

DEPARTAMENT DE FÍSICO-QUÍMICA

FACULTAT DE FARMÀCIA

UNIVERSITAT DE BARCELONA

DEPARTAMENT DE QUÍMICA FÍSICA

FACULTAT DE QUÍMICA

UNIVERSITAT DE BARCELONA

Estudio teórico de efectos del entorno en sistemas químicos

Carles E. Curutchet Barat

Noviembre 2004



UNIVERSITAT DE BARCELONA



Programa de Doctorado en Química Teórica y Computacional

Departamento de Química Física

Universidad de Barcelona

Bienio 2000-2002

Esta tesis ha sido realizada por el licenciado Carles E. Curutchet Barat bajo la dirección de los Drs. F. Javier Luque Garriga y Modesto Orozco López para optar al título de doctor por la Universidad de Barcelona

F. Javier Luque Garriga

Modesto Orozco López

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer muy especialmente a Javier y a Modesto la posibilidad que me han dado de integrarme en su grupo de investigación para realizar esta tesis doctoral. La posibilidad de aprender de sus conocimientos y el apoyo que me han prestado a lo largo de estos años me han permitido profundizar en el complejo y apasionante mundo de la modelización molecular, punto de encuentro de disciplinas tan diversas como la química, la biología, la física o la informática.

También quiero agradecer la ayuda que me prestaron los “veteranos”, especialmente Bego y Xavi, pero también Robert y Elena, cuando empecé a pelearme con los ordenadores hace ya unos años. Igualmente quiero agradecer a la gente del departamento en farmacia, el Jordi, el Ramón, el Axel, el Ignasi, el Oscar, el Damián, la Assumpta, la Yolanda... y los que me olvido o ya no están como la Nuria o la Susana, por el buen ambiente que reina siempre en el “palomar”. Extiendo también mis agradecimientos al resto de componentes del grupo del Parc Científic de Barcelona.

Quiero agradecer también al Dr. Christophe Chipot y del Dr. Manuel Ruiz-López por su hospitalidad al acogerme en el Laboratoire de Chimie et Biochimie Théoriques de Nancy, así como a los Dr. Jacopo Tomasi y Dra. Benedetta Menucci por acogerme en el Dipartimento di Chimica e Chimica Industriale de la Università di Pisa. De la misma forma, quiero manifestar mi satisfacción por haber tenido la suerte de trabajar con el Dr. Christopher Cramer durante su estancia en nuestro grupo.

Finalmente, quiero agradecer a mi madre y a mi hermano los ánimos que me dan en todo momento .

La presente tesis se ha realizado gracias a una beca predoctoral concedida por el Ministerio de Ciencia y Tecnología.

A la meva família i els meus amics

Índice

1. Introducción	13
1.1. Métodos teóricos en fases condensadas	14
1.2. Los efectos de polarización en simulaciones moleculares	16
1.3. Objetivos	18
1.3.1. Modelos continuos de solvatación	18
1.3.2. Fuerzas de polarización	19
Bibliografía	21
2. Metodología	25
2.1. Introducción	25
2.2. Química Cuántica	26
2.2.1. Introducción	26
2.2.2. Métodos <i>ab initio</i>	27
2.2.2.1. El operador Hamiltoniano	27
2.2.2.2. El principio variacional	28
2.2.2.3. Construcción de la función de onda	29
2.2.2.4. El método Hartree-Fock	32
2.2.2.5. Funciones de base	34
2.2.2.6. Métodos que incluyen la correlación electrónica	35
2.2.3. Métodos semiempíricos	37
2.2.4. Métodos basados en la teoría del funcional de la densidad	39
2.2.5. Análisis de descomposición de la energía de Kitaura-Morokuma	40
2.3. Métodos clásicos	42
2.3.1. Introducción	42
2.3.2. Campos de fuerzas	42
2.3.2.1. Términos enlazantes	43
2.3.2.2. Términos no enlazantes	45
2.3.2.3. Efectos de polarización	46
2.3.2.4. Parametrización	48

2.4. Modelos continuos de solvatación	50
2.4.1. Introducción	50
2.4.2. El modelo continuo	51
2.4.3. Contribución electrostática	53
2.4.3.1. Métodos de expansión multipolar (MPE)	53
2.4.3.2. Modelo generalizado de Born (GB)	55
2.4.3.3. Carga superficial aparente (ASC)	56
2.4.4. Contribuciones no electrostáticas	58
2.4.4.1. Cavitación	58
2.4.4.2. Dispersión-repulsión	59
2.4.5. La cavidad: parametrización de modelos continuos	61
2.4.5.1. El modelo MPE	62
2.4.5.2. El modelo SMx	62
2.4.5.3. El modelo MST	63
2.4.6. Hacia modelos de solvatación fraccionales	64
Bibliografía	69
3. Resultados sobre modelos continuos de solvatación	79
3.1. Solvation in octanol: Parametrization of the continuum MST model	81
3.2. MST continuum study of the hydration free energies of monovalent ionic species	97
3.3. Electrostatic component of solvation: Comparison of SCRF continuum models .	143
4. Resultados sobre solvatación fraccional	159
4.1. Transferability of fragmental contributions to the octanol/water partition coef- ficient: An NDDO-based MST study	161
4.2. MST study of group contributions for alkane derivatives: Effect of the charge normalization	177
5. Resultados sobre fuerzas de polarización	187
5.1. Energy decomposition in molecular complexes: Implications for the treatment of polarization in molecular simulations	189
5.2. Development of distributed models of atomic polarizability for molecular si- mulations	205
6. Discusión de los resultados	221
6.1. Modelos continuos de solvatación	221
6.2. Contribuciones fraccionales al coeficiente de partición	232
6.3. Fuerzas de polarización	238

<i>ÍNDICE</i>	11
Bibliografía	247
7. Conclusiones	249
A. Otros trabajos publicados	251
B. Abreviaciones	253

Capítulo 1

Introducción

La química teórica y computacional es un campo que ha experimentado un extraordinario avance en las últimas décadas [1, 2, 3]. Las bases teóricas sobre las que se asienta este campo, sin embargo, se remontan hasta principios del siglo pasado, con el desarrollo de la mecánica cuántica, pero los altos requerimientos computacionales demandados por los métodos mecanocuánticos tradicionalmente limitaban su aplicación a sistemas modelo de pequeño tamaño. En este contexto, no sólo el desarrollo de metodologías apropiadas, sino también el espectacular avance de la informática han sido decisivos en el paso que ha dado este campo para estudiar sistemas cada vez más complejos y más cercanos a la realidad. Tal como afirmaba la Academia Sueca al conceder en 1998 el premio Nobel de Química a los profesores W. Kohn y J.A. Popel, “la química ha dejado de ser una ciencia puramente experimental”. Actualmente disponemos de técnicas que nos permiten diseñar verdaderos experimentos numéricos, capaces de aportar información microscópica no accesible mediante técnicas experimentales. Esto es especialmente cierto para sistemas aislados, en fase gas, para los cuales disponemos de elaboradas metodologías basadas en cálculos cuánticos de alto nivel que alcanzan precisiones comparables a la experimental.

En el mundo de la química, sin embargo, lo más frecuente es trabajar en fases condensadas, y no en fase gas. En este tipo de sistemas, el medio juega un papel clave, alterando de forma decisiva las propiedades estructurales y reactivas de los compuestos químicos. Uno de los efectos más destacables del solvente es la polarización de la distribución electrónica del soluto, que produce un aumento en la polaridad del mismo reflejada, por ejemplo, en el momento dipolar. En solución acuosa, el aumento del momento dipolar respecto a su valor *in vacuo* es de aproximadamente el 20-30 % para solutos neutros [4, 5, 6], y se relaja a un 8-10 % en solventes apolares como el cloroformo [7]. El efecto del solvente también es importante en las propiedades espectroscópicas electrónicas de los solutos, debido a la solvatación diferencial entre el estado fundamental y los estados excitados. Este efecto puede dar lugar a efectos solvatocrómicos positivos (desplazamientos al rojo) o negativos (desplazamientos al

azul), dependiendo de la polaridad relativa de los estados fundamental y excitado y de la polaridad y polarizabilidad del solvente. También son notables los efectos sobre las transiciones vibracionales, que dependen de la magnitud de las interacciones específicas y no específicas soluto-solvente. Numerosos ejemplos de cambios en las propiedades espectroscópicas inducidos por el solvente se pueden encontrar en la literatura [8, 9]. También la influencia del solvente es importante en la reactividad química, ya que: i) introduce fricción en el movimiento de los átomos del sistema reactivo, ii) modula las características intrínsecas de los reactivos, y iii) modifica las superficies de energía potencial debido a la solvatación diferencial de reactivos, productos y estados de transición. Por ejemplo, el ataque nucleofílico de un grupo R-O^- sobre un grupo carbonilo es muy exotérmico y ocurre sin barrera de activación en fase gas, mientras en solución acuosa es claramente endotérmico con una barrera de activación notable [10, 11, 12, 13].

Además, los efectos del solvente también son importantes en la distribución nuclear. Por ejemplo, hay casos donde tenemos cambios conformacionales completos después de transferir una molécula desde fase gas a una fase condensada [9], debido a la tendencia de solventes polares a estabilizar estructuras con separación de carga. Igualmente importantes son los cambios inducidos por el solvente en equilibrios tautoméricos o ácido/base. Un ejemplo clásico es el equilibrio ceto/enol de β -dicetonas, en las cuales la forma enólica es la especie más poblada en solventes apolares mientras que la especie ceto es el tautómero más estable en solución acuosa [9]. Los equilibrios conformacionales, tautoméricos y ácido/base juegan un papel clave en la mayoría de procesos químicos y biológicos, ya que afectan la actividad de las biomoléculas de forma decisiva [14, 15, 16, 17, 18].

1.1. Métodos teóricos en fases condensadas

El estudio de sistemas en fases condensadas es mucho más complicado que el de sistemas aislados, principalmente debido al enorme número de moléculas a considerar, y al hecho de que sea preciso considerar el conjunto de entidades químicas (soluto, solvente) que constituyen el estado líquido desde un punto de vista dinámico. A pesar de estas dificultades, el desarrollo metodológico experimentado en las últimas décadas [19] ha permitido que la introducción de los efectos del solvente sea una práctica común en el campo.

Se han desarrollado una gran diversidad de estrategias para introducir el efecto del solvente. Ellas difieren principalmente en dos aspectos básicos: i) el nivel de detalle usado para describir el soluto, y ii) la manera en que se describe el solvente, principalmente mediante técnicas clásicas discretas o modelizándolo como un medio continuo caracterizado por la constante dieléctrica del solvente.

Modelos clásicos continuos Los modelos continuos, a pesar de su simplicidad formal y computacional, han demostrado ser una aproximación muy eficiente para el estudio de los efectos del solvente. Basados en la teoría de líquidos polarizables, estos modelos centran la atención en una parte del sistema, el soluto, mientras que el entorno se describe de forma aproximada como un medio continuo caracterizado por la constante dieléctrica del solvente. Al introducir el soluto, el medio se polariza generando un campo eléctrico que actúa sobre la distribución de carga del soluto (el campo de reacción). La componente electrostática a la solvatación se calcula, entonces, como la mitad de la energía de interacción electrostática entre la distribución de carga del soluto y el campo de reacción, de acuerdo con la teoría de respuesta lineal [20, 21]. Los primeros modelos clásicos continuos, que se desarrollaron en la primera mitad del siglo XX, usaban cavidades muy simples (esferas o elipsoides) y expansiones multipolares truncadas a diferentes niveles para describir la carga del soluto: monopolo en el modelo de Born [22], dipolo en los modelos de Bell [23] y Onsager [24], o a niveles de orden superior. Actualmente disponemos de modelos más elaborados como el modelo de Born generalizado para distribuciones de monopolos multicéntricas (GB, *Generalized Born*), o la resolución de la ecuación de Poisson-Boltzmann mediante métodos de diferencias finitas (FDPB, *Finite Difference Poisson-Boltzmann*), aunque también se han desarrollado muchos otros. La extensión de estos métodos para incorporar otras componentes a la energía libre de solvatación (dispersión-repulsión y cavitación) se puede realizar posteriormente mediante diversas estrategias, generalmente basadas en términos proporcionales a la superficie de área accesible al solvente (modelos PB/SA y GB/SA). Descripciones detalladas de estas metodologías pueden encontrarse en la referencia [19].

Modelos cuánticos continuos De entre los modelos de solvatación, los modelos cuánticos continuos han experimentado una de las evoluciones más destacables a lo largo de los últimos 30 años. La publicación del primer método continuo mecanocuántico se remonta a 1973 y fue realizada por el grupo de Nancy [25]. Unos años después, el grupo de Pisa presentó la formulación del método del continuo polarizable (PCM, *Polarizable Continuum Model*) [26]. Ambos métodos, que pertenecen a los denominados métodos del campo de reacción autocohérente (SCRf, *Self-Consistent Reaction Field*), han alcanzado una destacada popularidad y se encuentran implementados en diversos programas mecanocuánticos. En ellos, el soluto se describe a nivel cuántico, mientras que el solvente se describe de forma aproximada mediante un medio continuo caracterizado por la constante dieléctrica del medio. En estos modelos, el medio continuo se polariza a causa de la distribución de carga del soluto, pero también la densidad electrónica del soluto se polariza por la respuesta (campo de reacción) del solvente. Los modelos continuos SCRf, a pesar de su simplicidad, permiten obtener buenas descripciones de las propiedades de solvatación con un coste computacional moderado. En algunos casos, sin embargo, cuando existen interacciones soluto-solvente muy específicas, es necesario añadir al soluto alguna molécula de solvente de forma explícita para obtener así una descripción más

adecuada del sistema [27].

Modelos clásicos discretos Una alternativa lógica al uso de modelos continuos es la consideración explícita de las moléculas de solvente en el sistema. La gran cantidad de partículas que hay que considerar, sin embargo, nos limita en general a descripciones clásicas de éstas basadas en el uso de campos de fuerzas (*force field*), que permiten evaluar rápidamente la energía del sistema mediante expresiones analíticas clásicas y parámetros ajustados que sólo dependen de las coordenadas de los núcleos. A diferencia de los modelos continuos, en las técnicas discretas es necesario obtener representaciones promediadas del sistema soluto-solvente, que se obtienen por técnicas de dinámica molecular (MD) o de Monte Carlo (MC). La energía libre de solvatación nuevamente se obtiene a partir de la diferencia en el trabajo necesario para aniquilar la molécula de soluto en fase gas y en solución. Esto puede realizarse de acuerdo con la teoría de respuesta lineal, o mediante técnicas de mecánica estadística, bastante más costosas pero también más precisas, como los métodos de perturbación de la energía libre (FEP, *free energy perturbation*) o integración termodinámica (TI, *thermodynamic integration*).

Modelos mixtos QM/MM El tratamiento completo del sistema soluto-solvente a nivel cuántico es poco viable con los recursos computacionales actuales. No obstante, es posible el uso de modelos mixtos de mecánica cuántica/mecánica molecular (QM/MM, *Quantum Mechanics/Molecular Mechanics*), en los que la descripción cuántica se limita al soluto, mientras que el solvente se describe mediante un campo de fuerzas clásico. De esta forma, el Hamiltoniano total del sistema está formado por i) un Hamiltoniano QM para la descripción del soluto, ii) un Hamiltoniano clásico para la descripción de las moléculas de solvente, y iii) un término de acoplamiento que incluye las interacciones soluto-solvente. De forma análoga a las técnicas discretas clásicas, la exploración del espacio conformacional se realiza mediante técnicas de MD o MC, y el trabajo necesario para aniquilar el soluto en fase gas y en solución, necesarios para evaluar la energía libre de solvatación, se obtienen en base a la teoría de respuesta lineal o bien mediante métodos estadísticos de cálculo de la energía libre. Este tipo de métodos son conceptualmente muy atractivos y tienen un amplio espectro de potenciales aplicaciones. Sin embargo, su impacto real está limitado por su elevado coste computacional, que hace necesario el uso de campos de fuerzas no polarizables para el solvente y Hamiltonianos simples para el soluto.

1.2. Los efectos de polarización en simulaciones moleculares

Los métodos computacionales modernos permiten la simulación de sistemas que comprenden millares de átomos a partir de un amplio abanico de niveles de teoría [19, 2]. La elección de un nivel de teoría concreto depende del tamaño del sistema y de las propiedades de interés. De esta forma, los campos de fuerzas clásicos acoplados a algoritmos de Monte Carlo

o de dinámica molecular ofrecen una aproximación computacional eficiente para explorar los aspectos estructurales, dinámicos y energéticos de sistemas químicos y bioquímicos complejos. Sin embargo, en casos donde ocurren grandes cambios en la densidad electrónica, se tiene que recurrir a la mecánica cuántica, ya sea para simular el sistema entero o para simular parte de él, donde los efectos cuánticos son especialmente importantes. Los métodos híbridos de mecánica cuántica/mecánica molecular (QM/MM) [28, 29, 30, 31] son adecuados para este tipo de casos, donde la mayoría del sistema se describe a nivel clásico mientras que una pequeña parte requiere una representación mecanocuántica (por ejemplo, un soluto en un solvente polar, o el centro activo de un enzima).

La fiabilidad y el poder predictivo de las simulaciones basadas en descripciones clásicas dependen principalmente de la precisión del campo de fuerzas, que típicamente sólo incluye fuerzas electrostáticas y de dispersión-repulsión. Los efectos de inducción, necesarios para describir adecuadamente sistemas altamente polarizables como las interacciones catión- π [32, 33, 34], generalmente se incluyen a través de dipolos puntuales inducidos. No obstante, la aplicación de este método está limitado por factores como el uso de polarizabilidades isotrópicas, la transferibilidad de polarizabilidades atómicas (o de grupo) entre distintos entornos químicos, el tratamiento de los acoplamientos de corto alcance entre dipolos inducidos, y el coste computacional adicional requerido para determinar iterativamente los dipolos inducidos. Algunos de estos inconvenientes se pueden aliviar mediante procedimientos eficientes para derivar modelos distribuidos de polarizabilidades atómicas [35], y mediante el desarrollo de estrategias alternativas para el tratamiento de los efectos de inducción, como el modelo de fluctuación de carga [36] o del oscilador de Drude [37]. Sin embargo, a pesar de la existencia de diferentes métodos para tratar explícitamente la polarización, la gran mayoría de simulaciones de sistemas complejos aún se realiza mediante campos de fuerzas no polarizables basados en el uso de cargas puntuales, que incluyen la polarización implícitamente en su parametrización.

Para sistemas relativamente simples como líquidos puros, por ejemplo, los efectos de inducción se incluyen implícitamente ajustando los parámetros del campos de fuerzas empírico de forma que reproduzcan propiedades experimentales como la densidad o el calor de vaporización. Esta estrategia la ejemplifica el campo de fuerzas OPLS [38, 39, 40]. Para sistemas más complejos, donde el desarrollo de campos de fuerzas efectivos sólo a partir de información experimental es menos viable, la polaridad de las moléculas generalmente se exagera para compensar los efectos de inducción ausentes causados por las cargas del entorno. Estos efectos entonces se incluyen de forma promediada escalando la interacción electrostática entre la molécula y su entorno. Por ejemplo, una estrategia muy extendida consiste en ajustar las cargas atómicas al potencial electrostático calculado a nivel Hartree-Fock con un conjunto de base 6-31G(d), ya que este nivel de teoría tiende a sobreestimar (alrededor del 10-20 %) el momento dipolar de moléculas neutras en fase gas.

Dada la naturaleza crucial de la contribución electrostática a la energía total en la mayoría de sistemas químicos/bioquímicos, uno de los retos actuales de la química teórica y computacional que está siendo objeto de un gran esfuerzo investigador consiste en el desarrollo de campos de fuerzas polarizables que sean capaces de introducir explícitamente y de forma eficiente los efectos de inducción en simulaciones moleculares clásicas.

1.3. Objetivos

1.3.1. Modelos continuos de solvatación

La primera parte de esta tesis se sitúa en el contexto de los modelos continuos de solvatación SCRF, y más concretamente en el modelo continuo MST desarrollado en nuestro grupo. Los estudios que se presentan intentarán profundizar en varios aspectos.

Un aspecto central en el que se intentará profundizar es la definición de la cavidad en este tipo de modelos, ya que tanto la magnitud de la componente electrostática a la solvatación como la polarización inducida en la densidad electrónica del soluto por el solvente dependen en gran medida de la definición de la frontera soluto-solvente. En este sentido, el modelo MST usa cavidades adaptadas al solvente, con el fin de obtener un balance apropiado entre las distintas contribuciones a la energía libre de solvatación.

El otro aspecto central de estos estudios tiene como objetivo el desarrollo de metodologías apropiadas para el estudio de las contribuciones a la solvatación asociadas a átomos/grupos a partir de métodos continuos SCRF. La partición de la energía libre de solvatación en contribuciones fragmentales se ha usado para profundizar en las bases fisicoquímicas de la interacción soluto-solvente [41, 42, 43], y para explorar relaciones estructura-actividad en biomoléculas y en diseño de fármacos [44, 45, 46]. En particular, el proceso de desolvatación energéticamente desfavorable del ligando y el receptor debe ser compensado por la formación de contactos ligando-receptor favorables, lo cual implica un cierto grado de complementariedad (electrostática, estérica y hidrofóbica) entre el ligando y el receptor. De esta forma, no sólo la energía libre de solvatación total, también la distribución tridimensional de regiones polares y apolares a lo largo del esqueleto molecular son importantes para entender la actividad biológica de las moléculas. Este aspecto justifica la extensión del modelo MST al solvente *n*-octanol, que es claramente el solvente orgánico más popular para el estudio de las propiedades hidrofóbicas en sistemas biológicos y farmacológicos.

Concretamente, los objetivos de esta primera parte de la tesis son:

- La extensión del modelo continuo MST a la solvatación en n-octanol para la predicción de energías libres de solvatación y coeficientes de partición octanol/agua, prestando una atención especial a la definición de la cavidad para este solvente. También se llevará a cabo la revisión de las parametrizaciones existentes en agua, cloroformo y tetracloruro de carbono. Tales parametrizaciones se han llevado a cabo a nivel *ab initio* HF/6-31G(d) y semiempíricos AM1 y PM3.
- La revisión de la definición de la cavidad para el estudio de especies iónicas monovalentes, motivada por la aparición de nuevos y mejores valores experimentales de energías de solvatación de compuestos iónicos, y la aplicación del modelo revisado a la predicción de equilibrios ácido/base.
- La comparación de la componente electrostática a la solvatación y la polarización del soluto obtenida por diferentes métodos SCRF para distintas cavidades y diferentes dieléctricos.
- El estudio detallado de las contribuciones a la solvatación asociadas a átomos o grupos obtenidas por métodos SCRF, que permiten obtener una descripción tridimensional del concepto de solvatación. Se incidirá en la posible transferibilidad de estas contribuciones fraccionales.

1.3.2. Fuerzas de polarización

La segunda parte de esta tesis está orientada a profundizar en la naturaleza de las fuerzas de polarización y en el desarrollo de metodologías efectivas para introducir estos efectos en los campos de fuerzas clásicos, y tiene diversos objetivos:

- El estudio de la contribución de polarización a la energía de interacción para una serie de sistemas modelo, que caracterizan interacciones habituales presentes en los sistemas bioquímicos, tales como puentes de hidrógeno y catión- π .
- El desarrollo de una metodología destinada a introducir los efectos de polarización en campos de fuerzas clásicos basada en un doble conjunto de cargas, que describen la respuesta electrostática de las moléculas aisladas o bien frente a campos externos, respectivamente.
- El desarrollo de protocolos sistemáticos y robustos para la derivación de modelos distribuidos de polarizabilidades atómicas, que permiten tratar explícitamente los efectos de inducción a través de dipolos puntuales inducidos mediante la asignación de polarizabilidades a los distintos átomos del sistema. En este estudio se prestará especial atención, además, a la transferibilidad entre diferentes compuestos de las polarizabilidades atómicas obtenidas.

Bibliografía

- [1] A. R. Leach. *Molecular Modelling. Principles and Applications*. Longman, Singapore, 1996.
- [2] C. J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. Wiley, Chichester, 2002.
- [3] J. Andrés and J. Bertrán, editors. *Química Teórica y Computacional*. Publicaciones de la Universitat Jaume I, Castellón de la Plana, 2000.
- [4] C. J. Cramer and D. G. Truhlar. *Chem. Phys. Lett.*, 198:74, 1992.
- [5] F. J. Luque, M. J. Negre, and M. Orozco. *J. Phys. Chem.*, 97:4386, 1993.
- [6] J. Gao, F. J. Luque, and M. Orozco. *J. Phys. Chem.*, 98:2975, 1993.
- [7] F. J. Luque, Y. Zhang, C. Alemán, M. Bachs, J. Gao, and M. Orozco. *J. Phys. Chem.*, 100:4269, 1996.
- [8] C. Reichardt. *Chem. Rev.*, 94:2319, 1994.
- [9] C. Reichardt. *Solvent and solvent effects in organic chemistry*. VCH, Weinheim, 1998.
- [10] J. D. Madura and W. L. Jorgensen. *J. Am. Chem. Soc.*, 108:2517, 1986.
- [11] A. Howard and P. Kollman. *J. Am. Chem. Soc.*, 110:7195, 1988.
- [12] D. Arad, R. Langridge, and P. Kollman. *J. Am. Chem. Soc.*, 112:491, 1990.
- [13] M. Strajbl, J. Florian, and A. Warshel. *Int. J. Quant. Chem.*, 77:44, 2000.
- [14] S. Madurga and E. Vilaseca. *J. Phys. Chem. A*, 106:11822, 2002.
- [15] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán. *J. Mol. Struct. (TEOCHEM)*, 632:197, 2003.
- [16] B. Hernández, M. Orozco, and F. J. Luque. *J. Comput-Aided Mol. Design*, 11:153, 1997.
- [17] M. Orozco, B. Hernández, and F. J. Luque. *J. Phys. Chem. B*, 102:5228, 1998.

- [18] J. R. Blas, F. J. Luque, and M. Orozco. *J. Am. Chem. Soc.*, 126:154, 2004.
- [19] M. Orozco and F. J. Luque. *Chem. Rev.*, 100:4187, 2000.
- [20] C. J. Böttcher. *Theory of Electric Polarization*, volume 1. Elsevier, Amsterdam, 1973.
- [21] C. J. Böttcher and P. Bordewijk. *Theory of Electric Polarization*, volume 2. Elsevier, Amsterdam, 1978.
- [22] M. Z. Born. *Phys.*, 1:45, 1920.
- [23] R. P. Bell. *Trans. Faraday Soc.*, 27:797, 1931.
- [24] L. Onsager. *J. Am. Chem. Soc.*, 58:1486, 1936.
- [25] D. Rinaldi and J. L. Rivail. *Theor. Chim. Acta*, 32:57, 1973.
- [26] S. Miertus, E. Scrocco, and J. Tomasi. *Chem. Phys.*, 55:117, 1981.
- [27] J. Tomasi, R. Cammi, B. Menucci, C. Capelli, and S. Corni. *Phys. Chem. Chem. Phys.*, 4:5697, 2002.
- [28] J. Gao and M. A. Thompson, editors. *Combined Quantum Mechanical and Molecular Mechanical Methods*. ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998.
- [29] G. Monard and Jr. K. M. Merz. *Acc. Chem. Res.*, 32:904, 1999.
- [30] M. F. Ruiz-López and J. L. Rivail. *Encyclopedia of Computational Chemistry*, volume 1, page 437. Wiley, New York, 1998.
- [31] J. Aqvist and A. Warshel. *Chem. Rev.*, 93:2523, 1993.
- [32] J. Caldwell and P. A. Kollman. *J. Am. Chem. Soc.*, 117:4177, 1995.
- [33] H. Minoux and C. Chipot. *J. Am. Chem. Soc.*, 121:10366, 1999.
- [34] E. Cubero, F. J. Luque, and M. Orozco. *Proc. Natl. Acad. Sci. USA*, 95:5976, 1998.
- [35] C. Chipot, F. Dehez, J. Angyan, C. Millot, M. Orozco, and F. J. Luque. *J. Phys. Chem. A*, 105:11505, 2001.
- [36] S. W. Rick, S. J. Stuart, and B. J. Berne. *J. Chem. Phys.*, 101:6141, 1994.
- [37] G. Lamoureux and B. Roux. *J. Chem. Phys.*, 119:3025, 2003.
- [38] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. *J. Chem. Phys.*, 79:926, 1983.

- [39] N. A. McDonald and W. L. Jorgensen. *J. Phys. Chem. B*, 102:8049, 1998.
- [40] M. L. P. Price, D. Ostrovsky, and W. L. Jorgensen. *J. Comput. Chem.*, 22:1340, 2001.
- [41] J. Hine and P. K. Mookerjee. *J. Org. Chem.*, 40:292, 1975.
- [42] J. P. Guthrie. *Can. J. Chem.*, 55:3700, 1977.
- [43] S. Cabani, P. Gianni, V. Mollica, and L. Lepori. *J. Solut. Chem.*, 10:563, 1981.
- [44] Y. C. Martin. *Quantitative Drug Design. A Critical Introduction*. Marcel Dekker, New York, 1978.
- [45] A. L. Horvath. *Molecular Design*. Elsevier, Amsterdam, 1992.
- [46] C. Hansch and A. Leo. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*. American Chemical Society, Washington DC, 1995.

Capítulo 2

Metodología

2.1. Introducción

El campo de la química teórica y computacional comprende una amplia gama de metodologías que se aplican a sistemas químicos muy diversos [1, 2, 3]. Si bien las leyes físicas que rigen el comportamiento de los sistemas químicos son conocidas gracias al desarrollo de la mecánica cuántica durante el siglo pasado, la aplicación rigurosa de estas leyes se ve limitada por la enorme cantidad de cálculos que conlleva resolver las ecuaciones que plantean. Esta limitación, aunque aliviada por el espectacular aumento en la capacidad de cálculo de los ordenadores, ha hecho necesaria la introducción de aproximaciones, más o menos rigurosas, que permiten el estudio de sistemas de mayor tamaño al reducir el volumen de cálculo requerido. En general, la calidad de los resultados que se pueden obtener de un modelo dependerá del nivel de simplificaciones que se hayan introducido en él, pero también de la validez que tengan éstas en el sistema que se quiera estudiar. Por esta razón, es esencial conocer los fundamentos físicos y matemáticos en que se basan, y obtener así una visión crítica sobre los resultados que nos permita extraer información válida.

En este capítulo se describirán brevemente las técnicas de cálculo usadas en la presente tesis doctoral. El esquema del capítulo es el siguiente:

- Introducción a los métodos basados en la química cuántica.
- Descripción de los campos de fuerza clásicos para el cálculo de la energía de sistemas.
- Descripción de los modelos continuos para introducir el efecto del solvente en cálculos de química cuántica.

2.2. Química Cuántica

2.2.1. Introducción

El estudio riguroso de sistemas microscópicos, tales como los sistemas químicos, requiere del uso de métodos basados en la mecánica cuántica. Aunque se han desarrollado diversas teorías cuánticas para el estudio de sistemas moleculares, en esta sección se describirán principalmente los métodos basados en la teoría de orbitales moleculares, ya que son los más ampliamente difundidos y los que se han utilizado en el presente trabajo. Descripciones extensas de las dos teorías se pueden encontrar en diversas fuentes bibliográficas [4, 5].

El postulado fundamental de la mecánica cuántica establece que, para cualquier sistema, existe una función de onda tal que la aplicación sobre ella de unas funciones apropiadas llamadas operadores nos devuelve las propiedades observables del sistema. Para obtener esta función de onda debemos resolver la ecuación de Schrödinger, que en su forma independiente del tiempo, en la que nos centraremos, toma la forma siguiente :

$$H\Psi = E\Psi \quad (2.1)$$

donde H es el operador Hamiltoniano, que operado sobre la función de onda nos devuelve como valor propio la energía E del sistema.

Es importante señalar la naturaleza probabilística de Ψ , dado que el producto de la función de onda por su complejo conjugado ($|\Psi^*\Psi|$) refleja la densidad de probabilidad. Así, para el caso más común de funciones de onda reales, la probabilidad de que un sistema químico se encuentre en una región del espacio multidimensional será igual a la integral de $|\Psi|^2$ sobre esa región del espacio. Esto implica una restricción adicional a la función de onda. En concreto, para una partícula, la integral normalizada de $|\Psi|^2$ sobre todo el espacio debe ser igual a la unidad.

La ecuación de Schrödinger sólo se puede resolver de forma exacta para sistemas muy simples, como el átomo de hidrógeno, de forma que para aplicarla a sistemas de interés químico hace falta realizar diversas aproximaciones. En esta sección se describirán las características principales de algunas de estas aproximaciones y de los métodos a que han dado lugar. En primer lugar, se describirá la llamada metodología *ab initio* [6], que nos permite estudiar un sistema a partir de los postulados de la química cuántica sin disponer de información empírica adicional. No obstante, los altos requerimientos computacionales que implica el uso de esta metodología hacen que en la práctica no se pueda aplicar a todos los sistemas que son objeto de estudio. Este hecho ha impulsado el desarrollo de los llamados métodos semiempíricos [7, 8],

que se describirán posteriormente. La filosofía de estos métodos se basa en la reducción del coste computacional de los cálculos *ab initio* mediante la introducción de diversas aproximaciones, basadas en el uso de parámetros ajustados para reproducir diversas propiedades experimentales. A continuación se comentarán brevemente los métodos basados en la teoría del funcional de la densidad. Finalmente, se describirá el análisis de descomposición de la energía de Kitaura-Morokuma, que ha sido usado en uno de los estudios que se presentan.

2.2.2. Métodos *ab initio*

2.2.2.1. El operador Hamiltoniano

La resolución de la ecuación de Schrödinger independiente del tiempo (ec. 2.1), que nos permite obtener la función de onda, y a partir de ésta, propiedades del sistema como su energía, depende de la forma del operador Hamiltoniano. En ausencia de otros campos externos, de efectos relativistas, etc., este operador incluye típicamente cinco contribuciones a la energía total del sistema: la energía cinética de núcleos y electrones, la atracción entre núcleos y electrones, y la repulsión interelectrónica y internuclear. Para un sistema formado por N electrones y M núcleos, descritos por los vectores posición R_A y r_i , respectivamente, el Hamiltoniano en unidades atómicas toma entonces la siguiente forma:

$$H = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (2.2)$$

donde M_A es la masa del núcleo A relativa a la masa del electrón y Z_A es su número atómico. Los operadores Laplacianos ∇_i^2 y ∇_A^2 involucran diferenciales respecto a las coordenadas del electrón i -ésimo y el núcleo A -ésimo.

Aproximación Born-Oppenheimer

La obtención de buenas funciones de onda para sistemas moleculares es extremadamente difícil debido a los movimientos correlacionados de núcleos y electrones. A fin de simplificar el problema, se puede introducir la llamada aproximación Born-Oppenheimer, que es básica en los métodos de química cuántica. Esta aproximación asume que los núcleos son mucho más pesados que los electrones, y que, por lo tanto, se mueven mucho más lentamente. Así, en la práctica, podemos considerar la relajación electrónica respecto a los movimientos nucleares como instantánea. Esto permite desacoplar los dos movimientos y calcular energías electrónicas para posiciones nucleares fijas, de forma que el término de energía cinética de los núcleos se puede omitir, la correlación en el potencial atractivo electrón-núcleo también se elimina, y

la repulsión entre núcleos se puede considerar una constante para una geometría dada. Los términos que nos quedan de la ecuación 2.2 reciben el nombre de Hamiltoniano electrónico:

$$H_{el} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2.3)$$

De esta forma, la ecuación de Schrödinger electrónica se puede escribir como:

$$(H_{el} + V_N)\Psi_{el}(r_i; R_A) = E_{el}\Psi_{el}(r_i; R_A) \quad (2.4)$$

donde V_N es la energía de repulsión internuclear, y las coordenadas electrónicas r_i son variables independientes pero las coordenadas nucleares R_A son parámetros. El valor propio E_{el} de la ecuación de Schrödinger electrónica se llama la energía electrónica. Como el término V_N es una constante para una geometría dada, y las constantes añadidas al Hamiltoniano no afectan a las funciones de onda obtenidas, en la práctica se suele resolver la ecuación 2.4 sin incluir V_N , que se añade posteriormente al valor propio obtenido para la energía.

La aproximación Born-Oppenheimer está justificada en la gran mayoría de casos, y nos permite introducir conceptos importantes como el de superficie de energía potencial, que es la superficie definida por E_{el} sobre todas las posibles coordenadas nucleares. También permite definir conceptos muy útiles como las geometrías de equilibrio o de estados de transición, que son puntos críticos de la superficie de energía potencial del sistema.

2.2.2.2. El principio variacional

Dado un operador Hamiltoniano hermítico H , existe un conjunto de soluciones exactas a la ecuación de Schrödinger, infinitas en número, etiquetadas por el índice i :

$$H\Psi_i = E_i\Psi_i \quad i = 0, 1, \dots \quad (2.5)$$

donde

$$E_0 \leq E_1 \leq E_2 \leq \dots \leq E_i \leq \dots$$

Es decir, para una molécula existen muchas posibles funciones propias Ψ_i del operador H . Como H es un operador hermítico, los valores propios E_i son reales y las correspondientes funciones propias Ψ_i son ortonormales, es decir, son ortogonales entre sí y están normalizadas, de forma que cumplen:

$$\int \Psi_i \Psi_j d\tau = \delta_{ij} \quad (2.6)$$

donde la integral se realiza sobre todas las coordenadas del sistema, indicado por $d\tau$, y δ_{ij} es la delta de Krönecker (igual a uno si $i = j$, y cero en caso contrario). Además, asumiremos

que el conjunto de funciones propias Ψ_i de H , que podemos considerar discreto, forman un conjunto completo, de forma que cualquier función Φ que satisfaga las mismas condiciones de contorno que el conjunto $\{\Psi_i\}$ (normalmente que la función de onda desaparezca en el infinito), se puede escribir como una combinación lineal de las Ψ_i :

$$\Phi = \sum_i c_i \Psi_i \quad (2.7)$$

La ecuación de Schrödinger no nos indica como encontrar este conjunto completo de funciones de onda moleculares, sólo sabemos que este conjunto está acotado inferiormente por E_0 , la energía del estado fundamental. Sin embargo, se puede demostrar, a partir de las ecuaciones 2.1, 2.6 y 2.7, que el valor esperado del Hamiltoniano para la función Φ será siempre una energía igual o superior a la del estado fundamental:

$$E = \frac{\int \Phi H \Phi d\tau}{\int \Phi^2 d\tau} \geq E_0 \quad (2.8)$$

donde el valor esperado de la energía a la izquierda se obtiene a partir de la ecuación de Schrödinger multiplicando por Φ , integrando, y aislando la energía. Este teorema es el llamado principio variacional, y tiene implicaciones muy poderosas ya que permite evaluar la calidad de una función de onda de prueba para describir el estado fundamental de un sistema en base a su energía. Además, permite construir nuestra función de onda no a partir de una combinación lineal de funciones ortonormales Φ_i desconocidas, sino de funciones definidas dependientes de ciertos parámetros, y variar esos parámetros hasta que la integral en la ecuación 2.8 alcance un mínimo. Ese valor mínimo será entonces la estimación variacional a la energía exacta del estado fundamental.

2.2.2.3. Construcción de la función de onda

El principio variacional permite juzgar la calidad de una función de onda por la energía que tenga asociada, de forma que la función que mejor definirá el estado fundamental de un sistema será la que minimice su energía. Estas funciones de onda polielectrónicas se construyen a partir de orbitales moleculares, que son funciones que describen el comportamiento de un electrón. En la aproximación CLOA (Combinación Lineal de Orbitales Atómicos), los orbitales moleculares se expresan como una combinación lineal a partir de una base finita de orbitales atómicos:

$$\psi_i = \sum_{\mu=1}^N c_{\mu i} \varphi_{\mu} \quad (2.9)$$

donde el conjunto de N funciones φ_{μ} se llama el conjunto de base (*basis set*), y los coeficientes $c_{\mu i}$ nos dan la contribución de cada orbital atómico μ al orbital molecular i . Aunque podemos

escoger otro tipo de funciones matemáticas para el conjunto de base, la elección de orbitales atómicos, además de ser muy intuitiva desde el punto de vista químico, es una manera muy efectiva de describir los orbitales moleculares.

Spin electrónico y antisimetría

El Hamiltoniano electrónico (ec. 2.3) depende únicamente de las coordenadas espaciales de los electrones. Sin embargo, para describir completamente un electrón es necesario especificar también su spin. En un contexto no relativista, esto puede hacerse mediante la introducción de dos funciones de spin, $\alpha(w)$ y $\beta(w)$, ortonormales entre sí, y que forman un conjunto completo. Las funciones de spin son propias del operador S_Z , y sólo tienen dos valores propios asociados, spin alto (\uparrow) y spin bajo (\downarrow). Es necesario, entonces, introducir un requerimiento adicional a la función de onda, en concreto, ésta debe cambiar de signo si se intercambian las coordenadas tanto espaciales como de spin de dos electrones cualquiera. Este requerimiento es el llamado principio de antisimetría, y es un postulado independiente de la mecánica cuántica con consecuencias más generales que el conocido principio de exclusión de Pauli, que nos dice que no puede haber dos electrones caracterizados por el mismo conjunto de números cuánticos. Más adelante se mostrará como se pueden formar funciones de onda antisimétricas a partir de orbitales moleculares mediante el uso de determinantes de Slater.

Teniendo esto en cuenta, cada orbital molecular se escribe como el producto entre una función espacial, $\psi_i(r)$, y una función de spin, $\alpha(w)$ o $\beta(w)$, y recibe el nombre de spin orbital, $\chi_i(r, w)$. De esta forma, a partir de un orbital espacial se pueden construir dos spin orbitales diferentes, $\psi_i(r)\alpha(w)$ o $\psi_i(r)\beta(w)$.

Productos Hartree

Habiendo visto que las funciones apropiadas para describir un electrón son los spin orbitales, ahora consideraremos funciones de onda para un conjunto de N electrones. Antes de considerar la forma exacta de la función de onda, consideraremos un sistema más simple en el cual los electrones no interactúan entre sí, de forma que el Hamiltoniano toma la forma:

$$H = \sum_{i=1}^N h(i) \quad (2.10)$$

donde $h(i)$ es el operador que describe la energía cinética y la atracción nuclear del electrón i . Esta es la forma que toma el Hamiltoniano electrónico completo de la ecuación 2.3 si omitimos

la repulsión electrón-electrón. De esta forma, el operador $h(i)$ tendrá un conjunto de funciones propias que tomarán la forma de spin orbitales $\{\chi_j\}$,

$$h(i)\chi_j(x_i) = \varepsilon_j\chi_j(x_i) \quad (2.11)$$

donde x_i corresponde a las coordenadas espaciales y de spin del electrón i .

Dado que H es la suma de los Hamiltonianos monoeléctricos $h(i)$, la función de onda N -electrónica propia de H será simplemente el producto de spin orbitales para cada electrón:

$$\Psi_{PH}(x_1, x_2, \dots, x_N) = \chi_i(x_1)\chi_j(x_2) \cdots \chi_k(x_N) \quad (2.12)$$

La función de onda Ψ^{HP} es propia de H y tiene como valor propio la energía E , que corresponde a la suma de energías de cada spin orbital:

$$H\Psi_{PH} = E\Psi_{PH} = (\varepsilon_i + \varepsilon_j + \dots + \varepsilon_k)\Psi_{PH} \quad (2.13)$$

Este tipo de funciones de onda monoeléctricas se llaman productos Hartree, con el electrón uno siendo descrito por el spin orbital χ_i , el dos por el spin orbital χ_j , etc.

El producto Hartree presenta diversas deficiencias como función de onda. En primer lugar, no considera la correlación en el movimiento de los electrones, de forma que la probabilidad de encontrar un electrón en una región del espacio, a partir de su spin orbital, es independiente de la posición del resto de electrones. Además, hay aún otra deficiencia básica en este tipo de funciones, ya que no tienen en cuenta la indistinguibilidad de los electrones al situar cada uno en un orbital concreto. El principio de antisimetría no distingue entre electrones idénticos y requiere que la función de onda electrónica sea antisimétrica respecto al intercambio de coordenadas espaciales y de spin de dos electrones.

