J. Phys. Chem. B **1998,** *102,* 3637−3641

Comparison of a GB Solvation Model with Explicit Solvent Simulations: Potentials of Mean Force and Conformational Preferences of Alanine Dipeptide and 1,2-Dichloroethane

Marco Scarsi, Joannis Apostolakis, and Amedeo Caflisch*

Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received: December 11, 1997

The potential of mean force in aqueous solution for rotation around the two backbone dihedrals ϕ and ψ of the alanine dipeptide is computed in explicit water and in the continuum approximation by numerical integration of the self-energies and the generalized Born (GB) equation. The two models show good agreement. The experimentally observed increase in the gauche/trans population ratio for dichloroethane in going from the gas phase to the pure liquid is reproduced by the GB solvation model with a solvent dielectric constant of 10.5. This test case shows that the GB model gives accurate predictions also for solvents with much lower polarizability than water. For both test systems additional calculations with a finite difference Poisson equation solver yield somewhat more accurate results at a much higher computational cost than the GB solvation model.

1. Introduction

The accurate and efficient estimation of electrostatic energies of molecules in solution is required to solve a number of relevant problems in theoretical structural biology. These include the capability to discriminate between near-native and nonnative conformations of a protein as well as the predictions of molecular recognition events (docking, ligand design). When simulating molecules in a liquid environment, it is computationally convenient to approximate the electrostatic effects of the solvent by a continuum dielectric model.[1−10] The system is partitioned into solvent and solute regions and two different dielectric constants are assigned to each region. In this approximation only the intrasolute electrostatic interactions need to be evaluated. This strongly reduces the number of interactions with respect to an explicit treatment of the solvent. Moreover, the high computational cost due to the equilibration of the solvent molecules is avoided. The electrostatic effects of the solvent are often relevant, and it has been shown that the continuum dielectric model provides an accurate description of molecules in solution.[11] The problem of evaluating the electrostatic energy of a system with a spatial discontinuity in the dielectric constant can be solved at a low computational cost by calculating the interaction energies with the generalized Born (GB) equation and a numerical[5,10] or analytical[6−9] integration of the self-energies. For a large data set of peptide structures and conformations it has recently been shown that the numerical integration gives more accurate solvation free energies than the analytical.[12]

In this paper the implementation of the GB model described by Scarsi et al.[10] is shown to yield results in good agreement with explicit solvent simulations for two test systems. The first is the terminally blocked alanine dipeptide, which has been extensively used as a model for simulations of peptides and proteins.[13] The potential of mean force (PMF) for rotation around the ϕ and ψ backbone dihedrals is calculated by molecular dynamics (MD) simulations in explicit water and by a systematic search with evaluation of the solvation energy by the GB model. The nonelectrostatic contribution to solvation, if assumed to be proportional to the solvent accessible surface, is negligible.[11]

The second system is 1,2-dichloroethane (DCE) which was chosen for two reasons: it has been experimentally investigated with diffraction and spectroscopic methods,[14] and it has a much lower polarizability than water (dielectric constant value of 10.5, ref 15). An increase in the gauche/trans population ratio was observed experimentally in going from the gas to the liquid phase. This originates from the dielectric effect of liquid DCE and is a purely electrostatic effect.[16] These conformational preferences are correctly reproduced by MD in neat explicit DCE liquid and by a systematic search with an energy function based on the GB model.

For comparison, all calculations performed with the GB solvation model were repeated with a finite difference solution of the Poisson equation (FDP). The latter provides the exact continuum electrostatic energy (in the accuracy limits of a numerical method) in a nonhomogeneous dielectric, whereas the GB method is based on two main approximations: the Coulombic approximation, which is used in the evaluation of the self-energies, and the GB equation for the interactions between partial charges on the solute.[10] Although not as accurate as the results obtained with FDP, the GB results are good and require only a small fraction of the computational time needed by MD or FDP.