Determinantes de Slater

El producto Hartree no cumple el principio de antisimetría. Sin embargo, podemos realizar combinaciones lineales de diferentes productos Hartree, de manera que la función de onda no distinga entre electrones y sea antisimétrica. De manera general para un sistema N -electrónico, esta función de onda se puede expresar como un determinante de Slater [9] con la siguiente forma:

$$\Psi_{DS} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_N) & \chi_j(x_N) & \cdots & \chi_k(x_N) \end{vmatrix} \quad (2.14)$$

donde el factor $\frac{1}{\sqrt{N!}}$ es un factor de normalización. Este determinante tiene N electrones ocupando N spin orbitales $(\chi_i, \chi_j, \dots, \chi_k)$, sin especificar qué electrón está en cada orbital. Además, intercambiar las coordenadas de dos electrones corresponde a intercambiar dos filas del determinante de Slater, lo cual cambia de signo el determinante. De esta forma, la función de onda cumple el principio de antisimetría. Además, dos electrones no pueden ocupar el mismo orbital, ya que tendríamos dos filas iguales y el determinante valdría cero (principio de exclusión de Pauli). Si escogemos que las etiquetas de los electrones estén siempre en el orden x_1, x_2, \dots, x_N , estos determinantes se suelen escribir de forma compacta como:

$$\Psi_{DS}(x_1, x_2, \dots, x_N) = |\chi_i \chi_j \cdots \chi_k\rangle \quad (2.15)$$

donde el factor de normalización $\frac{1}{\sqrt{N!}}$ está incluido implícitamente.

2.2.2.4. El método Hartree-Fock

Encontrar soluciones aproximadas a la ecuación de Schrödinger para sistemas de muchos electrones es complicado debido a los movimientos correlacionados de los electrones. En este contexto, el método Hartree-Fock ha sido fundamental para el desarrollo de la química moderna, y es el punto de partida para tratamientos *ab initio* más exactos a este tipo de problemas. Este método se basa en el uso de una función de onda constituida por un solo determinante de Slater para describir el estado fundamental del sistema. Consideremos entonces una función de onda polielectrónica construida a partir de un determinante de Slater de n spin orbitales, χ_k ($k = 1, 2, \dots, n$), que se puede escribir como $\Psi = |\chi_1 \cdots \chi_k \chi_l \cdots \chi_n\rangle$. El Hamiltoniano electrónico completo H_{el} de la ecuación 2.3 nos permite calcular el valor esperado de la energía E del sistema, que toma la forma siguiente:

$$E = \sum_k \int \chi_k(1) h(1) \chi_k(1) dx_1 + \quad (2.16)$$

$$+ \sum_k \sum_{l>k} \left[\int \int \chi_k(1) \chi_l(2) \frac{1}{r_{12}} \chi_k(1) \chi_l(2) dx_1 dx_2 - \int \int \chi_k(1) \chi_l(2) \frac{1}{r_{12}} \chi_k(2) \chi_l(1) dx_1 dx_2 \right]$$

donde $h(1)$ es el Hamiltoniano de core relativo a un electrón, que comprende su energía cinética y su energía de atracción con los núcleos. La evaluación del segundo término de esta suma, donde aparecen integrales en que intervienen 4 spin orbitales, requiere de un gran esfuerzo computacional. La flexibilidad variacional de la función de onda está en la elección de los spin orbitales, de forma que minimizando la energía respecto a su elección se puede llegar a las ecuaciones de valores propios de Hartree-Fock, que determinan los mejores spin orbitales:

$$f(i) \chi_j(x_i) = \varepsilon_j \chi_j(x_i) \quad (2.17)$$

donde $f(i)$ es un operador efectivo monoeléctrico, llamado operador de Fock, que tiene la forma:

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{iA}} + v^{HF}(i) \quad (2.18)$$

El término $v^{HF}(i)$ es el potencial medio que experimenta el electrón i debido a la presencia de los otros electrones. Esta es la esencia del método Hartree-Fock, que reemplaza el complicado problema de tratar muchas partículas por un problema monoeléctrico en el que la repulsión electrónica se trata de forma promediada.

El potencial v^{HF} sobre un electrón, que escogemos como uno, tiene la forma siguiente:

$$v^{HF}(1) = \sum_{j=1}^N J_j(1) - K_j(1) \quad (2.19)$$

donde el operador J_j , llamado operador de Coulomb, introduce el potencial local medio generado en x_1 por un electrón descrito por χ_j , según:

$$J_j(1) = \int dx_2 |\chi_j(2)|^2 \frac{1}{r_{12}} \quad (2.20)$$

El término K_j se denomina operador de intercambio y aparece como consecuencia de la naturaleza antisimétrica del determinante de Slater. Este término, que es menos intuitivo y no tiene una interpretación clásica como el término de Coloumb, se define por su efecto cuando opera sobre un spin orbital $\chi_k(1)$ de la forma siguiente:

$$K_j(1)\chi_k(1) = \left[\int dx_2 \chi_j(2) \frac{1}{r_{12}} \chi_k(2) \right] \chi_j(1) \quad (2.21)$$

La ecuación de Hartree-Fock (ec. 2.17) es una ecuación de valores propios con los spin orbitales como funciones propias y las energías de los spin orbitales como valores propios. Los N spin orbitales de menor energía son entonces los llamados orbitales ocupados, mientras que el resto se denominan orbitales virtuales.

La resolución de las ecuaciones de Hartree-Fock generalmente se lleva a cabo introduciendo un conjunto de funciones de base para expandir los spin orbitales (aproximación CLOA, ver ecuación 2.9). Entonces el problema se reduce a resolver unas ecuaciones matriciales, cuya forma fue introducida por primera vez en 1951 por Roothaan [10]. Hay que señalar, además, que el potencial Hartree-Fock $v^{HF}(i)$ depende de los spin orbitales de los otros electrones, de forma que la ecuación 2.17 no es lineal y debe resolverse iterativamente. Este procedimiento se denomina el método del campo autoconistente (SCF, *Self-Consistent Field*).

Finalmente, hay dos posibilidades en cuanto al formalismo empleado. El formalismo RHF (*Restricted Hartree-Fock*) utiliza orbitales sin discriminación de spin y se aplica a sistemas de

capa cerrada, es decir, con todos los electrones apareados. Esta es la situación más general en sistemas bioquímicos, y es el método en el que se basan las técnicas de cálculo del presente trabajo. Sin embargo, existe también el formalismo UHF (*Unrestricted Hartree-Fock*), que se puede aplicar a sistemas con electrones desapareados, cuya parte espacial puede diferir según el spin sea α o β .

2.2.2.5. Funciones de base

El conjunto de base es el conjunto de funciones matemáticas a partir de las cuales construimos la función de onda. Como se ha explicado anteriormente, cada orbital molecular en la teoría Hartree-Fock se expresa como una combinación lineal de funciones de base (ec. 2.9), cuyos coeficientes se determinan a partir de la solución iterativa de las ecuaciones HF SCF. Entonces la función de onda electrónica completa se expresa como un determinante de Slater construido a partir de los orbitales moleculares ocupados. En teoría, se puede llegar al límite HF si usamos un conjunto de base infinito, que necesariamente permite una descripción óptima de la densidad de probabilidad electrónica. En la práctica, de todas formas, no podemos usar un conjunto de base infinito, hecho por el cual se ha dedicado un gran esfuerzo en identificar funciones que permitan acercarnos al límite HF de la manera más eficiente posible.

Generalmente, las funciones de base que se usan tiene formas análogas a las soluciones que se encuentran para los orbitales atómicos del átomo de hidrógeno (funciones hidrogenoides). Estas funciones se denominan orbitales de Slater [11] (STO, *Slater-type Orbital*), y tienen la forma $r^{n-1}e^{-\alpha r} \cdot Y_l^m(\theta, \varphi)$, donde $Y_l^m(\theta, \varphi)$ designa un armónico esférico. En la teoría *ab initio* HF, de todas formas, los STOs sufren una limitación importante. En concreto, no disponemos de soluciones analíticas para la evaluación de las integrales generales de cuatro índices de la ecuación 2.16 cuando las funciones de base son STOs. En 1950 Boys propuso resolver este problema cambiando la función radial de $e^{-\alpha r}$ a $e^{-\alpha r^2}$, de forma que los orbitales atómicos tuvieran forma de funciones gaussianas [12] (GTO, *Gaussian-type Orbital*).

Aunque son convenientes desde un punto de vista computacional, los GTOs tienen varios problemas que dificultan su uso como funciones de base. La primera es que para funciones tipo s, cuando r tiende a cero las funciones gaussianas no reproducen el pico que generan los STOs. Además, para valores grandes de r , la función gaussiana decae más rápidamente que la función de Slater. En la práctica, para resolver estos inconvenientes, las funciones de base usadas en los cálculos *ab initio* se construyen a partir de combinaciones lineales de GTOs, escogidas de forma que describan correctamente la forma de los STO:

$$\varphi_\mu = \sum_{a=1}^M c_a \phi_{a\mu} \quad (2.22)$$

donde M es el número de gaussianas usadas como combinación lineal.

Una función de base definida como combinación lineal de gaussianas se llama una función de base contraída, y las funciones gaussianas individuales que la forman se llaman primitivas. Pople y colaboradores [13] fueron los primeros en determinar de forma sistemática los coeficientes y exponentes de contracción óptimos para reproducir STOs a partir de diferentes cantidades de GTOs para un gran número de átomos de la tabla periódica.

Las bases STO-NG son ejemplos de bases mínimas, en la que sólo se usa una función de base para cada tipo de orbital, sea interno o de valencia. Las bases en que se permite dos funciones de base por orbital se denominan *doble zeta*, las que permiten tres *triple zeta*, etc. Cuando los orbitales internos se describen por una función y los de valencia por dos o más, las bases se llaman *split-valence*. Además, aparte de las funciones de base comentadas, se suelen añadir al conjunto funciones adicionales para lograr una mejor descripción de la densidad electrónica, tales como funciones de polarización y funciones difusas. Las primeras son funciones de alto momento angular (p para el hidrógeno, d y f para átomos pesados), que permiten una mayor anisotropía de la densidad electrónica; las segundas son funciones con exponentes muy bajos que permiten que la densidad electrónica se aleje de los núcleos.

Las bases split-valence más extendidas en la literatura son las desarrolladas por el grupo de Popel, entre las cuales se encuentran las bases M-NG. El primer número indica el número de funciones usadas en la contracción de cada orbital interno, mientras que los números después del guión indican el número de funciones usadas para cada orbital de valencia (si hay dos números es una base doble zeta para los orbitales de valencia, si hay tres, triple zeta). Un ejemplo es la base 6-31G(d), que es la que se ha usado más extensamente en este trabajo. Esta base usa un orbital formado por seis primitivas para describir cada orbital interno, y dos orbitales formados por tres y una primitivas para describir cada orbital de valencia. La d entre paréntesis indica además que se ha añadido una función de polarización a los átomos pesados. Finalmente comentaremos la base 6-31+G(d), también usada en el presente trabajo, en la cual el símbolo + indica que se ha añadido funciones difusas a los átomos pesados.

2.2.2.6. Métodos que incluyen la correlación electrónica

La calidad de un resultado Hartree-Fock depende en gran medida del tamaño de la base, pero incluso para una base infinitamente grande (base completa) hay una diferencia entre la energía límite Hartree-Fock y la energía exacta no relativista, esta diferencia se denomina energía de correlación:

$$E_{exacta} = E_{limite HF} + E_{correlacion} \quad (2.23)$$

La energía de correlación tiene su origen en que en la teoría HF el movimiento de los electrones no está correlacionado, ya que cada electrón interactúa con un campo promedio del resto de

los electrones y no con los demás electrones individualmente. Esto hace que HF sobreestime la repulsión entre electrones al no considerar la correlación electrónica. Existen diversos métodos para calcular este efecto, que engloban tratamientos variacionales y perturbacionales. A continuación se comentarán muy brevemente estos métodos, aunque no han sido usados en el presente trabajo.

Interacción de Configuraciones Los métodos de interacción de configuraciones se basan en la constatación de que la función de onda exacta no se puede expresar bajo la fórmula de un simple determinante, como supone el método HF. La técnica CI consiste, entonces, en construir nuevos determinantes sustituyendo uno o varios orbitales ocupados por orbitales virtuales para generar estados excitados (Ψ_1 , Ψ_2 , etc.) y la función de onda correlacionada (Ψ) se optimiza bajo la forma de una combinación lineal de todos estos determinantes de Slater y del determinante HF (Ψ_0):

$$\Psi = c_0\Psi_0 + c_1\Psi_1 + c_2\Psi_2 + \dots \quad (2.24)$$

Debido a que es prácticamente imposible para sistemas usuales llevar a cabo todas las permutaciones de todos los electrones en todos los orbitales (*Full CI*), es una práctica común limitar el número de estados excitados a monoexcitaciones (CIS), a excitaciones simples y dobles (CISD), o incluso términos superiores como excitaciones triples (CISDT) o cuádruples (CISDTQ). Usualmente, además, se limitan las excitaciones a los electrones de valencia (aproximación *frozen core*). En un CI tradicional, los diferentes determinantes de la ecuación 2.24 se obtienen por un cálculo HF y sólo se optimizan variacionalmente los coeficientes c_0 , c_1 , etc. Si se permite, además, que varíen los coeficientes de las funciones de base en cada determinante, se obtienen mejores resultados para la función de onda. Este es el denominado método multiconfiguracional (MCSCF *MultiConfigurational Self-Consistent Field*).

Un cálculo CI es variacional, lo que asegura que la energía obtenida será siempre superior a la energía exacta. Pero una desventaja de estos métodos, excepto el *full CI*, es que no son coherentes en tamaño. Es decir, la energía de N átomos o moléculas no interaccionantes no será igual a la suma de las energías de los N átomos o moléculas por separado. El método QCI (*Quadratic Configuration Interaction*) [14] se desarrolló para corregir esta deficiencia.

En general, una interacción de configuraciones permite obtener la descripción no relativista más completa posible de un sistema molecular, representando el estado fundamental de la molécula por una mezcla de todos los estados electrónicos accesibles. Por contra, esta aproximación es habitualmente costosa para un uso rutinario y resulta prohibitiva para sistemas de tamaño medio.

Teoría de perturbaciones de Møller-Plesset La teoría de perturbaciones proporciona una metodología para encontrar la energía de correlación. Para ello, el Hamiltoniano del sistema total se separa en dos partes: una de referencia H_0 , la cual tiene función de onda y valores propios conocidos, y una de perturbación V , siendo ésta considerada “pequeña” respecto a H_0 . Bajo esta premisa, los métodos perturbacionales, cuya base teórica se ha denominado MBPT (*Many-Body Perturbation Theory*), expresan la función de onda generada como una suma de contribuciones mediante una expansión en serie en función del parámetro perturbacional λ . La teoría de perturbaciones fue aplicada a la química cuántica para sistemas polielectrónicos por Møller y Plesset [15]. En éste caso, para obtener una expansión de la energía de correlación se elige el Hamiltoniano HF como H_0 . El Hamiltoniano completo (H), se escribe como la suma del Hamiltoniano de orden cero (H_0) y una perturbación (V) acoplada con un parámetro λ (que varía entre 0 y 1):

$$H = H_0 + \lambda V \quad (2.25)$$

Asumiendo que la perturbación es pequeña comparada con H_0 , la función de onda perturbada y la energía se pueden expresar como:

$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots \quad (2.26)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \quad (2.27)$$

H_0 se construye a partir de la suma de los operadores de Fock monoeléctricos para todos los orbitales ocupados y toma la función de onda HF como $\Psi^{(0)}$. Un tratamiento posterior muestra que $E^{(0)} + E^{(1)} = E_{HF}$, con lo que las correcciones para la correlación electrónica empiezan en el orden 2. En función del grado de corrección que se considere se denomina el método de estimación de la energía de correlación (MP2, MP3, MP4...).

La ventaja del método MP sobre CI limitado es que el primero sí es coherente en tamaño. La desventaja es que como no es variacional la energía que resulta no es un límite superior a la energía exacta. Además, en la práctica hay que comprobar que la serie de resultados (MP2, MP3, MP4...) es convergente.

2.2.3. Métodos semiempíricos

Los métodos semiempíricos [7, 8], partiendo de la teoría HF SCF, introducen una serie de aproximaciones a fin de reducir el esfuerzo computacional que requieren los cálculos. En un cálculo HF, la evaluación de las integrales bielectrónicas es el paso que más esfuerzo exige en términos de coste computacional, ya que escalan formalmente según N^4 , donde N es el número de funciones de base. Por esta razón, estas aproximaciones básicamente tratan de reducir el número de integrales a evaluar, permitiendo que en la práctica se puedan abordar sistemas

de mayor tamaño. Estos métodos introducen, entonces, diversas aproximaciones en la expresión del Hamiltoniano, que se compensan con la introducción de un conjunto de parámetros. Esto hace necesario el ajuste de dichos parámetros, bien mediante la ejecución de cálculos *ab initio* de alto nivel, bien de forma que dichos métodos reproduzcan diferentes propiedades experimentales, como geometrías moleculares, calores de formación, energías de ionización, etc.

La primera aproximación de estos métodos es que no tratan explícitamente los electrones internos, de forma que éstos y el núcleo forman el llamado core atómico. Por otro lado, los electrones de valencia se describen mediante una base mínima formada por orbitales de Slater (STO) tipo s, p, y a veces también d. Finalmente, estos métodos aplican en diferente medida la aproximación *zero differential overlap* (ZDO). En esta aproximación, el solapamiento entre pares de orbitales diferentes para todos los elementos de volumen dv se iguala a cero:

$$\phi_\mu\phi_\nu dv = 0 \quad (2.28)$$

donde ϕ_μ y ϕ_ν son dos funciones de base.

Este conjunto de aproximaciones hacen que se reduzcan de forma muy importante el número de integrales a evaluar para resolver las ecuaciones de Roothaan-Hall. La primera implementación de forma práctica de la aproximación ZDO dio lugar al método *complete neglect of differential overlap* (CNDO), desarrollado por Pople y colaboradores [16, 17]. Sin embargo, la aplicación estricta de la aproximación ZDO a todos los productos de funciones de base es demasiado drástica, y los métodos semiempíricos más usados actualmente se basan en la aproximación *neglect of diatomic differential overlap* (NDDO), en la cual solamente se anulan los productos entre funciones de base que pertenecen a átomos diferentes.

Basados en la aproximación NDDO, e introduciendo algunas variaciones, Dewar y Thiel [18] desarrollaron el método *modified neglect of differential overlap* (MNDO), que supuso un avance importante en el campo. Una de las principales limitaciones del método MNDO es su incapacidad para modelizar apropiadamente los puentes de hidrógeno. Esto se debe a que el método MNDO tiende a sobreestimar la repulsión entre átomos separados por distancias típicas de interacciones de van der Waals. Con el fin de eliminar este problema, posteriormente se desarrolló el método AM1 [19] (*Austin Model 1*), que introduce términos adicionales en la energía de repulsión nuclear. Basado en las mismas expresiones que AM1, pero adoptando una parametrización más automatizada, se desarrolló también el modelo PM3 [20] (*Parametrized Model 3*). Estos dos modelos, AM1 y PM3, son dos de los modelos semiempíricos más extendidos en la actualidad, y los que se han usado en el presente trabajo. Diversas variantes de estos dos métodos han sido publicadas en los últimos 10 años [2].

2.2.4. Métodos basados en la teoría del funcional de la densidad

Los métodos basados en la teoría del funcional de la densidad (DFT, *Density Functional Theory*) [5], que aunque no se han usado en este trabajo se comentarán brevemente, se basan en la idea de calcular el intercambio y la correlación electrónica a partir de los funcionales generales de la densidad electrónica. Esta aproximación resulta de los trabajos de los años 1920, principalmente del modelo de Thomas, Fermi y Dirac del átomo, y de los trabajos de Slater de los años 1950. Más recientemente, se popularizaron los métodos DFT después de la publicación de los teoremas de Hohenberg-Kohn [21] en 1964, que demuestran la existencia de un funcional de la densidad electrónica que determina la energía del estado fundamental. Sin embargo, el teorema no determina la forma del funcional.

La razón por la cual los cálculos DFT son fácilmente aplicables es que la densidad electrónica puede ser representada como la suma del cuadrado de N densidades de orbitales mono-electrónicos. Estos orbitales definen un simple determinante de Slater. A partir de esta idea la teoría de Kohn-Sham (KS) [22] calcula la energía electrónica como la suma de varios términos:

$$E = E^T + E^V + E^J + E^{XC} \quad (2.29)$$

donde E^T es la energía cinética, E^V contiene los términos de energía potencial atractiva electrón-núcleo y los términos repulsivos entre núcleos, E^J representa la energía de repulsión coulombiana entre electrones, y E^{xc} es la energía de intercambio-correlación.

Todos los términos arriba mencionados, excepto la repulsión núcleo-núcleo, son funciones de la densidad electrónica, ρ , y son a menudo aproximados por integrales implicando la densidad de spin y a veces su gradiente:

$$E^{XC}(\rho) = \int f[\rho_\alpha(r), \rho_\beta(r), \nabla\rho_\alpha(r), \nabla\rho_\beta(r)] d^3r \quad (2.30)$$

donde ρ_α designa la densidad electrónica de spin α , ρ_β la de spin β , y $\rho = \rho_\alpha + \rho_\beta$, la densidad electrónica total.

En general E^{XC} se divide en una parte de intercambio (E^X) y una parte de correlación (E^C):

$$E^{XC}(\rho) = E^X(\rho) + E^C(\rho) \quad (2.31)$$

Los funcionales de intercambio y de correlación pueden ser locales y no locales, y a veces también se combinan con el término de intercambio de Hartree-Fock, como en el funcional de Becke de tres parámetros [23]. El acoplamiento de un funcional de intercambio con un funcional de correlación da nombre a cada método DFT usado en la práctica. Una de las combinaciones más empleadas actualmente es la B3LYP, donde se usa el funcional híbrido de

Becke de tres parámetros para la parte de intercambio, y el funcional de Lee-Yang-Parr (LYP) [24] para la parte de correlación.

2.2.5. Análisis de descomposición de la energía de Kitaura-Morokuma

El análisis de las interacciones moleculares implica determinar la naturaleza de la energía de interacción. Para ello es preciso realizar una descomposición de la misma en términos cercanos a nuestra percepción física. El método de Kitaura-Morokuma [25, 26] es un método que permite descomponer la energía de interacción HF de un complejo intermolecular en cinco componentes: electrostático, polarización, repulsión-intercambio, transferencia de carga y un término final denominado *mixing*, que aparece porque las cuatro componentes no dan lugar a la energía de interacción completa. Afortunadamente, este término suele ser relativamente pequeño.

Supongamos que hemos realizado un cálculo *ab initio* SCF en dos moléculas, X e Y , y sobre el complejo intermolecular XY . Las funciones de onda obtenidas se pueden escribir como $A\Psi_X^0$, $A\Psi_Y^0$ y $A\Psi_{XY}^0$, donde "A" indica el uso de funciones de onda antisimétricas, como por ejemplo un determinante de Slater. La suma de las energías de las moléculas aisladas es E_0 y la energía del complejo es E_4 . La energía de interacción es entonces igual a $E_4 - E_0$.

La contribución electrostática E_{es} se calcula como la interacción clásica de Coulomb entre las funciones de onda sin perturbar de las dos especies aisladas, X e Y , y es igual a $E_1 - E_0$, donde E_1 es la energía asociada al producto de las dos funciones de onda individuales, Ψ_1 :

$$\Psi_1 = A\Psi_X^0 A\Psi_Y^0 \quad (2.32)$$

Sin embargo, la distribución electrónica tanto de X como de Y cambia por la presencia de la otra molécula, es decir, se polariza, de forma que aparecen multipolos inducidos en cada molécula. Para calcular la polarización hay que calcular primero las funciones de onda Ψ_X y Ψ_Y cuando están en presencia de la otra molécula. La energía de la función de onda Ψ_2 se determina como E_2 , donde Ψ_2 es:

$$\Psi_2 = A\Psi_X A\Psi_Y \quad (2.33)$$

La contribución de polarización E_{pol} equivale entonces a $E_2 - E_1$ y es siempre atractiva.

Al determinar Ψ_1 y Ψ_2 no se consideran interacciones de intercambio electrónico. El solapamiento entre las distribuciones electrónicas de X e Y a distancias cortas causa una repulsión por el hecho de acercar electrones con el mismo spin en la misma región del espacio, lo cual lleva a la violación del principio de exclusión de Pauli. La contribución de repulsión-intercambio

E_{ex} se calcula como $E_3 - E_1$, donde E_3 es la energía de la función de onda Ψ_3 :

$$\Psi_3 = A(\Psi_X^0 \Psi_Y^0) \quad (2.34)$$

Ψ_3 se deriva, pues, a partir de las funciones de onda sin distorsionar de X e Y pero permitiendo el intercambio de electrones. Este término es siempre repulsivo.

Al movernos desde E_3 a E_4 , el sistema gana en polarización E_{pol} y transferencia de carga o deslocalización electrónica E_{ct} entre las dos moléculas. De forma similar, $E_4 - E_2$ incluye E_{ex} y E_{ct} . De esta forma, la energía de transferencia de carga E_{ct} se puede calcular como:

$$E_{ct} = E_1 + E_4 - E_2 - E_3 \quad (2.35)$$

El análisis de KM presenta algunos problemas, por ejemplo, cuando se introducen funciones difusas hay una cantidad considerable de solapamiento intermolecular, lo cual puede hacer difícil el descomponer la energía de interacción en los diferentes términos. Tampoco es una técnica que sea naturalmente extrapolable a cálculos multideterminantales. Sin embargo, es una herramienta muy útil que permite analizar distintos tipos de interacciones y racionalizar así los principales factores que contribuyen a sus energías.

2.3. Métodos clásicos

2.3.1. Introducción

Los métodos clásicos o de mecánica molecular son una alternativa a las metodologías cuánticas para abordar el estudio de sistemas de gran tamaño, y su uso está muy extendido en campos como la bioquímica, donde se suele trabajar con macromoléculas como proteínas o ácidos nucleicos. En estos métodos, la energía de un sistema químico se expresa como una función que depende únicamente de las posiciones de los núcleos, evaluándose de forma muy rápida mediante el uso de expresiones analíticas clásicas y de parámetros ajustados que incorporan de forma promediada la información electrónica. El conjunto de expresiones y parámetros constituyen lo que se denomina un campo de fuerzas (*force field*). Estos campos de fuerzas clásicos se suelen usar en combinación con técnicas de mecánica estadística como la Dinámica Molecular o el método de Monte Carlo. Una de las principales limitaciones de estas técnicas, sin embargo, es que no permiten estudiar roturas ni formación de enlaces covalentes, si bien la aparición de métodos híbridos QM/MM, en los cuales una parte del sistema se describe cuánticamente y el resto clásicamente, permiten aprovechar las principales ventajas de ambos tipos de metodologías para realizar, por ejemplo, estudios de reactividad enzimática.

A continuación se explicarán brevemente los términos que componen estos campos de fuerzas, que generalmente tratan los efectos de polarización en las distribuciones de carga de las moléculas de forma implícita. Posteriormente, se comentarán algunas de las aproximaciones que se han desarrollado para tratar explícitamente los efectos de polarización.

2.3.2. Campos de fuerzas

Existe un gran número de campos de fuerzas, y cada uno de ellos difiere en cuanto a las expresiones y los parámetros que usan para evaluar la energía. Sin embargo, todos ellos siguen un esquema común, donde la energía se calcula como una suma de contribuciones que pueden calcularse de forma independiente. Se suelen diferenciar entre los términos que dan cuenta de las contribuciones de átomos enlazados entre sí (*bonded terms*) y los términos que contabilizan las interacciones entre átomos no enlazados (*nonbonded terms*):

$$E_{total} = E_{bonded} + E_{nonbonded} \quad (2.36)$$

Generalmente, los términos enlazantes se componen de tres contribuciones básicas, E_{str} , E_{angle} y $E_{torsion}$, y los términos no enlazantes de dos, E_{ele} y E_{vdW} . Estos términos están presentes en todos los campos de fuerzas, y se muestran de forma esquemática en la figura 2.1. También se

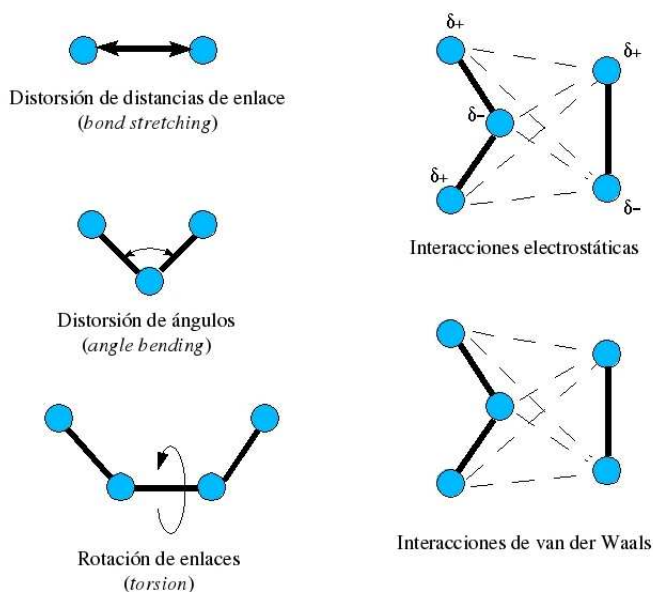


Figura 2.1: Contribuciones a la energía en mecánica molecular

han desarrollado campos de fuerzas que introducen los efectos de polarización explícitamente mediante un término E_{pol} adicional. Más adelante, después de describir los términos enlazantes y no enlazantes, se expondrán los principales métodos que se han propuesto para la inclusión de la polarización en los campos de fuerzas.

Adicionalmente, aunque no entraremos en ellos, podemos encontrar términos de puente de hidrógeno (E_{hbond}), términos cruzados (E_{cross}) y términos asociados a movimientos fuera del plano ($E_{out-of-plane}$). Los términos cruzados acoplan pares de contribuciones, es decir, introducen dependencias mutuas entre pares de contribuciones, por ejemplo $E_{str-str}$ o $E_{str-angle}$. Los términos $E_{out-of-plane}$, también llamados torsiones impropias, están asociados a movimientos fuera del plano y se usan generalmente para mantener la planaridad de átomos con hibridación sp^2 .

2.3.2.1. Términos enlazantes

Términos de *stretching*

Los términos de *stretching* (E_{str}) están relacionados con el coste energético que conlleva la distorsión de las distancias de enlace respecto a su valor de equilibrio. La curva que mejor

se ajusta al potencial de *stretching* es el potencial de Morse:

$$E_{str} = D_e \left[1 - e^{-a(l-l_0)} \right]^2 \quad (2.37)$$

donde l es la longitud del enlace, D_e es la profundidad del mínimo de energía potencial y $a = w\sqrt{(\mu/2D_e)}$, donde μ es la masa reducida y w es la frecuencia de vibración del enlace, que está relacionada con la constante de *stretching* del enlace, k , por $w = \sqrt{(k/\mu)}$. l_0 es el valor de referencia para la longitud del enlace.

Generalmente este tipo de potencial no se usa debido a que es costoso de evaluar, y en su lugar se considera un comportamiento armónico (ley de Hooke), que es mucho más eficiente computacionalmente:

$$E_{str} = \frac{k}{2}(l - l_0)^2 \quad (2.38)$$

Esta aproximación se comporta bien a distancias cercanas a las de equilibrio, que es donde se encuentran los enlaces en general, y es la forma usada, por ejemplo, en los campos de fuerza AMBER/parm94 [27], CHARMM22 [28, 29], GROMOS [30] y OPLS [31, 32]. Sin embargo, el acercamiento a los núcleos del potencial de Morse es más repulsivo que una parábola, mientras que el alejamiento que llevaría a una disociación del enlace es más suave. De esta forma, algunos campos de fuerzas, como MM3 [33, 34, 35], incluyen en la ecuación 2.38 términos cúbicos y de orden superior para reproducir mejor la curva de Morse.

Términos de *bending*

Los términos relacionados con los ángulos de enlace se denominan términos de *bending* (E_{bend}), y en general se calculan mediante potenciales armónicos (AMBER/parm94, CHARMM22, GROMOS, OPLS):

$$E_{bend} = \frac{k}{2}(\theta - \theta_0)^2 \quad (2.39)$$

donde θ es el ángulo de enlace y θ_0 denota el valor de equilibrio, y k es la constante de fuerza. Las constantes de fuerza de *bending* son menores que las asociadas a distancias de enlace, ya que se necesita menos energía para distorsionar un ángulo que un enlace. De la misma forma que con los términos de *stretching*, la precisión del campo de fuerzas se puede mejorar con la inclusión de términos de orden superior (MM3), o por el acoplamiento del *bending* con otros términos del campo de fuerzas (MM3).

Términos de torsión

Los términos de torsión (E_{tors}) describen las barreras de rotación de los enlaces químicos. Estos perfiles de energía presentan una periodicidad con diversos máximos y mínimos, que los

campos de fuerzas representan mediante series de Fourier:

$$E_{tors} = \sum_i \frac{V_i}{2} (1 + \cos(n\chi - \gamma_i)) \quad (2.40)$$

donde V_i representa la constante de fuerza para cada término de la función, n la periodicidad del término, χ el ángulo diedro y γ_i el ángulo de fase. Los ángulos de fase sirven para poder desplazar los máximos de cada función a voluntad. El número de términos en la serie de la ecuación 2.40 depende del tipo de torsión a considerar. Para torsiones muy sencillas basta con un sólo término, pero para torsiones más complejas es necesario añadir más términos. En general, se suelen usar como máximo términos de periodicidad 1, 2 y 3.

2.3.2.2. Términos no enlazantes

Término de van der Waals

El término de van der Waals (E_{vdW}) tiene en cuenta las interacciones de dispersión y de repulsión entre átomos no enlazados. Las fuerzas de dispersión son atractivas y aparecen a distancias generalmente entre 2 y 4 Å, debido fundamentalmente a reordenaciones e interacciones instantáneas de las nubes electrónicas de ambas moléculas o átomos. A distancias muy cortas, por otra parte, los núcleos y las capas internas de electrones dominan la interacción, de forma que se generan fuerzas repulsivas.

Como ocurre con otras contribuciones, existen varias expresiones para calcular este término. Las más populares son los denominados potenciales de *Lennard-Jones* con expresiones 6-12 como la siguiente:

$$E_{vdW} = 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.41)$$

donde σ_{ij} es el diámetro de colisión, para el cual la energía es cero, y ε_{ij} determina la profundidad del mínimo y se calcula a partir de la media geométrica de valores individuales $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$. En el campo de fuerzas OPLS, que adopta la ecuación 2.41, σ_{ij} se calcula como la media geométrica de valores individuales $\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$.

El potencial de Lennard-Jones se puede expresar también en función de la separación correspondiente al mínimo de energía, r_{ij}^* , que se calcula como la suma de los radios de van der Waals de los átomos implicados $r_{ij}^* = r_i + r_j$. Esta expresión es la que adopta, por ejemplo, el campo de fuerzas AMBER/parm94:

$$E_{vdW} = \varepsilon_{ij} \left[\left(\frac{r_{ij}^*}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^*}{r_{ij}} \right)^6 \right] \quad (2.42)$$

Otros campos de fuerzas como MM3 utilizan el potencial de Hill, derivado del potencial de Buckingham, donde se usa un término exponencial para describir las fuerzas de repulsión que es más preciso que el término r^{-12} , pero también más costoso de evaluar computacionalmente:

$$E_{vdW} = \varepsilon_{ij} \left[1,84 \cdot 10^5 e^{\left(-12,00 \frac{r_{ij}}{r_{ij}^*}\right)} - 2,25 \left(\frac{r_{ij}^*}{r_{ij}}\right)^6 \right] \quad (2.43)$$

Las fuerzas de van der Waals se suelen calcular para todos los pares de átomos, pertenezcan o no a la misma molécula, siempre que se encuentren separados por al menos tres enlaces. En este caso, las interacciones a tres enlaces (1-4) se escalan para disminuir la repulsión a lo largo de los perfiles de rotación de los enlaces químicos.

Término electrostático

El término electrostático (E_{ele}) describe las interacciones electrostáticas existentes entre las distribuciones de carga de las diferentes partes del sistema. La evaluación de esta contribución puede realizarse mediante diferentes tipos de expresiones. La estrategia más habitual consiste en asignar cargas puntuales a los átomos del sistema. Estas cargas se derivan de forma que reproduzcan las propiedades electrostáticas de la molécula. Entonces la contribución electrostática a la energía total del sistema se calcula como la suma de interacciones entre pares de cargas puntuales usando la ley de Coulomb:

$$E_{ele} = \sum_i \sum_{j>i} \frac{q_i q_j}{\varepsilon r_{ij}} \quad (2.44)$$

donde ε es la constante dieléctrica del medio y r_{ij} la distancia que separa a las cargas q_i y q_j . Generalmente las interacciones electrostáticas se calculan entre todos los pares de átomos separados por al menos tres enlaces. Como en el caso del término de van der Waals, la interacción a tres enlaces se suele escalar por un factor a fin de mejorar la energética del perfil de rotación.

2.3.2.3. Efectos de polarización

Hasta ahora hemos considerado que las distribuciones de carga son estáticas. En realidad, sin embargo, estas distribuciones de carga pueden ser perturbadas por la influencia de un campo eléctrico como el que generan el resto de cargas del sistema, es decir, las moléculas son polarizables. Los efectos de polarización se suelen introducir de forma promediada mediante el uso de cargas sobrepolarizadas respecto a las de la molécula aislada. Se han desarrollado, sin embargo, diversos modelos para introducir los efectos de polarización en los campos de fuerzas

[36], ya que una descripción adecuada de este tipo de efectos parece necesario en sistemas altamente polarizables, como las interacciones catión- π [37, 38, 39]. De estos métodos, los más extendidos son los basados en el modelo de dipolos puntuales, en modelos de fluctuación de carga, o en modelos basados en osciladores clásicos de Drude, que se describirán brevemente a continuación.

El modelo de dipolos puntuales La aproximación más extendida se basa en el modelo de dipolos puntuales formulado por Applequist [40], que consiste en asignar un tensor de polarizabilidad a cada átomo. Asumiendo un modelo de polarizabilidad atómica isotrópico, el campo eléctrico genera sobre cada átomo un dipolo inducido según:

$$\mu_{ind,i} = \alpha_i E_i \quad (2.45)$$

donde α_i es la polarizabilidad atómica.

El campo eléctrico E_i sobre el átomo i es la suma del campo eléctrico estático y el creado por los dipolos inducidos de los otros átomos del sistema. De esta forma, la determinación de los momentos dipolares inducidos sobre cada átomo debe resolverse iterativamente hasta alcanzar convergencia. Una vez alcanzada, la energía de interacción entre un dipolo y el campo eléctrico (la energía de inducción) se calcula según:

$$E_{ind} = -\frac{1}{2} \mu_{ind,i} \cdot E_{i,0} \quad (2.46)$$

donde $E_{i,0}$ es el campo eléctrico estático generado únicamente por las cargas puntuales. Esta contribución se calcula para todos los dipolos inducidos, y sumado al término de la ecuación 2.44 nos permite obtener la energía electrostática total del sistema. La evaluación de los efectos de polarización precisa de la autocoherencia en las distribuciones de dipolos inducidos para cada movimiento nuclear. Esto hace que el incremento en el tiempo de cálculo necesario para la evaluación de los dipolos inducidos sea muy elevado, y por ello su uso no se ha popularizado. A pesar de estas dificultades, el modelo de dipolos puntuales se encuentra ya implementado en el campo de fuerzas polarizable AMBER/parm02 [41].

Modelos de fluctuación de carga En estos métodos, las cargas atómicas fluctúan en respuesta al entorno según el principio de igualdad de electronegatividad, que implica que la carga fluye entre los átomos hasta que las electronegatividades instantáneas de los átomos son iguales. En esta aproximación, a las cargas que fluctúan se les asignan masas ficticias y son tratadas como grados de libertad adicionales en las ecuaciones de movimiento. En el contexto de la dinámica molecular, las ecuaciones resultantes se pueden resolver de forma eficiente con un coste computacional no muy superior al necesario para un campo de fuerzas

con cargas fijas [42]. Este tipo de modelos, sin embargo, pueden estar seriamente limitados para reproducir las diferentes componentes del tensor de polarizabilidad de algunas especies moleculares adecuadamente. Por ejemplo, una molécula lineal sólo podría polarizarse a lo largo de su eje. Recientemente, se ha publicado una versión del campo de fuerzas CHARMM para el estudio de proteínas que incorpora el modelo de fluctuación de carga para tratar las fuerzas de polarización [43].

Modelos basados en osciladores clásicos de Drude En este tipo de modelos, para modelizar la polarizabilidad α de un átomo con carga parcial q se asocia a éste un “oscilador de Drude”. Esto consiste en enlazar a cada átomo una partícula adicional mediante un enlace químico ficticio con distancia de enlace igual a 0 y constante de fuerza k_D . A cada partícula Drude, entonces, se le asigna una masa y una carga q_D que se sustraen de la masa y la carga q de su átomo correspondiente, de forma que la masa y carga totales se mantienen para la pareja Drude-átomo. De esta forma, en ausencia de campos externos la partícula Drude oscila alrededor de la posición del átomo (r) y por tanto la pareja Drude-átomo se comporta en promedio como una carga puntual q . Sin embargo, cuando actúa un campo eléctrico E , la partícula Drude se desplaza del átomo una distancia $d = \frac{qE}{k_D}$, de forma que oscila alrededor de $r + d$, y el dipolo inducido sobre el átomo toma un valor promedio de $\mu = \frac{q_D^2 E}{k_D}$. Entonces, la polarizabilidad isotrópica atómica tiene la simple expresión:

$$\alpha = \frac{q_D^2}{k_D} \quad (2.47)$$

Para una polarizabilidad α dada, entonces, se escoge la constante de fuerza k_D de forma que la partícula Drude se mantenga a una distancia d mucho menor que cualquier distancia interatómica. De esta forma, el dipolo inducido resultante es casi equivalente a un dipolo puntual.

Recientemente se han publicado algoritmos eficientes para la implementación de este modelo en simulaciones de dinámica molecular [44]. También se ha desarrollado un modelo de agua polarizable basado en el oscilador de Drude [45] como primer paso para la definición de un campo de fuerzas basado en este modelo.

2.3.2.4. Parametrización

Un aspecto fundamental que distingue los distintos campos de fuerzas que se encuentran en la bibliografía es el tipo de compuestos para el que han sido parametrizados. De esta forma, existen campos de fuerzas diseñados y parametrizados para moléculas orgánicas, como MM3, u otros diseñados para el estudio de biomoléculas, como AMBER/parm94, GROMOS y

CHARMM22. Así, la elección de uno u otro campo de fuerzas depende del problema concreto que se vaya a estudiar, y debe considerar siempre la parametrización de éstos.

El desarrollo de los parámetros de los campos de fuerzas es un trabajo costoso. Las ecuaciones y los parámetros deben estar cuidadosamente balanceados, y se ajustan a partir de datos experimentales o de cálculos cuánticos realizados sobre sistemas modelo usados como referencia. Es importante destacar, además, que aunque la magnitud de los parámetros individuales debe conservar cierto sentido químico, no tienen porqué corresponder con medidas experimentales directas, ya que un campo de fuerzas es un modelo matemático conjunto y debe mantener un balance adecuado entre las diversas contribuciones energéticas. En caso de que no se hayan descrito en la bibliografía los parámetros necesarios para abordar un nuevo sistema, estos deben desarrollarse como paso previo a la modelización.

2.4. Modelos continuos de solvatación

2.4.1. Introducción

El creciente número de estudios teóricos aparecidos en las últimas décadas demuestran la extraordinaria evolución que han sufrido los métodos quimicuánticos en su aplicación al estudio de sistemas químicos. El esfuerzo investigador en este campo ha cristalizado en una serie de elaborados métodos que han permitido el acercamiento de la química cuántica al límite de la precisión experimental en fase gas.

En el mundo de la química, sin embargo, lo más habitual es trabajar con sistemas en estado líquido, y no en fase gas. En este tipo de sistemas, el medio juega un papel clave, alterando de forma decisiva las propiedades reactivas y estructurales de los compuestos químicos. El estudio de tales sistemas en solución es, sin embargo, mucho más complicado que en fase gas [46]. La principal limitación consiste en el enorme número de moléculas que hay que considerar de forma dinámica para representar el conjunto de entidades químicas (solute, solvente) que constituyen el estado líquido. Esta complejidad ha dado lugar a una amplia variedad de aproximaciones computacionales, los cuales incluyen métodos basados en i) la elaboración de funciones físicas; ii) simulaciones moleculares clásicas de líquidos, donde las propiedades del sistema se obtienen a partir de un conjunto de configuraciones representativas del sistema soluto-solvente; iii) descripciones de la solución mediante el concepto de supermolécula, que proporciona información limitada pero detallada sobre las interacciones específicas soluto-solvente; iv) la combinación de tratamientos mecanocuánticos del soluto con descripciones discretas clásicas de las moléculas de solvente promediadas estadísticamente; y v) modelos continuos, en los cuales la atención se centra básicamente en una parte del sistema, el soluto, mientras el solvente se representa de forma muy simplificada como un medio continuo polarizable.

En las últimas décadas ha habido una evolución hacia métodos más elaborados para el estudio de fases condensadas. En este sentido han aparecido multitud de estrategias que combinan los diversos métodos antes mencionados para la descripción de las moléculas de solvente en conjunción con tratamientos clásicos o quimicuánticos para el soluto. La evolución que han experimentado estos métodos ha sido objeto de análisis de diversas revisiones [47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57], que proporcionan una evaluación crítica de su utilidad para el estudio de sistemas químicos y bioquímicos.

En esta sección se describirán los principales métodos teóricos que tratan el efecto del solvente mediante modelos continuos de solvatación en los cuales el soluto se describe a nivel mecanocuántico. Descripciones más detalladas de estos métodos pueden encontrarse en numerosas publicaciones [58, 59, 60, 61, 62, 63]. En primer lugar se introducirán los fundamentos

y las características básicas de los modelos continuos. Posteriormente se introducirán las expresiones más extendidas para el cálculo de las distintas contribuciones a la energía libre de solvatación, para después describir las distintas estrategias de parametrización que siguen los métodos usados en el presente trabajo. Finalmente, se describirán los fundamentos y las aplicaciones derivadas de modelos de solvatación fraccionales, los cuales permiten particionar la energía libre de solvatación en contribuciones fragmentales asociadas a átomos o grupos, y cómo esta información puede ser de gran utilidad en campos como el diseño de fármacos mediante el establecimiento de relaciones estructura-actividad.

2.4.2. El modelo continuo

Los modelos continuos de solvatación se basan en la definición de un Hamiltoniano efectivo (H_M) para el soluto M (ec. 2.48). Este Hamiltoniano se puede expresar como la suma del Hamiltoniano *in vacuo* (H_M^0) y un potencial de interacción (V_{int}). Dentro de la aproximación Born-Oppenheimer, el Hamiltoniano del soluto depende de las coordenadas de los electrones, q , y paramétricamente de las coordenadas de los núcleos, Q . El potencial de interacción depende además de las coordenadas del solvente, indicadas de forma colectiva por Ω , de forma que el Hamiltoniano electrónico puede expresarse como:

$$H_M(q, Q, \Omega) = H_M^0(q, Q) + V_{int}(q, Q, \Omega) \quad (2.48)$$

Siguiendo la definición de Ben-Naim para el proceso de solvatación [64], la introducción de un soluto en un solvente se puede definir como el proceso por el cual una partícula de soluto es transferida desde una posición fija en fase gas a una posición fija en solución a temperatura, presión y composición del solvente constantes. Bajo estas condiciones, la energía libre de Gibbs de solvatación, ΔG_{sol} , se puede relacionar con el trabajo reversible necesario para crear el soluto M dentro del solvente S a igual densidad numérica en fase gas y en solución. Tal trabajo incluye interacciones soluto-solvente directas, así como contribuciones debidas a cambios internos en el soluto y el solvente debidos a la solvatación:

$$\Delta G_{sol} = W(M/S) + RT \ln \left(\frac{(q_{rot} q_{vib})_{gas}}{(q_{rot} q_{vib})_{sol}} \right) \quad (2.49)$$

donde $W(M/S)$ es el trabajo de acoplamiento del soluto M en el solvente S , y q_{rot}, q_{vib} denotan la función de partición para la rotación y vibración de M .

La definición precedente omite la contribución translacional, que aparece debido a las diferencias en funciones de partición del momento del soluto en fase gas y en solución. En la práctica, la contribución a la solvatación derivada de las diferencias en la función de partición

rotacional y vibracional en fase gas y en solución es también ignorada, lo cual simplifica el cálculo de la energía libre de solvatación al trabajo de acoplamiento M/S.

El trabajo de acoplamiento entre M y S se puede determinar mediante el método del parámetro de carga (ec. 2.50) [65], donde la transición desde los estados correspondientes al soluto aislado + solvente puro al del soluto inmerso en el solvente (denotados por los índices 0 y 1, respectivamente) se lleva a cabo mediante un parámetro de carga, λ , de forma que la interacción entre el soluto y el solvente se activa progresivamente al variar λ entre 0 y 1:

$$W(M/S) = \int_0^1 d\lambda \int d\Omega [\rho_S(\lambda) V_{int}(q, Q, \Omega) g_s(\Omega; \lambda)] \quad (2.50)$$

donde $g_s(\Omega; \lambda)$ indica la función de distribución promediada térmicamente del conjunto infinito de moléculas de solvente correspondientes al potencial de interacción $\lambda V_{int}(q, Q, \Omega)$.

El trabajo de acoplamiento de M con S se puede descomponer en distintas contribuciones, típicamente cavitación, dispersión, repulsión y contribución electrostática, de forma que la ecuación 2.50 se puede resolver a partir de integraciones consecutivas en las que intervienen distintos parámetros de carga:

$$W(M/S) = \Delta G_{ele} + \Delta G_{cav} + \Delta G_{dis} + \Delta G_{rep} \quad (2.51)$$

ΔG_{cav} es un término positivo que corresponde al trabajo necesario para crear una cavidad suficientemente grande para acomodar el soluto dentro del solvente. Los términos ΔG_{dis} y ΔG_{rep} aparecen al activar las fuerzas de dispersión y repulsión entre el soluto y el solvente. Finalmente, el término ΔG_{ele} corresponde al trabajo necesario para generar la distribución de carga del soluto dentro del solvente. La contribución electrostática es la más relevante para solutos polares o cargados, especialmente en solventes polares, mientras que el resto de componentes, llamadas conjuntamente contribución no electrostática, son dominantes para solutos y solventes apolares.

En general, los modelos continuos usan una forma simplificada del potencial de interacción V_{int} , que se reduce a la componente electrostática, mientras los términos no electrostáticos se determinan usando diversos formalismos basados en parámetros geométricos del soluto y propiedades generales del solvente. Sin embargo, se han propuesto también diversos métodos para la introducción de los efectos de dispersión y repulsión en la función de onda, por ejemplo, el método propuesto por Amovilli y Menucci [66], que permite calcular la repulsión de Pauli y las contribuciones de dispersión de forma autoconistente como parte del operador de interacción.

2.4.3. Contribución electrostática

En el contexto de la aproximación de respuesta lineal, el tratamiento quimicuántico de un soluto dentro de un medio continuo se puede representar formalmente por la ecuación de Schrödinger que se indica a continuación (ec. 2.52), donde el operador perturbativo V_{ele} añadido al Hamiltoniano del soluto acopla la respuesta electrostática del solvente, caracterizado por una constante dieléctrica ε , a la distribución de carga del soluto ρ_M , y E es la energía electrostática del soluto más la contribución electrostática a la energía libre de solvatación. El factor $1/2$ de la ecuación 2.52 tiene en cuenta el trabajo necesario para polarizar el solvente, que de acuerdo con la teoría de respuesta lineal es igual a la mitad de la interacción soluto-solvente.

$$\left[H_M^0(q, Q) + \frac{1}{2} V_{ele}(q, Q, \rho_M, \varepsilon) \right] \Psi(q, Q) = E(Q) \Psi(q, Q) \quad (2.52)$$

La energía libre electrostática se puede determinar restando la energía del soluto en el vacío, E^0 , de la energía libre electrostática del soluto en solución, G_{ele} (ec. 2.53), que se puede obtener a partir de un cálculo variacional del funcional de la ecuación 2.54, donde se optimizan simultáneamente la función de onda del soluto y el campo de reacción, que dependen mutuamente entre sí.

$$\Delta G_{ele} = G_{ele} - E^0 \quad (2.53)$$

con

$$G_{ele} = \left\langle \Psi(q, Q) \left| H^0 + \frac{1}{2} V_{ele}(q, Q, \rho_M, \varepsilon) \right| \Psi(q, Q) \right\rangle \quad (2.54)$$

A nivel Hartree-Fock, esto se puede realizar mediante un matriz de Fock modificada, cuyos elementos se ven alterados, respecto a los del soluto en vacío, por la adición de la contribución V_{ele} al término monoeléctrico:

$$F_{\mu\nu} = F_{\mu\nu}^0 + \langle \mu | V_{ele} | \nu \rangle \quad (2.55)$$

Se han descrito diversos métodos en la literatura para describir el campo de reacción del solvente, y por tanto para resolver el problema electrostático. A continuación se describirán las tres aproximaciones más populares y extendidas, que han sido las aplicadas en la presente tesis.

2.4.3.1. Métodos de expansión multipolar (MPE)

En esta aproximación tanto la distribución de carga del soluto como la del campo de reacción del solvente se expresan como una expansión multipolar truncada. Esta definición engloba

y generaliza los modelos originalmente propuestos por Born, Kirkwood, Bell y Onsager. El modelo de Born [67] incluye un monopolo en una cavidad esférica, y fue desarrollado posteriormente por Kirkwood [68] para tener en cuenta los efectos de la fuerza iónica. El modelo de Bell [69] describe la energía libre de solvatación de un dipolo puntual en una cavidad esférica. El modelo de Onsager [70] es un refinamiento del modelo anterior, que añade al sistema un término de polarizabilidad dipolar isotrópica. Posteriores mejoras a estos métodos incluyen más términos en la expansión multipolar y usan cavidades más realistas.

En el contexto de métodos QM, los modelos MPE (*multipolar expansion methods*) fueron desarrollados por Hall *et al.* [71], Tapia y Goscinski [72], Rivail y Rinaldi [73, 74], Yomosa [75, 76], Karelson y Zerner [77], Mikkelsen y colaboradores [78] y Wiberg *et al.* [79]

El grupo de Nancy ha desarrollado una versión muy elaborada y elegante de modelo SCRF basado en expansiones multipolares. En las últimas versiones del modelo se consideran expansiones multipolares multicéntricas con términos hasta $l = 12$ [80], aunque truncando en $l = 6$ se obtiene una convergencia aceptable en la mayoría de casos. La cavidad esférica inicialmente adoptada se reemplazó por un elipsoide [81], y más adelante por cavidades adaptadas a la forma de la molécula [82]. Típicamente esta cavidad se construye a partir de la superposición de esferas centradas en los núcleos usando los radios de Bondi escalados por un factor de 1.3, que se derivó a partir de ajustes al volumen molecular experimental para un conjunto de líquidos de referencia.

Si indicamos por I, J los centros de la expansión multipolar, la componente electrostática a la energía libre de solvatación se obtiene introduciendo en los elementos de la matriz de Fock, $F_{\mu\nu}$, un término debido al campo de reacción del solvente:

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_I \sum_J \sum_{l,l'} \sum_{m,m'} \langle \mu | M_l^m(I) | \nu \rangle \cdot f_{ll'}^{mm'}(I, J) M_{l'}^{m'}(J) \quad (2.56)$$

donde

$$M_{l'}^{m'}(J) = \sum_{\lambda} \sum_{\eta} \langle \lambda | M_{l'}^{m'}(J) | \eta \rangle \quad (2.57)$$

donde $F_{\mu\nu}^0$ representa el elemento de la matriz de Fock para el soluto aislado, M_l^m denota el operador para la expansión multipolar de la distribución de carga del soluto, y $f_{ll'}^{mm'}$ son los factores del campo de reacción, que dependen de la constante dieléctrica del solvente y de la forma de la cavidad.

Este método ha sido extendido para incluir los efectos de correlación electrónica a nivel de teoría de perturbaciones Møller-Plesset o de la teoría del funcional de la densidad [83, 84, 85]. Además, también se han derivado expresiones analíticas para las primeras y segundas derivadas de la energía libre [86, 87].

2.4.3.2. Modelo generalizado de Born (GB)

El modelo generalizado de Born (GB, *Generalized Born*) se puede considerar un caso especial de los métodos MPE, ya que es una generalización multicéntrica de la ecuación de Born para iones monoatómicos. De acuerdo a este formulismo, la distribución de carga se describe a través de un conjunto de cargas puntuales centradas en los núcleos. La componente electrostática a la solvatación se calcula a partir de la solvatación individual de Born de cada átomo (carga), corregida por el efecto perturbativo de los otros átomos (cargas).

La versión QM más elaborada del modelo GB corresponde a las series de modelos SMx desarrollados por Cramer y Truhlar [63], que se implementaron originalmente en los métodos semiempíricos AM1 y PM3, y posteriormente a diversos niveles *ab initio* y DFT. Aquí los elementos de la matriz de Fock adoptan la forma indicada por la ecuación 2.58, donde sólo los elementos diagonales de la matriz de Fock se modifican como consecuencia de la función delta de Krönecker, $\delta_{\mu\nu}$:

$$F_{\mu\nu} = F_{\mu\nu}^0 + \delta_{\mu\nu} \frac{\varepsilon - 1}{\varepsilon} \sum_{k'} q_{k'} \gamma_{kk'} \quad \text{con } \mu \in k \quad (2.58)$$

donde

$$q_k = Z_k - \sum_{\mu \in k} P_{\mu\mu} \quad (2.59)$$

k, k' son los índices atómicos, Z_k es la carga nuclear del átomo k , P es la matriz densidad de primer orden, q_k es la carga atómica y $\gamma_{kk'}$ es la integral de Coulomb.

Las versiones originales del modelo se han refinado en distintas direcciones que han culminado con el último modelo SM5 [88]. El punto más destacable es el refinamiento de la distribución de monopolos. Las versiones originales usaban el análisis de población de Mulliken [89] para obtener las cargas atómicas, pero el modelo SM5 adopta cargas de clase IV [90], que se basan en un procedimiento destinado a reproducir observables dependientes de las cargas. En particular, el modelo SM5.42 se basa en el procedimiento CM2 (*Charge Model 2*) [91, 92, 93, 94], donde las cargas atómicas se determinan primero calculando las cargas de Löwdin [95], y después aplicando la expresión siguiente:

$$q_k^{CM2} = q_k^{Lwdin} + \sum_{k' \neq k} B_{kk'} (D_{Z_k Z_{k'}} + C_{Z_k Z_{k'} B_{kk'}}) \quad (2.60)$$

donde $B_{kk'}$ es el orden de enlace de Mayer entre los átomos k y k' , y los elementos de las matrices C y D , que dictan como se redistribuye la carga entre átomos que tienen diferentes números atómicos, se ajustan para reproducir momentos dipolares en fase gas experimentales.

Otro aspecto crítico del método GB es el cálculo de los radios efectivos de Born para cada

átomo, que dependen de los radios intrínsecos de Coulomb y de la geometría de la molécula. En versiones iniciales del modelo GB los radios intrínsecos de Coulomb se calculaban en función de la carga atómica. En el modelo SM5, sin embargo, adoptan valores constantes para todos los átomos excepto el hidrógeno, al cual se le asignan diferentes valores según el átomo al que esté enlazado.

2.4.3.3. Carga superficial aparente (ASC)

La aproximación de carga superficial aparente (ASC, *apparent surface charge*) nos permite obtener la componente electrostática a la solvatación a partir de otra estrategia. En particular, la formulación original del modelo del continuo polarizable (PCM, *Polarizable Continuum Model*) fue desarrollada por Miertus, Scrocco y Tomasi [96, 97].

Este método se basa en el cálculo de una carga superficial aparente que simula la respuesta del solvente a la distribución de carga del soluto. Para obtener esta superficie, la cavidad del soluto se divide en un conjunto de elementos de superficie llamados teseras suficientemente pequeños para que la densidad de carga del solvente sea constante. Entonces, la carga aparente se puede calcular a partir de la componente normal a la superficie del campo eléctrico:

$$q_j = -\frac{\varepsilon - 1}{\varepsilon} S_j \left(\frac{\partial V_T}{\partial n} \right)_j \quad (2.61)$$

donde V_T es el potencial electrostático total, que incluye ambas contribuciones del soluto y del solvente, n es el vector unitario normal al elemento de superficie j , S_j es el área del elemento de superficie j , y ε es la constante dieléctrica del solvente.

La componente electrostática a la energía libre de solvatación se obtiene añadiendo un término que describe la interacción electrostática con las cargas aparentes a los elementos de la matriz de Fock, $F_{\mu\nu}$:

$$F_{\mu\nu} = F_{\mu\nu}^0 + \left\langle \mu \left| \sum_k \frac{q_k}{|r - r_k|} \right| \nu \right\rangle \quad (2.62)$$

donde M es el número total de elementos de superficie en los que se ha dividido la cavidad y $\{q_k\}$ es el conjunto de cargas aparentes (localizadas en r_k) que representan la respuesta del solvente a la distribución de carga del soluto.

A lo largo de los años se han ido refinando diversos aspectos del modelo original PCM. Debido a la dependencia mutua entre las distribuciones de carga del soluto y del solvente, originalmente se propuso un proceso iterativo hasta alcanzar autoconherencia entre la función de onda del soluto y el campo de reacción del solvente. Este proceso iterativo se reformuló a través de otros procedimientos. En concreto, el método de inversión de matriz [98, 99] permitió

el desarrollo de expresiones analíticas para las primeras y segundas derivadas de la energía libre, que ha permitido la implementación de algoritmos de optimización de la geometría en solución eficientes [100, 101, 102].

Un aspecto de este método que ha sido objeto de una atención particular es el tratamiento de la compensación de carga [103, 104], un problema que se plantea a causa de i) las colas de densidad de carga del soluto que se encuentran fuera de la cavidad, y ii) los errores numéricos debidos a la división de la superficie de la cavidad en teseras. Originalmente este efecto se corregía definiendo un factor f^X , que satisface la relación de la ecuación 2.63, donde Q_M^X indica la carga neta del soluto:

$$f^X \sum_k q_k = -\frac{\varepsilon - 1}{\varepsilon} Q_M^X \quad (2.63)$$

Los tratamientos más elaborados para realizar la compensación de carga se basan en dos aspectos. El primero, en la introducción de factores diferentes para las cargas aparentes inducidas por los núcleos y por los electrones. En este caso Q_M^X representa la carga total nuclear ($X = N$) y el número total de electrones ($X = e$). En segundo lugar, en el tratamiento explícito de la densidad de carga que se encuentra fuera de la cavidad generando una carga aparente efectiva adicional. Esto se realiza calculando la carga escapada del soluto, Q_{out} , a partir de la diferencia entre el número total de electrones, N_{el} , y la carga de soluto dentro de la cavidad, Q_{in} , que se determina a partir del teorema de Gauss:

$$Q_{out} = -(N_{el} - Q_{in}) = -\left(N_{el} + \frac{1}{4\pi} \int_S E_M(s) n_s ds\right) \quad (2.64)$$

Una vez hemos calculado de forma precisa la carga exterior del soluto, Q_{out} , y hemos introducido sus efectos a través de un conjunto extra de cargas aparentes, la respuesta del solvente inducida por los núcleos o por la densidad electrónica sólo se ven afectados por los mismos tipo de errores numéricos.

El modelo original PCM se desarrolló para estudiar medios isotrópicos homogéneos, pero el grupo de Pisa ha desarrollado recientemente el llamado formalismo de la ecuación integral (IEF, *Integral Equation Formalism*) [105, 106, 107, 108], que permite estudiar medios dieléctricos anisotrópicos y soluciones iónicas. Este nuevo formalismo representa una mejora considerable respecto a su formulación original, y es menos sensible a los errores de compensación de carga que se indicaban para su versión tradicional.

La inclusión de los efectos de correlación electrónica en el método PCM se ha realizado a diferentes niveles [109, 110, 111, 112, 113, 114, 115], usando métodos CI, MC-SCF, MBPT y de la teoría del funcional de la densidad, y prestando especial atención a la inclusión de

los efectos del solvente sólo a nivel energético o a nivel de la energía y de la matriz densidad simultáneamente.

Otro método basado en la aproximación ASC similar al PCM es el modelo COSMO (*conductor-like screening model*) desarrollado por Klamt y Schuurman [116, 117]. La característica principal de COSMO es que el campo de reacción se trata asumiendo que el solvente se comporta como un conductor en vez de un dieléctrico. Los principales beneficios de este modelo son la simplificación del problema electrostático, que facilita el cálculo de gradientes analíticos. Una extensión a este modelo es el llamado método COSMO-RS (*real solvent*) [118, 119], el cual se ha parametrizado para diferentes solventes proporcionando resultados de notable calidad. No entraremos en los detalles de estos modelos ya que no se han usado en el presente trabajo.

2.4.4. Contribuciones no electrostáticas

2.4.4.1. Cavitación

El término de cavitación ΔG_{cav} corresponde al trabajo necesario para crear una cavidad de volumen y forma adecuados para acomodar el soluto dentro del solvente. Se han propuesto diferentes formalismos para el cálculo de la energía libre de cavitación basados en la forma y tamaño del soluto y en propiedades del solvente, como la tensión superficial, la compresibilidad isotérmica, o el radio molecular y la densidad numérica [120, 121, 122, 123]. Una formulación especialmente elaborada proviene de la teoría de partículas escaladas (SPT, *Scaled Particle Theory*) de Pierotti [124]. En este formalismo se asume que las moléculas se comportan como esferas rígidas, y la energía libre de cavitación se expande en potencias del radio de la esfera que excluye los centros de las moléculas de solvente (ec. 2.65). Los coeficientes K se expresan en términos de propiedades del solvente y de la solución (ec. 2.66-2.69), como el radio molecular (R_S), la densidad numérica del solvente (n_S), la presión (P) y la temperatura (T).

$$\Delta G_{cav} = K_0 + K_1 R_{MS} + K_2 R_{MS}^2 + K_3 R_{MS}^3 \quad (2.65)$$

$$K_0 = RT \left[-\ln(1-y) + \frac{9}{2} \left(\frac{y}{1-y} \right)^2 \right] - \frac{4\pi R_S^3 P}{3} \quad (2.66)$$

$$K_1 = -\frac{3RT}{R_S} \left[\frac{y}{1-y} + 3 \left(\frac{y}{1-y} \right)^2 \right] + 4\pi R_S^2 P \quad (2.67)$$

$$K_2 = \frac{3RT}{R_S} \left[\frac{y}{1-y} + \frac{3}{2} \left(\frac{y}{1-y} \right)^2 \right] - 4\pi R_S P \quad (2.68)$$

$$K_3 = \frac{4\pi P}{3} \quad (2.69)$$

donde $y = \frac{4\pi R_S^3 n_S}{3}$, y R_{MS} es la suma de los radios del solvente y del soluto, es decir, $R_{MS} = R_M + R_S$.

El formalismo de Pierotti se espera que funcione razonablemente bien para solutos y solventes de tamaño pequeño y esférico. Sin embargo, para solventes con forma no esférica la aplicación del modelo es más delicada, siendo necesario definir un radio efectivo para las moléculas de solvente. Por ejemplo, para la parametrización en n-octanol del modelo continuo MST descrita en el primer trabajo de esta tesis se adoptó un valor de 6.84 Å, determinado a partir del análisis de datos de solubilidad de Xe en octanol (medida como función de la temperatura en el intervalo 283-323 K) a partir de la teoría de partículas escaladas.

La extensión del formalismo de Pierotti a solutos con formas más complejas se puede realizar a partir de la expresión de Claverie [125], donde la energía libre de cavitación asociada a cada átomo de la molécula se pondera por un factor proporcional a la área expuesta al solvente de ese átomo:

$$\Delta G_{cav} = \sum_i \frac{S_i}{4\pi R_i^2} \Delta G_{cav}(R_i) \quad (2.70)$$

2.4.4.2. Dispersión-repulsión

En métodos basados en un tratamiento discreto de las moléculas estas contribuciones se evalúan típicamente a partir de potenciales de par que relacionan fragmentos químicos (átomos, enlaces, grupos) de las moléculas de soluto y solvente, tal como se ha explicado en la sección 2.3.2.2 para los campos de fuerzas clásicos. Para un soluto (M) rodeado de moléculas de solvente (S), la energía de dispersión-repulsión promediada se puede determinar a partir de la ecuación 2.71, donde N_S es el número de fragmentos de tipo s en la molécula de solvente, n_s es la densidad macroscópica del solvente, y g_{ms} es una función de correlación entre los fragmentos químicos m y s :

$$\langle E_{dis-rep} \rangle = n_s \sum_{s \in S} N_S \sum_{m \in M} \int U_{ms}(r_{ms}) g_{ms}(r_{ms}) dr_{ms}^3 \quad (2.71)$$

Los potenciales de par se suelen expresar como expansiones truncadas en potencias de $1/r$:

$$U_{ms} = \sum_k \alpha_{ms,k} r_{ms}^{-k} \quad (2.72)$$

donde $\alpha_{ms,k}$ son los coeficientes de la expansión en serie, k denota los diferentes términos considerados en la expansión truncada y r_{ms} es la distancia entre los fragmentos m y s . Los

términos considerados y sus respectivos coeficientes dependen del campo de fuerzas adoptado. Alternativamente, para el término repulsivo también se usan potenciales de par de tipo exponencial, más adecuados para describir esta interacción:

$$U_{rep,ms} = c_{ms}e^{-\gamma_{ms}r_{ms}} \quad (2.73)$$

Las integrales que aparecen en la ecuación 2.71 se pueden limitar a valores r_{ms} mayores que un valor mínimo, de forma que alrededor de cada centro m habrá una región donde no se pueden encontrar átomos de solvente s . Esta región se puede asimilar para cada fragmento de solvente con la cavidad σ_s , formada por esferas de van der Waals centradas en los núcleos del soluto y aumentadas para tener en cuenta el radio del fragmento de solvente s .

Una vez conocido el conjunto de cavidades σ_s , podemos reemplazar las integrales de la ecuación 2.71 por integrales de superficie. Para realizar esta transformación necesitamos definir unas funciones auxiliares vectoriales $A_{ms}(r_{ms})$ que cumplen:

$$\vec{\nabla} A_{ms}(r_{ms}) = U_{ms}g_{ms}(r_{ms}) \quad (2.74)$$

de forma que la ecuación 2.71 se convierte en:

$$\langle E_{dis-rep} \rangle = n_s \sum_{s \in S} N_S \sum_{m \in M} \int_{\sigma_s} A_{ms} n_\sigma d\sigma \quad (2.75)$$

donde n_σ es el vector normal exterior a la superficie σ_s en la posición σ .

Las funciones de distribución $g_{ms}(r_{ms})$ son en principio desconocidas y se han propuesto varias aproximaciones para resolver la ecuación 2.74. Una consiste en reemplazar las funciones $g_{ms}(r_{ms})$ por factores calibrados para un soluto a partir de medidas experimentales o derivadas de métodos teóricos, como simulaciones moleculares discretas. La manera más sencilla de definir estas funciones es, sin embargo, la basada en la “aproximación uniforme” [126], según la cual $g_{ms}(r_{ms})$ vale 0 dentro de la cavidad y 1 fuera de ella. Esto permite derivar expresiones sencillas analíticas para $A_{ms}(r_{ms})$.

Otra estrategia muy extendida y más sencilla que las vistas hasta ahora para calcular la contribución de dispersión-repulsión se basa en asumir que esta contribución tiene una dependencia lineal con la superficie de la cavidad (ec. 2.76), lo cual es un caso límite de las ecuaciones anteriores. De hecho, estudios previos han mostrado una relación lineal entre valores de dispersión-repulsión calculados a partir de la ecuación 2.75 dentro de la “aproximación uniforme” y la superficie de van der Waals de hidrocarburos [127]. También se han encontrado relaciones lineales entre las energías libres de solvatación de hidrocarburos y su área accesible

al solvente (SASA, *solvent accessible surface area*) [128].

$$\Delta G_{dis-rep} = \sum_k \zeta_k A_k \quad (2.76)$$

donde A_k denota la contribución del átomo k a la superficie de la cavidad y ζ_k su tensión superficial.

La expresión simplificada de la ecuación 2.76 se beneficia del hecho que las fuerzas de dispersión decaen según r^{-6} , lo que facilita que la contribución de un átomo sea menos dependiente de la naturaleza y la distribución espacial de átomos cercanos.

Finalmente, se debe distinguir entre las tensiones superficiales atómicas microscópicas usadas en modelos continuos y la tensión superficial macroscópica del solvente. Los parámetros ζ_k se derivan típicamente a partir de ajustes a energías libres de solvatación experimentales. De esta forma, además de las interacciones de dispersión-repulsión entre el soluto y el solvente, también tienen que corregir cualquier error sistemático en el tratamiento electrostático así como otros efectos relacionados con la interacción del soluto con la primera esfera de solvatación.

2.4.5. La cavidad: parametrización de modelos continuos

Un factor clave en cualquier modelo continuo es la definición de la interfase soluto-solvente, ya que esta modula de forma decisiva la contribución electrostática a la energía libre de solvatación. Generalmente, las cavidades se construyen a partir de la superposición de esferas centradas en los átomos cuyo tamaño se determina a partir de radios atómicos estándar [129, 130, 131, 132, 133]. Sin embargo, también se han propuesto otras estrategias, como el uso de radios atómicos variables, cuyo valor depende del entorno molecular o de la distribución de carga del soluto [134, 135, 136], el análisis de funciones de distribución radial obtenidas a partir de simulaciones discretas [137, 138] o la elección de un determinado contorno de isodensidad electrónica [139, 140, 141]. Finalmente, aunque la mayoría de métodos SCRF usan la misma cavidad para cualquier solvente, también se han considerado cavidades adaptadas al solvente [142].

El desarrollo de cualquier modelo continuo implica necesariamente un cierto grado de arbitrariedad, que principalmente incluye la elección de la cavidad del soluto y el tratamiento a las contribuciones no electrostáticas a la energía libre de solvatación. Este requiere un adecuado esfuerzo de parametrización, que se consigue principalmente mediante el ajuste entre energías libres de solvatación experimentales y calculadas, pero se extiende también a propiedades como efectos solvatocrómicos, variaciones en el equilibrio químico inducidas por el solvente, o

propiedades específicas de solutos en solución [46, 50, 53, 58, 63]. A continuación se describirán brevemente los aspectos principales que involucran la parametrización de los modelos continuos de solvatación que se han usado en el presente trabajo.

2.4.5.1. El modelo MPE

La versión utilizada en el presente trabajo del modelo MPE, parametrizada para tratar la solvatación en agua a nivel Hartree-Fock con bases 6-31G(d), utiliza cavidades construidas a partir de la superposición de esferas centradas en los núcleos usando los radios de Bondi escalados por un factor de 1.3, tal y como se ha explicado en la sección 2.4.3.1. Esta elección se basa en obtener un volumen coherente con la teoría de Onsager de los líquidos polares [143]. En el caso de líquidos puros, el volumen de Onsager es igual al volumen promedio de una molécula en el líquido. El factor 1.3 se obtuvo a partir de ajustes con el volumen molecular experimental para un conjunto de líquidos de referencia.

Los términos no electrostáticos, tanto cavitación como dispersión-repulsión, se describen conjuntamente mediante una relación lineal con la contribución atómica a la área superficial (ec. 2.77). Las tensiones superficiales atómicas γ_k se obtienen a partir del ajuste a valores experimentales de energías libres de solvatación para un conjunto de moléculas de referencia.

$$\Delta G_{n-ele} = \sum_k \gamma_k A_k \quad (2.77)$$

Finalmente cabe señalar que el modelo MPE, a diferencia de los modelos SMx y MST que se explicarán a continuación, adopta la misma cavidad tanto para los términos electrostáticos como no electrostáticos.

2.4.5.2. El modelo SMx

Los modelos SMx se han visto sometidos a un intenso proceso de parametrización por el grupo de Minnesota. Además de los Hamiltonianos semiempíricos AM1 y PM3, el método SM5.42R [88] (donde 42 indica el uso de cargas de clase IV CM2, explicadas en la sección 2.4.3.2, y R el uso de geometrías optimizadas en fase gas para el soluto) ha sido parametrizado tanto a nivel *ab initio* HF como para métodos de la teoría del funcional de la densidad para diversos conjuntos de funciones de base, considerando un gran número de solventes orgánicos (incluyendo agua, alcanos, cicloalcanos, alcoholes, cetonas, éteres, etc.). En total, la parametrización se realizó considerando 2135 energías libres de solvatación para 275 solutos neutros en 91 solventes, y para 49 iones en agua.

Los términos no electrostáticos, típicamente indicados por CDS (*cavitation-dispersion-solvent structure effects*), incluyen efectos de cavitación, dispersión y efectos en la estructura del solvente, y se determinan mediante una expresión que relaciona el área accesible al solvente de los átomos del soluto (ec. 2.78). Las tensiones superficiales atómicas, σ_k , se definen a partir de una función multilineal de coeficientes semiempíricos de tensión superficial, $\tilde{\sigma}_{km}$ (ec. 2.79), donde los valores de T_m dependen de las distancias interatómicas con los átomos a distancias enlazantes y a veces en el intervalo de distancias típicas de interacciones geminales. Además, dentro del llamado modelo de solvatación universal [144, 145, 146], también dependen de un pequeño número de descriptores del solvente.

$$\Delta G_{CDS} = \sum_k \sigma_k A_k \quad (2.78)$$

$$\sigma_k = \sum_m \tilde{\sigma}_{km} T_m \quad (2.79)$$

El modelo SM5, a diferencia del modelo MPE, adopta cavidades distintas para el término electrostático y no electrostático. Los radios de las esferas en la parte electrostática son constantes empíricas llamadas radios intrínsecos de Coulomb, mientras que para las fuerzas de corto alcance se usan los radios atómicos de Bondi [147].

2.4.5.3. El modelo MST

Este modelo usa la aproximación PCM [97, 96] descrita en la sección 2.4.3.3 para el cálculo de la componente electrostática a la energía libre de solvatación y se ha parametrizado para varios solventes a nivel *ab initio* HF/6-31G(d) y semiempíricos AM1 y PM3 [148, 149, 150, 151]. Anteriormente había sido parametrizado en agua, cloroformo y tetracloruro de carbono. En el presente trabajo, se han reparametrizado estas versiones y se ha extendido el método al solvente n-octanol.

El método MST calcula las contribuciones no electrostáticas a partir de la suma de un término de cavitación y otro de van der Waals. La componente de cavitación se calcula mediante la expresión de Claverie-Pierotti (ec. 2.70), y el término de van der Waals se calcula usando una relación lineal con la superficie atómica (ec. 2.76) definiendo tensiones superficiales para los átomos. Las cavidades se construyen a partir de esferas centradas en los núcleos con el algoritmo GEPOL [152, 153, 154]. En la parametrización que se presenta en este trabajo se ha adoptado una estrategia dual, de forma que las contribuciones no electrostáticas se calculan a partir de superficies de van der Waals construidas a partir de radios de Pauling, mientras que el término electrostático se determina a partir de una cavidad creada escalando los radios atómicos por un factor dependiente del solvente. El uso de cavidades dependientes del solvente

se justifica por la necesidad de tener en cuenta la posición diferente en que se encuentra la primera esfera de solvatación en solventes distintos [148, 149, 150, 151], determinada a partir de la comparación de resultados MST con otros obtenidos a partir de simulaciones discretas del soluto en solución usando campos de fuerzas clásicos y a partir de cálculos mixtos de mecánica cuántica-mecánica molecular. Para iones, tal como se describe en uno de los trabajos que se presentan, se ha encontrado necesario reducir el tamaño de la cavidad por un factor de aproximadamente 0.9 [155], debido al hecho que las moléculas de solvente se acercan más a solutos cargados que neutros.

2.4.6. Hacia modelos de solvatación fraccionales

La partición de energías libres de solvatación en contribuciones fragmentales asociadas a átomos o grupos ha sido objeto de un amplio esfuerzo de investigación debido a dos razones principales. En primer lugar, esto permite obtener una visión más profunda de las bases fisicoquímicas de la interacción soluto-solvente, que puede ser valiosa para analizar procesos tales como el “efecto hidrofóbico” [156, 157, 158]. En segundo lugar, nos permite obtener una descripción compacta de las propiedades de solvatación de moléculas para establecer relaciones estructura-actividad [159, 160, 161].

Las contribuciones fraccionales a la energía libre de solvatación se pueden estimar cuantitativamente usando métodos sencillos, como el cálculo de superficies hidrofóbicas y hidrofílicas, o el uso de contribuciones fragmentales empíricas. En particular, esta última aproximación a dado lugar a una larga serie de métodos desarrollados para predecir propiedades fisicoquímicas relevantes como el coeficiente de partición octanol-agua [162, 163, 164]. Típicamente, estos métodos explotan bases de datos de contribuciones asignadas a átomos o grupos específicos, y suelen incluir una serie de reglas que permiten corregir los efectos debido a la naturaleza no aditiva de las contribuciones fragmentales. En este contexto, el desarrollo de métodos de partición fraccional basados en modelos físicos pueden ser extremadamente útiles para el refinamiento de esos modelos empíricos, porque: i) las contribuciones se calculan considerando de forma explícita el entorno molecular, la conformación, y los estados de ionización y tautoméricos, y ii) las contribuciones fragmentales se extraen directamente a partir de propiedades físicas reales. En el contexto de los modelos continuos de solvatación se han desarrollado métodos fraccionales para modelos GB por Cramer y Truhlar [165, 166, 167] y por nuestro grupo en el modelo MST [168, 169, 170, 171].

En el modelo MST, la partición de los términos no electrostáticos es directa, ya que están relacionadas con la superficie accesible al solvente de los átomos (ec. 2.70 y 2.76). La partición del término no electrostático es más complicada debido a problemas relacionados con la partición de la distribución de carga del soluto. De todas formas, este problema se puede resolver

a través de la teoría de perturbaciones, que permite reescribir la componente electrostática a la energía libre de solvatación como [172, 173, 174, 175]:

$$\Delta G_{ele} = \left\langle \Psi^0 \left| \frac{1}{2} V^{sol} \right| \Psi^0 \right\rangle \quad (2.80)$$

donde V^{sol} es el campo de reacción generado por el soluto totalmente polarizado en solución, cuya función de onda se indicaría por Ψ^{sol} .

La ecuación 2.80 nos permite determinar la componente electrostática a partir de la adición de contribuciones fraccionales (ec. 2.81). Cabe señalar, además, que mediante este esquema de partición, la contribución del átomo i ($\Delta G_{ele,i}$) incluye la interacción de toda la molécula con los elementos de superficie j que pertenecen a la superficie expuesta al solvente del átomo i .

$$\Delta G_{ele} = \sum_{i=1}^N \Delta G_{ele,i} = \sum_{i=1}^N \sum_{j \in i} \left\langle \Psi^0 \left| \frac{1}{2} \frac{q_j^{sol}}{|r_j - r|} \right| \Psi^0 \right\rangle \quad (2.81)$$

El uso de métodos semiempíricos basados en la aproximación NDDO permiten además subdividir la contribución de un átomo i entre una componente propia (*self*) y otra cruzada (*cross*) (ec. 2.82-2.84), tal y como se propone en uno de los estudios presentados en esta tesis. El término propio incluye la interacción electrostática de la carga del core (Z_i) y los elementos $\phi_\mu \phi_\nu$ de la distribución de carga electrónica del átomo i con las cargas del campo de reacción del solvente localizadas en la superficie del átomo i (ec. 2.83). El término cruzado incluye la interacción de las cargas de core (Z_k) y los elementos $\phi_\mu \phi_\nu$ de la distribución de carga electrónica del resto de átomos con las cargas aparentes del solvente en la superficie del átomo i (ec. 2.84).

$$\Delta G_{ele,i} = \Delta G_{ele,i}^{self} + \Delta G_{ele,i}^{cross} \quad (2.82)$$

$$\Delta G_{ele,i}^{self} = \frac{1}{2} \sum_{j \in i} \left[\left(\sum_{\mu \in i} \sum_{\nu \in i} P_{\mu\nu} \left\langle \chi_\mu \left| \frac{q_j}{|r_j - r|} \right| \chi_\nu \right\rangle \right) + Z_i \frac{q_j}{|r_j - r|} \right] \quad (2.83)$$

$$\Delta G_{ele,i}^{cross} = \frac{1}{2} \sum_{j \in i} \sum_{k \neq i} \left[\left(\sum_{\mu \in k} \sum_{\nu \in k} P_{\mu\nu} \left\langle \chi_\mu \left| \frac{q_j}{|r_j - r|} \right| \chi_\nu \right\rangle \right) + Z_k \frac{q_j}{|r_j - r|} \right] \quad (2.84)$$

donde χ representa la base de orbitales atómicos, y P es la matriz densidad monoeléctronica.

A pesar de la arbitrariedad implícita en cualquier método fraccional, el presente esquema descrito en esta sección proporciona resultados similares y consistentes con otros esquemas de partición alternativos definidos en el contexto de métodos continuos y discretos [176]. Esto nos da confianza en las contribuciones fraccionales obtenidas a partir de este método y hace que sea útil para derivar parámetros destinados a estudios de relaciones estructura-actividad. Estos esquemas de particionamiento se han aplicado al estudio de la transferibilidad de las

contribuciones fragmentales al coeficiente de partición octanol-agua, $\log P_{ow}$, en dos de los trabajos presentados en esta tesis.

La visualización gráfica de las contribuciones fraccionales a la energía libre de solvatación permite obtener un dibujo 3D de las propiedades de solvatación de las moléculas [168, 169, 170, 171], y esto facilita la comparación de compuestos basándonos en la distribución de sus regiones hidrofóbicas/hidrofílicas. Por ejemplo, se ha encontrado muy útil para estudios de relaciones estructura-actividad el uso de “dipolos de solvatación” [169, 177, 178] como descriptores simplificados del perfil de solvatación de una molécula:

$$\mu_{sol} = \sum_i^N \Delta G_{sol,i} (r_i - r_0) - \frac{\Delta G_{sol}}{N} \sum_{i=1}^N (r_i - r_0) \quad (2.85)$$

donde r_0 es un origen arbitrario, típicamente el centro de masas de la molécula.

La comparación entre dos dipolos de solvatación se puede hacer fácilmente a partir del producto escalar de los vectores unitarios (ec. 2.86), obteniéndose así un índice de similitud muy compacto que varía entre -1 y 1.

$$\vartheta_{wz} = \frac{\mu_{sol}^w}{|\mu_{sol}^w|} \cdot \frac{\mu_{sol}^z}{|\mu_{sol}^z|} \quad (2.86)$$

Una comparación más detallada entre las distribuciones fraccionales de las dos moléculas (w con N_w átomos, y z con N_z átomos) se puede obtener a partir de funciones de similitud como la mostrada en la ecuación 2.87 [171]. Como la función Λ_{wz} depende de la orientación relativa de w y z , la orientación óptima se puede obtener maximizando Λ_{wz} a través de rutinas de optimización del gradiente o mediante otros procedimientos basados, por ejemplo, en simulaciones de Monte Carlo.

$$\Lambda_{wz}(R_{wz}) = \sum_{i=1}^{N_w} \sum_{l=1}^{N_z} \frac{\Delta G_{sol,i} \Delta G_{sol,l}}{(r_{il}^n + \delta)} \quad (2.87)$$

donde n es un parámetro ajustable que controla la forma de la función de similitud y δ es una constante que evita la aparición de singularidades en la función.

El índice de similitud se puede determinar usando, por ejemplo, un índice de tipo Carbó [179] como el indicado en la ecuación 2.88, donde Λ_{ww} y Λ_{zz} indican los índices de autosimilitud determinados a partir de la ecuación 2.87 cuando $w = z$ y $R_{ww} = R_{zz} = 0$.

$$\gamma_{wz} = \frac{\Lambda_{wz}}{\sqrt{\Lambda_{ww} \Lambda_{zz}}} \quad (2.88)$$

La utilidad del índice de similitud hidrofóbico en el contexto de las relaciones estructura-actividad, biodisponibilidad de fármacos y reconocimiento molecular se ha investigado recientemente [169, 170, 171]. Los resultados sugieren que esta técnica puede ser una herramienta potente para el desarrollo racional de moléculas bioactivas.

Bibliografía

- [1] A. R. Leach. *Molecular Modelling. Principles and Applications*. Longman, Singapore, 1996.
- [2] C. J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. Wiley, Chichester, 2002.
- [3] J. Andrés and J. Bertrán, editors. *Química Teórica y Computacional*. Publicaciones de la Universitat Jaume I, Castellón de la Plana, 2000.
- [4] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. McGraw-Hill, New York, 1982.
- [5] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, Oxford, 1989.
- [6] W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople. *Ab Initio Molecular Orbital Theory*. Wiley, New York, 1986.
- [7] J. J. P. Stewart. *Reviews in Computational Chemistry*, volume 1, page 45. VCH, New York, 1990.
- [8] M. C. Zerner. *Reviews in Computational Chemistry*, volume 2, page 313. VCH, New York, 1991.
- [9] J. C. Slater. *Quantum Theory of Matter*. McGraw-Hill, New York, 1968.
- [10] C. C. J. Roothaan. *Rev. Mod. Phys.*, 23:69, 1951.
- [11] J. C. Slater. *Phys. Rev.*, 36:57, 1930.
- [12] S. F. Boys. *Proc. R. Soc. (London)*, 200:542, 1950.
- [13] W. J. Hehre, R. F. Stewart, and J. A. Pople. *J. Chem. Phys.*, 51:2657, 1969.
- [14] J. A. Popel, M. Head-Gordon, and K. Raghavachari. *J. Chem. Phys.*, 87:5968, 1987.
- [15] C. Møller and M. S. Plesset. *Phys. Rev.*, 46:618, 1934.