2. Methods

The implementation of the GB approach described in ref 10 was used to calculate the solvation energy. The dielectric boundary was described by the molecular surface of the solute. A dielectric constant of 1 was assigned to the volume occupied by the solute, while the continuum solvent region was assigned a dielectric constant of 78.5 (alanine dipeptide) and 10.5 (liquid DCE). A grid size of 0.4 Å was used for the numerical integration over the solute volume to calculate the self-energies. The total solvation energy (self-energies plus pairwise screening)

---

* Corresponding author. Tel: (411) 635 55 21. Email: caflisch@bioc.unizh.ch.

S1089-5647(98)00506-9 CCC: $15.00

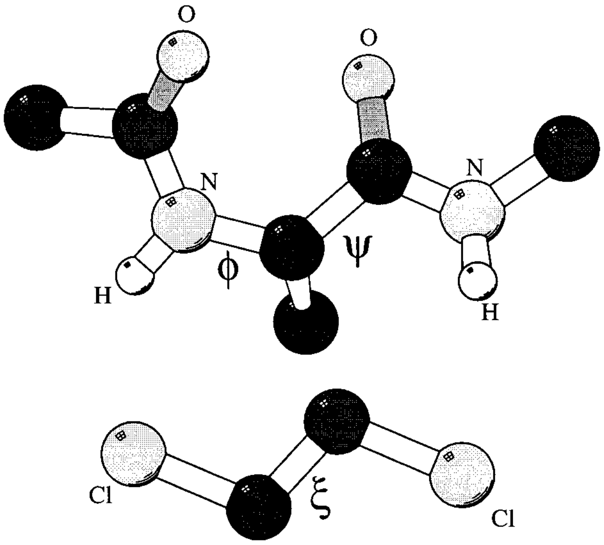

Figure 1. 3D representation of alanine dipeptide (a, top) and 1,2-dichloroethane (b, bottom). Carbon atoms are black and without labels. Single bonds are white and double bonds are gray. Rotatable bonds are labeled with Greek letters. This figure was made with the program MOLSCRIPT.[30]

was added to the energy in vacuo calculated by the CHARMM force field[17] to obtain the energy in solution.

All the FDP calculations were performed with the program UHBD.[18−20] The solute boundary and dielectric constants were the same as for the GB calculations. Three FDP calculations with decreasing grid size (focusing) were performed for every structure. The spacing of the final grid was 0.4 Å.

All minimizations and MD simulations with explicit water molecules were performed with CHARMM.[17]

2.1. Alanine Dipeptide. The alanine dipeptide was modeled according to the polar hydrogen parametrization of CHARMM (param19).[17] A 3D picture of the molecule is shown in Figure 1a.

Implicit Solvent. The PMF profiles were calculated as described by Marrone et al.[11] First a total of $36 \times 36 = 1296$ conformations of the alanine dipeptide were generated by incrementing in 10° steps the two backbone dihedrals ϕ and ψ from $-180°$ to 180°. These conformations were subjected to conjugate gradient minimization[21] in vacuo with harmonic force restraints on ϕ and ψ to maintain the initial structure. The harmonic force constant was 72 kcal/(mol rad^2). The minimizations were assumed to have converged when the root-mean-square of the gradient reached a value smaller than 0.01 kcal/(mol Å). The vacuum energy and the continuum solvation energies (GB and FDP) were calculated for each relaxed conformation on the ϕ, ψ map. Defining $E(\phi,\psi)$ as the energy in vacuo or in solution (CHARMM vacuum energy plus continuum solvation term) of a conformation with given ϕ and ψ dihedrals, the PMF for rotation around the ϕ dihedral angle at temperature T is

$$PMF(\phi) = -RT \ln\left(\frac{\sum_{\psi}\exp(-E(\phi,\psi)/RT)}{\sum_{\phi'\psi'}\exp(-E(\phi',\psi')/RT)}\right) \qquad (1)$$

where R is the gas constant. The argument of the natural logarithm is the probability of finding the alanine dipeptide in a conformation with the dihedral angle ϕ at temperature T. For rotation around the dihedral ψ the same formula applies with substitution of ϕ by ψ and vice versa.

Explicit Solvent. The PMF for the alanine dipeptide in explicit water was calculated by MD simulations and the umbrella sampling procedure over ϕ and ψ. The force constant for the quadratic umbrella potential was 125 kcal/(mol rad^2). The alanine dipeptide was simulated in a cubic box of 18.86 Å on a side containing 207 water molecules with periodic boundary conditions. To scan adequately all the conformational space, the minimum of the umbrella potential was moved by 10° from $-180°$ to 180° along the dihedral angle of interest. An equilibration run of 10 ps and a data collection run of 40 ps were performed for every position of the restraining potential. Thus, a total of 1.44 ns of production runs were performed for each dihedral. Fast bond oscillations were frozen by the SHAKE algorithm[22] and a time step of 2 fs was used. The production runs were performed at constant temperature (300 K) by weak coupling of the system to a heat bath (time constant of 5 ps).[23]

2.2. 1,2-Dichloroethane. DCE was modeled in the united-atom representation (Figure 1b). The dihedral energy term for rotation around the dihedral angle ξ was parametrized as described by Jorgensen et al.[16] to reproduce the experimental data on the energy barriers and differences between the trans and the two gauche conformers in the gas phase.[24−26]

Implicit Solvent. The energy in the gas phase is assumed to consist only of the dihedral potential.[16] Values of the partial charges and atomic radii needed for the evaluation of the GB solvation energy were taken from partial charges and van der Waals radii of Jorgensen et al.[16] A dielectric constant of 10.5 was assigned to the continuum solvent (liquid DCE).[27] A systematic search over the conformational space was performed by incrementing the dihedral angle in 10° steps from $-180°$ to 180°. No minimization was required since the dihedral potential was the only energy term. The vacuum energy and the solvation energy (with the GB and FDP models) were evaluated for every conformation.

Explicit Solvent. MD of the pure liquid phase was performed by simulating 216 DCE molecules in a periodic cubic box of 30.54 Å on a side. The dimensions of the box reproduce the experimental density of DCE, 1.246 g/cm^3.[28] The intramolecular energy includes bond and bond angle energy terms in addition to the dihedral potential. The minima and force constants for the first two terms are the same as the ones used by Gilson et al.[27] To compare the MD simulation results with a previous Monte Carlo simulation of Jorgensen et al.,[16] four different MD schemes were tested. They differed in the long-range truncation option of the nonbonded interactions and in the use of SHAKE to constrain the covalent bonds. In the first truncation scheme the "shift" function[17] was applied with a cutoff of 12.5 Å. In the second truncation scheme the nonbonded forces were multiplied by a "switching" function[29] acting between 10.75 and 11.25 A. This small interval is expected to reproduce better the Monte Carlo simulation scheme, where a cutoff at 11 Å was used for the nonbonded interactions.[16] Two simulations, with and without SHAKE, were performed for every truncation option. When SHAKE was invoked, additional bonds were added between the two 1−3 atom pairs C−Cl. In this way the bonds and angles were constrained by invoking SHAKE only on the bonds. This allows for a better comparison with the Monte Carlo simulation.[16] The following acronyms will be used henceforth: MD/SH-SK for "shift" on the nonbonded energies and SHAKE, MD/SH for "shift" on the nonbonded energies and no intramolecular constraints, MD/FSW-SK for "switch" on the nonbonded forces and SHAKE, and MD/FSW for "switch" on the nonbonded forces and no