- [16] J. A. Pople, D. P. Santry, and G. A. Segal. *J. Chem. Phys.*, 43:S129, 1965.
- [17] J. A. Pople and G. A. Segal. *J. Chem. Phys.*, 43:S136, 1965.
- [18] M. J. S. Dewar and W. Thiel. *J. Am. Chem. Soc.*, 99:4899, 1977.
- [19] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart. *J. Am. Chem. Soc.*, 107:3902, 1985.
- [20] J. J. P. Stewart. *J. Comput. Chem.*, 10:209, 1989.
- [21] P. Hohenberg and W. Kohn. *Phys. Rev. B*, 136:864, 1964.
- [22] W. Kohn and L. J. Sham. *Phys. Rev. A*, 140:1133, 1965.
- [23] A. D. Beck. *J. Chem. Phys.*, 98:5648, 1993.
- [24] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785, 1988.
- [25] K. Morokuma. *J. Chem. Phys.*, 55:1236, 1971.
- [26] K. Kitaura and K. Morokuma. *Int. J. Quant. Chem.*, 10:325, 1976.
- [27] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, M. D. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. Kollman. *J. Am. Chem. Soc.*, 117:5179, 1995.
- [28] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. *J. Phys. Chem. B*, 102:3586, 1998.
- [29] A. D. MacKerell, J. Wiórkiewicz-Kuczera, and M. Karplus. *J. Am. Chem. Soc.*, 117:11946, 1995.
- [30] W. F. van Gunsteren, H. J. C. Berendsen, J. Hermans, W. G. J. Hol, and J. P. M. Postma. *Proc. Natl. Acad. Sci. USA*, 80:4315, 1983.
- [31] W. L. Jorgensen and J. Tirado-Rives. *J. Am. Chem. Soc.*, 110:1657, 1988.
- [32] J. Pranata, S. G. Wierschke, and W. L. Jorgensen. *J. Am. Chem. Soc.*, 113:2810, 1991.
- [33] N. L. Allinger, Y. H. Yuh, and J. H. Lii. *J. Am. Chem. Soc.*, 111:8551, 1989.
- [34] J. H. Lii and N. L. Allinger. *J. Am. Chem. Soc.*, 111:8566, 1989.
- [35] J. H. Lii and N. L. Allinger. *J. Am. Chem. Soc.*, 111:8576, 1989.

- [36] T. A. Halgren and W. Damm. *Curr. Op. Struct. Biol.*, 11:236, 2001.
- [37] J. Caldwell and P. A. Kollman. *J. Am. Chem. Soc.*, 117:4177, 1995.
- [38] H. Minoux and C. Chipot. *J. Am. Chem. Soc.*, 121:10366, 1999.
- [39] E. Cubero, F. J. Luque, and M. Orozco. *Proc. Natl. Acad. Sci. USA*, 95:5976, 1998.
- [40] J. Applequist, J. R. Carl, and K. Fung. *J. Am. Chem. Soc.*, 94:2952, 1972.
- [41] P. Cieplak, J. Caldwell, and P. Kollman. *J. Comput. Chem.*, 22:1048, 2001.
- [42] S. W. Rick, S. J. Stuart, and B. J. Berne. *J. Chem. Phys.*, 101:6141, 1994.
- [43] S. Patel, A. D. MacKerell Jr., and C. L. Brooks III. *J. Comput. Chem.*, 25:1504, 2004.
- [44] G. Lamoureux and B. Roux. *J. Chem. Phys.*, 119:3025, 2003.
- [45] G. Lamoureux, Jr. A. D. MacKerell, and B. Roux. *J. Chem. Phys.*, 119:5185, 2003.
- [46] M. Orozco and F. J. Luque. *Chem. Rev.*, 100:4187, 2000.
- [47] W. L. Jorgensen. *Acc. Chem. Res.*, 22:184, 1989.
- [48] A. Warshel. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. Wiley, New York, 1991.
- [49] W. F. van Gunsteren, F. J. Luque, D. Timms, and A. E. Torda. *Annu. Rev. Biophys. Biomol. Struct.*, 23:847, 1994.
- [50] C. J. Cramer and D. G. Truhlar, editors. *Structure and Reactivity in Aqueous Solution*, page 24. ACS Symposium Series 568, American Chemical Society, Washington, D.C., 1994.
- [51] P. A. Kollman. *Acc. Chem. Res.*, 29:461, 1996.
- [52] J. Gao. *Acc. Chem. Res.*, 29:298, 1996.
- [53] O. Tapia and J. Bertrán, editors. *Solvent Effects and Chemical Reactivity*. Kluwer, Dordrecht, 1996.
- [54] M. F. Ruiz-López and J. L. Rivail. *Encyclopedia of Computational Chemistry*, volume 1, page 437. Wiley, New York, 1998.
- [55] G. Monard and K. M. Merz. *Acc. Chem. Res.*, 32:904, 1999.
- [56] S. Ten-no, F. Hirata, and S. Kato. *Chem. Phys. Lett.*, 214:391, 1993.
- [57] L. Shao, H. Yu, and J. Gao. *J. Phys. Chem. A*, 102:10366, 1998.

- [58] J. Tomasi and M. Perisco. *Chem. Rev.*, 94:2027, 1994.
- [59] J. L. Rivail and D. Rinaldi. *Computational Chemistry. Reviews of Current Trends*, page 139. World Scientific, Singapore, 1995.
- [60] C. J. Cramer and D. G. Truhlar. *Reviews in Computational Chemistry*, volume 6, page 1. VCH, New York, 1995.
- [61] M. Orozco, C. Alhambra, X. Barril, J. M. López, M. A. Busquets, and F. J. Luque. *J. Mol. Model.*, 2:1, 1996.
- [62] J. Tomasi, B. Menucci, R. Cammi, and M. Cossi. *Theoretical Aspects of Biochemical Reactivity*. Kluwer, Dordrecht, 1997.
- [63] C. J. Cramer and D. G. Truhlar. *Chem. Rev.*, 99:2161, 1999.
- [64] A. Ben-Naim. *Solvation Thermodynamics*. Plenum, New York, 1987.
- [65] J. Tomasi and B. Menucci. *Encyclopedia of Computational Chemistry*, page 53. Wiley, Chichester, 1998.
- [66] C. Amovilli and B. Menucci. *J. Phys. Chem. B*, 101:1051, 1997.
- [67] M. Z. Born. *Phys.*, 1:45, 1920.
- [68] G. Tanford and J. G. Kirkwood. *J. Am. Chem. Soc.*, 79:533, 1957.
- [69] R. P. Bell. *Trans. Faraday Soc.*, 27:797, 1931.
- [70] L. Onsager. *J. Am. Chem. Soc.*, 58:1486, 1936.
- [71] J. Hylton, R. E. Christoffersen, and G. G. Hall. *Chem. Phys. Lett.*, 26:501, 1974.
- [72] O. Tapia and O. Goscinski. *Mol. Phys.*, 29:1653, 1975.
- [73] D. Rinaldi and J. L. Rivail. *Theor. Chim. Acta*, 32:57, 1973.
- [74] J. L. Rivail and D. Rinaldi. *Chem. Phys.*, 18:233, 1976.
- [75] S. Yomosa. *J. Phys. Soc. Japan*, 35:1738, 1973.
- [76] S. Yomosa. *J. Phys. Soc. Japan*, 36:1655, 1974.
- [77] A. R. Katritzky, M. C. Zerner, and M. M. Karelson. *J. Am. Chem. Soc.*, 108:7213, 1986.
- [78] H. Agren, C. Medina-Llanos, and K. V. Mikkelsen. *Chem. Phys.*, 115:43, 1987.
- [79] M. W. Wong, M. J. Frisch, and K. B. Wiberg. *J. Am. Chem. Soc.*, 113:4776, 1991.

- [80] C. Curutchet, C. J. Cramer, D. G. Truhlar, M. F. Ruiz-López, D. Rinaldi, M. Orozco, and F. J. Luque. *J. Comput. Chem.*, 24:284, 2003.
- [81] D. Rinaldi, M. F. Ruiz-López, and J. L. Rivail. *J. Chem. Phys.*, 78:834, 1983.
- [82] V. Dillet, D. Rinaldi, J. G. Angyan, and J. L. Rivail. *Chem. Phys. Lett.*, 202:18, 1993.
- [83] J. L. Rivail. *CR Acad. Sci. Paris*, 307:311, 1990.
- [84] J. G. Angyan. *Chem. Phys. Lett.*, 241:51, 1995.
- [85] M. F. Ruiz-López, F. Bohr, M. T. C. Martins Costa, and D. Rinaldi. *Chem. Phys. Lett.*, 221:109, 1994.
- [86] V. Dillet, D. Rinaldi, and J. L. Rivail. *J. Phys. Chem.*, 98:5034, 1994.
- [87] V. Dillet, D. Rinaldi, J. Bertrán, and J. L. Rivail. *J. Chem. Phys.*, 104:9437, 1996.
- [88] C. C. Chamber, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem.*, 100:16385, 1996.
- [89] R. S. Mulliken. *J. Chem. Phys.*, 23:1933, 1955.
- [90] J. W. Storer, D. J. Giesen, C. J. Cramer, and D. G. Truhlar. *J. Comput-Aided Mol. Des.*, 9:87, 1995.
- [91] J. Li, T. Zhu, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. A*, 102:1820, 1998.
- [92] J. Li, B. Williams, C. J. Cramer, and D. G. Truhlar. *J. Chem. Phys.*, 110:724, 1999.
- [93] J. Li, J. Xing, C. J. Cramer, and D. G. Truhlar. *J. Chem. Phys.*, 111:885, 1999.
- [94] J. Li, T. Zhu, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. B*, 104:2178, 2000.
- [95] P.O. Löwdin. *J. Chem. Phys.*, 18:365, 1950.
- [96] S. Miertus, E. Scrocco, and J. Tomasi. *Chem. Phys.*, 55:117, 1981.
- [97] S. Miertus and J. Tomasi. *Chem. Phys.*, 65:239, 1982.
- [98] R. Cammi and J. Tomasi. *J. Comput. Chem.*, 16:1449, 1995.
- [99] R. J. Zauhar. *J. Comput. Chem.*, 12:575, 1991.
- [100] R. Cammi, M. Cossi, and J. Tomasi. *J. Chem. Phys.*, 104:4611, 1996.
- [101] R. Cammi, M. Cossi, B. Menucci, C. S. Pomelli, and J. Tomasi. *Int. J. Quant. Chem.*, 60:1165, 1996.

- [102] M. Cossi, J. Tomasi, and R. Cammi. *Int. J. Quant. Chem.*, 29:695, 1995.
- [103] B. Menucci and J. Tomasi. *J. Chem. Phys.*, 106:5151, 1997.
- [104] M. Cossi, B. Menucci, J. Pitarch, and J. Tomasi. *J. Comput. Chem.*, 19:833, 1998.
- [105] E. Cancés, B. Menucci, and J. Tomasi. *J. Chem. Phys.*, 107:3032, 1997.
- [106] B. Menucci, E. Cancés, and J. Tomasi. *J. Phys. Chem. B*, 101:10506, 1997.
- [107] B. Menucci, M. Cossi, and J. Tomasi. *J. Phys. Chem.*, 100:1807, 1996.
- [108] B. Menucci, M. Cossi, and J. Tomasi. *J. Chem. Phys.*, 102:6837, 1995.
- [109] F. J. Olivares del Valle and J. Tomasi. *Chem. Phys.*, 150:139, 1991.
- [110] F. J. Olivares del Valle and M. A. Aguilar. *J. Comput. Chem.*, 13:115, 1992.
- [111] F. J. Olivares del Valle, M. A. Aguilar, and S. Tolosa. *J. Mol. Struct. (TEOCHEM)*, 279:223, 1993.
- [112] M. A. Aguilar, F. J. Olivares del Valle, and J. Tomasi. *J. Chem. Phys.*, 98:7375, 1993.
- [113] A. Fortunelli. *J. Mol. Struct. (TEOCHEM)*, 357:117, 1995.
- [114] J. C. Contador, M. A. Aguilar, and F. J. Olivares del Valle. *Chem. Phys.*, 214:113, 1997.
- [115] R. Cammi, B. Menucci, and J. Tomasi. *J. Phys. Chem. A*, 103:9100, 1999.
- [116] A. Klamt and G. J. Schüürmann. *J. Chem. Soc., Perkin Trans 2*, 5:799, 1993.
- [117] A. Klamt and V. Jonas. *J. Chem. Phys.*, 92:9972, 1996.
- [118] A. Klamt. *J. Phys. Chem.*, 99:2224, 1995.
- [119] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz. *J. Phys. Chem. A*, 102:5074, 1998.
- [120] I. Tuñón, E. Silla, and J. L. Pascual-Ahuir. *Chem. Phys. Lett.*, 203:289, 1993.
- [121] T. Simonson and A. T. Brünger. *J. Phys. Chem.*, 98:4683, 1994.
- [122] O. J. Sinanoglu. *Chem. Phys.*, 75:463, 1981.
- [123] V. Gogonea and E. J. Osawa. *Mol. struct. (TEOCHEM)*, 311:305, 1994.
- [124] R. A. Pierotti. *Chem. Rev.*, 76:717, 1976.
- [125] P. Claverie. *Intermolecular Interactions: From Diatomics to Biomolecules*. Wiley, Chichester, 1978.

- [126] F. M. Floris and J. Tomasi. *J. Comput. Chem.*, 10:616, 1989.
- [127] F. M. Floris, J. Tomasi, and J. L. J. Pascual-Ahuir. *J. Comput. Chem.*, 12:784, 1991.
- [128] D. Sitkoff, K. A. Sharp, and B. Honig. *J. Phys. Chem.*, 98:1978, 1994.
- [129] A. A. Rashin and K. Namboodiri. *J. Phys. Chem.*, 91:6003, 1987.
- [130] G. E. Chudinov, D. V. Napolov, and M. V. Basilevsky. *Chem. Phys.*, 160:41, 1992.
- [131] A. A. Rashin, M. A. Bukatin, J. Andzelm, and A. T. Hagler. *Biophys. Chem.*, 51:375, 1994.
- [132] E. V. Stefanovich and T. N. Truong. *Chem. Phys. Lett.*, 244:65, 1995.
- [133] J. R. Pliego Jr. and J. M. Riveros. *Chem. Phys. Lett.*, 355:543, 2002.
- [134] M. A. Aguilar, M. A. Martín, S. Tolosa, and F. J. Olivares del Valle. *J. Mol. Struct (TEOCHEM)*, 43:313, 1988.
- [135] M. A. Aguilar and F. J. Olivares del Valle. *Chem. Phys.*, 129:439, 1989.
- [136] P. F. B. Gonçalves and P. R. Livotto. *Chem. Phys. Lett.*, 304:438, 1999.
- [137] M. Orozco, F. J. Luque, D. Habibollahzadeh, and J. Gao. *J. Chem. Phys.*, 102:6145, 1995.
- [138] P. F. B. Gonçalves and H. Stassen. *J. Comput. Chem.*, 23:706, 2002.
- [139] J. B. Foresman, T. A. Keith, K. B. Wiberg, and M. J. Frisch. *J. Phys. Chem.*, 100:16098, 1996.
- [140] C. Kölle and K. Jug. *J. Comput. Chem.*, 18:1, 1997.
- [141] K. Jug, C. Kölle, and F. Neumann. *J. Comput. Chem.*, 20:301, 1999.
- [142] M. Orozco, M. Bachs, and F. J. Luque. *J. Comput. Chem.*, 16:563, 1995.
- [143] L. Onsager. *J. Am. Chem. Soc.*, 58:1486, 1936.
- [144] J. Li, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. *Chem. Phys. Lett.*, 288:293, 1998.
- [145] T. Zhu, J. Li, G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. *J. Chem. Phys.*, 109:9117, 1998.
- [146] J. Li, T. Zhu, G. D. Hawkins, P. Winget, D. A. Liotard, C. J. Cramer, and D. G. Truhlar. *Theor. Chem. Acc.*, 103:9, 1999.

- [147] A. Bondi. *J. Phys. Chem.*, 68:441, 1964.
- [148] M. Bachs, F. J. Luque, and M. Orozco. *J. Comput. Chem.*, 15:446, 1994.
- [149] F. J. Luque, C. Alemán, M. Bachs, and M. Orozco. *J. Comput. Chem.*, 17:806, 1996.
- [150] F. J. Luque, Y. Zhang, C. Alemán, M. Bachs, J. Gao, and M. Orozco. *J. Phys. Chem.*, 100:4269, 1996.
- [151] C. Curutchet, M. Orozco, and F. J. Luque. *J. Comput. Chem.*, 22:1180, 2001.
- [152] J. L. Pascual-Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi. *J. Comput. Chem.*, 8:778, 1987.
- [153] J. L. Pascual-Ahuir and E. Silla. *J. Comput. Chem.*, 11:1047, 1990.
- [154] E. Silla, I. Tuñón, and J. L. Pascual-Ahuir. *J. Comput. Chem.*, 9:1077, 1991.
- [155] M. Orozco and F. J. Luque. *Chem. Phys.*, 182:237, 1994.
- [156] C. Tanford. *The Hydrophobic Effect*. Wiley, New York, 1980.
- [157] T. P. Straatsma, H. J. C. Berendsen, and J. P. M. Postma. *J. Chem. Phys.*, 85:6720, 1986.
- [158] N. T. Southall, K. A. Dill, and A. D. J. Haymet. *J. Phys. Chem. B*, 106:521, 2002.
- [159] Y. C. Martin. *Quantitative Drug Design. A Critical Introduction*. Marcel Dekker, New York, 1978.
- [160] A. L. Horvath. *Molecular Design*. Elsevier, Amsterdam, 1992.
- [161] C. Hansch and A. Leo. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*. American Chemical Society, Washington DC, 1995.
- [162] R. Mannhold and K. Dross. *Quant. Struct. Act. Relat.*, 15:403, 1996.
- [163] R. Manhold, G. Cruciani, K. Dross, and R. J. Rekker. *J. Comput-Aided Mol. Des.*, 12:573, 1998.
- [164] R. Manhold and H. J. van de Waterbeemd. *J. Comput-Aided Mol. Des.*, 15:337, 2001.
- [165] C. J. Cramer and D. G. Truhlar. *Chem. Phys. Lett.*, 198:74, 1992.
- [166] D. J. Giesen, C. C. Chamber, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. B*, 101:5084, 1997.
- [167] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. B*, 102:3257, 1998.

- [168] F. J. Luque, X. Barril, and M. Orozco. *J. Comput-Aided Mol. Des.*, 13:139, 1999.
- [169] X. Barril, J. Muñoz, F. J. Luque, and M. Orozco. *Phys. Chem. Chem. Phys.*, 2:4897, 2000.
- [170] J. Muñoz, X. Barril, F. J. Luque, J. L. Gelpí, and M. Orozco. *Fundamentals of Molecular Similarity*, page 143. Kluwer Academic/Plenum, New York, 2001.
- [171] J. Muñoz, X. Barril, B. Hernández, M. Orozco, and F. J. Luque. *J. Comput. Chem.*, 23:554, 2002.
- [172] F. J. Luque, J. M. Bofill, and M. Orozco. *J. Chem. Phys.*, 103:10183, 1995.
- [173] J. G. Angyán. *J. Chem. Phys.*, 107:1291, 1997.
- [174] F. J. Luque, J. M. Bofill, and M. Orozco. *J. Chem. Phys.*, 107:1293, 1997.
- [175] F. J. Luque and M. Orozco. *J. Phys. Chem. B*, 101:5573, 1997.
- [176] A. Morreale, J. L. Gelpí, F. J. Luque, and M. Orozco. *J. Comput. Chem.*, 24:1610, 2003.
- [177] D. Eisenberg and A. D. McLachlan. *Nature*, 319:199, 1986.
- [178] D. Eisenberg, E. Schware, M. Komaromy, and R. J. Wall. *Mol. Biol.*, 179:125, 1984.
- [179] R. Carbó, L. Leyda, and M. Arnau. *Int. J. Quant. Chem.*, 17:1185, 1980.

Capítulo 3

Resultados sobre modelos continuos de solvatación

3.1.

**Solvation in octanol: Parametrization
of the continuum MST model**

Carles Curutchet, Modesto Orozco and F. Javier Luque

Journal of Computational Chemistry **2001**, 22, 1180

Texto del artículo disponible en / Text de l'article disponible a / Available online at:

<http://www3.interscience.wiley.com/cgi-bin/abstract/83002114/ABSTRACT>

3.2.

MST continuum study of the hydration free energies of monovalent ionic species

Carles Curutchet, Axel Bidon-Chanal, Ignacio Soteras, Modesto Orozco and F. Javier Luque

Journal of Physical Chemistry B, accepted

MST Continuum Study of the Hydration Free Energies of Monovalent Ionic Species

Carles Curutchet,¹ Axel Bidon-Chanal,¹ Ignaci Soteras,¹ Modesto Orozco,^{2*}
and F. Javier Luque^{1*}

¹ Departament de Físicoquímica. Facultat de Farmàcia. Universitat de Barcelona. Avgda. Diagonal 643. Barcelona 08028. Spain.

² Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, c/. Martí i Franqués 1, 08028, Barcelona, and Unitat de Modelització Molecular i Bioinformàtica, Institut de Recerca Biomèdica. Parc Científic de Barcelona, c/. Josep Samitier 1, 08028 Barcelona, Spain.

Send correspondence to FJL (e-mail: javier@far1.far.ub.es) or MO (e-mail: modesto@mmb.pcb.ub.es)

Abstract

In this study we revisit the protocol previously proposed within the framework of the Miertus-Scrocco-Tomasi (MST) continuum model to define the cavity between solute and solvent for predicting hydration free energies of univalent ions [Orozco, M.; Luque, F. J. *Chem. Phys.* **1994**, *182*, 237]. The protocol relies on the use of a reduced cavity (around 10-15% smaller than the cavity used for neutral compounds) around the atom(s) bearing the formal charge. The suitability of this approach is examined here for a series of 47 univalent ions for which accurate experimental hydration free energies are available. Attention is also paid to the effect of the charge renormalization protocol used to correct uncertainties arising from the electron density located outside the solute cavity. The method presented here provides, with a minimum number of fitted parameters, reasonable estimates of the hydration free energy of ions (average relative error of 4.7%) and is able to reproduce solvation in water of both small and large ions.

Introduction

The theoretical study of chemical systems in condensed phases has experienced an extraordinary evolution in the last decades, leading to a number of computational approaches that combine quantum mechanical (QM) or classical descriptions of the solute with continuum or discrete representations of the solvent molecules.¹ One of these approaches is the self-consistent reaction field (SCRF) continuum model, where the solute resides in a cavity carved into a continuum polarizable medium that simulates the solvent.² In spite of the rather crude description of the solvent, QM SCRF methods have found a widespread acceptance due to several reasons. First, the availability of well-defined formalisms developed to account for the mutual coupling between the solute's charge distribution and the solvent's reaction field (see ref. 3 for a recent review). Second, the ease of gaining insight into the solvent-induced changes in the properties of the solute upon transfer from the gas phase to solution.^{2,4} Third, its low computational cost, which is only slightly larger than the corresponding QM calculation in the gas phase.

Though the preceding features have made QM SCRF methods very popular for the study of molecules in solution, their *chemical* accuracy has mainly been restricted by the uncertainty in the location of the boundary between solute and solvent.² In the last decade a large research effort has been spent to parametrize the solute cavity in different SCRF methods, such as the multipolar expansion method developed at Nancy,⁵ the generalized Born SMx solvation models from Cramer and Truhlar,⁶ the Polarizable Continuum model from Tomasi and coworkers⁷ and the related Miertus-Scrocco-Tomasi (MST) version developed at Barcelona,⁸ or the conductor-like screening model from Klamt and coworkers.⁹ As a result, the most refined QM SCRF methods predict

the free energy of solvation of neutral molecules in a variety of solvents with an average error clearly below 1 kcal/mol when solute cavities adjusted to the specific electrostatic model are used.¹⁰

For ionic solutes, the uncertainty in the hydration free energies determined from QM SCRF methods is larger (around 4-5 kcal/mol). The lower predictive accuracy for ions arises not only for the intrinsic problems in representing the strong solute-solvent interactions existing between water and ions, but also to the poor quality of most experimental data. For neutral solutes the solvation free energy can be directly obtained by measuring partition coefficients of solutes between gas phase and dilute aqueous solutions in equilibrium. In this way the uncertainty of the measured partition coefficient increases with increasing solvation free energy, and eventually the concentration of solute molecules in the gas phase falls below the experimental detection limits when the solvation free energy (in absolute terms) exceeds ~12 kcal/mol.¹¹ This makes it necessary to determine the solvation free energy of ionic species by indirect approaches based on the use of suitable thermodynamic cycles. For example, Pearson¹² determined the experimental solvation free energies for a number of univalent ions using the relationships given by Eqs.1 and 2 for cations and anions, respectively.

$$\Delta G_{sol}(AH^+) = B_g(A) + \Delta G_{sol}(H^+) + \Delta G_{sol}(A) - 2.303RTpK_a(AH^+) \quad (1)$$

$$\Delta G_{sol}(B^-) = -B_g(B^-) - \Delta G_{sol}(H^+) + \Delta G_{sol}(BH) + 2.303RTpK_a(BH) \quad (2)$$

where B_g is the gas phase basicity and ΔG_{sol} denotes the free energy of solvation.

Note that according to Eqs. 1 and 2 the quality of the “experimental” solvation free energies of cations, $\Delta G_{sol}(AH^+)$, and anions, $\Delta G_{sol}(B^-)$, depend on the accuracy

of different thermodynamic quantities, such as the gas phase basicity, the solvation free energy of the conjugate neutral base, $\Delta G_{sol}(A)$, or acid, $\Delta G_{sol}(BH)$, the pK_a constant, and the solvation free energy of the proton, $\Delta G_{sol}(H^+)$. It is also worth noting that correction terms (not explicitly indicated in Eqs. 1 and 2) arising from the standard state definition of the chemical species in the gas phase and in solution might be necessary. Thus, solvation free energies can be defined for the hypothetical process of transferring a species from a fixed position in the gas phase to a fixed position in the solution phase, which is equivalent to using the same standard states (typically 1 M concentration) for gas and solution phases.¹³ However, experimental data in the gas phase is generally referred to an ideal gas at 1 atm. of pressure, so that a correction term of $RT\ln(RT)$ must be considered.

In his original work Pearson used an estimated $\Delta G_{sol}(H^+)$ value of -259.5 kcal/mol, which was also adopted by Florian and Warshel in the parametrization of their Langevin dipole solvation model.¹⁴ However, the values of $\Delta G_{sol}(H^+)$ used by other authors vary considerably (from -252 to -263 kcal/mol)¹⁵. This range of error is similar to that determined from theoretical calculations¹⁶ and introduces a noticeable range of uncertainty in the “experimental” hydration free energy of ions. More recently, Tissandier et al.,¹⁷ using a cluster-pair approximation and thermodynamic properties of ion-water clusters, derived a value of $\Delta G_{sol}(H^+) = -264$ kcal/mol. A very similar value has recently been derived by Tuttle et al and by Coe and coworkers.¹⁸ Because this new $\Delta G_{sol}(H^+)$ value is considerably more negative than the previous ones, it has a significant effect on the absolute solvation free energies of ions and the corresponding acidity/basicity constants.

Pliego and Riveros have recently derived new experimental values for the hydration free energies of 47 univalent ions using as reference the solvation free energy of the proton derived by Tissandier et al.¹⁹ In light of the new reported values, the aim of this work is to revisit the original parametrization of the MST model²⁰ to predict the corrected experimental hydration free energies for ions. In turn, this makes it necessary to solve additional fundamental problems for a continuum description of solvation. First, how should solute cavities be defined for ionic compounds, which are generally characterized by a large anisotropy in the charge distribution? Second, how can errors in the solvation free energy arising from the solute charge density lying outside the cavity be corrected?

In our previous MST parametrization for ions,²⁰ the van der Waals radii used to build up the solute/solvent interface were scaled by a factor of 1.10-1.15, which is around 10% smaller than that recommended for the hydration of neutral solutes.⁸ Moreover, the inaccuracies in the apparent surface charges due to the solute electron density lying outside the cavity was treated by means of a simple uniform correction procedure. However, owing to the small size of the ions examined in that study, the strategy of reducing the whole cavity might not be appropriate for ionic compounds having larger, more complex chemical structures, specially when there is high anisotropy in the charge distribution along the molecule. Then, we explore here the suitability of our original strategy, but limiting the reduction of the cavity to the atom(s) that bear the formal charge in the ion. Furthermore, to reduce artefacts arising from the solute charge density penetration, different charge renormalization procedures are tested to determine their consistency in predicting the hydration free energies, particularly for anions having diffuse charge distributions

Methods

The MST continuum method. In the MST model the free energy of solvation is determined as the sum of three contributions: electrostatic, cavitation, and van der Waals. Following the formalism of the Polarizable Continuum model (PCM) originally developed by Miertus, Scrocco and Tomasi,⁷ the electrostatic term (ΔG_{ele}) is determined by using a set of apparent charges spread over the cavity surface (Eq. 3), which account for the reaction field created in the solvent by the presence of the solute's charge distribution. The apparent charges are obtained by solving the Laplace equation with suitable boundary conditions (Eq. 4).

$$V_R = \sum_{j=1}^M \frac{q_j}{|r_j - r|} \quad (3)$$

where V_R is the perturbation operator used to couple the solute's charge distribution and the solvent reaction field, M is the total number of surface elements, j , into which the solute/solvent boundary is divided, and q_j is the apparent charge at the j surface element, which is centered at r_j .

$$q_j = - \frac{\varepsilon - 1}{4 \pi \varepsilon} S_j \left(\frac{\partial V_T}{\partial n} \right)_j \quad (4)$$

where V_T is the total electrostatic potential, which includes both solute and solvent contributions, n is the unit vector normal to the surface element j , S_j is the area of the surface element j , and ε is the solvent dielectric constant.

The cavitation free energy is determined following Pierotti's scaled particle theory²¹ adapted to molecular shaped cavities by using the procedure proposed by Claverie (Eq. 5).²² Thus, the cavitation free energy of atom i , $\Delta G_{C-P,i}$ is determined weighting the contribution of the isolated atom, $\Delta G_{P,i}$, by the ratio between the solvent-exposed surface of such an atom, S_i , and the total surface of the molecule, S_T (Eq. 5). Finally, the van der Waals term (ΔG_{vdw}) is computed using a linear relationship to the solvent-exposed surface of each atom (Eq. 6).

$$\Delta G_{cav} = \sum_{i=1}^N \frac{S_i}{S_T} \Delta G_{P,i} \quad (5)$$

where N is the number of atoms

$$\Delta G_{vdw} = \sum_{i=1}^N \xi_i S_i \quad (6)$$

where ξ_i is the atomic surface tension, which is determined by fitting to the experimental free energy of solvation.

Molecular-shaped GEPOL cavities²³ are used to define the solute/solvent interface for the calculation of electrostatic and non-electrostatic (cavitation, van der Waals) terms. In the last parametrization of the MST model, a dual-cavity strategy was used,²⁴ so that nonelectrostatic contributions are determined by using a van der Waals surface mainly built up from Pauling's radii (in Å; polar hydrogen: 0.9, apolar hydrogen: 1.2, C and N: 1.5, O: 1.4, F: 1.35, S: 1.75, and Cl: 1.80), whereas the electrostatic term is determined by using a solvent-exposed surface created by scaling the atomic radii by a solvent-dependent factor, λ , which adopts values of 1.25, 1.50, 1.60 and 1.80 for the solvation of neutral compounds in water, octanol, chloroform and

carbon tetrachloride. Based on the comparison of the ΔG_{ele} values determined from MST and from Monte Carlo-Free Energy Perturbation (MC-FEP) calculations for seven small ionic compounds,²⁰ it was found necessary to further reduce the solute cavity by around 10% to compute the hydration free energy of univalent ions. This reduction was justified by the fact that due to strong electrostatic interactions solvent molecules are in general closest to the atoms of charged molecules than in those of the neutral ones.

Charge normalization. Owing to the solute charge density that lies outside the cavity, and to the numerical errors due to tessellation of the cavity surface (Eq. 4), the sum of the apparent surface charges, $Q_{\sigma} = \sum_j q_j$, in the MST (PCM) model do not satisfy the relationship given by Eq. 7,^{7a,25} where Q_M denotes the total charge of the solute, and the error in the charges is given by $\Delta_{\sigma} = Q_{\text{theor}} - Q_{\sigma}$, where $Q_{\text{theor}} = -\frac{\epsilon-1}{\epsilon} Q_M$.

$$Q_{\sigma} = -\frac{\epsilon-1}{\epsilon} Q_M \quad (7)$$

The MST model⁸ takes into account the charge normalization by using the expressions given in the original formalism^{7a} (denoted in the following as the *standard correction* (SC) approach), as noted in Eqs. 8a-b, where Q_{σ}^{+} and Q_{σ}^{-} denote the sum of positive and negative apparent charges, respectively.

$$q'_j = q_j \left[1 + \Delta_{\sigma} / (2Q_{\sigma}^{+}) \right] \quad \text{for } q_j > 0 \quad (8a)$$

$$q'_j = q_j \left[1 + \Delta_{\sigma} / (2Q_{\sigma}^{-}) \right] \quad \text{for } q_j < 0 \quad (8b)$$

More elaborate treatments of the charge compensation have been developed by the Pisa group.²⁶ The most accurate procedure (denoted *density-corrected* (DC) approach hereafter) relies on two main features. First, the introduction of separate factors to correct the apparent charges induced by nuclei (f^Z) and electrons (f^e), which makes possible to correct the numerical errors on the nuclear charges, which are only affected by the discretization of the cavity (Eq. 9). Second, the explicit calculation of the charge density outside the cavity (Q_{out}), which is determined from the flux of the solute electric field through the cavity surface (Eq. 10). Using this new approach the solvent response induced by the electron charge density inside the cavity (Q_{in}) is then only affected by the numerical errors due to tessellation of the cavity (Eq. 11).

$$f^Z \sum_j q_j^Z = -\frac{\varepsilon - 1}{\varepsilon} Q_M^Z \quad (9)$$

$$Q_{out} = -(Q_M^e - Q_{in}) = -(Q_M^e + \frac{1}{4\pi} \int_S E_M(s) \cdot n_s \cdot ds) \quad (10)$$

where $E_M(s)$ is the electric field created at each tessera of the cavity.

$$f^e \sum_j q_j^e = -\frac{\varepsilon - 1}{\varepsilon} Q_{in,M}^e \quad (11)$$

For the sake of completeness, let us note that an alternative procedure used by other groups to alleviate the effects due to the charge density that lies outside the cavity, where the continuous charge distribution obtained from quantum mechanical or density functional calculations is replaced by a set of atomic point charges, often supplemented by point dipoles, which are then used to evaluate the solvent reaction field by means of finite-difference Poisson-Boltzmann or boundary element techniques.²⁷ This approach,

nevertheless, demands that the set of point charges accurately reproduces the charge distribution of the solute, which then raises the question about the best procedure to derive the atomic charges.²⁸

Computational details. The set of compounds compiled by Pliego and Riveros¹⁹ has been used to examine the suitability of our previous parametrization of the MST model to predict the hydration free energy of univalent ions. In order to be consistent with our previous work,²⁰ all calculations have been performed at the HF/6-31+G(d) level. Following the standard MST protocol,^{8,20} the gas phase geometry of the charged molecules was fully optimized and subsequently used for calculations in solution. The MST hydration free energies were determined by using a locally modified version of Monstergauss.²⁹ For the sake of comparison, computations were also performed by using the Integral Equation Formalism³⁰ of the Polarizable Continuum Model (IEF-PCM), which has been claimed to be less sensitive to the inaccuracies arising in the calculation of the electrostatic term. IEF-PCM calculations were performed by using the code implemented in Gaussian03.³¹

Results and discussion

About the cavity definition for ionic species. Though the choice of the solute/solvent boundary is a crucial parameter in the development of any QM SCRF method, the definition of the solute cavity is even more delicate for ionic compounds, whose electrostatic field strongly perturbs the solvent, thus making the solvent molecules in the first hydration shells to exhibit properties clearly different from those of the bulk solvent. In turn, hydration also changes the electron density distribution of ions,³² leading to differential trends in the electron rearrangement that takes place for cations

and anions. Therefore, it can be questioned whether the hydration of ions can be described in an effective way by means of a suitable refinement of the solute cavity in continuum models, and whether or not such a refinement is well suited to reproduce the solvent-induced change in different molecular properties of the solute.³³

An attempt often considered to describe the solvation of ions consists of the addition of an arbitrary number of solvent molecules to the solute, while the rest of the solvent is treated as a continuum dielectric.³⁴ This approach, however, raises some problems, such as the increase in the computational task, the need to obtain appropriate averages of the thermally accessible configurations or the treatment of the librational motions originated from the weak interactions between solute and solvent molecules. Moreover, by adding few solvent molecules, the conceptual problem of the solvation of the solute alone is transformed into the solvation of the solute-solvent cluster. Finally, the analysis of solute-solvent clusters for medium and large size molecules becomes even more problematic and computationally more expensive. Because of these considerations, we prefer here to limit the QM discrete description exclusively to the solute. This choice, nevertheless, raises the delicate question about the suitability of transferring the cavity definition optimized for the solvation of neutral solutes to ionic compounds.

In the original MST parametrization,²⁰ the optimum cavity for the hydration of ions was defined by scaling the atomic radii by a factor of 1.10-1.15, which implies a reduction with regard to the scaling factor of 1.25, i. e. the value optimized for the hydration of neutral solutes.⁸ A similar correction was also adopted in the UAHF parametrization of the PCM model,^{7c} where the atomic radii are reduced depending on the formal charge carried by atoms, though different correction factors are used for cations and anions. Others authors have also proposed to exploit alternative

relationships between the atomic charge and the effective atomic radius,³⁵ or to use modified set of atomic radii.³⁶ An alternative approach for cavity size comes from the use of isodensity contours of the electron density, which can be adjusted during the SCRF calculation as the charge distribution of the solute is polarized by the solvent reaction field.³⁷ Here the electron density contour that leads to better agreement with the experimental solvation free energy for ions is larger than that required for neutral compounds.³⁸

For our purposes here, we have decided to retain the original idea of reducing the scaling factor (λ) that multiplies the atomic radii in order to build up the solute cavity used to calculate the ΔG_{ele} component. However, instead of reducing the whole cavity of the solute, which might be appropriate for small size ions, the scaling down is here limited to the heavy atom (and the hydrogen atoms attached to it) that bears the formal charge. Thus, for alcohol cations and anions the reduction in the cavity only affected the $-\text{OH}_2^+$ and $-\text{O}^-$ groups. In the case of ionized carboxylic acids, such a reduction only affected the oxygen atoms in the $-\text{COO}^-$ group. Besides retaining the simplicity of the original approach, this strategy allows us to exploit the same set of scaling factors that modulate the cavity definition for the calculation of the electrostatic component of the solvation free energy (see **Methods**), while introducing a minimum number of parameters to describe the hydration of ionic species.

Refinement of the MST model for ions. The adjustment of the scaling factor λ used to determine ΔG_{ele} was performed with the aim to minimize the deviation of the calculated hydration free energy, ΔG_{hyd} , from the experimental value for the series of monovalent ions. Following our previous study,²⁰ MST-HF/6-31+G(d) calculations were performed by reducing progressively the scaling factor from $\lambda=1.25$ to $\lambda=1.00$. The ΔG_{ele} values

were calculated by using both SC (Eqs. 8a,b) and DC (Eqs. 10 and 11) charge renormalization schemes. The cavity used to compute the ΔG_{ele} term was built up by using the set of atomic radii typically used in MST calculations (see **Methods**). Finally, the non-electrostatic (cavitation and van der Waals) component, which was estimated by using the standard atomic surface tensions determined for neutral molecules, was added to the ΔG_{ele} term to calculate ΔG_{hyd} . This computational strategy benefits from the fact that the ΔG_{ele} term is the dominant contribution to the hydration of ionic solutes, and from the partial cancellation between cavitation and van der Waals components.

Figure 1 shows the profiles of the root mean square deviation (rmsd) and relative error (in percentage) between experimental and calculated ΔG_{hyd} values determined for different scaling factors. The smallest deviation is found for a scaling factor λ ranging between 1.10 and 1.15, which implies a reduction in the size of the cavity around 90% relative to that used for neutral molecules. It is also worth noting that such a range of values is in agreement with the results found previously by comparing the ΔG_{ele} values determined from MST and MC-FEP calculations for a subset of small ions.²⁰ The results in Figure 1 also point out that both the rmsd and the relative error are further reduced by near 1 kcal/mol when the DC charge renormalization scheme is considered. This finding can be attributed to the larger efficiency of this method to account for the anisotropies in the charge distribution of the ions.

Table 1 reports the electrostatic component of the hydration free energy determined from MST calculations by using a scaling factor of 1.15 and the two charge compensation schemes. For the sake of comparison, it also reports the values determined by using the IEF-PCM formalism and the same cavity definition for the univalent ions. Comparison of the MST and IEF results is shown in Figure 2. The regression equations indicate that there is close agreement between MST and IEF

results. Thus, the MST-SC ΔG_{ele} values (Eq. 12) are around 3% larger than the IEF ones, and such a difference is reduced to only 2% when the MST-DC results are considered (Eq. 13). The most notable difference between the two charge normalization schemes, nevertheless, is reflected in the correlation coefficient of the regression equations, which is significantly larger when the MST-DC results are considered. This finding points out not only the superiority of the DC charge normalization scheme, but also the lower sensitivity of the IEF formalism to the errors arising from the penetration of the charge density outside the solute cavity. Moreover, these results suggest that the MST-DC optimized cavities can be considered to be transferable to the IEF-PCM framework.

$$\Delta G_{ele}(IEF) = 0.970(\pm 0.009)\Delta G_{ele}(MST - SC) \quad r = 0.90 \quad F = 191.2 \quad (12)$$

$$\Delta G_{ele}(IEF) = 0.979(\pm 0.005)\Delta G_{ele}(MST - DC) \quad r = 0.97 \quad F = 748.8 \quad (13)$$

where r is the Pearson's correlation coefficient and F stands for the Snedecor's distribution parameter.

Considering the MST-DC results shown in Table 1, the rmsd between calculated and experimental hydration free energies amounts to 5.8 kcal/mol (relative error of 5.9%). For the subset of 25 anions the rmsd is 4.7 kcal/mol, but it amounts to 6.9 kcal/mol for the 22 cations. Based on this finding, one might be tempted to use different scaling factors to build up the solute cavity for anions and cations. However, the larger rmsd found for cations mainly arises from deviations in protonated R-O-R' (methanol, ethanol, dimethyl ether, diethylether) and R-CO-R' (acetone, methyl phenyl ketone, methyl ethyl ester) groups, which have relative errors $\geq 10\%$. When these compounds are excluded, the rmsd for the subset of cations (3.6 kcal/mol) is similar to that found

for anions. It is not clear whether the notable deviation found for these compounds might stem from short-range interactions with hydrating water molecules not properly accounted for in the continuum model. Keeping in mind the limited availability of accurate ΔG_{hyd} values for ionic compounds and the assumptions intrinsic to the continuum method, it seems to us not justified, at least on the basis of present results, to use different scaling factors for anions and cations, a procedure already used in the context of solvation calculations of ion hydration carried out by using the Born equation.³⁹ Likewise, attempts to reduce these errors by distinct reshaping of solute cavities seems not justified and might even yield to non-physical over-trained models.

An alternative strategy to reduce the difference between computed and experimental data involves the adjustment of the atomic surface tensions, ξ_i (see Eq. 6), which in the MST model are typically determined by fitting the experimental free energies of solvation. Since steric terms are used in continuum models to account implicitly for effects other than pure bulk electrostatic effects, the fitted parameters for neutral and charged molecules might not be equal. Accordingly, one can expect that restricting the refinement of the surface tensions only to the atom(s) bearing the formal charge would suffice to improve the agreement between calculated and experimental ΔG_{hyd} values. From a practical point of view, caution is required owing to the limited availability of data for certain atom types, which can lead to overtraining of the method. Thus, the only statistically significant improvement was obtained when the surface tension of the OH^+ group was adjusted to $-0.51 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, which is near 4-fold more negative than the corresponding value for neutral molecules ($-0.14 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$).^{8d} With the only correction of this surface tension, the rmsd between calculated and experimental ΔG_{hyd} values is 4.6 kcal/mol.