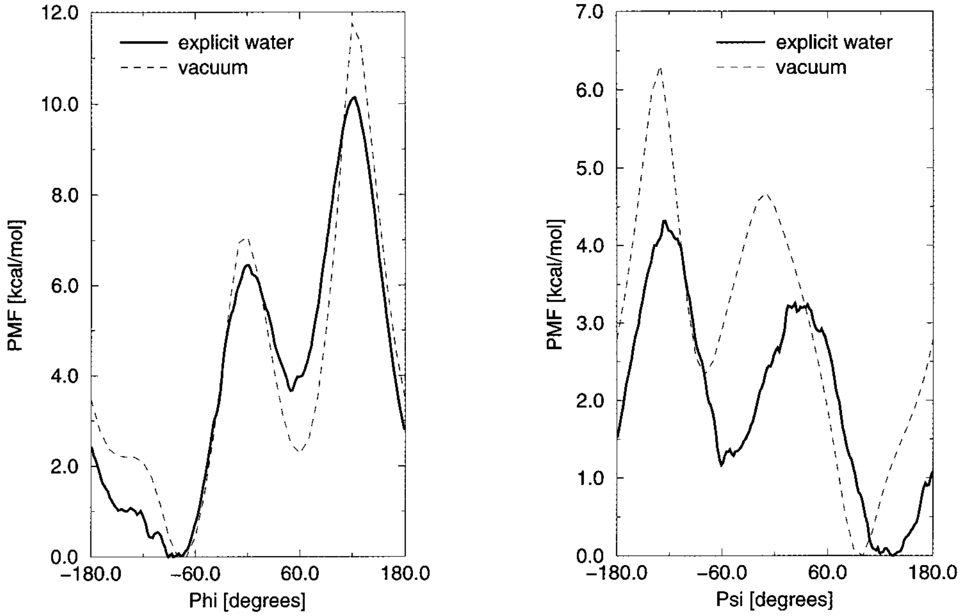

Figure 2. PMF of alanine dipeptide in vacuo and in solution (MD in explicit water) along the ϕ (a, left) and the ψ (b, right) coordinates.

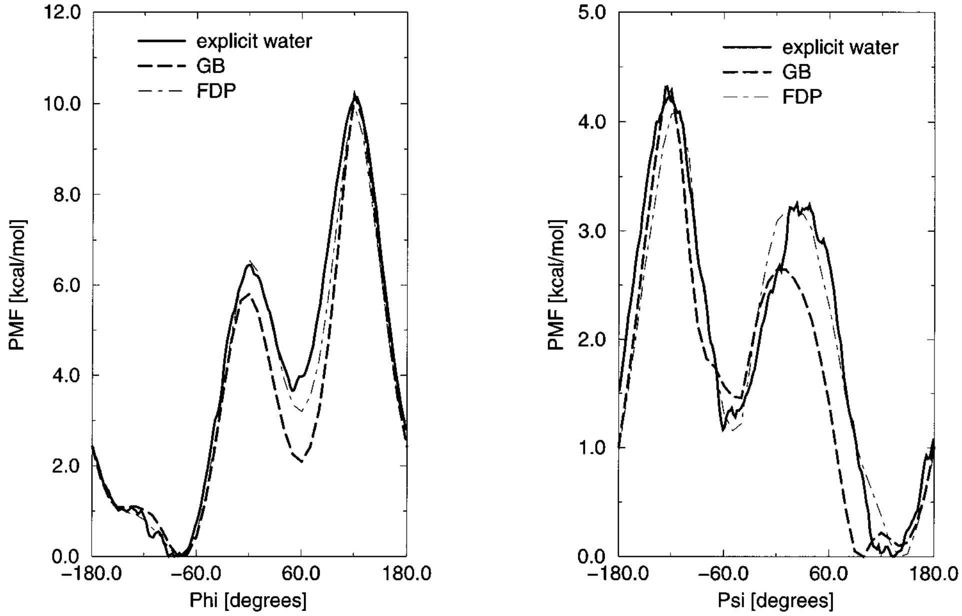

Figure 3. PMF of alanine dipeptide in solution along the ϕ (a, left) and the ψ (b, right) coordinates calculated with three different solvation models (MD in explicit water, GB and FDP continuum approaches).