The calculated ΔG_{hyd} and its electrostatic and non-electrostatic components determined for the univalent anions are given in Table 2. With the exception of the protonated R-O-R' and R-CO-R' groups (see above), the non-electrostatic term is generally positive and only amounts, on average, to 3.9% of the hydration free energy. Overall, there is satisfactory agreement between experimental and calculated ΔG_{hyd} values, as noted in a relative error of 4.7%, and in the comparison of the regression equation given in Eq. 14 (see Figure 3), specially keeping in mind the variety of functional groups and sizes of the ionic compounds.

$$\Delta G_{\text{hyd}}(\text{exp}) = 1.019(\pm 0.008) \Delta G_{\text{hyd}}(\text{MST} - \text{DC}) \quad r = 0.94 \quad F = 351.0 \quad (14)$$

Prediction of pK_a values. Since the acidity/basicity of molecules is crucial to understand chemical reactivity and intermolecular interactions, an intense research effort has been devoted to check the reliability of theoretical methods, particularly QM SCRF continuum methods, to predict accurately the pK_a. Thus, Chen et al. were able to predict the pK_a values for several molecules by using a model where the reaction potential was solved via finite-difference solutions to the Poisson-Boltzmann equation.^{25c} For a small series of imidazole derivatives, Topol et al. predicted pK_a values with an average absolute deviation of 0.8 units from experimental data by combining results obtained from G2 or density functional calculations with SCRF estimates of the hydration free energies.^{25d} Schüürmann et al. estimated the acidities of carboxylic acids by using the PCM-UAHF model at the SCF and MP2 levels, obtaining good correlations for relative pK_a values but not for the absolute ones, arguing that the major sources of error stem from gas-phase calculations.⁴⁰ Recently Shields and co-workers predicted accurate absolute pK_a's for carboxylic acids⁴¹ and phenols⁴² using the

C-PCM model along with *state-of-the-art* gas phase *ab initio* calculations. Chen and MacKerell compared pK_a differences in substituted pyridines by using semiempirical and *ab initio* gas-phase free energies along with IPCM, SCIPCM and SM2 solvation calculations, and relevant differences between the pK_a values obtained at different levels of theory were found.⁴³ By using the SSC(V)PE model, Chipman checked different isodensity thresholds to build up cavities in order to predict the pK_a 's for alcohols, acids and amines.^{38b} This study showed that no single isodensity contour can be applied to neutral compounds, anions and cations, and that good correlations can be obtained only for compounds sharing the same acidic functional group. Other studies also have pointed out the difficulty to predict absolute pK_a values, which demand highly accurate estimates of both the gas phase free energy difference between acid and base partners and of the corresponding solvation free energies.⁴⁴

To assess the performance of the MST parametrization for ionic compounds, we have determined the pK_a values for a series of carboxylic acids and substituted pyridines by using the thermodynamic cycle shown in Figure 4 and Eq. 15. These compounds were chosen because of the availability of experimental data for the gas phase basicities.^{45,46} For consistency with the MST parametrization, the experimental $\Delta G_{sol}(H^+)$ value determined by Tissandier¹⁷ was used. The hydration free energies of the neutral species were calculated by using the HF/6-31G(d) version of the MST model.^{8d} Finally, to further compare the effect of the charge normalization scheme both MST-SC and MST-DC calculations were performed for ionic compounds.

$$pK_a = -\log K_a = \Delta G_{aq} / 2.303RT \quad (15)$$

Table 3 shows the results obtained for the two series of compounds, which encompasses a range of pK_a values larger than 5 units. Comparison of the results obtained by using the two charge renormalization schemes supports the better performance of the DC procedure. Thus, whereas both MST-SC and MST-DC methods gives rise to similar relative pK_a variations within the series of compounds (the rmsd between calculated and experimental pK_a differences varies between 0.4 and 1.2 pK_a units), the absolute pK_a values determined from MST-DC computations agree better (in absolute terms) with the experimental values, as noted in a rmsd (in pK_a units) of 1.0 (MST-DC) and 2.4 (MST-SC) for carboxylic acids and of 1.9 (MST-DC) and 3.2 (MST-SC) for substituted pyridines.

Taking together the two series of compounds, the rmsd of the pK_a values predicted from MST-DC calculations amount to 1.6 pK_a units. This uncertainty can be mainly ascribed to the error in the hydration free energy of the ionic species, since the difference between calculated and experimental hydration free energy for neutral compounds (experimental ΔG_{hyd} available for acetic acid: -6.8 kcal/mol, pyridine: -4.7 kcal/mol, and 4-methylpyridine: -4.9 kcal/mol) is less than 0.5 kcal/mol. Chen and McKerell reported rmsd values ranging from 1.22 to 4.14 pK_a units depending on the method used to compute the solvation free energy (see above). For carboxylic acids alone, standard deviations of 0.4 pK_a units have been reported by using Complete Basis Set models in conjunction with C-PCM calculations.⁴¹ However, accurate estimates of pK_a values (i. e., with an error of half pK_a unit) are typically obtained by means of an empirical fitting of the experimental pK_a values.^{44b-d} In the absence of those empirical adjustments, the pK_a values determined from MST-DC calculations compare satisfactorily with the experimental ones, thus giving support to the computational strategy used here to refine the MST method for the hydration of ionic compounds.

Zwitterion-neutral equilibria. Because of the high anisotropy in the charge distribution of zwitterions, the equilibrium between neutral and charge-separated forms is particularly well suited to examine the effect of the charge renormalization scheme within the framework of the MST model. To this end, we have examined the equilibrium between neutral and zwitterionic forms of nicotinic and isonicotinic acids (see Figure 5), a system that was previously studied by Nagy et al.⁴⁷ In their study the free energy of tautomerization in water was determined by adding the relative free energies of hydration to the gas phase free energy difference between neutral and zwitterionic species. This latter quantity was estimated from calculations carried out at the MP2/6-311++G(d,p) level and subsequently corrected from zero-point energy, thermal and entropy corrections determined at the MP2/6-31G(d). At this level, the neutral forms are predicted to be favoured by 33-35 kcal/mol in the gas phase. To determine the relative stability of tautomers in water, we adopted the relative gas phase stabilities determined by Nagy et al., the hydration free energies of the neutral species were determined from MST/HF6-31G(d) calculations,^{8d} and those of the zwitterionic compounds were obtained from both MST-SC and MST-DC calculations.

Table 4 reports the hydration free energies of the neutral and zwitterionic species as well as the predicted tautomerization free energy (ΔG_t) in water. As was found for the calculated pK_a values, the results obtained from MST-DC computations agree better with the experimental data than those determined from MST-SC calculations. Thus, whereas both neutral and zwitterionic forms of nicotinic acid are predicted to have similar stability in water from MST-SC calculations, the neutral form is predicted to be destabilized by around 2.8 kcal/mol according to MST-DC results, which agrees with the observed preference of the zwitterion in aqueous solution by around 2.0 kcal/mol.

With regard to the isonicotinic acid, the MST-SC method predicts again the neutral form to be preferred in water, whereas the opposite trend is found from MST-DC calculations, which predict the zwitterion to be more stable by 1.7 kcal/mol, in agreement also with the experimental data.

Solvent reaction field in SC and DC schemes. The whole of the preceding results support the better performance of the DC charge renormalization scheme for predicting the hydration free energy of ionic compounds. This finding agrees with the results obtained from the analysis of the fractional contributions of specific groups to the hydration free energy within series of congeneric alkane derivatives,⁴⁸ since the DC scheme was found to yield consistently more transferable group contributions than the SC scheme.

To gain further insight into the origin of the differences noticed between both SC and DC charge renormalization schemes, we have compared a series of electrostatic properties determined by using the wavefunction of the fully hydrated solute for a small subset of charged monosubstituted alkane derivatives (CH_3X and $\text{CH}_3(\text{CH}_2)_6\text{X}$, with X being NH_3^+ , OH_2^+ , NH^- and O^-). This series of compounds is well suited to our purposes because i) the enlargement of the side chain is expected to play a relatively small perturbing effect on the hydration of the X unit, and ii) this allows us to concentrate onto the differences in the solvent reaction field around the X group. To this end, we have computed the rmsd between the electrostatic potential at selected points in four layers around the molecule. For our purposes here, such a comparison was performed by including exclusively those points generated by scaling the atomic radii of atoms in the group X. Moreover, we have determined the well depth (E_{min}) of the molecular interaction potential⁴⁹ in the vicinity of the group X for the solutes interacting with

classical probe particles O^- (cations) and H^+ (anions). Finally, the results are examined in light of the solvent's net apparent charge (Q_σ) induced on the surface elements spread on the portion of the cavity created from the group X. The results are shown in Table 5.

Not unexpectedly, there is close similarity in the results obtained from the two charge renormalization schemes for $CH_3NH_3^+$ and $CH_3OH_2^+$. However, there are notable differences when the largest members of the congeneric series, $CH_3(CH_2)_6NH_3^+$ and $CH_3(CH_2)_6OH_2^+$, are considered. Thus, the electrostatic potential around the group X is more positive when the DC scheme is used, which reflects the more negative apparent charge induced on the cavity surface around the NH_3^+ and OH_2^+ groups. In fact, by enlarging the size of the side chain there is a reduction of 25-32% in $Q_\sigma(X)$ when the SC scheme is used, which is notably larger than the 8-10% variation found for the DC results.

For anions the differences between the two charge renormalization schemes are clear even for the parent anions, CH_3NH^- and CH_3O^- , and tend to decrease as the size of the ionic compound is enlarged. This is particularly true for the former anion, CH_3NH^- , where the solvent apparent charge determined in the SC scheme ($Q_\sigma(NH^-)=+0.91$) is sensibly larger than that obtained in the DC formalism ($Q_\sigma(NH^-)=+0.78$). A detailed inspection of the results reveals that such a difference stems from the solvent reaction field created in the portion of the cavity associated to the hydrogen atom. Thus, when the DC scheme is used, the solvent reaction field is mainly concentrated in the surface elements pertaining to the nitrogen atom, giving rise to an apparent charge of +0.80, whereas the apparent charge on the portion of the cavity created by the hydrogen atom is only -0.02. In contrast, when the SC scheme is considered, both nitrogen and hydrogen atoms contribute markedly to the solvent reaction field, as noted in apparent charges of +0.69 and +0.22, respectively. This different distribution of the reaction

field in the surroundings of the NH group gives rise to differences in the electrostatic properties of the solute (see Table 5), leading in turn to different solute-solvent interaction energy.

In summary, the preceding results clearly point out that the explicit treatment of anisotropy in the charge distribution leads to subtle, but nonnegligible differences in the solvent reaction field, which in turn might modulate the calculation of the solvent-induced change in solute's properties, as noted in the case of the group contributions to the solvation free energy.⁴⁸

Concluding remarks

A satisfactory agreement between predicted and experimental hydration free energies of univalent ions can be obtained by means of simple, conservative adjustment in the definition of the cavity used to obtain the electrostatic contribution within the framework of the MST model. Such a refinement implies a reduction in the cavity of 8% relative to the standard scaling used for neutral species, and is limited to the atom(s) that bear the formal charge. This allows us to retain the original simplicity of the approach and to exploit the same set of scaling factors that modulate the cavity definition for the calculation of the electrostatic term in the MST model.

The results point out the need to account explicitly for the anisotropies in the charge distribution of ionic species in order to correct the effects due to the solute's charge density that lies outside the cavity. The explicit consideration of the anisotropy in the charge distribution of the ionic compounds leads to a subtle, but nonnegligible improvement in the calculated electrostatic component of the hydration free energy. Moreover, it also modulates the magnitude of the solvent reaction field, which in turn influences the solvent-induced change in the solute's properties. On the other hand, the

IEF values are very close to and exhibit high correlation with the MST-DC results, which are on average around 2% larger (in absolute value) than the hydration free energies determined by using the IEF formalism. Therefore, these findings point out that the optimized cavities should be largely transferable between the two computational schemes.

The optimized MST-DC method leads to an average relative error of 4.7% in the hydration free energy for the series of univalent ions containing a variety of functional groups and molecular sizes. The basic strategy proposed here is still open to further refinement, at least for the prediction of the hydration free energy, through the adjustment of the atomic surface tensions used for the charged atoms. However, this refinement would require the existence of a larger database of experimental values to avoid overtrained models that can fail when applied to compounds very different to those used in the parametrization of the model. In any case the robustness of the strategy is supported by the results obtained for the absolute and relative pK_a s of carboxylic acids and substituted pyridines, as well as for the tautomeric equilibrium between neutral and zwitterionic forms of nicotinic and isonicotinic acids. These examples suffice to point out that the MST model can be used confidently to measure the magnitude of the solvent influence of chemically relevant properties and to use this continuum method as an interpretative tool able to gain insight into the influence of solvation on chemical processes of ionic species.

Acknowledgments

We are grateful to Prof. J. Tomasi for providing us with his original code of the PCM model, which was modified by us to carry out the MST calculations. This work has been supported by the Ministerio de Ciencia y Tecnología (BIO2003-06848 and

SAF2002-04282), and the Centre de Supercomputació de Catalunya. C.C. is fellowship from the Ministerio de Ciencia y Tecnología.

Table 1. Electrostatic component of the free energy of hydration (kcal/mol) determined at the HF/6-31+G(d) level from MST calculations using both SC and DC charge normalization procedures and from IEF calculations.

Ion	MST-SC	MST-DC	IEF
F ⁻	-101.7	-101.7	-98.4
Cl ⁻	-75.4	-75.4	-72.5
OCl ⁻	-74.6	-76.1	-73.5
OH ⁻	-106.4	-104.1	-98.7
HO ₂ ⁻	-92.6	-92.5	-89.4
NH ₂ ⁻	-102.5	-96.8	-90.3
CN ⁻	-74.8	-75.0	-71.6
HCC ⁻	-81.7	-81.9	-75.8
CH ₃ O ⁻	-90.3	-88.8	-84.7
CH ₃ CH ₂ O ⁻	-89.8	-86.1	-81.4
C ₆ H ₅ O ⁻	-77.0	-70.9	-66.1
HCOCH ₂ ⁻	-76.5	-73.5	-69.4
CH ₃ COCH ₂ ⁻	-77.3	-73.0	-68.9
HCOO ⁻	-80.3	-79.8	-76.6
CH ₃ COO ⁻	-82.8	-80.7	-76.9
CH ₂ COOEt ⁻	-74.0	-70.0	-66.7
CH ₃ CONH ⁻	-79.3	-76.4	-72.7
N ₃ ⁻	-72.6	-71.9	-68.9
CH ₂ CN ⁻	-70.3	-69.6	-65.8
CH ₃ CHCN ⁻	-69.2	-67.1	-64.1
CH ₂ NO ₂ ⁻	-76.5	-73.5	-70.7
HS ⁻	-80.3	-76.5	-71.8
CH ₃ S ⁻	-78.4	-76.2	-71.4
C ₆ H ₅ S ⁻	-74.9	-67.4	-63.2
CH ₃ SO ₃ ⁻	-77.8	-76.4	-72.9
H ₃ O ⁺	-105.2	-106.3	-107.6
CH ₃ OH ₂ ⁺	-83.8	-84.7	-86.3
CH ₃ CH ₂ OH ₂ ⁺	-77.3	-79.1	-80.3
(CH ₃) ₂ OH ⁺	-68.5	-68.9	-69.6
(Et) ₂ OH ⁺	-59.2	-60.1	-61.4
NH ₄ ⁺	-93.5	-93.7	-94.9
CH ₃ NH ₃ ⁺	-81.2	-81.3	-82.9
(CH ₃) ₂ NH ₂ ⁺	-70.7	-70.6	-72.4
(CH ₃) ₃ NH ⁺	-60.9	-60.7	-62.3
CH ₃ CH ₂ NH ₃ ⁺	-76.1	-77.2	-78.9
(Et) ₃ NH ⁺	-52.4	-53.1	-54.1
<i>n</i> -PrNH ₃ ⁺	-73.1	-75.6	-77.5
<i>n</i> -BuNH ₃ ⁺	-71.2	-75.2	-76.7
C ₆ H ₅ NH ₃ ⁺	-69.1	-76.1	-77.8
(CH ₃ COCH ₃)H ⁺	-66.5	-67.6	-68.4
(PhCOCH ₃)H ⁺	-53.6	-56.8	-58.6

$(\text{CH}_3\text{COOEt})\text{H}^+$	-61.2	-63.1	-63.4
HCONH_3^+	-78.4	-79.4	-80.4
$\text{CH}_3\text{CONH}_3^+$	-74.2	-74.7	-75.7
$\text{C}_3\text{H}_5\text{NH}^+$	-59.5	-60.5	-61.8
$(\text{CH}_3)_2\text{SH}^+$	-63.7	-63.5	-64.7
$(\text{CH}_3\text{SOCH}_3)\text{H}^+$	-63.3	-64.9	-65.7

Table 2. Electrostatic and non-electrostatic contributions (kcal/mol) to the free energy of hydration determined from MST-HF/6-31+G* calculations for the series of univalent ions used in the parametrization. The electrostatic term was determined by using a scaling factor of 1.15 and the DC charge renormalization scheme. The experimental hydration free energies are also given.

Ion	ΔG_{ele}	$\Delta G_{\text{n-ele}}$	$\Delta G_{\text{hyd}}(\text{MST-DC})$	Expt. ^a
F ⁻	-101.7	1.7	-99.9	-105.0
Cl ⁻	-75.4	0.5	-74.8	-74.6
OCl ⁻	-76.1	2.4	-73.7	-80.7
OH ⁻	-104.1	0.1	-104.0	-105.0
HO ₂ ⁻	-92.5	2.3	-90.2	-97.3
NH ₂ ⁻	-96.8	0.3	-96.5	-92.2
CN ⁻	-75.0	2.9	-72.1	-67.6
HCC ⁻	-81.9	-0.8	-82.7	-76.1
CH ₃ O ⁻	-88.8	2.7	-86.0	-95.2
CH ₃ CH ₂ O ⁻	-86.1	3.0	-83.1	-91.1
C ₆ H ₅ O ⁻	-70.9	2.8	-68.1	-71.3
HCOCH ₂ ⁻	-73.5	2.5	-71.0	-75.7
CH ₃ COCH ₂ ⁻	-73.0	2.6	-70.4	-75.6
HCOO ⁻	-79.8	4.1	-75.7	-76.2
CH ₃ COO ⁻	-80.7	4.2	-76.5	-77.3
CH ₂ COOEt ⁻	-70.0	4.0	-66.0	-66.4
CH ₃ CONH ⁻	-76.4	2.5	-73.9	-80.1
N ₃ ⁻	-71.9	0.3	-71.6	-70.7
CH ₂ CN ⁻	-69.6	3.6	-66.0	-65.7
CH ₃ CHCN ⁻	-67.1	4.3	-62.8	-65.6
CH ₂ NO ₂ ⁻	-73.5	4.2	-69.3	-75.9
HS ⁻	-76.5	1.8	-74.8	-71.6
CH ₃ S ⁻	-76.2	0.9	-75.3	-73.7
C ₆ H ₅ S ⁻	-67.4	1.1	-66.3	-63.3
CH ₃ SO ₃ ⁻	-76.4	6.1	-70.3	-71.7
H ₃ O ⁺	-106.3	0.4	-105.9	-110.2
CH ₃ OH ₂ ⁺	-84.7	-6.9	-91.5	-93.1
CH ₃ CH ₂ OH ₂ ⁺	-79.1	-6.6	-85.7	-88.4
(CH ₃) ₂ OH ⁺	-68.9	-3.2	-72.0	-79.8
(Et) ₂ OH ⁺	-60.1	-2.6	-62.6	-71.5
NH ₄ ⁺	-93.7	0.6	-93.1	-85.2
CH ₃ NH ₃ ⁺	-81.3	1.2	-80.2	-76.5
(CH ₃) ₂ NH ₂ ⁺	-70.6	1.7	-68.9	-68.6
(CH ₃) ₃ NH ⁺	-60.7	2.3	-58.4	-61.2
CH ₃ CH ₂ NH ₃ ⁺	-77.2	1.5	-75.8	-73.0
(Et) ₃ NH ⁺	-53.1	3.1	-50.0	-54.7
<i>n</i> -PrNH ₃ ⁺	-75.2	2.0	-73.2	-70.9
<i>n</i> -BuNH ₃ ⁺	-75.6	1.7	-73.9	-71.5

$C_6H_5NH_3^+$	-76.1	1.4	-74.7	-72.8
$(CH_3COCH_3)H^+$	-67.6	-5.6	-73.2	-76.8
$(PhCOCH_3)H^+$	-56.8	-5.2	-61.9	-65.1
$(CH_3COOEt)H^+$	-63.1	-4.0	-67.1	-68.5
$HCONH_3^+$	-79.4	0.6	-78.7	-82.5
$CH_3CONH_3^+$	-74.7	0.8	-73.9	-73.8
$C_5H_5NH^+$	-60.5	1.3	-59.2	-61.1
$(CH_3)_2SH^+$	-63.5	2.6	-60.9	-64.5
$(CH_3SOCH_3)H^+$	-64.9	1.6	-63.3	-68.2

^a Data taken from ref.19.

Table 3. Calculated and experimental pK_a values for series of substituted pyridines and carboxylic acids. MST calculations have been performed by using both SC and DC charge renormalization methods. Absolute and relative pK_a values are given in plain and in italics, respectively.

Compound	ΔG_{gas}^a	$\Delta G_{\text{sol}}(A)$	$\Delta G_{\text{sol}}(AH^+)$ (SC)	pK_a^b (SC)	$\Delta G_{\text{sol}}(AH^+)$ (DC)	pK_a^b (DC)	pK_a (exp) ^c
Pyridine derivatives							
H	213.7	-4.9	-58.2	2.2 <i>0.0</i>	-59.2	2.9 <i>0.0</i>	5.2 <i>0.0</i>
<i>p</i> -CH ₃	218.0	-5.3	-54.9	2.6 <i>0.4</i>	-56.2	3.6 <i>0.7</i>	6.0 <i>0.8</i>
<i>m</i> -CH ₃	216.6	-5.2	-55.1	1.8 <i>-0.4</i>	-56.7	3.0 <i>0.1</i>	5.7 <i>0.5</i>
<i>m</i> -F	206.7	-3.5	-61.2	0.3 <i>-1.9</i>	-62.3	1.1 <i>-1.8</i>	3.0 <i>-2.2</i>
<i>p</i> -Cl	210.4	-3.9	-59.4	1.4 <i>-0.9</i>	-60.8	2.4 <i>-0.5</i>	3.8 <i>-1.4</i>
<i>m</i> -Cl	207.5	-4.0	-60.0	-0.3 <i>-2.5</i>	-61.4	0.7 <i>-2.3</i>	2.8 <i>-2.4</i>
<i>o</i> -OCH ₃	220.9	-4.9	-52.9	3.6 <i>1.4</i>	-54.5	4.8 <i>-1.9</i>	6.6 <i>1.4</i>
<i>m</i> -OCH ₃	216.7	-4.6	-54.2	1.7 <i>-0.5</i>	-55.9	2.9 <i>0.0</i>	4.8 <i>-0.4</i>
<i>p</i> -CN	202.5	-5.5	-64.4	-1.9 <i>-4.1</i>	-67.8	0.6 <i>-2.3</i>	1.9 <i>-3.4</i>
<i>m</i> -CN	201.7	-6.1	-65.5	-2.2 <i>-4.4</i>	-69.0	0.4 <i>-2.5</i>	1.4 <i>-3.9</i>
<i>m</i> -NO ₂	200.2	-5.4	-66.4	-2.1 <i>-4.3</i>	-69.8	0.4 <i>-2.5</i>	1.2 <i>-4.0</i>
mse ^d				-3.2 <i>-0.2</i>		-1.8 <i>0.5</i>	
rmsd				3.2 <i>0.4</i>		1.9 <i>0.8</i>	
Carboxylic acids							
Formic	336.5	-7.1	-76.2	2.5 <i>0.0</i>	-75.7	2.9 <i>0.0</i>	3.8 <i>0.0</i>
Acetic	339.5	-6.8	-78.6	2.7 <i>0.2</i>	-76.5	4.3 <i>1.4</i>	4.8 <i>1.0</i>
Cyanoacetic	321.8	-9.4	-67.1	0.1 <i>-2.4</i>	-66.1	0.9 <i>-2.0</i>	2.5 <i>-1.3</i>
Chloroacetic	327.0	-8.3	-68.7	1.9 <i>-0.6</i>	-69.0	1.7 <i>-1.2</i>	2.9 <i>-0.9</i>
Pivalic	335.9	-4.7	-75.5	0.8 <i>-1.7</i>	-69.8	5.0 <i>2.1</i>	5.0 <i>1.2</i>

mse				-2.2		-0.8	
rmsd				-0.2		-0.3	
				2.4		1.0	
				1.2		0.6	
All compounds							
mse				-2.9		-1.5	
				-0.2		0.3	
rmsd				3.0		1.6	
				0.7		0.7	

^a Data taken from refs. 45 and 46.

^b A correction term of 1.9 kcal/mol was added to the gas phase free energy difference to account for the change in standard states at 298.15 K.

^c Values taken from ref. 50.

^d rms: root-mean square deviation; mse: mean signed deviation.

Table 4. Calculated and experimental free energies of tautomerization between neutral and zwitterionic forms of nicotinic (NA) and isonicotinic (INA) acids. MST calculations were performed by using both SC and DC charge renormalization methods.

Compound	$\Delta G_{\text{gas}}^{\text{a}}$	$\Delta G_{\text{sol}}(\text{A})$	$\Delta G_{\text{sol}}(\text{AH}^+)$ (SC)	ΔG_{aq} (SC)	$\Delta G_{\text{sol}}(\text{AH}^+)$ (DC)	ΔG_{aq} (DC)	ΔG_{aq} (exp.) ^a
NA (A)	33.3	-9.3	-42.6	0.0	-45.3	-2.7	-2.0
(B)	33.5	-9.0	-42.6	0.0	-45.3	-2.7	
INA	35.1	-8.6	-42.1	1.6	-45.4	-1.7	-2.5

^a Data taken taken from ref. 47

Table 5. Differences in the electrostatic potential computed in layers around the molecule, the well depth of the molecular interaction potential, and the net apparent charge induced on the cavity surface for series of congeneric ionic compounds.

Compound	X	rmsd ^a in layer				E _{min} ^b		Q _e (X) ^c	
		x 1.4	x 1.6	x 1.8	x 2.0	MST-SC	MST-DC	MST-SC	MST-DC
CH ₃ NH ₃ ⁺	NH ₃ ⁺	0.0	0.0	0.0	0.0	-123.8	-123.8	-0.71	-0.71
CH ₃ (CH ₂) ₆ NH ₃ ⁺	NH ₃ ⁺	2.1	1.8	1.6	1.4	-118.6	-121.5	-0.53	-0.65
CH ₃ OH ₂ ⁺	OH ₂ ⁺	0.5	0.4	0.3	0.3	-130.1	-129.7	-0.57	-0.59
CH ₃ (CH ₂) ₆ OH ₂ ⁺	OH ₂ ⁺	3.6	3.0	2.5	2.2	-127.1	-127.5	-0.39	-0.53
CH ₃ NH ⁻	NH ⁻	6.6	5.5	4.4	3.7	-187.6	-192.3	+0.91	+0.78
CH ₃ (CH ₂) ₆ NH ⁻	NH ⁻	1.6	1.3	1.1	0.9	-190.6	-190.6	+0.79	+0.72
CH ₃ O ⁻	O ⁻	0.9	0.8	0.6	0.5	-169.6	-168.5	+0.81	+0.77
CH ₃ (CH ₂) ₆ O ⁻	O ⁻	0.6	0.5	0.5	0.4	-168.2	-168.2	+0.76	+0.75

^a rmsd between the electrostatic potentials computed from the wavefunctions determined from MST-SC and MST-DC calculations for the ionic compounds. Layers defined by scaling the van der Waals radii by 1.4, 1.6, 1.8 and 2.0.

^b Minimum energy of the molecular interaction potential for the ionic compounds interacting with classical probe particles O^- (anions) and H^+ (anions).

^c Net apparent charge spread on the cavity surface of the group X determined from both SC and DC charge renormalization scheme.

Legends to Figures

Figure 1. Representation of the change in the root-mean-square deviation (rmsd; solid line) and the relative error (in percentage; dashed line) between the experimental hydration free energies and the MST values obtained by varying the scaling factor used to correct the atomic radii of charged atoms in the calculation of the electrostatic component of the hydration free energy (see text for details). The electrostatic free energy of solvation was determined by using both SC (black) and DC (white) charge renormalization schemes (see text for details). The rmsd values are in kcal/mol.

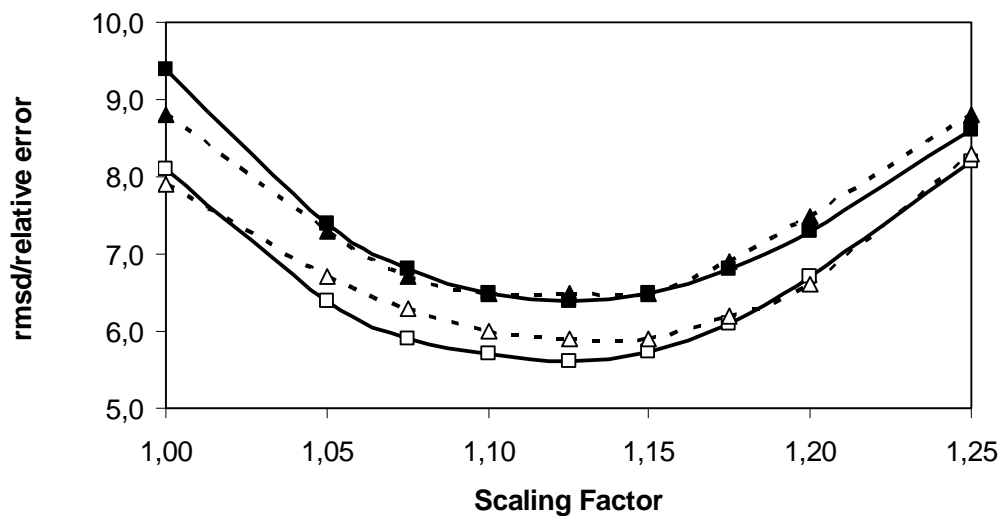


Figure 2. Comparison of the electrostatic component of the free energy of hydration (kcal/mol) determined from MST calculations by using a scaling factor of 1.15 and both SC (black) and DC (white) charge renormalization schemes and IEF calculations at the HF/6-31+G(d) level. The plots displays the ideal regression line $y=cx$ with slope c equal to unity.

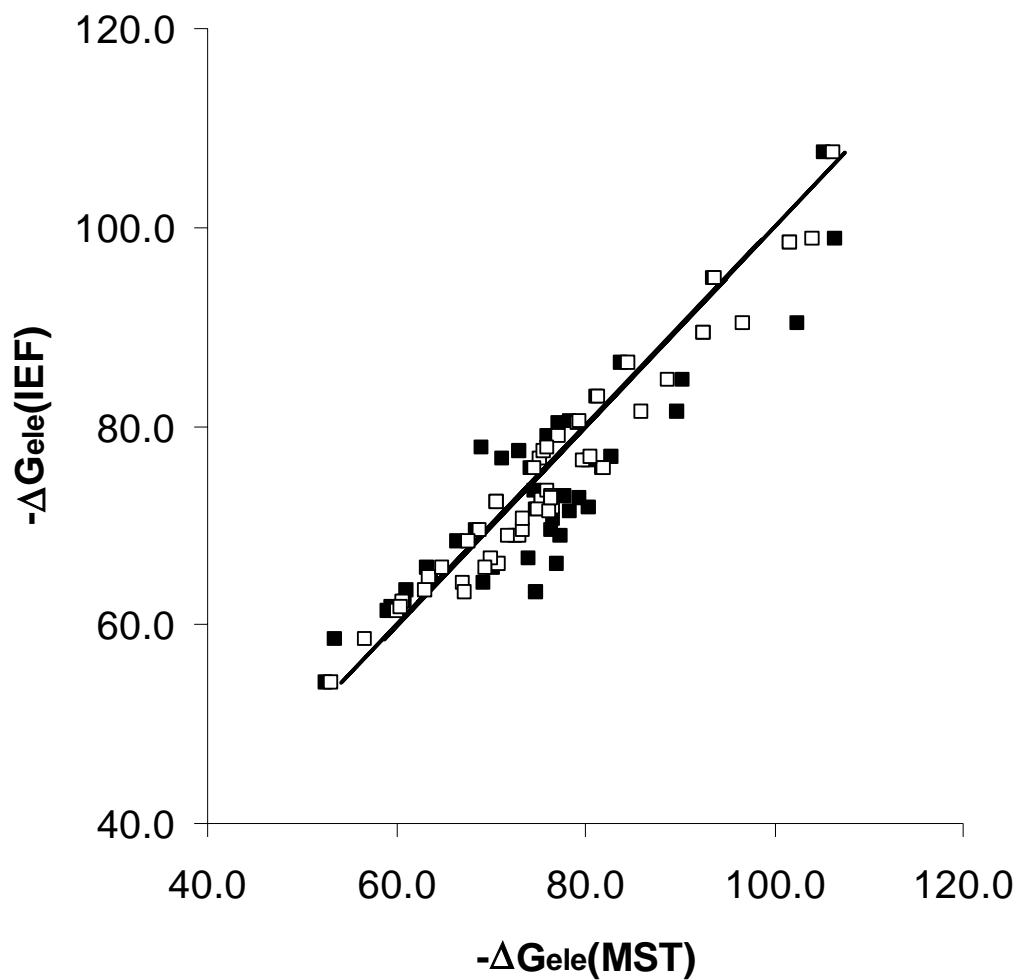


Figure 3. Comparison of the experimental and calculated hydration free energies (kcal/mol) for the series of univalent ions. Theoretical estimates were obtained from MST-HF/6-31+G(d) calculations carried out by using a scaling factor of 1.15 and the DC charge renormalization schemes. The plots displays the ideal regression line $y=cx$ with slope c equal to unity.

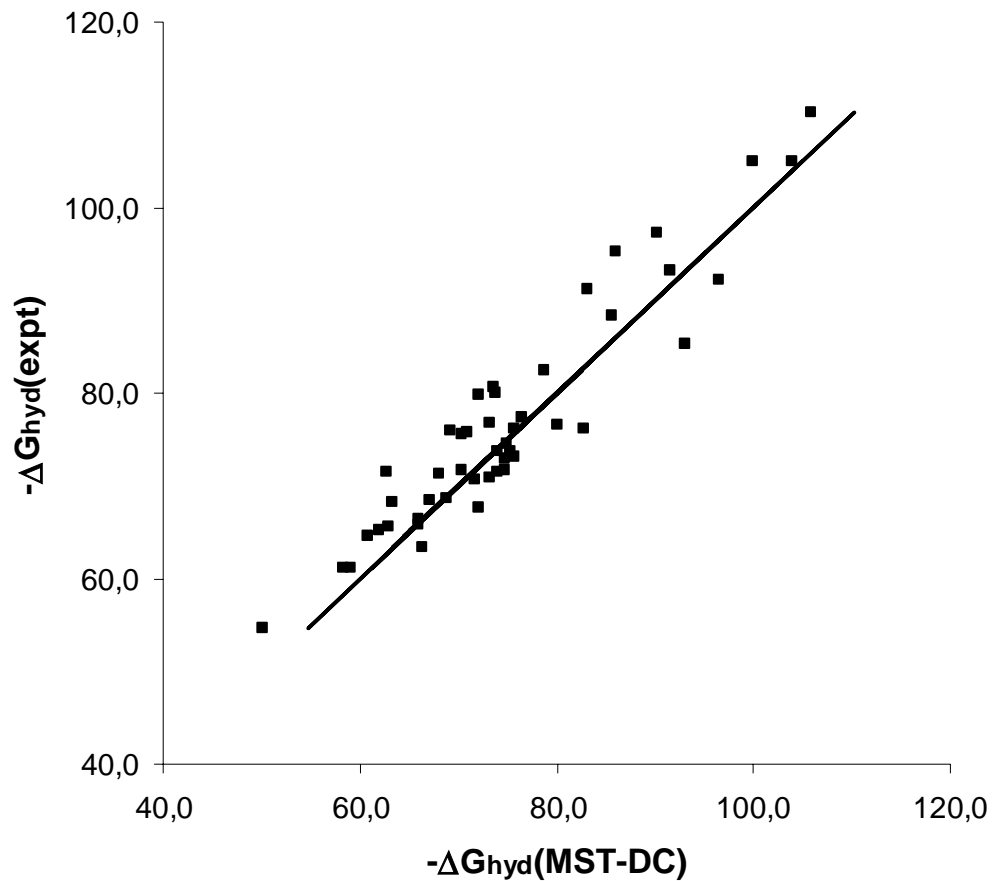


Figure 4. Thermodynamic cycle used for the calculation of the pK_a values for the series of substituted pyridines and carboxylic acids.

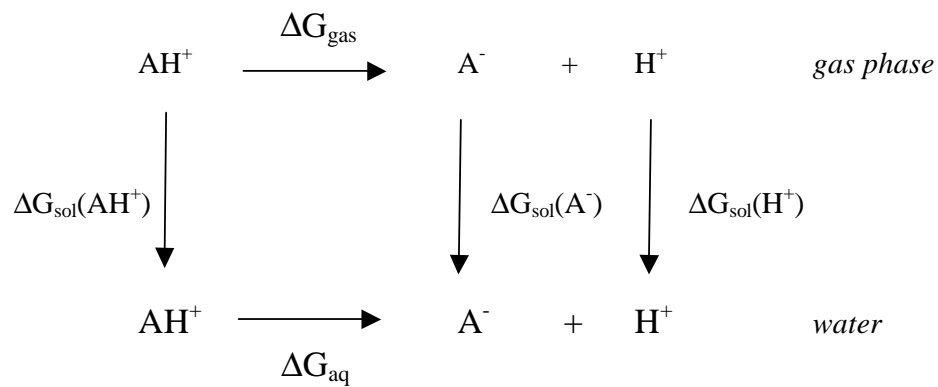
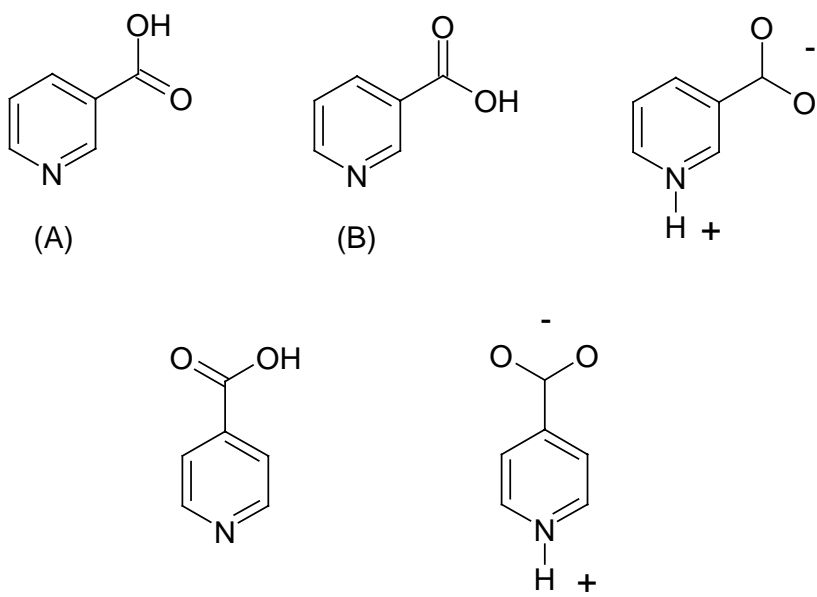


Figure 5. Representation of the neutral and zwitterionic species of nicotinic (top) and isonicotinic (bottom) acids.



References

-
1. (a) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. Wiley: New York, 1991. (b) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395. (b) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184. (c) Cramer, C. J.; Truhlar, D. G. *Structure and Reactivity in Aqueous Solution*. American Chemical Society: Washington, 1994, Vol. 568. (c) *Computer Simulation in Biomolecular Systems*. Theoretical and Experimental Applications. van Gunsteren, W.F.; Weiner, P.K.; Wilkinson, A.J. (eds.) Kluwer: Dordrecht, 1997, Vol. 3. (d) Tomasi, J.; Mennucci, B.; Cammi, R.; Cossi, M. In *Computational Approaches to Biochemical Reactivity*. Náray-Szabó, G., Warshel, A, Eds.; Kluwer: Dordrecht, 1997, p. 1. (e) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187.
 2. (a) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027. (b) Rivail, J. L.; Rinaldi, D. In *Computational Chemistry, Reviews of Current Trends*, Leszczynski, J, ed. World Scientific: Singapore, 1995, p. 139. (c) Cramer, C. J.; Truhlar, D.G. *Chem. Rev.* **1999**, *99*, 2161.
 3. Luque, F. J.; Curutchet, C.; Muñoz-Muriedas, J.; Bidon-Chanal, A.; Soteras, I.; Morreale, A.; Gelpí, J. L. and Orozco, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3827.
 4. (a) Tomasi, J.; Cammi, R.; Mennucci, B. *Int. J. Quantum Chem.* **1999**, *75*, 783, (b) Tomasi, J. ; Cammi, R. ; Mennucci, B. ; Capelli, C. ; Corni, S. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5967.
 5. (a) Rinaldi, D. ; Ruiz-López, M. F. ; Rivail, J. L. *J. Chem. Phys.* **1983**, *78*, 834. (b) Dillet, V.; Rinaldi, D.; Rivail, J. L. *J. Phys. Chem.* **1994**, *98*, 5034. Tuñón, I.; Ruiz-López, M. F.; Rinaldi, D. ; Bertrán, J. J. *Comput. Chem.* **1996**, *17*, 148.

-
6. (a) Giesen, D. J.; Gu, M. Z.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **1997**, *101*, 2061. (b) Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1998**, *109*, 9117. (c) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **199**, *103*, 9.
7. (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (b) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239. (c) Barone, V.; Cossi, M. and Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210.
8. (a) Orozco, M.; Bachs, M.; Luque, F. J. *J. Comput. Chem.* **1995**, *16*, 563. (b) Luque, F. J.; Zhang, Y.; Aleman, C.; Bachs, M.; Gao, J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 4269. (c) Luque, F. J.; Aleman, C.; Bachs, M.; Orozco, M. *J. Comput. Chem.* **1996**, *17*, 806. (d) Curutchet, C.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2001**, *22*, 1180.
9. (a) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799. (b) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224. (b) Klamt, A.; Jonas, V.; Bürger, T. and Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074.
10. Curutchet, C.; Cramer, C. J.; Truhlar, D. G.; Ruiz-López, M. F.; Rinaldi, D.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2003**, *24*, 284.
11. Wolfenden, R. *Science* **1983**, *222*, 1087.
12. Pearson, R. G. *J. Am. Chem. Soc.* **1986**, *108*, 6109.
13. Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 782.
14. Florian, J.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 5583.
15. (a) Marcus, Y. *Ion Solvation*. Wiley: New York, 1985. (b) Gomer, R.; Tryson, G. J. *Chem. Phys.* **1977**, *66*, 4413. (c) Conway, R. E. *J. Sol. Chem.* **1978**, *7*, 721. (d) Klots, C. E. *J. Phys. Chem.* **1981**, *85*, 3585.
16. (a) Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. *J. Chem. Phys.* **1998**, *109*, 4852. (b) Mejias, J. A.; Lago, S. *J. Chem. Phys.* **2000**, *113*, 7306. (c)

-
- Zhan, C.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 11534 (d) Grabowski, P.; Riccardi, D.; Gomez, M.A.; Asthagiri, D.; Pratt, L. R. *J. Phys. Chem. A* **2002**, *106*, 9147.
17. Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V. and Tuttle Jr, T. R. *J. Phys. Chem. A* **1998**, *102*, 7787.
18. (a) Tuttle, T. R., Jr.; Malaxos, S.; Coe, J. V. *J. Phys. Chem. A* 2002, *106*, 925. (b) Coe, J. V. *Int. Rev. Phys. Chem.* **2001**, *20*, 33.
19. Pliego Jr., J. R.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622.
20. Orozco, M.; Luque, F. J. *Chem. Phys.* **1994**, *182*, 237.
21. Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717.
22. Claverie, P. In *Intermolecular Interactions: From Diatomics to Biomolecules*, B. Pullman (Ed.), Wiley, Chichester, 1978.
23. (a) Pascual-Ahuir, J. L.; Silla, E. ; Tomasi, J. ; Bonaccorsi, R. *J. Comput. Chem.* **1987**, *8*, 778. (b) Pascual-Ahuir, J. L.; Silla, E. *J. Comput. Chem.* **1990**, *11*, 1047. (c) Silla, E. ; Tuñón, I. ; Pascual-Ahuir, J. L. *J. Comput. Chem.* **1991**, *9*, 1077.
24. Colominas, C.; Luque, F. J. ; Teixidó, J. ; Orozco, M. *Chem. Phys.* **1999**, *240*, 253.
25. Rashin, A. A.; Namboodiri, K. *J. Phys. Chem.* **1987**, *91*, 6003.
26. (a) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151. (b) Cossi, M.; Mennucci, B.; Pitarch, J.; Tomasi, J. *J. Comput. Chem.* **1998**, *8*, 833.
27. (a) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Rignalda, M.; Goddard, W. A., III; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875. (b) Marten, B.; Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Rignalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775. (c) Chen, J. L.; Noodleman, L.; Case, D. A.; Bashford, D. *J. Phys. Chem.* **1994**, *98*, 11059. (d) Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Phys. Chem. A* **1997**, *101*, 10075.