intramolecular constraints. In the four simulations the system was heated and equilibrated at 298 K for a total of 140 ps. Subsequently 1 ns of data collection was run. The time step was 2 fs in MD/SH and MD/FSW and 4 fs in MD/SH-SK and MD/FSW-SK. Snapshots of the system saved every 20 time steps were used for the statistical analysis of the conformational preferences of DCE in the liquid phase. Given *prob* (ξ) as the probability of occurrence of the dihedral ξ, the potential V for rotation around ξ was obtained in the liquid phase from

$$V(\xi) = -RT\ln(prob\,(\xi)) \qquad (2)$$

3. Results

3.1. Alanine Dipeptide. Figure 2 shows a comparison between the PMF along the ϕ and ψ coordinates in explicit water and in vacuo. As a convention the global minimum of a PMF will be set to zero, since it is undetermined to an additive constant. Along the ψ coordinate the solvation strongly affects the shape of the potential, whereas the solvent effect is smaller along ϕ. Figure 3 shows the PMF obtained with MD in explicit water and the PMF calculated by the systematic search supplemented by the continuum approaches. The two continuum models reproduce correctly the main solvent effects, despite the differences in the conformational sampling adopted for the explicit and implicit solvent. For the GB model the largest deviations from the explicit water results are of the order of $3RT$ at $\phi \simeq 60°$ and $2RT$ at $\psi \simeq 30°$. These discrepancies are smaller in the FDP approach, where the largest deviations are about $1.5RT$ at $\phi \simeq 60°$ and $1RT$ at $\psi \simeq 120°$. The lower accuracy originates from the Coulombic approximation to evaluate the self-energies and the use of the GB equation for the interactions between partial charges in the solute.[10] For the alanine dipeptide PMF the discrepancies between the GB solvation model and the full solution of the Poisson equation

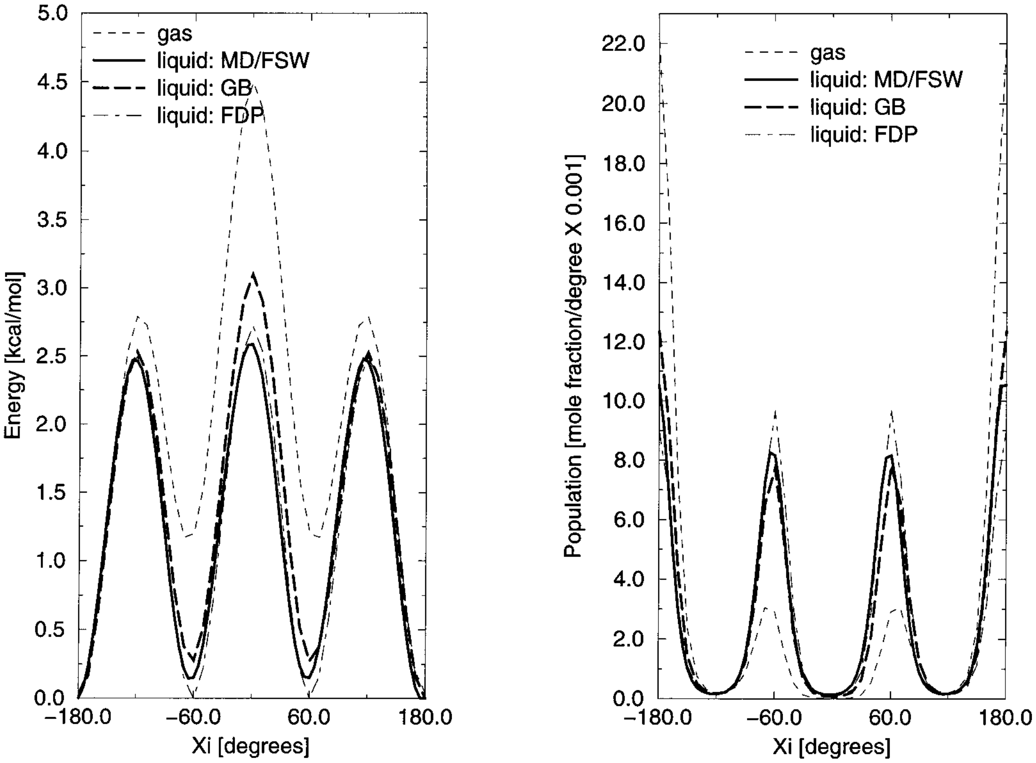

Figure 4. (a, left) Potential energy for rotation around the C−C bond of DCE and (b, right) dihedral population distribution in the gas phase and in the neat liquid phase calculated with three different solvation models (MD in explicit liquid DCE, GB, and FDP continuum approaches).

TABLE 1: Conformational Preferences of DCE at 298 K

	$(g/t)_{\mathrm{gas}}$	$(g/t)_{\mathrm{liquid}}$	ΔΔG ([kcal/mol])[a]
experiments[26]	0.30 ± 0.03	1.86 ± 0.3	1.08
GB	0.31	1.25	0.85
FDP	0.31	2.04	1.12
MD/FSW	0.31	1.57	0.98
MD/FSW-SK	0.31	2.22	1.18
MD/SH	0.31	2.38	1.20
MD/SH-SK	0.31	3.30	1.40

[a] Change in the free energy difference between the gauche and trans conformations in going from the gas to the liquid phase: $\Delta\Delta G = -RT \ln ((g/t)_{\mathrm{gas}}/(g/t)_{\mathrm{liquid}})$.

are somewhat larger than the error originating from the continuum dielectric approximation, in contrast with previous assumptions.[6,9] The better accuracy of the FDP approach comes at a higher computational cost. The evaluation of the PMF along the two coordinates took 3 min with the GB model compared to about 10 h for the FDP approach on an SGI workstation (processor R4400, clock frequency 175 MHz). On the same processor 1 ns of MD took about 64 h.

3.2. 1,2-Dichloroethane. Figure 4a shows the potential for rotation around the C−C bond in the gas phase and in the liquid phase calculated by the GB and FDP approaches and the MD/FSW simulation. In the liquid phase the energy of the two gauche conformations is lower than in the gas phase. The consequences of the dielectric effects in the neat liquid are evident in the population distribution (Figure 4b). To calculate the ratio gauche/trans *(g/t)*, all the conformations having a dihedral smaller than −120° or larger than 120° are considered trans, while all the other ones are defined gauche.[16] Integration of the population distribution according to this assumption yields the ratios shown in Table 1. The GB solvation model gives a $(g/t)_{\mathrm{liquid}}$ of 1.25, to be compared with the experimental value of 1.86 ± 0.3 obtained from analysis of infrared intensities.[26] $(g/t)_{\mathrm{liquid}}$ from the FDP approach is 2.04. The four MD trajectories yield values of $(g/t)_{\mathrm{liquid}}$ ranging from 1.57 (MD/FSW) to 3.30 (MD/SH-SK). As a basis of comparison, a Monte Carlo simulation in a periodic box containing 128 monomers gave a $(g/t)_{\mathrm{liquid}}$ of 1.30.[16] Surprisingly, this deviates less from the MD results obtained without SHAKE than from the ones obtained with SHAKE.

The accuracy of the continuum approximation is remarkable if one considers the simplicity of the model and the fact that no parameter has been optimized. Table 1 shows the change in the free energy difference between the gauche and trans conformations in going from the gas to the liquid phase. It is calculated as $\Delta\Delta G = -RT \ln ((g/t)_{\mathrm{gas}}/(g/t)_{\mathrm{liquid}})$. The present implementation of the GB model gives a ΔΔG of 0.85 kcal/mol to be compared with 1.08 kcal/mol from experimental data. The GB implicit description of the solvent, up to now tested only with aqueous solvent (dielectric constant of about 80), approximates correctly also the liquid phase of DCE which has a much lower polarizability (10.5). The calculation of the potential in solution required 11 s with the GB model compared to 20 min with the FDP model and about 4 h for 100 ps of MD on an SGI workstation (processor R4400, clock frequency 175 MHz).