-
28. (a) Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1997**, *101*, 5573. (b) Gogonea, V.; Merz, K. M., Jr. *J. Phys. Chem. A* **1999**, *103*, 5171.
29. Peterson, M.; Poirier, R. MonsterGauss; Department of Biochemistry, Univ. of Toronto, Canada. Version modified by Cammi, R.; Tomasi, J. (1987); and by Curutchet, C.; Orozco, M.; Luque, F.J. (2003).
30. (a) Cancés, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032. (b) Mennucci, B.; Cancés, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, *101*, 10506. (c) Cancés, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1998**, *109*, 260.
31. Gaussian 98 (Revision A.11), 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, 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. 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. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle; J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2001.
32. (a) Luque, F. J.; Orozco, M.; Bhadane, P.K.; Gadre, S.K. *J. Chem. Phys.* **1994**, *100*, 6718. (b) Luque, F. J.; Gadre, S.R.; Bhadane, P.K.; Orozco, M. *Chem. Phys. Letters* **1995**, *232*, 509. (c) Mestres, J.; Solà, M.; Carbó, R.; Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 606.
33. Cappelli, B.; Mennucci, B.; Silva, C. O. D. ; Tomasi, J. *J. Chem. Phys.* **2000**, *112*, 5382.

-
34. (a) Topol, I. A.; Tawa, G. J.; Burt, S. K.; Rashin, A. A. *J. Chem. Phys.* **1999**, *111*, 10998. (b) Pliego, J. R.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241. (c) Pliego, J. R.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *116*, 7434. (d) Zhan, C.-G.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 9737. (e) Leontyev, I. V.; Vener, M. V.; Rostov, I. V.; Basilevsky, M. V.; Newton, M. D. *J. Chem. Phys.* **2003**, *119*, 8024. (f) Vener, M. V.; Leontyev, I. V.; Basilevsky, M. V. *J. Chem. Phys.* **2003**, *119*, 8038. (g) Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. *J. Am. Chem. Soc.* **2004**, *126*, 1285.
35. (a) Aguilar, M. A.; Olivares del Valle, F. J. *Chem. Phys.* **1989**, *120*, 439. (b) Olivares del Valle, F. J.; Aguilar, M. A. *J. Comput. Chem.* **1992**, *13*, 115. (c) Takahashi, O.; Sawahata, H.; Ogawa, Y.; Kikuchi, O. *J. Mol. Struct. (THEOCHEM)* **1997**, **393**, 141. (d) Camaioni, D. M.; Dupuis, M.; Bentley, J. J. *J. Phys. Chem. A* **2003**, *107*, 5778.
36. Stefanovich, V.; Truong, T. N. *Chem. Phys. Lett.* **1995**, *240*, 65.
37. (a) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098.
38. (a) Chipman, D. M. *J. Chem. Phys.* **2002**, *116*, 10129. (b) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413.
39. (a) Latimer, W. M.; Pitzer, K. S.; Slansky, C. M. *J. Chem. Phys.* **1939**, *7*, 108. (b) Rashin, A. A.; Honig, B. *J. Phys. Chem.* **1985**, *89*, 5588.
40. Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. *J. Phys. Chem. A* **1998**, *102*, 6706.
41. Liptak, M.D.; Shields, G.C. *J. Am. Chem. Soc.* **2001**, *123*, 7314.
42. Liptak, M.D.; Gross, K.C.; Seybold, P.G.; Feldgus, S.; Shields, G.C. *J. Am. Chem. Soc.* **2002**, *124*, 6421.
43. Chen, I.; MacKerell Jr, A.D. *Theor. Chem. Acc.* **2000**, *103*, 483.

-
44. (a) Colominas, C.; Orozco, M.; Luque, F. J.; Borrell, J. I.; Teixidó, J. *J. Org. Chem.* **1998**, *63*, 4947. (b) Kličić, J.J.; Friesner, R.A.; Liu, S.; Guida, W.C. *J. Phys. Chem. A* **2002**, *106*, 1327. (c) Adam, K. R. *J. Phys. Chem. A* **2002**, *106*, 11963. (d) Pliego, J. R., Jr. *Chem. Phys. Letters* **2003**, *367*, 145. (e) Klamt, A.; Eckert, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380.
45. (a) Cumming, J. P.; Kebarle, P. *Can. J. Chem.* **1978**, *56*, 1. (b) Fujio, M.; McIver, R. T., Jr.; Taft, R. W. *J. Am. Chem. Soc.* **1981**, *103*, 4017. (c) Caldwell, G.; Renneboog, R.; Kebarle, P. *Can. J. Chem.* **1989**, *67*, 611. (d) Cai, J.; Topsom, R. D.; Headley, A. D.; Koppel, I.; Mishima, M.; Taft, R. W.; Veji, S. *J. Mol. Struct. (THEOCHEM)* **1988**, *168*, 141. (e) Taft, R. W.; Topsom, R. D. *Prog. Phys. Org. Chem.* **1987**, *16*, 1.
46. (a) Taagepera, M.; Summerhays, K. D.; Hehre, W. J.; Topsom, R. D.; Pross, A.; Radom, L.; Taft, R. W. *J. Org. Chem.* **1981**, *46*, 891. (b) Taft, R. W. *Prog. Phys. Org. Chem.* **1983**, *14*, 247.
47. Nagy, P. I.; Takács-Novák, K. *J. Am. Chem. Soc.* **1997**, *119*, 4999.
48. Curutchet, C.; Bidon-Chanal, A.; Orozco, M.; Luque, F. J. *Chem. Phys. Letters.* **2004**, *384*, 299.
49. Orozco, M.; Luque, F. J. *J. Comput. Chem.* **1993**, *14*, 587.
50. (a) Fischer, A.; Galloway, W. J.; Vaughan, J. *J. Chem. Soc.* **1964**, 3591. (b) Sargeant, E.P. and Dempsey, B. *Ionisation Constants of Organic Acids in Aqueous Solution*; IUPAC no. 23; Pergamon Press: New York, 1979.

3.3.

**Electrostatic component of solvation:
Comparison of SCRF continuum models**

Carles Curutchet, Cristopher J. Cramer, Donald G. Truhlar,
Manuel F. Ruiz-Lopez, Daniel Rinaldi, Modesto Orozco
and F. Javier Luque

Journal of Computational Chemistry **2003**, 24, 284

Texto del artículo disponible en / Text de l'article disponible a / Available online at:

<http://www3.interscience.wiley.com/cgi-bin/abstract/102524599/ABSTRACT>

Capítulo 4

Resultados sobre solvatación fraccional

4.1.

Transferability of fragmental contributions to the octanol/water partition coefficient: An NDDO-based MST study

Carles Curutchet, Albert Salichs, Xavier Barril, Modesto Orozco and F. Javier Luque

Journal of Computational Chemistry **2003**, 24, 32

Texto del artículo disponible en / Text de l'article disponible a / Available online at:
<http://www3.interscience.wiley.com/cgi-bin/abstract/101521414/ABSTRACT>

4.2.

MST study of group contributions for alkane derivatives: Effect of the charge normalization

Carles Curutchet, Axel Bidon-Chanal, Modesto Orozco
and F. Javier Luque

Chemical Physics Letters **2004**, 384, 299

Texto del artículo disponible en / Text de l'article disponible a / Available online at:

http://www.sciencedirect.com/science?_ob=ArticleURL&_aset=V-WA-A-W-A-MsSAYWW-UUW-U-AAVDYCCAZ-AAAWBZZBAZ-ACDADAZZ-A-U&_rdoc=1&_fmt=summary&_udi=B6TFN-4BBHDG1-N&_coverDate=01%2F26%2F2004&_cdi=5231&_orig=search&_st=13&_sort=d&view=c&_acct=C000043460&_version=1&_urlVersion=0&_userid=793840&md5=a45aaa9c8596e13aac95d08c2e0af4bf

Capítulo 5

Resultados sobre fuerzas de polarización

5.1.

**Energy decomposition in molecular complexes:
Implications for the treatment of polarization in
molecular simulations**

Carles Curutchet, Josep María Bofill, Begoña Hernandez,
Modesto Orozco and F. Javier Luque

Journal of Computational Chemistry **2003**, 24, 1263

Texto del artículo disponible en / Text de l'article disponible a / Available online at:
[http://www3.interscience.wiley.com/cgi-bin/abstract/10453794
2/ABSTRACT](http://www3.interscience.wiley.com/cgi-bin/abstract/104537942/ABSTRACT)

5.2.

Development of distributed models of atomic polarizability for molecular simulations

Carles Curutchet, Ignacio Soteras, Axel Bidon-Chanal, Christophe Chipot, Modesto Orozco and F. Javier Luque

Development of distributed models of atomic polarizability for molecular simulations

Carles Curutchet,¹ Ignacio Soteras,¹ Axel Bidon-Chanal,¹ Cristophe Chipot,^{2*} Modesto Orozco,^{3*} and F. Javier Luque^{1*}

¹ Departament de Fisicoquímica. Facultat de Farmàcia. Universitat de Barcelona. Avgda. Diagonal 643. Barcelona 08028. Spain.

² Equipe de Dynamique des Assemblages Membranaires. Unité Mixte de Recherche CNRS/UHP 7565. Institut Nancéien de Chimie Moléculaire. Université Henri Poincaré – Nancy I. BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France.

³ Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, c/. Martí i Franqués 1, 08028, Barcelona, and Unitat de Modelització Molecular i Bioinformàtica, Institut de Recerca Biomèdica. Parc Científic de Barcelona, c/. Josep Samitier 1, 08028 Barcelona, Spain.

Send correspondence to FJL (e-mail: javier@far1.far.ub.es) or MO (e-mail: modesto@mmb.pcb.ub.es)

Abstract

This study is focused on the development of atomic distributed polarizability models for the treatment of induction effects in molecular simulations. Different models of dipolar polarizabilities are derived for a series of aromatic compounds from the OPEP (Optimal Partitioning of Electrical Properties) suite of programs. These models are obtained by fitting the induction energies computed in a grid surrounding the molecule by using a fast perturbational approach which permits to reproduce the induction energies determined at the MP2/Sadlej level of theory. Isotropic atomic dipole polarizability models are considered by imposing suitable restraints in the fitting procedure, and by exploring all-atom and united-atom models. The goodness of the distributed models is assessed by comparing the molecular polarizabilities recovered from the models and those obtained from MP2/Sadlej calculations. The behaviour of the models is further examined in the context of cation- π interactions by computing the polarization interaction energy of the aromatic systems with a sodium cation represented by a non-polarizable point charge, and by comparing them to MP2/Sadlej QM/MM calculations. In general, excellent results are obtained for isotropic all-atom models, with deviations around 1 kcal/mol. Energy profiles that characterize the approach of the sodium cation towards the center of the aromatic ring of benzene are also compared to QM/MM calculations and demonstrate the goodness of the models for different orientations. Finally, the transferability of the atomic polarizabilities within the series of benzene derivatives is also examined.

Introduction

The assumption that induction effects can be accounted for in an average sense by means of a simple, appropriate parameterization is at the origin of the durable success of pairwise, additive potential energy functions for cost-effective statistical simulations of liquids. An important ingredient in the development of such nonpolarizable, effective force fields consists of increasing artificially the polarity of the participating molecules to compensate for missing, through-space intermolecular induction phenomena.¹ A popular approach for implicit polarization is based upon the observation that, compared to the experimental gas-phase quantities, molecular dipole moments computed at the Hartree-Fock (HF) level of theory, using a split-valence 6-31G(d) basis set, are systematically exaggerated.² In a number of instances where explicit nonadditivity effects can be neglected, intermolecular potentials parametrized with net atomic charges derived from the HF/6-31G(d) electrostatic potential provide a reasonable description of the properties of the liquid,³ although explicit, mutual polarization of the constituent molecules and their environment is clearly absent. The advantage of representing the system using pairwise, additive potential energy functions lies in the cost-effectiveness of the statistical simulation, avoiding the computationally demanding estimation of induced moments at the expense of sampling of the phase space. Such an implicit polarization approach is, however, not equivalent to a rigorous, atomic-level description of the molecular response to a nonuniform, external electric field, thus constituting the inherent limitation of the approach. This is particularly relevant in statistical simulations of ionic species or highly polarizing systems interacting with polarizable ones. Furthermore, explicit polarization has proven to be pivotal in those cases where intramolecular induction effects are sizable.^{4,5}

The massive increase of computational resources witnessed in recent years, pushing back the limitations of statistical simulations, has allowed molecular assemblies of appreciable complexity to be tackled over realistic time-scales, using sophisticated potential energy functions. The generalized availability of such resources has inevitably contributed to the recent interest of nonadditive force fields, opening the way to the development of alternative approaches for modelling induction effects in chemical systems. To a large extent, the success of molecular simulations relying on polarizable potential energy functions is closely coupled to the ability of the modeller to design such functions. Arguably, the most popular scheme for constructing models of atomic polarizabilities is the one proposed by Applequist.⁶ In essence, his method is based on a self-consistent determination of parameters characterizing the atomic polarizability that see each other through screened dipole-dipole interactions. Just like its refined version devised by Thole,⁷ this heuristic approach implies a substantial component of arbitrariness in the parameterization of the model. In contrast, the partitioning scheme put forward by Stone,⁸⁻⁹ combining an earlier formulation of the susceptibility function of the charge density¹⁰ with the distributed multipole analysis (DMA) method,¹¹ is clearly more rigorous. Rational models of distributed polarizabilities are derived from high-quality quantum mechanical (QM) calculations of the response of an isolated molecule to an external perturbation, thereby providing one and two-center charge-charge, charge-dipole, dipole-dipole, etc. parameters. A pivotal aspect of this approach concerns the partitioning of the response charge density into atomic and nonlocal contributions. Closely related, an alternative procedure supplying reasonably transferable, distributed polarizabilities¹² relies on a topological partitioning of the

molecular space into atomic regions, according to the theory of “atoms in molecules” (AIM) devised by Bader.¹³ The derived polarizability parameters have been shown to reproduce accurately, in a given region, the induced moments due to a local electrostatic potential and its successive derivatives experienced at another site. Inherently nonlocal,^{8,9,13} the resulting models include a plethora of terms that rapidly become cumbersome to handle, especially in the context of statistical simulations. For the latter, the number of polarizability components should be reduced significantly to yield compact and tractable sets of parameters, limited to low-order terms, e.g., charge flow and one-center dipole polarizabilities, capable of describing induction phenomena at a minimal cost. As a result, the above methods are not necessarily compatible with such requirements, and the quality of the generated models strongly depends on their level of sophistication. For instance, limiting the description to atomic dipole polarizabilities, accurate reproduction of molecular quantities beyond the order of the dipole is hopeless.

Atomic point charges derived from the quantum mechanical electrostatic potential computed on a grid of points around the molecule¹⁴⁻¹⁶ are largely utilized for parameterizing the Coulomb part of all-purpose force fields. In the spirit of this approach, a number of schemes targeted at the construction of models of distributed polarizabilities have been devised, based on a least-squares fitting procedure to the induction energy.¹⁷⁻²⁰ Their strength resides in the possibility to generate easily compact and flexible sets of polarizability parameters at any given order. Contrasting with the electrostatic potential, directly available from the wave function, the induction energy is somewhat more difficult to attain. In its most straightforward formulation, referred to as finite perturbation (FP),¹⁹ it can be estimated by considering the interaction of the molecule with a nonpolarizable point charge. Whereas potential derived charges imply a single QM calculation to map the entire space around the molecule, discretized in the form of a grid of N_p points, evaluating induction energies for the same grid via FP would require N_p distinct QM calculations, hence constituting the computational bottleneck of the approach. Clearly, the significant computational effort involved in the determination of induction energies, including usually intramolecular electron correlation and employing large basis sets to warrant an appropriate reproduction of the molecular polarizabilities, limits the applicability of FP to sufficiently small, prototypical systems. In light of these conclusions, two alternatives have been developed concurrently, aimed at a faster estimation of induction energies and relying essentially upon a single QM calculation – a *sine qua non* condition for the development of nonheuristic models of distributed polarizabilities. The first scheme, based on second-order perturbation theory (PT), allows the generation of large grids of induction energies derived from one QM calculation at the HF level.^{21,22} In essence, it relies upon an uncoupled form of the HF equations and provides only an approximation of the exact induction energy. Using adequately chosen scaling factors, this perturbative approach has been shown to yield induction energies in good agreement with the FP quantities.²² The second, more elaborate scheme consists of computing induction energies from fully distributed models of polarizabilities obtained within the framework of the AIM theory.¹² The key feature of this approach lies in its ability to map grids of induction energies from a single QM calculation at a sophisticated level of approximation, followed by the topological partitioning of the electron density response into atomic regions.²³

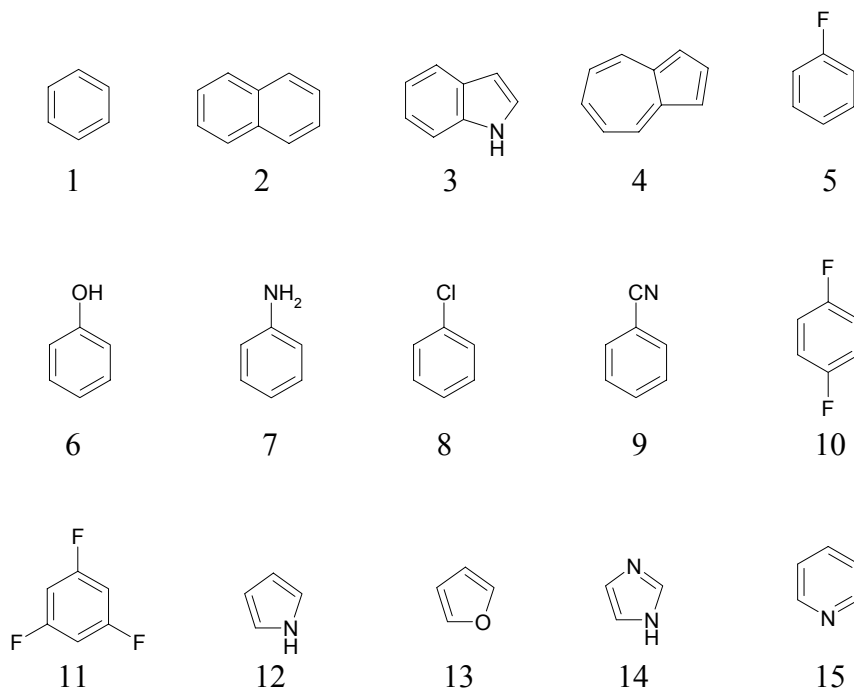
These two approaches have been compared in the evaluation of induction energy grids and the derivation of distributed polarizability models at the MP2 level of theory

and considering different choices of basis set²⁴. The results pointed out that both strategies supply induction energies of comparable quality in terms of error with respect to the reference FP values for neutral molecules, provided that the PT HF/Sadlej induction energy estimates are scaled appropriately to reproduce the MP2 quantities, this approach being more difficult to apply in the case of anions. Nevertheless, the strategy based on second-order perturbation theory constitutes a very appealing alternative due to its limited computational cost and ease of implementation. Moreover, this study showed the importance of sampling the space surrounding the molecule sufficiently far from the nuclei in the computed induction energy grids in order to obtain appropriate distributed models, and the suitability of the Sadlej²⁶ basis set at the MP2 level of theory as an optimal compromise in terms of computational cost and agreement with experimental molecular polarizabilities.

The purpose of this study is to explore the strategy relying on the PT calculation of induction energies for the derivation of distributed polarizability models able to introduce polarization effects in classical force fields in a systematic and well-defined way. As a test case, distributed models are derived for a series of aromatic compounds in order to explore the behaviour of this strategy in the context of cation- π interactions, where polarization effects play an important role²⁵. The models are obtained by considering isotropic dipole atomic polarizabilities, and by exploring both united-atom and all-atom models, and the quality of the fitted parameters is assessed by comparing the molecular polarizabilities recovered from the models and those obtained from MP2/Sadlej calculations. The behaviour of the models are then examined in the context of cation- π interactions by computing the polarization interaction energy of the aromatic systems with a sodium cation represented by a non-polarizable point charge, and by comparing them to MP2/Sadlej QM/MM calculations. In addition, energy profiles that characterize the approach of the sodium cation towards the centre of the aromatic ring of benzene are also compared to QM/MM calculations to explore the goodness of the models for different orientations. Finally, the transferability of the atomic polarizabilities within the series of benzene derivatives is also examined.

Methods

A series of aromatic systems shown in Figure 1 were chosen in order to test the performance of the atomic distributed polarizability models in the prediction of polarization energies in cation- π complexes. The benzene derivatives were optimized at the MP2/6-31G(d,p) level of theory. Then, induction energy grids were generated by using the approach based on second-order perturbation theory. This approach permits the evaluation of the whole grid of induction energies from a single quantum chemical calculation, which was performed at the HF/Sadlej level of theory. Then, the induction energy estimates were scaled by a distance-dependent function in order to reproduce MP2/Sadlej estimates²⁴. This choice is justified as the Sadlej²⁶ basis set used at the MP2 level of theory offers a good compromise between computational cost and agreement with experimental molecular polarizabilities. The grids were constructed by taking about 1000 points going from 2 to 5 times the van der Waals atomic radii of the nearest atom, in order to sample the space surrounding the molecule sufficiently far from the nuclei. Molecular polarizabilities were also computed at the MP2/Sadlej level in order to assess the goodness of the distributed models.

Figure 1. Aromatic structures considered in this study.

Isotropic dipole atomic components were fitted following different criteria in order to derive the distributed models, either by adopting an all-atom (AA) or united-atom (UA) approach. At the UA models, atomic polarizabilities were fitted for all atoms except for hydrogens. In order to test the transferability of the models, three different schemes were used for either AA or UA schemes. First, fitting to the induction energy grids was carried out without any restraint (no restrains - NR). Second, symmetric atoms were forced to share the same polarizability components (symmetry - S). Finally, all carbons and hydrogens pertaining to aromatic CH groups were forced to have the same components at each different system (atom type - AT). The atomic polarizabilities obtained by using these three procedures were subsequently compared and used to discuss the transferability of these values between different aromatic systems.

The final behaviour of the models was analyzed by comparing the fitted molecular polarizabilities with the pure QM estimates. In addition, polarization interaction energies between a sodium cation and the aromatic systems were computed both from the distributed models and from QM/MM calculations, where the sodium cation was simulated as a nonpolarizable point charge. The sodium cation was added at 2.47 Å above the aromatic ring for all systems (see ref. 25). The polarization energies derived from the isotropic distributed models were computed following the procedure proposed first by Applequist²⁷, and subsequently refined by Thole⁷. On the other hand, QM/MM estimates were determined by two different approaches. First, by computing the full variational polarization interaction energy of the point charge with the aromatic system, and second, by using the second-order perturbational approach used to build up the induction energy grids, i.e., by computing perturbationally the HF/Sadlej polarization energy and then scaling it to reproduce MP2/Sadlej values. This final calculation also permits to analyze the performance of the fast second-order perturbational approach in comparison with the more costly variational calculation.

Finally, polarization energy profiles determined for the approach of the sodium cation to the aromatic ring of benzene were calculated using the isotropic all-atom symmetry (AA-S) model and from QM/MM MP2/Sadlej calculations, with the aim to investigate the behaviour of the model at different orientations of the cation.

Results and discussion

Table 1 shows the polarization energies obtained variationally for the cation- π complexes at the MP2/Sadlej level and the values obtained using the perturbational approach mentioned before. The polarization energies determined perturbatively are able to reproduce the variational results with a deviation of only 9% ($r^2=0.97$) and a much lower computational cost, thus giving support to its use as a general strategy for the derivation of atomic polarizabilities adjusted to a given molecule.

Table 1. Variational and perturbational QM/MM MP2/Sadlej polarization interaction energies (kcal/mol) between the aromatic systems and the Na⁺ cation.

Compound	Variational	Perturbational
1	-21.4	-23.8
2	-25.2	-28.3
3	-25.1	-27.6
4	-28.7	-30.4
5	-21.2	-23.7
6	-22.4	-24.9
7	-24.3	-26.5
8	-22.3	-25.2
9	-22.2	-25.3
10	-21.0	-23.5
11	-21.0	-23.7
12	-22.4	-23.4
13	-20.5	-21.3
14	-18.9	-20.9
15	-19.8	-22.3
mse ^a		2.3
mue		2.3
rms		2.5

^a mse: mean signed deviation (kcal/mol); mue: mean unsigned deviation (kcal/mol); rms: root-mean square deviation (kcal/mol)

The comparison of molecular polarizabilities obtained from the different models with the values determined variationally at the MP2/Sadlej level are shown in Table 2. A very close agreement is found when an all-atom description is used, as noted in mean deviations less than 1% when atom-type restraints are imposed in the fitting of atomic polarizabilities. When a united-atom description is used, the agreement between molecular polarizabilities recovered from the distributed models and computed variationally is slightly worse, even though the mean deviation is only 4%. At this point, it must be pointed out that the molecular polarizability was not used as an adjustable parameter in the fitting procedure to derive the distributed models, and so the agreement obtained gives strong support to the method.

Table 2. Trace of the molecular polarizability (a.u.) obtained from the quantum mechanical calculation (MP2/Sadlej) and from different isotropic distributed models (All-atom/United-atom) obtained by imposing different kind of restrains: no restrains (NR), symmetry (S), atom type (AT).

Compound	QM	All-atom			United-atom		
		NR	S	AT	NR	S	AT
1	208.3	214.2	214.2	214.2	230.1	230.6	230.6
2	360.5	345.4	346.3	345.7	355.7	356.3	362.4
3	308.4	301.8	301.8	301.3	309.6	309.6	314.8
4	399.2	366.2	366.5	364.1	375.4	375.5	383.8
5	208.1	212.8	212.9	212.9	219.6	219.9	222.9
6	225.2	230.1	230.1	230.1	237.6	238.1	241.3
7	247.7	253.7	253.7	253.5	261.5	261.8	265.0
8	251.9	256.5	256.6	256.7	263.0	263.1	266.2
9	259.2	257.5	257.6	257.6	265.3	265.6	268.3
10	207.8	211.2	211.3	211.3	213.8	213.9	213.9
11	209.9	214.1	214.6	214.6	214.6	214.9	214.9
12	165.1	175.4	175.6	175.3	191.3	191.3	191.2
13	145.9	154.5	154.7	155.9	170.4	170.4	170.2
14	147.7	156.8	156.8	158.2	171.1	171.1	171.2
15	192.2	196.0	196.0	196.6	213.4	213.4	213.8
mse ^a		-0.6	-0.8	-0.7	-10.3	-10.6	-12.9
mue		8.1	8.1	8.5	14.2	14.3	14.9
rms		11.4	11.3	12.0	17.0	17.1	17.3

^a mse: mean signed deviation (kcal/mol); mue: mean unsigned deviation (kcal/mol); rms: root-mean square deviation (kcal/mol)

Another parameter considered to assess the goodness of the fittings is the deviation found between the induction energies computed quantum-mechanically and the values recovered from the distributed models along the grid points (see Table 3). In general, a good agreement is found for all the models, with small deviations ranging from $0.12 \cdot 10^{-3}$ to $0.27 \cdot 10^{-3}$ (in a.u.).

Table 3. Root-mean square deviation (10^{-3} a.u.) between the quantum mechanically and the recovered induction energies computed at the grids from different isotropic distributed models (All-atom/United-atom) obtained by imposing different kind of restrains: no restrains (NR), symmetry (S), atom type (AT).

Compound	All-atom			United-atom		
	NR	S	AT	NR	S	AT
1	0.11	0.11	0.11	0.36	0.36	0.36
2	0.13	0.14	0.14	0.21	0.22	0.30
3	0.12	0.12	0.13	0.20	0.20	0.27
4	0.17	0.17	0.26	0.24	0.25	0.40
5	0.11	0.11	0.11	0.24	0.24	0.28
6	0.12	0.12	0.12	0.26	0.27	0.31
7	0.13	0.13	0.13	0.26	0.26	0.30
8	0.11	0.11	0.12	0.21	0.21	0.24
9	0.13	0.13	0.13	0.22	0.22	0.25
10	0.10	0.10	0.10	0.17	0.17	0.17
11	0.11	0.11	0.11	0.12	0.12	0.12
12	0.11	0.11	0.11	0.25	0.25	0.25
13	0.10	0.10	0.11	0.23	0.23	0.23
14	0.11	0.11	0.13	0.25	0.25	0.25
15	0.11	0.11	0.12	0.36	0.36	0.36
Mean	0.12	0.12	0.13	0.24	0.24	0.27

In order to further investigate the behaviour of the models, the polarization energy of the aromatic systems interacting with the Na⁺ cation (placed at its optimum position to establish a cation- π interaction) was computed from the distributed models and compared to the variational MP2/Sadlej QM/MM calculations. Table 4 shows the results computed by neglecting the coupling between atomic induced dipoles, i. e., by considering only the electric field generated by the point charge of the sodium cation. On the other hand, Table 5 shows the results determined by computing the electric field generated by both the sodium cation and by the atomic induced dipoles following an iterative procedure, which usually converges in one or two cycles. The polarization energies of the different models with the Na⁺ cation show a good agreement, especially when an all-atom description is used. In this case, the differences between the polarization energies obtained at the QM and classical levels are lower than 2%, while this difference increases to 15% when united-atom models are considered. The results also indicate that the mutual coupling between the induced dipoles for the different cation- π complexes have a relatively small effect in the computed induction energies. These results suggest, then, that the mutual coupling could be neglected, thus avoiding the self-consistent procedure that must be carried out to reach convergence between the induced dipoles, with the corresponding saving in computational cost specially for simulations of complex systems.

Table 4. Polarization interaction energy (kcal/mol) between the Na⁺ cation and the aromatic systems determined from the variational quantum mechanical calculation and from different isotropic distributed models (All-atom/United-atom) obtained by imposing different kind of restrains (no restrains, symmetry, atom type). Energies computed without coupling between atomic induced dipoles.

Compound	QM	All-atom			United-atom		
		NR	S	AT	NR	S	AT
1	-21.4	-22.2	-22.3	-22.3	-29.1	-29.2	-29.2
2	-25.2	-23.2	-22.9	-23.6	-20.9	-20.7	-24.3
3	-25.1	-21.7	-21.7	-23.7	-21.8	-21.8	-24.0
4	-28.7	-24.1	-24.0	-23.0	-19.1	-19.1	-22.0
5	-21.2	-21.4	-21.5	-21.5	-24.2	-24.2	-25.6
6	-22.4	-22.5	-22.6	-22.6	-25.6	-25.7	-27.0
7	-24.3	-23.8	-23.9	-24.0	-26.5	-26.5	-27.6
8	-22.3	-22.4	-22.4	-22.4	-25.0	-25.0	-25.9
9	-22.2	-23.4	-23.4	-23.4	-23.5	-23.5	-24.1
10	-21.0	-20.6	-20.7	-20.7	-21.7	-21.7	-21.7
11	-21.0	-20.4	-20.5	-20.5	-20.4	-20.5	-20.5
12	-22.4	-21.6	-21.7	-21.5	-27.9	-27.9	-27.9
13	-20.5	-19.0	-19.1	-19.5	-25.0	-25.0	-25.0
14	-18.9	-19.3	-19.3	-19.8	-25.3	-25.3	-25.4
15	-19.8	-20.5	-20.5	-20.8	-27.4	-27.4	-27.4
mse ^a		0.7	0.7	0.5	-1.8	-1.8	-2.7
mue		1.2	1.2	1.1	4.2	4.2	4.0
rms		1.8	1.8	1.8	5.1	5.1	4.8

^a mse: mean signed deviation (kcal/mol); mue: mean unsigned deviation (kcal/mol); rms: root-mean square deviation (kcal/mol)

Table 5. Polarization interaction energy (kcal/mol) between the Na⁺ cation and the aromatic systems determined from the variational quantum mechanical calculation and from different isotropic distributed models (All-atom/United-atom) obtained by imposing different kind of restrains (no restrains, symmetry, atom type). Energies computed coupling atomic induced dipoles.

Compound	All-atom			United-atom			
	QM	NR	S	AT	NR	S	AT
1	-21.4	-22.5	-22.6	-22.6	-29.1	-29.2	-29.2
2	-25.2	-23.2	-22.7	-23.4	-20.9	-20.7	-24.3
3	-25.1	-21.8	-21.8	-23.8	-21.8	-21.8	-24.0
4	-28.7	-24.2	-24.1	-23.4	-19.1	-19.1	-22.0
5	-21.2	-22.0	-22.1	-21.9	-26.9	-26.9	-26.7
6	-22.4	-22.7	-22.8	-22.7	-25.6	-25.7	-27.0
7	-24.3	-23.9	-24.0	-24.1	-26.5	-26.5	-27.6
8	-22.3	-23.5	-23.5	-23.2	-28.3	-28.3	-27.7
9	-22.2	-23.7	-23.7	-23.8	-22.2	-22.1	-22.7
10	-21.0	-21.1	-21.2	-21.2	-23.0	-23.2	-23.2
11	-21.0	-21.0	-21.1	-21.1	-21.6	-21.7	-21.7
12	-22.4	-22.2	-22.2	-22.0	-27.9	-27.9	-27.9
13	-20.5	-19.6	-19.7	-19.9	-25.0	-25.0	-25.0
14	-18.9	-23.7	-23.7	-26.0	-33.0	-33.0	-33.0
15	-19.8	-21.2	-21.3	-21.2	-27.4	-27.4	-27.4
mse ^a		0.0	0.0	-0.3	-2.8	-2.8	-3.5
mue		1.5	1.6	1.5	5.1	5.2	4.7
rms		2.2	2.3	2.6	6.4	6.5	6.1

^a mse: mean signed deviation (kcal/mol); mue: mean unsigned deviation (kcal/mol); rms: root-mean square deviation (kcal/mol)

Finally, the energy profiles that characterize the approach of the sodium cation to benzene were calculated from the isotropic all-atom symmetry (AA-S) model and from QM/MM calculations. As shown in Figure 2, approximation profiles were computed along three orientations: (i) along the middle of a C-C bond (x orientation), (ii) along the C-H bond (y orientation), and (iii) perpendicular to the center of the aromatic ring (z orientation). In all cases, polarization energies at the “hard particle radius” (defined as the point where dispersion and repulsion cancel each other, see ref. 25) and in points closer to and further away from the molecule were computed. The results given in Figure 3 show the good behaviour encountered for the model along these three orientations.

Figure 2. Orientations of cation-aromatic interactions in this study. The z axis corresponds to the cation- π interaction.

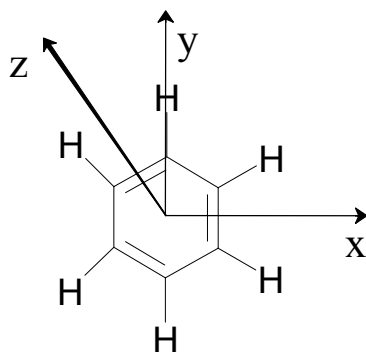
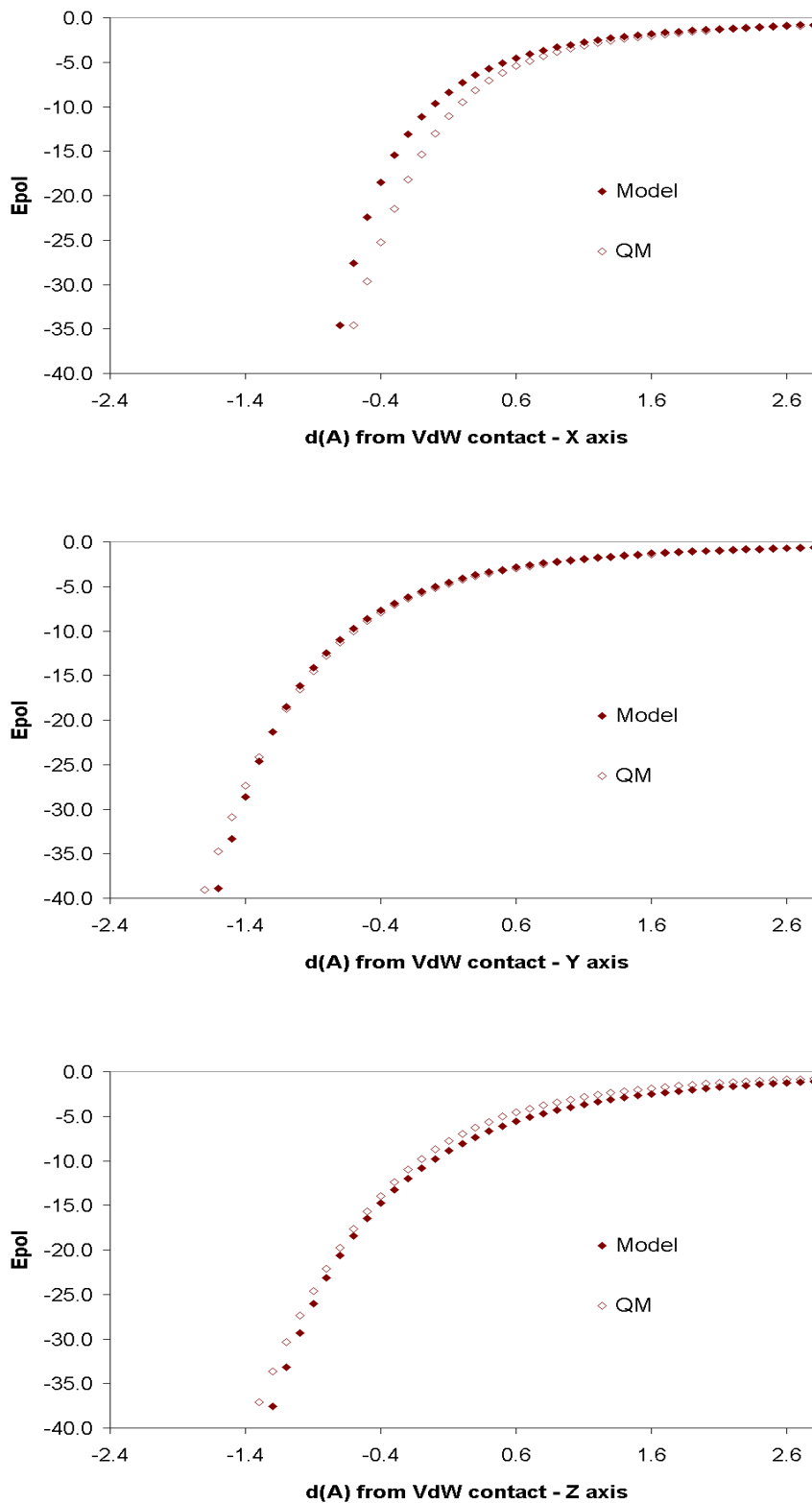


Figure 3. Polarization energy profiles that characterize the approach of Na⁺ to benzene determined from QM/MM MP2/Sadlej calculations and from the isotropic all-atom symmetry model of distributed polarizabilities. Distances (in Å) are referred in all cases to the hard sphere radius, the point where dispersion and repulsion cancel each other. Hard sphere radii are: 4.58 Å (x axis), 4.97 Å (y axis), and 3.24 Å (z axis). Profiles computed without coupling between atomic induced dipoles.



In order to further investigate the transferability of atomic polarizability components between the different systems considered, the atomic polarizabilities obtained for the C and H atoms (all-atom) and the C-H group (united-atom) in the different distributed models are shown in Table 6.

Table 6. Isotropic polarizability components (a.u.) for C and H pertaining to aromatic CH groups derived from different isotropic models. Mean values and standard deviations are given for the whole series of benzene derivatives considered in the study.

All-atom	α_C	NR	10.2	6.8
		S	10.4	5.9
		AT	9.6	2.3
	α_H	NR	3.5	1.1
		S	3.4	0.9
		AT	3.6	0.4
United-atom	α_C	NR	19.1	8.8
		S	19.1	8.4
		AT	16.7	8.1

Inspection of the mean values obtained for the series of complexes shows that no relevant differences are found in the atomic polarizabilities obtained either without restraints or by imposing symmetry or atom-type equivalences. The atomic polarizability of C and H oscillates (in a.u.) between 9.6 and 10.4, and between 3.4 and 3.6, and for the C-H group adopts values between 16.7 and 19.1. However, it must be noted that notably large standard deviations are found, which suggests the existence of marked differences in the atomic polarizabilities obtained for these atoms in different molecules. This is clearly reflected in the results shown in Tables 7 (all-atom) and 8 (united-atom), where the mean atomic polarizabilities obtained for each compound are given. For instance, the C polarizability obtained without restraints oscillates between 4.0 and 18.2 (a.u.), while the C-H polarizability oscillates between 12.7 and 31.5. These results seriously question the assumption that distributed models of atomic polarizabilities can be transferable between structurally related molecules.

Table 7. Isotropic polarizability components (a.u.) for C and H pertaining to aromatic CH groups derived from different all-atom isotropic models. For the atom-type model (AT) components derived for each aromatic system are given. In cases where different values can be found at each aromatic system, i.e. no restraints (NR) and symmetry (S) models, mean, maximum and minimum values, along with standard deviations are given for each compound.

	Atom-type	No restraints				Symmetry			
	α_C	α_C				α_C			
		Mean	Max	Min	Dev	Mean	Max	Min	Dev
1	8.2	8.1	8.9	5.8	1.1	8.2	8.2	8.2	0.0
2	10.8	11.5	23.8	-0.6	9.1	12.5	13.6	11.5	1.1
3	11.0	11.6	21.3	4.2	6.5	11.6	21.3	4.2	6.5
4	11.5	16.1	29.9	2.3	9.3	16.5	30.2	7.8	8.8
5	8.6	9.8	13.7	4.4	3.9	9.7	12.1	7.5	2.3
6	8.6	9.9	13.8	3.8	4.7	9.9	12.4	6.7	3.0
7	9.5	9.9	14.0	5.5	3.4	9.9	12.4	8.2	1.7
8	8.6	10.0	15.9	1.2	6.2	10.0	13.5	4.8	4.7
9	8.6	9.2	12.7	4.0	4.2	9.2	11.0	6.8	2.2
10	10.9	10.6	12.3	8.5	1.7	10.9	10.9	10.9	0.0
11	18.5	18.2	19.7	16.4	1.6	18.5	18.5	18.5	0.0
12	8.4	8.4	12.7	4.4	3.6	8.4	9.9	6.8	1.8
13	7.6	6.5	13.2	-0.7	6.8	6.6	12.3	0.9	6.6
14	7.2	4.0	6.9	-1.8	5.0	4.0	6.9	-1.8	5.0
15	6.8	5.3	19.2	-4.9	12.2	5.3	18.5	-4.0	12.0

	α_H	α_H				α_H			
		Mean	Max	Min	Dev	Mean	Max	Min	Dev
1	3.7	3.8	4.7	3.5	0.5	3.7	3.7	3.7	0.0
2	3.5	3.5	6.1	1.4	1.6	3.1	3.2	2.9	0.2
3	3.4	3.3	5.3	0.9	1.5	3.3	5.3	0.9	1.5
4	3.9	2.6	5.0	0.4	1.4	2.5	3.6	0.4	1.3
5	3.6	3.4	4.5	3.0	0.6	3.4	3.6	3.2	0.2
6	3.8	3.6	4.6	2.9	0.6	3.6	3.7	3.2	0.2
7	3.7	3.7	4.5	3.0	0.6	3.7	4.0	3.2	0.3
8	3.7	3.5	4.9	2.7	0.9	3.5	4.0	2.9	0.5
9	3.6	3.6	4.7	3.0	0.7	3.5	3.7	3.3	0.2
10	3.2	3.3	3.8	2.9	0.4	3.2	3.2	3.2	0.0
11	2.1	2.2	2.8	1.6	0.6	2.1	2.1	2.1	0.0
12	3.7	3.6	4.7	2.8	0.8	3.6	4.2	3.1	0.6
13	3.5	3.7	5.6	2.1	1.7	3.7	5.0	2.3	1.6
14	3.8	4.4	5.6	3.4	1.2	4.4	5.6	3.4	1.2
15	4.0	4.1	5.6	2.2	1.6	4.1	5.4	2.3	1.6

Table 8. Isotropic polarizability components (a.u.) for C pertaining to aromatic CH groups derived from different united-atom isotropic models. For the atom-type model (AT) components derived for each aromatic system are given. In cases where different values can be found at each aromatic system, i.e. no restrains (NR) and symmetry (S) models, mean, maximum and minimum values, along with standard deviations are given for each compound.