4. Conclusion

In the continuum dielectric approximation the combination of the numerical integration of the self-energies with the GB equation for the interaction energies allows for the efficient and accurate evaluation of solvation energies. Two systems have been analyzed in this paper, the alanine dipeptide and DCE. The calculations were performed with explicit solvent molecules and in the continuum approximation with the GB and FDP approaches. For the PMF of the alanine dipeptide the GB solvation model provided a correct description of the main effects of the aqueous solvent. For liquid DCE, which has a much lower polarizability than water, the GB model yielded an increase in the gauche/trans population ratio in going from the gas phase to the neat liquid. This is in qualitative agreement with experimental data. No parametrization was performed for the continuum models. The FDP approach showed a somewhat higher accuracy than the GB model at the expense of much higher CPU requirements.

Acknowledgment. We thank Prof. J. A. McCammon for providing the UHBD program, which was used for all the finite difference calculations. This work was supported by the Swiss National Science Foundation (Nationalfonds, Grant 3139-43423.95), the Swiss Federal Office for Public Health (Nationales AIDS-Forschungs-Programm, Grant 3139-043652.95), the Hartmann Müller Foundation, and the Jubiläumspende of the University of Zürich.

References and Notes

(1) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157,* 671−679.

(2) Gilson, M. K.; Honig, B. H. *Proteins: Struct., Funct., Genet.* **1988**, *4,* 7−18.

(3) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29,* 10219−10225.

(4) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Comm.* **1991**, *62, 187*−197.

(5) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112,* 6127−6129.

(6) Hawkins, G. D.; Cramer, C. J.; Trulhar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.

(7) Hawkins, G. D.; Cramer, C. J.; Trulhar, D. G. *J. Phys. Chem.* **1996**, *100,* 19824−19839.

(8) Schaefer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100, 1578*−1599.

(9) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101,* 3005−3014.

(10) Scarsi, M.; Apostolakis, J.; Caflisch, A. *J. Phys. Chem. A* **1997**, *101,* 8098−8106.

(11) Marrone, T. J.; Gilson, M. K.; McCammon, J. A. *J. Phys. Chem.* **1996**, *100,* 1439−1441.

(12) Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101,* 1190−1197.

(13) Brooks III, C. L.; Case, D. A. *Chem. Rev.* **1993**, *93*, 2487−2502.

(14) Abraham, R. J.; Bretschneider, E. *Internal Rotation in Molecules;* Orville-Thomas, W. J., Ed.; Wiley: London, 1974; Chapter 13.

(15) Lide, D. R. *Handbook of Chemistry and Physics;* CRC Press: Boca Raton, FL, 1994; Section 6.

(16) Jorgensen, W. L.; Binning, Jr., R. C.; Bigot, B. *J. Am. Chem. Soc.* **1981**, *103*, 4393−4399.

(17) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(18) Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1989**, *10,* 386−391.

(19) Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1990,** *11,* 401−409.

(20) Davis, M. E.; McCammon, J. A. *J. Comput. Chem.* **1991,** *12,* 909−912.

(21) Fletcher, R.; Reeves, C. M. *Comput. J.* **1964**, *7,* 149−154.

(22) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1977**, *23,* 327−341.

(23) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81,* 3684−3690.

(24) Kveseth, A. *Acta Chem. Scand., Ser. A* **1975**, *A29,* 307.

(25) Lowe, J. P. *Prog. Phys. Org. Chem.* **1968**, *6*, 1.

(26) Tanabe, K. *Spectrochim. Acta, Part A* **1972**, *28A*, 407.

(27) Gilson, M. K.; McCammon, J. A.; Madura, J. D. *J. Comput. Chem.* **1994**, *16*, 1081−1095.

(28) Wilhelm, E.; Schano, R.; Becker, G; Findenegg, G. H.; F., Kohler *Trans. Faraday Soc.* **1969**, *65*, 1443.

(29) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15,* 667−683.

(30) Kraulis, P. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.