	Atom-type	No restrains				Symmetry			
	α_C	α_C				α_C			
		Mean	Max	Min	Dev	Mean	Max	Min	Dev
1	12.8	12.8	15.9	10.5	2.1	12.8	12.8	12.8	0.0
2	18.9	22.0	33.6	10.1	8.4	22.2	29.0	15.3	7.3
3	18.4	21.4	32.6	10.5	8.9	21.4	32.6	10.5	8.9
4	19.9	24.2	36.5	16.1	8.0	24.1	34.8	16.3	7.7
5	14.9	19.0	31.3	6.0	11.2	19.0	29.7	8.8	10.5
6	15.1	19.7	36.1	3.2	13.4	19.6	31.8	7.6	12.1
7	16.2	20.7	34.5	5.3	13.3	20.7	33.1	8.1	12.5
8	15.0	18.8	29.6	7.6	10.5	18.8	29.3	8.5	10.4
9	15.9	19.1	28.8	9.7	8.4	19.1	27.4	12.0	7.8
10	21.9	21.8	25.7	17.7	3.9	21.9	21.9	21.9	0.0
11	31.5	31.5	31.9	31.1	0.4	31.5	31.5	31.5	0.0
12	13.5	14.0	15.5	12.2	1.5	14.0	15.1	12.8	1.4
13	12.8	13.0	13.8	12.4	0.7	13.0	13.6	12.4	0.7
14	12.9	12.6	13.3	12.1	0.6	12.6	13.3	12.1	0.6
15	12.7	12.4	17.3	6.8	4.5	12.4	17.1	6.8	4.6

Concluding remarks

The procedure described in this study constitutes a promising alternative for the derivation of distributed models of atomic polarizabilities. The models obtained are able to reproduce either the polarization energies computed variationally or the global polarizability of the molecule. No relevant differences are found in the magnitude of the atomic polarizabilities due to the use of symmetry or atom type-based restraints compared to the values derived without imposing any restraint. On the other hand, marked differences are found in the values obtained for different molecules, thus difficulting the definition of a single set of atomic values transferable between molecules. These findings suggest that the procedure adopted can be an efficient alternative for the definition of distributed models of atomic polarizabilities.

References

1. Berendsen, H. J. C.; Crigera, J. R. and Straatsma, T. P. *J. Phys. Chem.* **1987**, 91, 6269
2. Hehre, W. J.; Radom, L.; Schleyer, P. v. R. and Pople, J. A. *Ab initio molecular orbital theory*; Wiley Interscience: New York, **1986**
3. Carlson, H. A.; Nguyen, T. B.; Orozco, M. and Jorgensen, W. L. *J. Comput. Chem.* **1993**, 14, 1240
4. Colonna, F.; Evleth, E. *Chem. Phys. Lett.* **1993**, 212, 665
5. Chipot, C.; Ángyán, J. G.; Maigret, B. and Scheraga, H. A. *J. Phys. Chem.* **1993**, 97, 9797
6. Applequist, J. *Acc. Chem. Res.* **1977**, 10, 79
7. Thole, B. T. *J. Chem. Phys.* **1981**, 59, 341
8. Stone, A. J. *Mol. Phys.* **1985**, 56, 1065
9. Le Sueur, C. R. and Stone, A. J. *Mol. Phys.* **1994**, 83, 293
10. Maaskant, W. J. A. and Oosterhof, L. J. *Mol. Phys.* **1964**, 8, 319
11. Stone, A. J. *Chem. Phys. Lett.* **1981**, 83, 233
12. Ángyán, J. G.; Jansen, G.; Loos, M.; Hättig, C. and Hess, B. A. *Chem. Phys. Lett.* **1994**, 219, 267
13. Bader, R. F. W. *Atoms in Molecules – A Quantum Theory*; Oxford University Press: London **1990**.
14. Bonacorsi, R.; Petrongolo, C. and Scrocco, J. *Theor. Chim. Acta* **1971**, 20, 331
15. Momany, F. A. *J. Phys. Chem.* **1978**, 82, 592
16. Cox, S. R. and Williams, D. E. *J. Comput. Chem.* **1981**, 2, 304
17. Nakagawa, S. and Kosugi, N. *Chem. Phys. Lett.* **1993**, 210, 180
18. Alkorta, I.; Bachs, M. and Perez, J. J. *Chem. Phys. Lett.* **1994**, 224, 160
19. Celebi, N.; Ángyán, J. G.; Dehez, F.; Millot, C. and Chipot, C. *J. Chem. Phys.* **2000**, 112, 2709
20. Dehez, F.; Soetens, J. C.; Chipot, C.; Ángyán, J. G. and Millot, C. *J. Phys. Chem. A* **2000**, 104, 1293
21. Luque, F. J. and Orozco, M. *J. Comput. Chem.* **1998**, 19, 866
22. Chipot, C. and Luque, F. J. *Chem. Phys. Lett.* **2000**, 332, 190
23. Dehez, F.; Chipot, C.; Millot, C. and Ángyán, J. G. *Chem. Phys. Lett.* **2001**, 338, 180

24. Chipot, C.; Dehez, F.; Ángyán, J.; Millot, C.; Orozco, M. and Luque, F. J. *J. Phys. Chem. A* **2001**, 105, 11505
25. Cubero, E.; Luque, F. J. and Orozco, M. *Proc. Natl. Acad. Sci. USA* **1998**, 95, 5976
26. Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, 53, 1995
27. Applequist, J.; Carl, J. R. and Fung, K. *J. Am. Chem. Soc.* **1972**, 94, 2952

Capítulo 6

Discusión de los resultados

6.1. Modelos continuos de solvatación

Los modelos continuos SCRF de solvatación [1, 2] constituyen una herramienta computacional particularmente interesante para el estudio de procesos químicos en solución. Ello obedece esencialmente a tres factores. En primer lugar, su reducido coste computacional en relación a técnicas que utilizan una descripción discreta del solvente. En segundo lugar, los diversos formalismos SCRF son formalmente simples, están bien establecidos en la actualidad, y han sido ampliamente validados en su aplicación a diversos procesos químicos. Finalmente, el uso de estas técnicas permite no únicamente estimar la energía libre de solvatación, sino asimismo predecir la influencia del solvente sobre la densidad electrónica y nuclear del soluto y por tanto el cambio inducido por el solvente en sus propiedades [3].

A pesar del potencial de las técnicas SCRF, su uso en la predicción e interpretación de procesos químicos en solución debe llevarse a cabo con cautela. Ello es debido a las simplificada descripción de la interacción soluto-solvente que subyace en un modelo continuo, a las particularidades formales del modelo físico y matemático utilizado para evaluar el acoplamiento entre soluto y solvente, y finalmente a las características del proceso de parametrización implícito al uso de este tipo de modelos.

Aprovechando la naturaleza de función de estado de la energía libre, los métodos continuos calculan la energía libre de solvatación partiendo el proceso de solvatación del soluto en tres subprocesos a los que asignan naturaleza física: un término de cavitación, otro electrostático y finalmente uno que comprende dispersión y repulsión. En muchos algoritmos continuos, los términos de cavitación y dispersión-repulsión se suelen agrupar como "términos no-electrostáticos". Cabe señalar que si bien la partición del proceso de solvatación en tres subprocesos es perfectamente rigurosa, la asignación de significado físico a los mismos no lo

es, al no tener presente el acoplamiento entre las distintas contribuciones la energía libre de solvatación. El uso pues de esta aproximación aditiva es una de las primeras simplificaciones inherentes a los métodos continuos.

La evaluación de los términos electrostático y no-electrostático conlleva implícitamente la definición arbitraria de la frontera entre soluto y solvente. Asociada a dicha definición cabe plantear una primera cuestión relativa a la conveniencia de definir una cavidad común a ambos términos, o bien asignar cavidades diferentes para cada uno de ellos. Por otra parte, dado que el término no-electrostático incluye componentes ligados a interacciones soluto-solvente de corta distancia, se asume generalmente una relación proporcional entre dichos componentes y la superficie de la cavidad que encierra al soluto. En cuanto al término electrostático, se considera que la respuesta del solvente viene descrita por la constante dieléctrica del solvente puro, lo cual sólo sería estrictamente correcto a distancias suficientemente grandes con respecto al soluto, pues la perturbación creada éste modifica la propiedades de las moléculas de solvente asociadas a las capas de hidratación más próximas al soluto, especialmente en el caso de especies iónicas. Además, sólo el término electrostático es evaluado de forma autocoherente en base al acoplamiento entre el campo de reacción del solvente y la densidad de carga del soluto, mientras que dicho proceso autocoherente no suele aplicarse al término no-electrostático. Por ello, la parametrización de un modelo continuo necesariamente implica un balance entre los términos electrostático y no-electrostático, de forma que éste último debe corregir necesariamente aquellas *imprecisiones* asociadas en la definición y evaluación de la interacción electrostática entre soluto y solvente. Esta conexión implícita entre términos electrostáticos y no-electrostáticos introducida en el proceso de parametrización del método puede introducir sesgos que conduzcan a un desbalance físico entre ambas contribuciones. Por ello, si bien los métodos continuos de última generación garantizan buenas aproximaciones a la energía libre de solvatación, no siempre está garantizado el peso relativo de las contribuciones electrostáticas y no-electrostáticas.

Los trabajos recogidos en esta primera parte de la discusión (artículos 1-3) abordan precisamente gran parte de la problemática expuesta anteriormente. En ellos se describe el proceso adoptado para la parametrización del método continuo MST elaborado en el grupo de Barcelona a la descripción de solutos en n-octanol, y la revisión de la parametrización efectuada previamente para la solvatación en agua, cloroformo y tetracloruro de carbono. Tomados en su conjunto, dicha parametrización permite cubrir una gama diversa de solventes, donde cabe esperar diferencias significativas en la contribución relativa de los términos electrostático y no electrostático a la energía libre de solvatación. Por otra parte, se abordará también la adaptación del formalismo MST a la hidratación de especies iónicas monovalentes. Finalmente, se comentarán los resultados obtenidos en el análisis comparativo del término electrostático obtenido a partir de tres formalismos diferentes, que comprenden la expresión de Born gen-

eralizada (modelo SMX elaborado en Minnesota), la expansión multipolar (modelo MPE de Nancy), y el formalismo de carga aparente (modelo MST).

Artículo 1. Parametrización del modelo MST en n-octanol

En el primer trabajo se presenta la parametrización de la versión HF/6-31G(d) del modelo continuo MST para la predicción de energías libres de solvatación en n-octanol, así como la predicción de coeficientes de partición octanol/agua. También se muestra el refinado de la parametrización de las versiones del modelo MST para agua, cloroformo y tetracloruro de carbono.

Debido a que la componente electrostática (ΔG_{ele}) a la energía libre de solvatación depende de forma muy importante de la definición de la cavidad, se ha seguido la misma estrategia que en las anteriores parametrizaciones para definir la interfase soluto/solvente óptima para n-octanol. Así, se ha realizado una comparación de los resultados obtenidos con el modelo continuo MST y aquellos proporcionados a partir de simulaciones discretas de solutos en n-octanol. Concretamente, se han realizado dos tipos diferentes de análisis. En primer lugar, se han efectuado cálculos de energía libre acoplados a simulaciones de Monte Carlo con el fin de obtener una estimación de la contribución ΔG_{ele} en octanol para una serie de compuestos representativos de diversos grupos funcionales. En segundo lugar, se han estimado los aumentos en el momento dipolar ($\Delta\mu$) inducidos por el solvente mediante cálculos QM/MM a nivel HF/6-31G(d) a partir de muestreos de equilibrio obtenidos por cálculos de Monte Carlo. La comparación de estos resultados discretos de ΔG_{ele} y de $\Delta\mu$ con los correspondientes resultados MST han permitido determinar la mejor interfase soluto/solvente electrostática.

Tal como se muestra en la Figura 1 del primer artículo, el factor de escalado (κ) que multiplica los radios de van der Waals para construir la cavidad electrostática toma un valor óptimo para n-octanol alrededor de 1.5. Este valor ajustado empíricamente es razonable, ya que representa un compromiso entre los valores parametrizados anteriormente para agua (1.25) y para cloroformo (1.6) y tetracloruro de carbono (1.8), valores obtenidos utilizando el mismo protocolo de parametrización descrito anteriormente.

La parametrización del término de van der Waals (ΔG_{vdW}) se ha realizado a continuación mediante ajuste a los valores experimentales de energías de solvatación en octanol para una amplia serie de 58 compuestos. Adicionalmente, los parámetros obtenidos se han refinado mediante la comparación de coeficientes de partición octanol/agua experimentales y calculados de 75 compuestos. Respecto al refinado del modelo en agua, cloroformo y tetracloruro de carbono, se han usado series de compuestos más amplias que las utilizadas en las parametrizaciones

originales del método MST (69, 54 y 47, respectivamente). Asimismo, se han introducido algunas variaciones respecto a las versiones anteriores, como la ampliación del número de tipos de átomos usados en el término de van der Waals, o la introducción de un término de corrección para los compuestos con grupos nitrilo. Estas modificaciones han sido suficientes para reproducir satisfactoriamente las energías de solvatación en cloroformo, octanol y tetracloruro de carbono de una amplia serie de moléculas. Para representar la energía libre de hidratación ha sido preciso además introducir radios diferentes para nitrógeno, azufre y oxígeno, dependiendo del número de hidrógenos a que están enlazados.

El conjunto de radios atómicos utilizados para construir la cavidad en cada uno de dichos solventes, así como las correspondientes tensiones superficiales atómicas, están recogidas en las Tablas I y II del artículo. La estrategia de parametrización definida en el marco del modelo MST implica un número reducido de parámetros. Así, aparte de los radios de van der Waals asignados a cada tipo de átomo (12 valores para el agua y 8 valores para el resto de los solventes), la definición de la cavidad empleada en la evaluación del término electrostático hace uso de un único parámetro, correspondiente al factor de escala, cuyo valor depende de la naturaleza del solvente. Por otra parte, la evaluación del término no-electrostático utiliza la densidad numérica y el diámetro molecular de las moléculas en cuanto al componente de cavitación, calculado de acuerdo con el formalismo de Pierotti, y un total de 12 tensiones superficiales atómicas más un término de corrección para el grupo nitrilo en cuanto concierne al componente de van der Waals.

El análisis estadístico de las energías libres de solvatación calculadas a partir de los cálculos MST-HF/6-31G(d) para las series de compuestos se muestran en la Tabla III del primer artículo. En conjunto, se obtienen resultados muy satisfactorios para todos los solventes. La desviación cuadrática media entre los valores experimentales y calculados es de 0.7, 0.6, 0.4 y 0.3 kcal/mol para la solvatación en agua, octanol, cloroformo y tetracloruro de carbono, respectivamente. En octanol, la contribución electrostática representa en promedio aproximadamente el 50 % de la energía de solvatación de moléculas polares. Este valor es claramente inferior a la contribución en agua, donde es el término dominante (ΔG_{ele} contribuye alrededor del 130 % de ΔG_{sol}), pero notablemente superior a la contribución en cloroformo y tetracloruro de carbono, donde representan cerca del 35 % y 15 % del total, respectivamente. Por otra parte, la variación del momento dipolar también refleja claramente las diferencias en ΔG_{ele} obtenidas para los distintos solventes. Así, la solvatación de solutos polares en octanol conlleva un incremento en el momento dipolar correspondiente a un 16 % con respecto al valor en fase gas, que contrasta con el cambio de 30 % detectado en agua y con las modestas variaciones del 12 y 6 % reportadas en cloroformo y tetracloruro de carbono.

La contribución no electrostática (cavitación y van der Waals) a la solvatación en octanol es negativa, tal como sucede en cloroformo y tetracloruro de carbono, y a diferencia de lo que

ocurre en agua. Esto indica que el trabajo de cavitación es menor (en valor absoluto) a la energía asociada a interacciones dispersivas entre el soluto y las moléculas de n-octanol. Estos resultados evidencian el balance entre las características hidrofílicas y lipofílicas del octanol como solvente, ya que los términos electrostático y no electrostático tienen contribuciones similares a las energías libres de solvatación de solutos polares.

Por otra parte, el modelo MST parametrizado permite predecir los coeficientes de partición octanol/agua de los 75 compuestos con una desviación cuadrática media de 0.4 unidades de logP, y sin errores sistemáticos (véase la Figura 3 del primer artículo). Finalmente, el modelo se ha verificado calculando los coeficientes de partición para un conjunto de 27 fármacos que no habían sido considerados en la parametrización, obteniéndose una desviación cuadrática media de 0.75 unidades de logP sin desviaciones sistemáticas. Estos resultados son altamente satisfactorios, teniendo en cuenta la diversidad estructural de los fármacos considerados y la naturaleza del logP, calculado a partir de las energías libres de solvatación en los dos solventes.

Debido a la diferente naturaleza química de las moléculas de agua, octanol, cloroformo y tetracloruro de carbono, la parametrización del modelo MST a estos cuatro solventes abre un abanico de posibilidades para poder explorar la influencia de la naturaleza del solvente en procesos químicos en solución. En particular, es de gran interés poder conocer la solvatación diferencial en agua y octanol, ya que ello permite profundizar en el origen del reparto de solutos entre ambos solventes, y así poder incidir en una mayor comprensión de procesos relacionados con la biodistribución de fármacos y el reconocimiento entre ligando y receptor.

Artículo 2. Definición de la cavidad MST para compuestos iónicos monovalentes

La energía libre de solvatación de solutos neutros puede ser obtenida experimentalmente a partir de mediciones de la concentración de soluto en fase gas y en disolución, una vez se ha alcanzado el punto de equilibrio entre ambas fases. En este procedimiento, la precisión del coeficiente de partición medido experimentalmente disminuye a medida que se incrementa (en valor absoluto) la energía libre de solvatación, y eventualmente la concentración de soluto en fase gas cae por debajo de los límites de detección experimentales cuando la energía libre de solvatación excede aproximadamente -12 kcal/mol. En el caso de especies iónicas, dicha dificultad hace necesario determinar la energía libre de solvatación mediante métodos indirectos basados en el uso de ciclos termodinámicos. Así, por ejemplo, a partir de los valores experimentales de constantes de acidez o basicidad, que pueden determinarse con gran precisión, es posible derivar energías libres de solvatación de las correspondientes especies iónicas (ácido o base) conociendo la correspondiente basicidad en fase gas, la energía libre de solvatación de la especie neutra correspondiente a la base o ácido conjugado, y la energía libre de solvatación

del protón. Este última cantidad ha sido objeto de un amplio debate en los últimos años, proponiéndose valores que variaban desde -252 hasta -263 kcal/mol, si bien recientemente se ha aceptado un nuevo valor consenso de -264 kcal/mol determinado tanto experimentalmente como a partir de cálculos teóricos de alto nivel [4, 5, 6]. Está claro que esta incertidumbre en los valores experimentales para la energía de hidratación del protón introduce un ruido importante en los valores aceptados de energía libre de hidratación de los iones.

Dentro del marco de modelos continuos SCRF, como se ha comentado anteriormente, la determinación del término electrostático es muy dependiente de la definición de la cavidad, siendo este hecho mucho más marcado en la solvatación de especies iónicas que en la de compuestos neutros. En el caso de los aniones, además, una fuente adicional de error se deriva de las incertidumbres numéricas asociadas a las colas de densidad electrónica situadas fuera de la cavidad. Estos dos hechos, pues, se suman a la problemática relacionada con la determinación precisa de la energía libre de hidratación del protón comentada antes, la cual puede introducir un ruido importante en los valores aceptados de energías libres de hidratación de iones, y ponen de relieve la cuestión relativa a la transferibilidad de la cavidad definida para solutos neutros al caso de especies iónicas, y de las posibles modificaciones que deben introducirse de cara a predecir las energías libres de solvatación de iones con una precisión comparable al error experimental, que generalmente se asume que puede ser de 4-5 kcal/mol.

En ocasiones, de cara a describir la solvatación de iones se ha considerado conveniente introducir un número arbitrario de moléculas discretas de solvente, tratando el resto del solvente como un continuo polarizable. Si bien con ello se persigue describir mejor la perturbación creada por la carga del soluto sobre las moléculas que forman la primera capa de solvatación, dicha aproximación presenta una serie de dificultades. Por un lado, hay que indicar el incremento en el coste computacional, pues la introducción de dichas moléculas discretas hace necesario evaluar los valores promedio del conjunto de configuraciones térmicamente accesibles y ponderar explícitamente la contribución de los movimientos libracionales originados a partir de las interacciones débiles entre soluto y las moléculas de solvente. Todo ello implica acoplar el cálculo continuo a un cálculo discreto de dinámica molecular o Monte Carlo. Por otro lado, la adición de dichas moléculas de solvente cambia el problema relativo a la energía libre de solvatación del ión a la solvatación del agregado soluto-solvente. Finalmente, el análisis de dichos agregados para especies iónicas se hace mucho más difícil y costoso a medida que se incrementa la complejidad de la especie iónica. Por todo ello, en nuestro estudio se ha preferido limitar la descripción mecanocuántica únicamente al soluto iónico.

En una parametrización previa del método MST para iones, la comparación entre resultados MST con aquéllos derivados a partir técnicas discretas acopladas a mecánica estadística (MC-FEP) para una serie de pequeños iones sugirió una reducción de la cavidad del 10% aproximadamente respecto a la cavidad estándar usada para solutos neutros. Dicha reducción

cabía justificarla en base a la necesidad de considerar la localización más próxima al soluto de la primera esfera de solvatación en este tipo de sistemas. En el trabajo que se presenta, motivado por la aparición reciente de valores experimentales más precisos de energías libres de solvatación de iones (véase más arriba), se ha revisado este protocolo y se ha ampliado significativamente la serie de iones considerados, prestando además especial atención al efecto que tiene el método usado para corregir los errores originados por las colas de densidad electrónica situada en el exterior de la cavidad. En particular, se ha retenido la idea original de reducir el tamaño de la cavidad usada para evaluar la energía libre electrostática a la solvatación para iones respecto a la definición estándar de cavidad para solutos neutros. Sin embargo, en vez de reducir la cavidad completa del soluto, aproximación que puede ser apropiada para iones pequeños, se ha optado sólo por los átomos pesados (y los hidrógenos enlazados a él) que soportan la carga formal.

De acuerdo con el protocolo estándar adoptado en el marco del método MST, el ajuste de la cavidad para especies iónicas monovalentes se realizó con el objetivo de reducir la desviación entre las energías libres de hidratación (ΔG_{hid}) calculadas y los valores experimentales recientemente determinados para una serie de 47 iones monovalentes. De acuerdo con la anterior parametrización, los cálculos se realizaron a nivel MST-HF/6-31+G(d) reduciendo el factor de escalado que multiplica los radios atómicos para construir la cavidad electrostática para los átomos que soportan la carga formal desde $\lambda=1.25$ a $\lambda=1.00$, y utilizando las componentes no electrostáticas antes parametrizadas para solutos neutros. Además, los cálculos se han realizado mediante dos métodos de renormalización de carga distintos, denotados en este trabajo por SC (*standard correction*) y DC (*density-corrected*). La principal diferencia entre ellos es que, a diferencia del método estándar (SC), el método corregido por densidad (DC) tiene en cuenta la anisotropía en la distribución de carga, de forma que la corrección es proporcional a la densidad electrónica situada fuera de cada región de la cavidad.

Los resultados obtenidos se muestran en la Figura 1 del trabajo y muestran la desviación entre los valores experimentales y calculados de ΔG_{hid} para diferentes factores de escalado. La desviación mínima se obtiene para un factor situado entre 1.10 y 1.15, lo cual implica una reducción del 10% de la cavidad con respecto a la usada para solutos neutros. Este resultado está de acuerdo con los resultados antes comentados obtenidos de la comparación de valores ΔG_{ele} MST y derivados de cálculos MC-FEP. Los resultados además apuntan que los errores relativos se reducen alrededor de 1 kcal/mol cuando se usa el esquema de renormalización de carga DC. Esto puede atribuirse a la mejor eficiencia de este método para tratar las anisotropías en la distribución de carga de iones. Los resultados obtenidos para $\lambda=1.15$ se han comparado además con valores obtenidos mediante el formalismo IEF-PCM, hallándose una buena correlación entre ellos. Sin embargo, la correlación es significativamente mejor cuando se considera el esquema DC de renormalización de carga. Este hecho, además de

apuntar la superioridad del método DC, también destaca la menor sensibilidad del formalismo IEF a los errores debidos a la densidad de carga situada fuera de la cavidad. Por último, estos resultados sugieren que las cavidades MST-DC optimizadas se pueden considerar transferibles en gran medida al formalismo IEF-PCM.

Los resultados finales MST-DC obtenidos se muestran en la Tabla 2. La desviación cuadrática media entre los valores ΔG_{hid} calculados y experimentales es de 5.8 kcal/mol. Para el subconjunto de 25 aniones esta desviación es de 4.7 kcal/mol, mientras que asciende a 6.9 kcal/mol para cationes. Basándose en estos resultados, podría parecer conveniente introducir diferentes factores de escalado para cationes y aniones. Sin embargo, las mayores desviaciones encontradas para cationes provienen principalmente de alcoholes y cetonas protonados (excluyendo estos compuestos, la desviación disminuye hasta 3.6 kcal/mol.) No está claro si las desviaciones notables encontradas para estos compuestos provienen de interacciones de corto alcance con moléculas de agua no descritas adecuadamente por el modelo continuo u otros efectos que no quedan bien representados en la formulación continua. Dada la disponibilidad limitada de valores precisos de ΔG_{hid} para iones y las aproximaciones intrínsecas al modelo continuo, no nos ha parecido justificado, al menos en base a estos resultados, el uso de factores de escalados distintos para aniones y cationes. Otra estrategia alternativa a la definición adecuada de la cavidad para rebajar las desviaciones encontradas se basa en el ajuste de las tensiones atómicas superficiales que caracterizan el término de van der Waals. Esto requiere no obstante mucha prudencia dada la disponibilidad limitada de datos para algunos grupos funcionales, y que pueden llevar a una sobreparametrización del modelo. Así, la única mejora estadísticamente significativa se ha obtenido ajustando la tensión del grupo OH^+ a $-0.51 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, que es casi cuatro veces más negativa que el valor correspondiente a moléculas neutras ($-0.14 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$). Corrigiendo únicamente esta tensión superficial, la desviación cuadrática media entre los valores ΔG_{hid} calculados y experimentales es de 4.6 kcal/mol. En general, el acuerdo con los valores experimentales es satisfactorio, tal como ilustra la ecuación de regresión que se muestra en la Figura 3 del trabajo y por el error relativo obtenido (4.7%).

Para realizar una evaluación adicional del comportamiento de la parametrización MST para compuestos iónicos, se determinaron los valores de pK_a para una serie de ácidos carboxílicos y piridinas sustituidas, escogidos dada la disponibilidad de basicidades en fase gas experimentales para estos compuestos. La determinación del pK_a se realizó mediante el ciclo termodinámico que se muestra en la Figura 4 del trabajo, y usando el mismo valor experimental $\Delta G_{hid}(\text{H}^+)$ determinado recientemente a partir del cual se derivaron los valores experimentales de ΔG_{hid} usados en la parametrización MST para iones. La Tabla 3 del trabajo muestra los resultados obtenidos mediante el esquema DC y SC, que abarcan un rango de valores de pK_a de más de 5 unidades. Los resultados vuelven a apuntar a la superioridad del método DC. Así, mientras

los dos métodos dan lugar a variaciones de pK_a relativas parecidas que oscilan entre 0.4 y 1.2 unidades de pK_a , los valores absolutos determinados a partir de cálculos MST-DC se ajustan mejor a los resultados experimentales, como indica la desviación cuadrática media (en unidades de pK_a) de 1.0 (MST-DC) y 2.4 (MST-SC) para ácidos carboxílicos y de 1.9 (MST-DC) y 3.2 (MST-SC) para las piridinas sustituidas. Los resultados comparan satisfactoriamente con los valores experimentales, y apoyan la estrategia computacional usada para refinar el modelo MST para la hidratación de compuestos iónicos.

Una última prueba de la calidad del modelo parametrizado consistió en el cálculo de la energía de tautomerización del equilibrio zwitteriónico de los ácidos nicotínico y isonicotínico, dado que la alta anisotropía en la distribución de carga de los zwitteriones hace este tipo de sistemas muy adecuados para investigar el efecto de los esquemas de renormalización de carga en el contexto del modelo MST. Los resultados obtenidos demuestran nuevamente la necesidad de tener en cuenta la anisotropía en la distribución de carga mediante el esquema de renormalización DC, cuyos resultados concuerdan satisfactoriamente con los valores experimentales. También se analizó el efecto del esquema de renormalización de carga sobre propiedades como el potencial electrostático para un pequeño conjunto de alcanos monosustituídos cargados, poniendo de relieve las diferencias sutiles pero no negligibles en el campo de reacción generado a partir de los dos esquemas.

En conjunto, los resultados obtenidos apoyan el procedimiento explorado en estudios previos relativo a una reducción limitada de la cavidad como un procedimiento computacionalmente simple y conservativo para el estudio de especies iónicas en solución mediante modelos continuos. En particular, en el caso del modelo MST dicho procedimiento conlleva una reducción de aproximadamente un 8 % de la cavidad creada por aquellos átomos que soportan la carga formal de la molécula, manteniendo inalterado el resto del protocolo de definición de la cavidad desarrollado para solutos neutros. Asimismo, los resultados también ponen en evidencia la necesidad de considerar de forma explícita las anisotropías en la distribución de carga de las especies iónicas de cara a corregir los efectos debidos a las colas de densidad electrónica, si bien el tratamiento IEF-PCM proporciona resultados muy similares a los obtenidos mediante un tratamiento explícito de dichos efectos. La estrategia básica presentada en este trabajo para el estudio de especies iónicas, si bien es susceptible de refinamientos adicionales, es suficiente para abordar el estudio del efecto del solvente en procesos que involucren la participación de especies iónicas en solución y la caracterización del efecto del solvente sobre los mismos.

Artículo 3. Comparación del término ΔG_{ele} entre modelos continuos SCRF

Tal como se comentó previamente, es conveniente particionar la energía libre de solvatación en componentes electrostática y no-electrostática. Sin embargo, debe hacerse énfasis en remar-

car que sólo la energía libre de solvatación es determinable experimentalmente, de forma que no es posible validar experimentalmente las componentes electrostática y no-electrostática. Ello hace necesario ser prudente en la interpretación del papel que juegan dichas componentes en el proceso de solvatación de solutos, siendo conveniente discutir su papel mediante el análisis de cambios inducidos por el solvente en propiedades moleculares experimentales o bien mediante la comparación con resultados obtenidos a partir de otras metodologías teóricas. Esta última filosofía es precisamente la que se ha adoptado en las distintas parametrizaciones del modelo MST, tal como se ha descrito para el solvente octanol en el artículo 1.

En la práctica, la problemática expuesta anteriormente implica que modelos SCRF distintos, aunque sean capaces de estimar con precisiones similares la energía libre de solvatación de una gama diversa de solutos, proporcionen un balance diferente entre sus contribuciones electrostáticas y no-electrostáticas, reflejo de posibles diferencias en el formalismo utilizado para acoplar soluto y solvente, la definición de la cavidad empleada en la evaluación de dichas contribuciones, o bien los detalles específicos utilizados en la parametrización del modelo. En el artículo 3 de la presente tesis nos planteamos una comparación sistemática de la componente electrostática en tres modelos SCRF de última generación, que proporcionan en todos los casos valores correctos de la energía de solvatación total. En concreto, se han comparado tres modelos que reflejan diferentes formalismos SCRF: i) el modelo de Born generalizado implementado en el modelo SM5.42R, ii) el método de expansión multipolar desarrollado en Nancy, y iii) la versión MST del modelo del continuo polarizable (PCM). Todos los cálculos se han realizado a nivel HF/6-31G(d) para una serie de 18 moléculas que contienen grupos polares prototípicos, considerando tres valores diferentes de permitividad dieléctrica $\epsilon = 2, 5$ y 80, que corresponden aproximadamente a los valores del tetracloruro de carbono, el cloroforno y el agua. El análisis se ha centrado tanto en el valor de la componente electrostática de la energía libre de solvatación, como en la variación en el momento dipolar inducido por el solvente.

Los resultados mostrados en la Tabla 2 ilustran las diferencias entre las energías libres electrostáticas obtenidas a partir de las implementaciones estándar de los tres modelos para $\epsilon=80$. La correlación entre los tres modelos es de $r=0.97$. Sin embargo, hay diferencias muy marcadas en la magnitud de la respuesta electrostática. Comparado con el modelo SM5.42R, las energías libres electrostáticas determinadas a partir del modelo MPE son en promedio menores por un factor de 0.87, mientras que las MST son mayores por un factor de 1.45. En contraste con estos resultados, la Tabla 3 muestra valores muy parecidos para las energías libres de hidratación calculadas con los tres modelos, todos ellos además muy próximos a los valores experimentales. Dadas las diferencias en los términos ΔG_{ele} , el buen acuerdo con las energías libres de hidratación experimentales se puede atribuir a la parametrización del término no electrostático intrínseco a cada modelo SCRF.

Por otra parte, se han determinado sistemáticamente para cada método las contribuciones ΔG_{ele} usando las mismas definiciones de cavidades, en concreto, las cavidades estándar definidas para cada modelo. Usando las cavidades definidas para el modelo SM5.42R, los aumentos relativos para cada modelo al pasar de $\epsilon=2$ a 5 y de $\epsilon=2$ a 80 varían entre 1.8 y 2.0 y entre 2.3 y 2.8, respectivamente. Estos factores son de 1.7-1.9 y 2.1-2.6 si se usan las cavidades MPE, y de 2.3-2.9 y 5.8-8.0 para las cavidades MST. Para las cavidades MST, las diferencias en el aumento relativo del término ΔG_{ele} reflejan parcialmente la pequeña magnitud de los valores ΔG_{ele} obtenidos para $\epsilon=2$. En este punto cabe recordar, además, que el modelo MST usa diferentes cavidades para cada solvente, de forma que el volumen de la cavidad se reduce a medida que la constante dieléctrica del medio aumenta, característica que es única entre los distintos métodos continuos.

El conjunto de resultados sugiere que los tres formalismos SCRF presentan similitudes en su capacidad para capturar los *cambios* en la energía libre electrostática obtenida a partir de diferentes constantes dieléctricas. De todas formas, hay diferencias relevantes en cuanto a las *magnitudes absolutas* de las energías libres electrostáticas obtenidas a partir de los tres formalismos, tal como se puede apreciar en la Tabla 7 de este trabajo. En todos los casos el modelo MST da lugar a las contribuciones electrostáticas más grandes, pero que son sólo ligeramente superiores a las obtenidas por el modelo MPE. Dichas diferencias cabe atribuir las principalmente a efectos relacionados con la corrección de las colas de densidad electrónica que se extienden más allá de la cavidad, ya que cuando la distribución de carga se representa por un conjunto de cargas puntuales CM2 los métodos MPE y MST dan lugar prácticamente a la misma respuesta electrostática. Por otra parte, las contribuciones electrostáticas obtenidas con el modelo SM5.42R son claramente inferiores a las estimaciones proporcionadas por los modelos MPE y MST (véanse Tablas 7 y 8), lo que fuerza a usar cavidades muy pequeñas para maximizar la respuesta electrostática a la solvatación. Este hecho cabe atribuirlo al menos en parte al uso de un modelo demasiado simplificado de cargas puntuales para representar la densidad de carga real de la molécula. Las diferencias mencionadas anteriormente en la contribución electrostática a la energía libre de solvatación quedan asimismo cualitativamente reflejadas en los momentos dipolares en solución obtenidos a partir de los tres modelos (véanse Tablas 10-12). No obstante, cuantitativamente dichas diferencias son menos marcadas que las obtenidas para la componente ΔG_{ele} .

En conjunto, los resultados obtenidos en el presente análisis indican que los formalismos de expansión multipolar y de carga superficial aparente empleados en los métodos MPE y MST dan formalmente la misma respuesta electrostática, salvo para cavidades excesivamente pequeñas, cuyo tamaño es inferior al recomendado en las parametrizaciones estándar de ambos métodos. De hecho, la similitud observada en la respuesta electrostática entre los métodos MPE y MST indica que la definición de la cavidad realizada en el marco del modelo MST

para diferentes solventes sería completamente transferible al modelo MPE. Por otra parte, el formalismo de Born generalizado proporciona una respuesta electrostática menor cuando se utiliza la misma cavidad comparada con los métodos MPE y MST, siendo por tanto necesario recurrir a cavidades más pequeñas con el fin de obtener una contribución electrostática comparable.

6.2. Contribuciones fraccionales al coeficiente de partición

La solvatación diferencial de las moléculas en medios polares y apolares, medida típicamente a partir de coeficientes de partición, permite profundizar en diversas propiedades bioquímicas, farmacológicas y toxicológicas de las moléculas [7, 8, 9, 10]. En general el agua se considera como el solvente polar de referencia, mientras que diversas moléculas orgánicas se emplean como modelos de la fase apolar. De ellas, el n-octanol es probablemente el solvente más extendido, al ser un buen modelo de los lípidos que constituyen las membranas biológicas.

El coeficiente de partición octanol/agua ($\log P_{OW}$) es una propiedad extensamente empleada para caracterizar la hidrofobicidad de las moléculas. De hecho, en el campo biofarmacéutico el $\log P_{OW}$ actúa como un autentico discriminador sobre el posible carácter terapéutico de una molécula, tal como se apuntó en el trabajo desarrollado por Lipinski y colaboradores que indicaron como eran muy raros los fármacos que tenían un coeficiente de partición octanol/agua inferior a 1 o superior a 5 [11]. Desafortunadamente, la medición experimental de $\log P_{OW}$ a través de la síntesis del compuesto y de su posterior determinación requiere una inversión de tiempo y esfuerzo elevada. Esto resulta difícilmente asumible en la actualidad, donde las técnicas de química combinatoria y de *screening* virtual hacen necesario disponer de métodos que permitan evaluar de forma rápida el valor de $\log P_{OW}$ para una amplia y diversa serie de compuestos. Como consecuencia de esta necesidad se han desarrollado toda una gama de métodos empíricos que explotan contribuciones asociadas a tipos de átomos y/o fragmentos a fin de poder estimar con rapidez el coeficiente de partición octanol/agua [12, 13, 14]. También se han reseñado algunos algoritmos que basándose en cálculos mecanocuánticos permiten aproximar, sin necesidad de introducir parametrizaciones específicas, el coeficiente de partición [15, 16, 17].

A pesar de su incuestionable importancia y capacidad descriptiva, no podemos olvidar que el coeficiente de partición es un parámetro escalar y como tal no aporta información tridimensional sobre la distribución hidrofóbica/hidrofílica de la molécula. Así, el análisis del coeficiente de partición permitirá conocer si en promedio una molécula tiene características tales que favorezcan su paso por membranas biológicas, pero no si la distribución hidrofóbica/hidrofílica

de la misma la hacen adecuada para complementar las pautas que marca el centro de unión del receptor biológico. Es por lo tanto interesante ir una etapa más allá del $\log P_{ow}$, y obtener una distribución tridimensional de las características hidrofóbicas/hidrofilicas de la molécula. Disponer de esta información, que requiere fraccionar en grupos el coeficiente de partición (o las energías libres de solvatación o transferencia) permitiría diseñar modificaciones conservativas que aumenten el perfil farmacológico del fármaco diana.

La evaluación del coeficiente de partición octanol/agua a partir de contribuciones fragmentales implícitamente asume la transferibilidad de las contribuciones asociadas a átomos y fragmentos. Dicha fragmentación es arbitraria, lo cual da lugar a diferencias en cuanto a la naturaleza y tipo de átomos o fragmentos considerados en los distintos métodos empíricos. Por otro lado, la transferibilidad de las contribuciones fragmentales no puede asumirse en un sentido estricto, pues cabe pensar que el entorno inmediato de un átomo o fragmento específico presente en una molécula influirá en su contribución al valor de $\log P_{ow}$. Todo ello hace necesario definir asimismo reglas empíricas para corregir la contribución de un determinado grupo en función de su conectividad a otros fragmentos. No obstante, aún estas reglas correctoras son incapaces de subsanar los errores intrínsecos de la aproximación fraccional empírica en casos donde existen fuertes efectos electrónicos y estéricos ejercidos por otros fragmentos de la molécula, o en casos donde cambios conformacionales o tautoméricos jueguen un papel importante.

La implementación de métodos que permitan la determinación de contribuciones fraccionales a la energía libre de solvatación en modelos continuos SCRF constituye un procedimiento para poder examinar el grado de transferibilidad de las contribuciones fragmentales al coeficiente de partición octanol/agua. Al estar basados en un cálculo mecanocuántico y no en parametrizaciones empíricas, estos métodos recogen de manera natural los efectos electrónicos y estéricos debidos al entorno molecular, permitiendo determinar la influencia de factores como cambios conformacionales, tautoméricos o de ionización en la contribución de grupo.

De hecho, dentro del marco del modelo MST se ha definido anteriormente un procedimiento fraccional que permite determinar contribuciones de grupo basado en el campo de reacción del solvente sobre la porción de la cavidad asignada a un átomo del soluto [18]. En base a dicho procedimiento, la contribución del átomo y al $\log P_{ow}$ se calcula a partir de la interacción de la densidad de carga total del soluto con aquellas cargas aparentes distribuidas en la superficie de la cavidad perteneciente a dicho átomo. El método, denominado *contribución fraccional por superficie*, ya ha sido implementado tanto en las versiones QM, semiclásica y clásica del modelo MST.

Los dos trabajos que se recogen en este apartado inciden precisamente en el análisis de la transferibilidad de las contribuciones de grupo al $\log P_{ow}$. En el primero de ellos se presenta

los métodos de partición basados en superficie y en átomo dentro de la versión semiempírica del modelo MST, y se analiza la transferibilidad de grupo para diversas series de compuestos. En el siguiente trabajo, se presta especial atención a la influencia ejercida por el esquema de normalización de carga en la determinación del grado de transferibilidad de las contribuciones de grupo a la energía libre de hidratación.

Artículo 4. Transferibilidad de contribuciones fraccionales al coeficiente de partición octanol/agua

Este trabajo tiene dos objetivos. En primera instancia, se ha extendido la parametrización previamente descrita del modelo MST en octanol a nivel semiempírico AM1 y PM3, así como el refinamiento de las versiones anteriores para agua, cloroformo y tetracloruro de carbono. A continuación, se ha presentado un procedimiento basado en la aproximación NDDO que permite descomponer de manera rigurosa el $\log P_{OW}$ en *contribuciones fraccionales por átomo*, y descomponer éstas a su vez en componentes intrínsecas y cruzadas para un fragmento/átomo dado. En base a este esquema de partición, se ha discutido entonces la transferibilidad de las contribuciones fragmentales para diversos sistemas moleculares.

La parametrización del método MST para los Hamiltonianos semiempíricos AM1 y PM3 se ha realizado siguiendo el mismo procedimiento antes descrito para la versión HF/6-31G(d), y adoptando la misma definición para la cavidad. Así, en la parametrización se han usado 58 compuestos de los cuales se conoce la energía libre de solvatación en octanol conjuntamente con 75 moléculas con coeficiente de partición octanol/agua conocido. También los modelos en agua, cloroformo y tetracloruro de carbono se han reparametrizado usando un mayor número de compuestos (69, 54 y 47, respectivamente). Asimismo, se ha usado el mismo número de tipos de átomos para el término de van der Waals y el término de corrección para compuestos con grupos funcionales nitrilo introducido en la versión HF/6-31G(d). De esta forma, y por coherencia con la versión *ab initio* del método, en el presente trabajo sólo se han optimizado las tensiones superficiales del término de van der Waals a partir de regresiones lineales múltiples a fin de minimizar las diferencias respecto a los valores experimentales. En general, los resultados obtenidos son similares a los obtenidos para la versión *ab initio*. El análisis estadístico de los resultados se muestra en la Tabla 2 del trabajo. Las desviaciones cuadráticas medias entre las energías libres de solvatación experimentales y calculadas en agua, n-octanol, cloroformo y tetracloruro de carbono son de 0.7, 0.7, 0.4 y 0.3 kcal/mol para AM1, y de 0.9, 0.8, 0.5 y 0.3 kcal/mol para PM3, respectivamente (resultados de calidad similar a los obtenidos *ab initio*). El coeficiente de escalado en la ecuación de regresión entre los valores experimentales y teóricos es próximo a la unidad, lo cual revela la ausencia de errores sistemáticos para la amplia serie de compuestos. El acuerdo obtenido entre las energías libres de solvatación *ab*

initio y semiempíricas, sus componentes electrostáticas y no electrostáticas, y los cambios inducidos en el momento dipolar confirman la transferibilidad de los parámetros definidos en el modelo MST para la solvatación en octanol. De forma similar, los valores calculados de $\log P_{OW}$ obtenidos a partir de cálculos MST en agua y en octanol presentan una desviación cuadrática media de 0.5-0.7 unidades de $\log P$, sólo ligeramente superior a la obtenida en la versión *ab initio*. A partir de estos resultados, la versión semiempírica del modelo MST se presenta como una alternativa útil y eficiente para examinar las propiedades de solvatación de moléculas grandes.

Como se ha descrito anteriormente, los métodos fragmentales estiman el $\log P_{OW}$ asumiendo la naturaleza aditiva y la transferibilidad de las contribuciones al $\log P_{OW}$ de fragmentos específicos de la molécula. El método fraccional basado en la aproximación NDDO que se ha presentado aquí permite examinar la transferibilidad de estas contribuciones, y cuantificar el término de corrección atribuible al efecto del entorno sobre la contribución de un fragmento dado, al dividir las contribuciones fragmentales en una parte intrínseca y otra cruzada. Así, al aplicar esta metodología al estudio de la solvatación de una serie de derivados bencénicos monosustituídos (que incluyen el benceno, el fenol, el benzonitrilo, el nitrobenzono, el tiofenol, el fluorobenceno y el clorobenceno) se puede ver como la contribución del benceno no es constante, como se asume en los modelos aditivos, y aumenta en 0.6-1.0 unidades de $\log P$ al reemplazar el átomo de hidrógeno por los diferentes grupos funcionales. A su vez, estas diferencias cabe atribuir las principalmente a la componente intrínseca de la contribución fraccional por átomo, y se originan debido a las diferencias en la solvatación en agua, mientras que la contribución a la solvatación en octanol es relativamente menos variable. Además, las contribuciones obtenidas por el método NDDO presentan diferencias significativas con respecto a las obtenidas experimentalmente asumiendo que la contribución del anillo bencénico es constante dentro de una serie de derivados bencénicos. Por ejemplo, mientras que la comparación de los valores de $\log P_{OW}$ del benceno y el fluorobenceno indican que el fluoro es ligeramente lipofílico [$\log P_{OW}(F)=+0.1$], el método de descomposición NDDO predice el fluoro como hidrofílico [$\log P_{OW}(F)=-0.6$] (véase Tabla 5). Estos resultados indican que la variación en la lipofilidad del fluorobenceno no se deben exclusivamente al átomo de fluoro, sino que provienen de un balance en la variación de las contribuciones asociadas a los fragmentos C_6H_5 y F.

Por otra parte, se ha calculado la contribución fragmental del sustituyente X en las series de derivados $CH_3(CH_2)_nX$ (X=CH₃, F, OH, COOH; n=0-7). En el caso de los alcanos, los resultados (véase Tabla 6) apuntan a que el grupo metilo tiene una contribución constante de 0.8 unidades de $\log P$, así como una contribución constante a la solvatación en agua y en octanol. Sin embargo, las contribuciones fragmentales de los grupos F, OH y COOH dependen de la longitud de la cadena alquímica. Así, sus contribuciones varían desde -0.5 hasta -1.6 (F), -1.6 hasta -2.8 (OH), y -1.7 hasta -2.7 (COOH) unidades de $\log P$ cuando el número de

metilenos aumenta de 0 a 7. En contraste, las contribuciones del grupo metilo y metileno son prácticamente constantes a lo largo de cada serie homóloga. El análisis de las contribuciones a la solvatación en agua y en octanol indican claramente que la dependencia de las contribuciones a la longitud de la cadena alquílica se deben principalmente a las diferencias en las contribuciones a la solvatación en agua, y que el término intrínseco es el principal responsable de estos cambios.

La influencia de los efectos conformacionales en las contribuciones fragmentales se puede examinar a partir de los resultados dados en la Tabla 7, que muestran las contribuciones fragmentales para el butano, el 1-propanol y el 1,2-etandiol a valores escogidos del ángulo de torsión ($\theta = 0, 60, 120, 180^\circ$) alrededor del enlace $\text{CH}_2\text{-CH}_2$. Los valores teóricos de $\log P_{\text{OW}}$ muestran una dependencia marcada en el ángulo dihedro. Así, varían desde 2.4 a 2.9, 1.6 a 0.3, y -0.7 a -1.4 para el butano, el 1-propanol, y el 1,2-etandiol, respectivamente, cuando el ángulo dihedro varía desde 0 hasta 180° , poniendo de relieve la sensibilidad de esta propiedad a los cambios conformacionales. En el caso del butano, la contribución fragmental del grupo metilo (0.8 unidades de $\log P$) es constante para los cuatro ángulos dihedros considerados. Sin embargo, en el caso del 1-propanol, la contribución del grupo hidroxilo es de -0.8 unidades de $\log P$ en la conformación eclipsada y aumenta hasta -2.0 unidades de $\log P$ en la conformación extendida. Por el contrario, la contribución del grupo metilo es de nuevo independiente de la conformación de la cadena. Para el 1,2-etandiol se obtiene resultados en la misma dirección que para el 1-propanol, donde la menor exposición al solvente de los grupos hidroxilo explican el aumento de la lipofilia de estos grupos en las conformaciones eclipsadas.

Como última prueba, se ha examinado la influencia de los equilibrios tautoméricos sobre las contribuciones fragmentales al $\log P_{\text{OW}}$ para una serie de derivados del imidazol. La Tabla 8 de este trabajo muestra las contribuciones fraccionales de la unidad de imidazol. En el caso del 2-metil,5-nitrobenzoimidazol y el 2-trifluorometil,5nitrobenzoimidazol, no solo el $\log P_{\text{OW}}$ total, también las contribuciones fragmentales del fragmento imidazol son idénticas en sus respectivas formas tautoméricas. La única diferencia entre los dos compuestos es que la unidad de imidazol es alrededor de 0.5 unidades de $\log P$ más lipofílica en el 2-trifluorometil,5nitrobenzoimidazol, contribuyendo así a la mayor lipofilia de este compuesto. En el caso del 2-trifluorometil,4-clorobenzoimidazol, sin embargo, la unidad de imidazol es 0.6 unidades de $\log P$ más lipofílica en el tautómero N3-H que en el N1-H. Finalmente, se han examinado las contribuciones del anillo imidazol para los tautómeros N1-H y N3-H de la histamina, en sus conformaciones extendidas, hallándose diferencias de 0.6 unidades de $\log P$ en la contribución del anillo de imidazol entre los diferentes tautómeros.

En conjunto, los resultados obtenidos ilustran la dependencia que experimentan las contribuciones fragmentales al $\log P_{\text{OW}}$ respecto a efectos electrónicos, conformacionales y tau-

toméricos, y apuntan a la prudencia con que se debe considerar la transferibilidad de estas contribuciones fragmentales entre distintos sistemas.

Artículo 5. Transferibilidad de grupo y renormalización de carga en cálculos MST

Como se comentó en la sección correspondiente a la parametrización de modelos continuos para iones (véase más arriba), los resultados derivados de cálculos MST son sensibles a artefactos generados por las colas de densidad electrónica situadas fuera de la cavidad molecular. Este efecto se corrige mediante procedimientos de renormalización como los descritos más arriba. Estos métodos permiten obtener resultados globales de energía libre de solvatación correctos, pero no está claro cual es su calidad para reproducir los parámetros fraccionales y hasta que punto un método de renormalización mal balanceado puede sesgar las contribuciones fraccionales desbalanceando las regiones hidrofóbicas e hidrofílicas. En este punto cabe hacer notar que, si bien la parametrización del modelo MST permite reproducir la energía libre de solvatación, ello no asegura que un tratamiento inadecuado de los efectos de normalización de carga no induzca errores en la magnitud de las contribuciones fraccionales a la solvatación, que previsiblemente serán más sensibles a la anisotropía en la distribución de carga de la molécula y por tanto al método utilizado en la normalización de carga. Constatar esto ha sido el objetivo de nuestro siguiente trabajo.

Para llevar a cabo este estudio se ha escogido una serie de alcanos monosustituidos análoga a la utilizada en el cuarto trabajo ($\text{CH}_3(\text{CH}_2)_n\text{X}$, $n=0-7$), pero considerando una serie más amplia de grupos funcionales (CH_3 , OH , F , NH_2 , CHO , OCH_3 , COOH , COOCH_3 , CONH_2 , NO_2 y SH), y considerando tanto formas neutras como iónicas. Los cálculos se han realizado a nivel MST-HF/6-31G(d) para moléculas neutras y a nivel MST-HF/6-31+G(d) para los compuestos cargados, de acuerdo con la parametrización del modelo MST antes descrita. Los resultados obtenidos a nivel *ab initio* usando esquemas de normalización de carga que consideran explícitamente la anisotropía molecular apoyan la transferibilidad de las contribuciones fragmentales de los grupos funcionales X entre los diferentes compuestos, contribuciones que convergen rápidamente a medida que aumentamos la longitud de la cadena alquílica. En cambio, no se obtienen iguales buenos resultados si se emplean métodos de renormalización que no tienen presente la anisotropía en la distribución de carga del soluto.

Finalmente, estos estudios abren la puerta a la aplicación de la metodología MST como un procedimiento alternativo para derivar contribuciones fragmentales al coeficiente de partición octanol/agua y a su aplicación en la optimización del perfil farmacológico de compuestos con potencial acción terapéutica. Ejemplos de dicha aplicación son la definición del índice hidrofóbico de similitud [19, 20], y más recientemente del farmacóforo hidrofóbico, propiedades

determinadas en ambos casos usando la información obtenida acerca de la distribución 3D de las contribuciones hidrofóbicas/hidrofílicas de una molécula.

6.3. Fuerzas de polarización

La comprensión de las fuerzas intermoleculares es fundamental para conocer la estructura íntima de la materia y poder explicar sus propiedades. Por ello, no es de extrañar el esfuerzo dedicado en las últimas décadas al desarrollo de modelos teóricos que permitan expresar la interacción entre moléculas como una combinación de fuerzas de distinta naturaleza. De acuerdo con la clasificación propuesta por Stone [21], cabe diferenciar entre fuerzas de *largo alcance* y de *corto alcance*. En las primeras, la energía es inversamente proporcional a la distancia entre las partículas interaccionantes, mientras que en las segundas la energía decrece exponencialmente con la distancia.

Dentro de las fuerzas de *largo alcance* cabe incluir esencialmente las fuerzas electrostáticas, de inducción o polarización y dispersivas. Las interacciones electrostáticas son fuerzas de par, aditivas, originadas por la interacción coulombica entre las distribuciones de carga permanente de las moléculas, y pueden ser atractivas o repulsivas. Las contribuciones de inducción son atractivas y surgen por efecto de la distorsión de la distribución de carga de una molécula a raíz de la influencia de un campo eléctrico externo. Dado que el campo eléctrico ejercido por diversas moléculas puede reforzarse o cancelarse entre sí, son fuerzas de naturaleza no aditiva. Finalmente, la contribución de dispersión, si bien no tiene una descripción precisa en términos clásicos, reflejaría las fluctuaciones instantáneas en las distribuciones de carga de las moléculas. Su carácter es atractivo, y en una primera aproximación cabe considerar que es aditiva. Con respecto a las fuerzas de corto alcance, éstas surgen como resultado de un recubrimiento significativo entre las distribuciones de carga de las moléculas interaccionantes, siendo de carácter no aditivo. Dichas fuerzas comprenden contribuciones de tipo atractivo, como transferencia de carga y de intercambio, y repulsivas.

Gran parte de los modelos teóricos desarrollados en el marco de la mecánica cuántica para descomponer la energía de interacción en componentes se basan en la teoría de perturbación de fuerzas intermoleculares a larga distancia. Dado que las fuerzas intermoleculares son relativamente débiles, es posible describirlas usando teoría de perturbaciones, obteniéndose así de forma natural diversos términos de interacción a los cuales se les puede atribuir un claro significado físico. Alternativamente, la energía de interacción se puede calcular directamente a partir del principio variacional dentro de la aproximación de la supermolécula. El cálculo variacional es simple y proporciona excelentes aproximaciones a la energía de interacción total, pero particionar dicha energía en componentes no es directo, siendo necesario desarrollar métodos

de partición adecuados. Sin duda alguna, el procedimiento que ha tenido mayor impacto es el método de partición propuesto por Kitaura y Morokuma [22, 23], donde la energía de interacción se expresa como la suma de términos electrostático, de polarización, de intercambio y de transferencia de carga, junto con un término de mezcla resultante de acoplamientos cruzados entre los diversos términos.

El elevado coste computacional asociado a métodos de partición implementados a nivel de mecánica cuántica limita su aplicabilidad al estudio de sistemas de pequeño tamaño, dificultando el estudio de procesos de interés químico y biológico. Por ello, se han desarrollado métodos alternativos que expresan la energía de interacción de un sistema complejo como la suma de diversas contribuciones energéticas, de las cuales parte (correspondiente a un pequeño subconjunto de átomos del sistema) se tratan a nivel cuántico, mientras que el resto se describen mediante aproximaciones clásicas. Esta aproximación constituye la base de los métodos híbridos QM/MM, en los que la energía de interacción se expresa como la adición de un término electrostático, otro de polarización y finalmente un tercero de dispersión-repulsión. El primero refleja la interacción entre la distribución de carga del subsistema cuántico con el conjunto de cargas atómicas puntuales empleado para describir la distribución de carga del subsistema clásico. La determinación del término de polarización se lleva a cabo introduciendo las cargas puntuales del subsistema clásico en el Hamiltoniano del subsistema cuántico, permitiendo así que su distribución de carga se relaje bajo la influencia del campo eléctrico creado por el subsistema clásico. Finalmente, la contribución de dispersión-repulsión viene descrita generalmente a partir de una expresión clásica del tipo Lennard-Jones.

Los métodos híbridos QM/MM constituyen la elección natural para el estudio de procesos químicos que implican cambios importantes en la distribución de carga del sistema, como rotura y formación de enlaces. En casos en los que no se producen cambios destacados en la distribución de carga, el sistema completo se suele representar a nivel de mecánica clásica utilizando funciones potenciales (*campos de fuerza*) que describen la energía de interacción del sistema combinando generalmente un término electrostático y un término de van der Waals. En este nivel, los efectos de polarización no son tratados explícitamente, sino que se incorporan de forma implícita en el término electrostático aumentando la magnitud de las cargas puntuales asignadas a los átomos del sistema, tal como ocurre en campos de fuerza como AMBER/parm94 [24] o CHARMM27 [25]. Estos y otros campos de fuerza relacionados han permitido un avance muy importante en la simulación de grandes sistemas moleculares. No obstante, está claro que la limitación inherente al uso de una distribución de carga fija limita su precisión en casos en los que el entorno molecular de los átomos cambie durante la simulación, o en casos en los que una pequeña parte del sistema genere fuertes efectos inductivos sobre el resto. Casos típicos son la solvatación de compuestos iónicos, especialmente aniones, o bien complejos sistemas biológicos, como canales iónicos o metaloproteínas.

La representación de efectos de polarización precisa dos operaciones básicas: i) determinar como una distribución de carga se adapta a la presencia de otra distribución en su cercanía y ii) evaluar el impacto que esa relajación de carga tiene en la energética de la interacción molecular. En la práctica, se han descrito cuatro procedimientos básicos para incluir efectos de polarización en campos de fuerza clásicos: i) el modelo de dipolos puntuales inducidos, ii) el modelo de fluctuación de carga, iii) el modelo de cargas duales y iv) el modelo de oscilador de Drude clásico (véase Metodología).

Uno de los objetivos planteados en esta tesis ha sido profundizar en la naturaleza de las fuerzas de polarización y en el desarrollo de metodologías efectivas para describir sus efectos en campos de fuerzas clásicos. Este objetivo se ha materializado en los dos trabajos que se presentan a continuación, donde se compara la energía de polarización obtenida a partir del análisis de Kitaura-Morokuma con los resultados obtenidos a partir de una descripción mixta QM/MM. Asimismo, se examinan dos procedimientos clásicos para representar efectos de polarización. Uno de ellos se basa en un modelo de carga dual para representar la polarización de una molécula por el campo eléctrico creado por moléculas vecinas. En el otro caso se considera el modelo de dipolos inducidos, si bien nuestro interés ha sido evaluar un procedimiento para derivar modelos distribuidos de polarizabilidades atómicas.

Artículo 6. Polarización en complejos de puente de hidrógeno y catión- π

Con el objeto de explorar la contribución de las fuerzas de polarización en la interacción de complejos moleculares, se han escogido dos series de sistemas intermoleculares correspondientes a complejos de puentes de hidrógeno y catión- π . La elección de dichos sistemas no es gratuita, sino que viene determinada por dos razones básicas. En primer lugar, ambos sistemas reflejan interacciones fundamentales en el reconocimiento entre moléculas en el ámbito bioquímico. Así, la formación de interacciones por puente de hidrógeno es clave para el mantenimiento del código genético, mientras que la interacción catión- π es clave en numerosos procesos de reconocimiento ligando-proteína, como es el caso de la unión de acetilcolina al supuesto *centro aniónico* del enzima acetilcolinesterasa [26]. En segundo lugar, cabe esperar que la contribución de las fuerzas de polarización a la estabilidad de ambos sistemas presente notables diferencias, pues mientras el término electrostático debe ejercer una contribución claramente superior a la energía de polarización en complejos de puente de hidrógeno estándar, la contribución de las fuerzas de polarización puede ser mucho más significativa en complejos catión- π . Ambos sistemas son por tanto adecuados para examinar la bondad de una aproximación mixta QM/MM de cara a describir la energía de polarización en tales sistemas.

En el trabajo se planteó inicialmente realizar una comparación entre las contribuciones electrostática y de polarización obtenidas a partir del método de descomposición de la energía

de Kitaura-Morokuma (KM) con las correspondientes contribuciones determinadas en el contexto de un tratamiento QM/MM. Todos los cálculos se llevaron a cabo utilizando el método RHF y una base 6-31G(d). Aunque la energía electrostática no es extremadamente sensible a la inclusión de efectos de correlación electrónica o al uso de conjuntos de base amplios y flexibles, la energía de polarización está muy influenciada por estos factores. Sin embargo, la elección de este nivel de teoría moderado se debe a varias razones. En primer lugar, nuestro interés es realizar un análisis comparativo de diferentes esquemas computacionales, pero no la derivación de estimaciones precisas de la polarización. En segundo lugar, el análisis de KM puede sobreestimar la polarización cuando se usan conjuntos de base extendidos que no obedecen el principio de exclusión de Pauli, mientras que da lugar a valores similares a los obtenidos mediante otros métodos de descomposición basados en funciones de onda antisimetrizadas para conjuntos de base *split-valence*. Finalmente, el método KM ofrece la posibilidad de dividir la energía de polarización total en contribuciones fragmentales, las cuales se pueden comparar directamente con las energías de polarización de los monómeros obtenidas a partir de otras metodologías.

Los resultados mostrados en la Tabla 1 muestran los valores obtenidos a partir del análisis de descomposición de la energía de KM. Para la serie de complejos de puente de hidrógeno, la energía de polarización representa en promedio el 10 % de la energía electrostática, y a su vez estas dos componentes representan conjuntamente el 55 % de la energía de interacción. En contraste con estos resultados, los complejos correspondientes a la interacción de etileno y acetileno con Na^+ presentan una energía de polarización cuya magnitud es de un 40 % con respecto a la electrostática, y por otra parte la suma de ambas contribuciones refleja prácticamente toda la energía de interacción en dichos sistemas. Por otra parte, los resultados de la Tabla 2 sugieren que, para las geometrías de equilibrio de los complejos considerados, la energía de polarización total del complejo se puede obtener a partir de la adición de las energías de polarización de cada monómero individual, siendo generalmente despreciable la magnitud del término de mezcla que se origina a partir de acoplamiento mutuo entre monómeros.

Los resultados obtenidos a partir del análisis de Kitaura-Morokuma han sido empleados con el fin de examinar la capacidad de un tratamiento mixto QM/MM para reproducir la magnitud de la energía de polarización en ambos tipos de complejos. Para ello, se han estimado tanto la energía electrostática como la de polarización a partir de una descripción mixta de las moléculas interaccionantes. Así, mientras uno de los monómeros se ha descrito a nivel QM, el otro ha sido descrito clásicamente reemplazando la distribución de carga por un conjunto de cargas atómicas puntuales derivadas a partir de las metodologías ESP y RESP. Con respecto a la energía electrostática, dado que puede llegar a diferir en 1 kcal/mol según cual sea el monómero descrito a nivel QM, su valor en el tratamiento QM/MM se ha determinado a partir del promedio entre los dos sistemas obtenidos alternando el monómero descrito a nivel QM. Por

otro lado, la energía de polarización de cada monómero se ha obtenido introduciendo en el Hamiltoniano del monómero QM las cargas puntuales correspondientes al monómero clásico.

La comparación entre los resultados QM/MM y los obtenidos a partir del esquema KM se muestran en las Figuras 2 y 3. Respecto a la contribución electrostática, hay una correlación muy buena entre los dos métodos ($r^2=0.98$), aunque los resultados QM/MM infraestiman en un 24% los valores KM. En contraste con estos resultados, hay un acuerdo muy bueno entre las energías de polarización obtenidas por los dos métodos, tanto en la determinación de las energías de polarización relativas a un sólo monómero ($r^2=0.99$) como en las energías de polarización totales del complejo ($r^2=0.99$). Por otra parte, dado que la contribución del término de mezcla es muy pequeña (véase más arriba), la ligera tendencia del método QM/MM a infraestimar la energía de polarización a nivel de monómero (alrededor del 2%) es básicamente aditiva (5% a nivel de dímero). Los resultados obtenidos para los trímeros confirman los resultados obtenidos para los dímeros, sugiriendo que la polarización total para pequeños agregados moleculares en sus geometrías de equilibrio se puede estimar razonablemente como la adición de las contribuciones de polarización a nivel de monómero en el complejo molecular.

Finalmente, se ha examinado una estrategia destinada a introducir los efectos de inducción de forma implícita en los campos de fuerzas clásicos, comparándose los resultados obtenidos por dicha metodología con los resultados KM y QM/MM. En esta estrategia las propiedades electrostáticas de una molécula se definen a partir de la asignación de dos conjuntos de cargas. Uno de ellos refleja el campo electrostático permanente generado por la molécula, mientras que el segundo se utiliza para describir la distorsión electrostática inducida por la presencia de una distribución electrostática externa que polariza la molécula. El primer conjunto de cargas, representado por $\{q^0\}$, se obtiene ajustando el potencial electrostático creado por la molécula en su entorno. El segundo conjunto de cargas, representado por $\{q^1\}$, se obtiene ajustando las energías de interacción (electrostático + polarización) con una carga puntual calculada en una serie de puntos alrededor de la molécula. De esta forma, se espera que las cargas $\{q^1\}$ representen un límite superior a la polarización para la mayoría de moléculas, y que para sistemas reales las cargas efectivas $\{q^{eff}\}$ se puedan determinar por interpolación a partir de la ecuación $\{q^1\}=\lambda\{q^1\}+(1-\lambda)\{q^0\}$. En concreto, en este trabajo se ha adoptado un valor óptimo de $\lambda=1/2$ para moléculas neutras y de $\lambda=1$ para sistemas cargados. En promedio, el conjunto de carga $\{q^1\}$ obtenido para los diferentes monómeros considerados reflejan un aumento del 37% con respecto a las cargas $\{q^0\}$, y representan un aumento del momento dipolar alrededor del 40% relativo al valor determinado a nivel QM en fase gas.

Los Figura 6 de este trabajo muestra la comparación entre las componentes electrostática, de polarización, y electrostática+polarización obtenidas a partir de los métodos KM y

QM/MM y las expresiones clásicas propuestas para el modelo de doble carga. Los resultados indican que la energía electrostática clásica correlaciona muy bien con los valores QM/MM ($r^2=0.99$), si bien es necesario escalar los primeros por un factor de 1.08 con el fin de mejorar la concordancia cuantitativa con los resultados QM/MM. Con respecto a los valores KM también se encuentra una buena correlación ($r^2=0.94$), pero hay una infraestimación de esta contribución en general, lo cual refleja la misma tendencia mencionada anteriormente al comparar los resultados KM con QM/MM. Por otra parte, las energías de polarización obtenidas a partir del modelo clásico de doble carga reproducen razonablemente bien los cálculos QM/MM y KM. Los coeficientes de correlación son buenos ($r^2=0.92$ y $r^2=0.95$, respectivamente), y la pendiente de la recta de regresión no muestra desviaciones claras entre el modelo clásico de doble carga y las energías de polarización obtenidas con los otros métodos. Ello, pues, apoya la estrategia adoptada para introducir los efectos de polarización a través del modelo de doble carga propuesto y la conveniencia del factor λ utilizado en este caso para la serie de complejos. Finalmente, el modelo de doble carga reproduce la energía electrostática+polarización determinada a nivel QM/MM y a nivel KM, con coeficientes de correlación de 0.99 y 0.96, respectivamente.

Como validación final del comportamiento del modelo clásico de doble carga, se han considerado una serie de configuraciones extraídas de simulaciones de dinámica molecular de varios solutos en solución acuosa. La comparación de los resultados obtenidos (véase Figura 8) a partir del modelo clásico de doble carga y de resultados QM/MM muestran una elevada similitud cuantitativa. Teniendo en cuenta la simplicidad inherente al modelo de doble carga y su reducido coste computacional, estos resultados son particularmente esperanzadores.

Artículo 7. Determinación de modelos distribuidos de polarizabilidades atómicas

Tal como se comentó anteriormente, la utilización de momentos dipolares inducidos es el procedimiento más extendido para la introducción de efectos de polarización en campos de fuerzas. Su uso, sin embargo, requiere como paso previo la asignación de polarizabilidades a cada uno de los centros asociados a los momentos dipolares inducidos. Generalmente, dichos parámetros han sido obtenidos mediante ajuste a las polarizabilidades moleculares determinadas experimentalmente para una serie de moléculas pequeñas representativas de diversos grupos funcionales, procedimiento originalmente propuesto por Applequist [27] y posteriormente extendido por otros autores [28]. No obstante, dicho procedimiento asume la transferibilidad de las polarizabilidades y conlleva cierta arbitrariedad en cuanto concierne a la definición de tipos de átomos. El objetivo de este trabajo fue proponer un procedimiento para derivar modelos distribuidos de polarizabilidades atómicas y examinar su transferibilidad para series homólogas de compuestos.

El procedimiento de obtención de los modelos distribuidos de polarizabilidades atómicas se basa en la evaluación de la energía de polarización utilizando una expresión perturbativa examinada previamente en nuestro grupo de trabajo [29]. De acuerdo con este procedimiento, cada molécula se coloca en el interior de una malla tridimensional, y para aquellos puntos situados a una distancia entre 2 y 5 veces el radio de van der Waals del átomo más próximo se evalúa perturbativamente la energía de polarización ejercida por la unidad de carga positiva colocada en cada uno de dichos puntos a nivel HF usando la base de Sadlej [30]. Las energías de inducción obtenidas de esta manera se escalan posteriormente utilizando una función dependiente de distancia, que fue parametrizada para reproducir los valores de energías de polarización determinados a nivel MP2 con dicha base. El conjunto de polarizabilidades atómicas se obtiene finalmente ajustando las energías de polarización obtenidas con un modelo de dipolos inducidos a las energías de polarización determinadas según el procedimiento descrito anteriormente. Para ello, se utiliza el programa OPEP (Optimal Partitioning of Electrical Properties) desarrollado por el grupo de Nancy [31].

Los modelos distribuidos de polarizabilidades atómicas se han determinado para una serie de compuestos aromáticos, generalmente derivados bencénicos (Figura 1). La elección de dichos compuestos obedece básicamente a dos motivos: i) la disponibilidad de datos previos sobre la energía de polarización resultante de la interacción con el ión sodio [32]; ii) al tratarse de una serie de compuestos estructuralmente relacionada, es posible examinar el grado de transferibilidad de las polarizabilidades atómicas entre diversos compuestos. Para los derivados seleccionados (Figura 1) se han determinado polarizabilidades atómicas isotrópicas considerando tanto una descripción *all-atom* como *united-atom* (asignando en este caso una única polarizabilidad al grupo C-H). Asimismo, para discutir la transferibilidad de las polarizabilidades, se han derivado dichos modelos sin imponer ningún tipo de restricción, o bien imponiendo igualdad de polarizabilidades atómicas para aquellos átomos relacionados por simetría, o finalmente imponiendo igualdad de polarizabilidades atómicas a aquellos centros englobados en el mismo tipo de átomo.

La Tabla 1 muestra las energías de polarización obtenidas variacionalmente para los complejos catión- π a nivel MP2/Sadlej y los valores obtenidos empleando el procedimiento perturbativo mencionado anteriormente. Tal como se aprecia en la Tabla 1, las estimaciones perturbativas reproducen los valores variacionales con una desviación de únicamente el 9% ($r^2=0.97$) y un coste computacional muy inferior, lo que apoya su uso como estrategia general para la derivación de polarizabilidades atómicas ajustadas a la molécula de interés.

La comparación de las polarizabilidades moleculares obtenidas a partir de los diversos modelos con los valores calculados variacionalmente a nivel MP2/Sadlej se muestran en la Tabla 2. La concordancia obtenida es muy satisfactoria cuando se utiliza una descripción *all-atom*, pues la desviación promedio es inferior al 1% aún cuando se imponen restricciones

ligadas al tipo de átomo en el ajuste de las polarizabilidades atómicas. En el caso de utilizar una descripción *united-atom*, la concordancia entre las polarizabilidades moleculares obtenidas por los modelos y variacionalmente empeora ligeramente, aunque la desviación promedio es sólo del 4%. Cabe hacer notar en este punto que en el proceso de obtención de los modelos distribuidos no se utilizó la polarizabilidad molecular como un parámetro a ajustar, por lo que la concordancia obtenida representa un fuerte soporte al método.

Asimismo, las energías de polarización de los distintos modelos con el Na^+ (puesto en su posición óptima para establecer interacción catión- π) muestran una elevada concordancia, especialmente cuando se utiliza una descripción *all-atom*. Así, en este caso la diferencia entre las energías de polarización obtenidas a nivel QM y clásicamente difieren menos de 2%, si bien dicha diferencia llega a ser del 15% cuando se consideran modelos *united-atom*. Los resultados también ponen de relieve que el acoplamiento mutuo entre los dipolos inducidos para los diversos complejos catión- π tiene un efecto relativamente pequeño sobre las energías de inducción. Ello, pues, sugiere que dicho acoplamiento podría obviarse, lo cual permitiría suprimir el proceso de autoconherencia que debe llevarse a cabo hasta conseguir la convergencia entre dipolos inducidos, con el consiguiente ahorro de coste computacional en simulaciones de sistemas complejos.

Las polarizabilidades atómicas obtenidas para los átomos C e H (*all-atom*) y el grupo atómico C-H (*united-atom*) en los diversos modelos distribuidos están indicadas en la Tabla 6. En promedio, no se aprecian diferencias relevantes en cuanto a las polarizabilidades atómicas obtenidas sin restricciones, o bien imponiendo simetría o equivalencias por tipo de átomo. Así, la polarizabilidad atómica de C oscila (en u.a.) entre 9.6 y 10.4, para el H varía entre 3.4 y 3.6, y para el grupo C-H adopta valores comprendidos entre 16.7 y 19.1. No obstante, cabe hacer notar que las desviaciones estándar son elevadas, lo cual sugiere que deben existir diferencias marcadas en las polarizabilidades atómicas obtenidas para dichos átomos en las diferentes moléculas. Ello queda claramente reflejado en los resultados expuestos en las Tablas 7 (*all-atom*) y 8 (*united-atom*), donde se muestran las polarizabilidades atómicas promedio obtenidas en cada uno de los compuestos. Así, por ejemplo, la polarizabilidad de C obtenida sin imponer restricciones varía desde 4.0 hasta 18.2 (en u.a.), mientras que la polarizabilidad de C-H está comprendida entre 12.7 y 31.5. Así, pues, los resultados obtenidos cuestionan seriamente que los modelos distribuidos de polarizabilidades atómicas sean transferibles entre moléculas estructuralmente relacionadas.

En conjunto, cabe apuntar que el procedimiento descrito en este trabajo constituye una alternativa prometedora para poder derivar modelos distribuidos de polarizabilidades atómicas. Así, los modelos obtenidos permiten reproducir tanto las energías de polarización obtenidas variacionalmente como la polarizabilidad global de la molécula. No se aprecian diferencias

relevantes en la magnitud de las polarizabilidades atómicas debidas al hecho de imponer restricciones por simetría o por tipo de átomo respecto a los valores obtenidos sin imponer ningún tipo de restricción. Por el contrario, si se observan diferencias marcadas en cuanto a los valores obtenidos para moléculas diferentes, lo cual dificulta la definición de un único conjunto de valores atómicos transferibles entre moléculas. Por ello, cabe pensar que este procedimiento pueda ser una alternativa eficaz para la definición de modelos distribuidos de polarizabilidades atómicas.

Bibliografía

- [1] J. Tomasi and M. Perisco. *Chem. Rev.*, 94:2027, 1994.
- [2] C. J. Cramer and D. G. Truhlar. *Chem. Rev.*, 99:2161, 1999.
- [3] J. Tomasi, R. Cammi, B. Menucci, C. Capelli, and S. Corni. *Phys. Chem. Chem. Phys.*, 4:5697, 2002.
- [4] M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart, J. V. Coe, and T. R. Tuttle Jr. *J. Phys. Chem. A*, 102:7787, 1998.
- [5] T. R. Tuttle Jr, S. Malaxos, and J. V. Coe. *J. Phys. Chem. A*, 106:925, 2002.
- [6] J. V. Coe. *Int. Rev. Phys. Chem.*, 20:33, 2001.
- [7] P. Carrupt, B. Testa, and P. Gaillard. *Reviews in Computational Chemistry*, volume 11, page 241. Wiley-VCH, New York, 1997.
- [8] J. Sangster. *Octanol-water Partition Coefficients: Fundamentals and Physical Chemistry*. Wiley, Chichester, 1997.
- [9] C. Hansch and A. Leo. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*. American Chemical Society, Washington DC, 1995.
- [10] V. Pliska, B. Testa, and H. van de Waterbeemd, editors. *Lipophilicity in Drug Action and Toxicology*. VCH, Weinheim, 1996.
- [11] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. *Adv. Drug Deliv. Rev.*, 23:3, 1997.
- [12] R. Mannhold and K. Dross. *Quant. Struct. Act. Relat.*, 15:403, 1996.
- [13] R. Manhold, G. Cruciani, K. Dross, and R. J. Rekker. *J. Comput-Aided Mol. Des.*, 12:573, 1998.
- [14] R. Manhold and H. J. van de Waterbeemd. *J. Comput-Aided Mol. Des.*, 15:337, 2001.
- [15] J. D. Thompson, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. A*, 108:6532, 2004.

- [16] A. Klamt. *J. Phys. Chem.*, 99:2224, 1995.
- [17] A. Klamt, V. Jonas, T. Bürger, and J. C. W. Lohrenz. *J. Phys. Chem. A*, 102:5074, 1998.
- [18] F. J. Luque, X. Barril, and M. Orozco. *J. Comput-Aided Mol. Des.*, 13:139, 1999.
- [19] J. Muñoz, X. Barril, F. J. Luque, J. L. Gelpí, and M. Orozco. *Fundamentals of Molecular Similarity*, page 143. Kluwer Academic/Plenum, New York, 2001.
- [20] J. Muñoz, X. Barril, B. Hernández, M. Orozco, and F. J. Luque. *J. Comput. Chem.*, 23:554, 2002.
- [21] A. J. Stone. *The Theory of Intermolecular Forces*. Clarendon Press, Oxford, 1996.
- [22] K. Morokuma. *J. Chem. Phys.*, 55:1236, 1971.
- [23] K. Kitaura and K. Morokuma. *Int. J. Quant. Chem.*, 10:325, 1976.
- [24] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, M. D. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. Kollman. *J. Am. Chem. Soc.*, 117:5179, 1995.
- [25] N. Foloppe and A. D. MacKerell. *J. Comput. Chem.*, 21:86, 2000.
- [26] X. Barril, S. G. Kalko, M. Orozco, and F. J. Luque. *Mini Rev. Med. Chem.*, 2:27, 2002.
- [27] J. Applequist, J. R. Carl, and K. Fung. *J. Am. Chem. Soc.*, 94:2952, 1972.
- [28] J. M. Stout and C. E. Dykstra. *J. Am. Chem. Soc.*, 117:5127, 1995.
- [29] C. Chipot, F. Dehez, J. Angyan, C. Millot, M. Orozco, and F. J. Luque. *J. Phys. Chem. A*, 105:11505, 2001.
- [30] A. J. Sadlej. *Collect. Czech. Chem. Commun.*, 53:1995, 1988.
- [31] J. G. Ángyán, C. Chipot, F. Dehez, C. Hättig, G. Jansen, and C. Millot. *J. Comput. Chem.*, 24:997, 2003.
- [32] E. Cubero, F. J. Luque, and M. Orozco. *Proc. Natl. Acad. Sci. USA*, 95:5976, 1998.

Capítulo 7

Conclusiones

- Se ha presentado la parametrización de las versiones *ab initio* HF/6-31G(d) y semiempíricas AM1 y PM3 del modelo continuo MST para la solvatación en octanol, así como el refinamiento de las parametrizaciones previas realizadas para agua, cloroformo y tetracloruro de carbono. El modelo MST final permite reproducir de forma precisa las energías libres de solvatación en octanol para una variedad de solutos neutros, así como los coeficientes de partición octanol/agua para un amplio conjunto de compuestos incluyendo fármacos.
- La reducción local de un 8 % en la cavidad electrostática con respecto a la empleada para solutos neutros es suficiente para reproducir satisfactoriamente la energía libre de hidratación de iones monovalentes, siempre que los efectos derivados de la densidad electrónica localizada fuera de la cavidad se hayan corregido. En este último aspecto es necesario que la corrección por renormalización de carga sea realizada mediante algoritmos que consideren explícitamente la anisotropía en la distribución de carga del soluto.
- El estudio comparativo llevado a cabo sobre tres modelos continuos SCRF (SMx, MPE y MST) ha permitido caracterizar la respuesta electrostática proporcionada por dichos métodos. Todos ellos dan un perfil similar de energía libre electrostática versus permitividad, pero los valores absolutos de las energías libre electrostática para una permitividad dada pueden ser diferentes. Para cavidades equivalentes las mayores discrepancias se encuentran para el método SMx que da una respuesta electrostática claramente inferior a la que proporcionan los métodos MPE y MST. Esto sugiere que mientras los métodos de Nancy y Pisa-Barcelona pueden emplear cavidades similares, el método de Minnesota debe emplear cavidades necesariamente menores a fin de obtener una respuesta electrostática similar.
- Las características intrínsecas a la aproximación NDDO permite definir un modelo de partición de la energía libre de solvatación basado en contribuciones por átomo. Esta

partición es muy útil para examinar el grado de transferibilidad de las contribuciones fragmentales a la solvatación, típicamente asumidas como aditivas en los métodos de partición experimentales. Empleando esta metodología hemos sido capaces de determinar la dependencia de las contribuciones fragmentales al logPow respecto a efectos electrónicos, de puente de hidrógeno, estéricos y tautoméricos.

- Estudios sistemáticos nos han permitido determinar la sensibilidad de las contribuciones fraccionales a la energía libre de solvatación con respecto al método de renormalización de carga empleado en el contexto de modelos SCRF de carga aparente (como MST). En concreto, para compuestos con una fuerte separación de carga como es el caso de los derivados de alcanos sustituidos, el uso de métodos de renormalización de carga que consideren explícitamente la anisotropía de los solutos conduce a una mayor transferibilidad de las contribuciones fraccionales.
- Los resultados obtenidos para complejos de puente de hidrógeno y catión- π indican que la energía de polarización electrónica se puede obtener a partir de la adición de las energías de polarización de los monómeros interaccionantes. Por otra parte, los métodos QM/MM son capaces de reproducir satisfactoriamente los resultados obtenidos por los métodos QM/MM, donde el monómero polarizado es tratado a nivel QM y el polarizante por un conjunto de cargas puntuales. Finalmente, un modelo de doble carga, basado en el ajuste simultáneo de energías electrostáticas y de polarización, permiten evaluar de forma satisfactoria la energía de polarización en dichos complejos. Teniendo en cuenta la simplicidad inherente al modelo de doble carga y su reducido coste computacional, el modelo puede ser refinado y empleado para tratar explícitamente efectos de polarización en simulaciones clásicas.
- Los modelos distribuidos de polarizabilidades atómicas permiten reproducir las energías de polarización obtenidas variacionalmente, así como la polarizabilidad global de la molécula. Las polarizabilidades atómicas aparecen difíciles de transferir entre átomos de un mismo tipo en moléculas diferentes, por lo que parece recomendable realizar una parametrización específica para cada molécula. Por ello, el procedimiento descrito para derivar modelos distribuidos de polarizabilidades atómicas constituye una herramienta prometedora para poder evaluar con precisión la energía de polarización en el marco de campos de fuerza clásico basados en el modelo de momentos dipolares inducidos.

Apéndice A

Otros trabajos publicados

1. *Prediction of conformational free energy differences of solutes in solution: an MC-MST study.* Hernández, B.; Curutchet, C.; Colominas, C.; Orozco, M. and Luque, F. J. *Molecular Simulations* **2002**, 28, 153
2. *Continuum solvation models: Dissecting the free energy of solvation.* Luque, F. J.; Curutchet, C.; Muñoz-Muriedas, J.; Bidon-Chanal, A.; Soteras, I.; Morreale, A.; Gelpí, J. L. and Orozco, M. *Physical Chemistry Chemical Physics* **2003**, 18, 3827
3. *Solute-solvent interactions from QM SCRF methods: Analysis of group contributions to solvation.* Luque, F. J.; Bidon-Chanal, A.; Muñoz-Muriedas, J.; Soteras, I.; Curutchet, C.; Morreale, A. and Orozco, M. in *Fundamental World of Quantum Chemistry: A tribute volume to the memory of Per-Olov Löwdin, Volume 3.* Eds: E. Brandas, E. Kryachko. Kluwer Academic Publishers, **2004**
4. *Extension of the MST model to the IEF formalism: HF and B3LYP parametrizations.* Soteras, I.; Curutchet, C.; Bidon-Chanal, A.; Orozco, M. and Luque, F. J., enviado

Apéndice B

Abreviaciones

Å	Angstrom
AA	All Atom
AM1	Austin Model 1
AMBER	Assisted Model Building with Energy Refinement
ASC	Apparent Surface Charge
CHARMM	Chemistry Harvard Molecular Mechanics
CI	Configuration Interaction
CM2	Charge Model 2
CNDO	Complete Neglect of Differential Overlap
COSMO	Conductor-like Screening Model
COSMO-RS	Conductor-like Screening Model for Real Solvents
DFT	Density Functional Theory
ESP	Electrostatic Potential
FEP	Free Energy Perturbation
FDPB	Finite Difference Poisson-Boltzmann
FF	Force Field
GB	Generalized Born
GROMOS	Groningen Molecular Simulation
GTO	Gaussian-Type Orbital
HF	Hartree-Fock
IEF	Integral Equation Formalism
IPCM	Isodensity Polarizable Continuum Model
KM	Kitaura Morokuma
MBPT	Many-Body Perturbation Theory
MC	Monte Carlo
MCSCF	Multi-Configurational Self-Consistent Field
MD	Molecular Dynamics

MEP Molecular Electrostatic Potential
MM Molecular Mechanics
MM3 Molecular Mechanics 3
MNDO Modified Neglect of Differential Overlap
MO Molecular Orbital
MOPAC Molecular Orbital Package
MP2 Møller-Plesset 2
MPE Multipolar Expansion
MSE Mean Signed Error
MST Miertus Scrocco Tomasi
MUE Mean Unsigned Error
NDDO Neglect of Diatomic Differential Overlap
OPLS Optimized Potential for Liquid Simulations
PB Poisson-Boltzmann
PCM Polarizable Continuum Model
PM3 Parametrized Model 3
QCI Quadratic Configuration Interaction
QM Quantum Mechanics
QSAR Quantitative Structure-Activity Relationship
RHF Restricted Hartree-Fock
RESP Restrained Electrostatic Potential
RMSd Root Mean Square deviation
SASA Solvent Accessible Surface Area
SCF Self-Consistent Field
SCIPCM Self-Consistent Isodensity Polarizable Continuum Model
SCRf Self-Consistent Reaction Field
SMx Solvation Model x
SPT Scaled Particle Theory
STO Slater-Type Orbital
TI Thermodynamic Integration
UA United Atom
UAHF United Atom model for Hartree-Fock
UHF Unrestricted Hartree-Fock
ZDO Zero Differential Overlap