*J. Phys. Chem. A* **2005,** *109,* 3425−3432

# Simulation of Liquid Water Using Semiempirical Hamiltonians and the Divide and Conquer Approach

**G. Monard,*,[†] M. I. Bernal-Uruchurtu,[‡] A. van der Vaart,[§,||] K. M. Merz Jr.,[§] and M. F. Ruiz-López[†]**

*Equipe de Chimie et Biochimie Théoriques, UMR CNRS-UHP No. 7565, Université Henri Poincaré-Nancy I, BP 239, 54506 Vandoeuvre-lès-Nancy, France, Centro de Investigaciones Químicas, Universidad Autónoma del Estado de Morelos, 62210 Cuernavaca, México, Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, and Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138*

*Received: September 10, 2004; In Final Form: January 19, 2005*

This work examines the ability of semiempirical methods to describe the structure of liquid water. Particularly, the standard AM1 and PM3 methods together with recently developed PM3-PIF and PM3-MAIS parametrizations have been considered. We perform molecular dynamics simulations for a system consisting of 64 or 216 water molecules in a periodic cubic box. The whole system is described quantum mechanically. Calculations with 64 molecules have been carried out using standard SCF techniques whereas calculations with 216 molecules have been done using the divide and conquer approach. This method has also been used in one simulation case with 64 molecules for test purposes. Within this scope, the molecular dynamics program ROAR have been coupled with a linear scaling semiempirical code (DivCon) implemented in a parallel MPI version. The predicted liquid water structure using either AM1 or PM3 is shown to be very poor due to well-known limitations of these methods describing hydrogen bonds. In contrast, PM3-PIF and PM3-MAIS calculations lead to results in reasonably good agreement with experimental data. The best results for the heat of vaporization are obtained with the PM3-PIF method. The average induced dipole moment of the water molecule in the liquid is underestimated by all semiempirical techniques, which seems to be related to the NDDO approximation and to the use of minimal basis sets. A brief discussion on charge-transfer effects in liquid water is also presented.

## Introduction

The simulation of chemical processes in biological systems, in solution or in other complex media is a major challenge for computational chemistry and justifies the large methodological efforts that have been expended. On one hand, the Car-Parrinello molecular dynamics method[1] (CP-MD) has opened up significant opportunities and has general applicability. Unfortunately, the corresponding computational cost is high and applications are still limited to simple systems and/or short simulation times. On the other hand, MD simulations using combined quantum mechanics and molecular mechanics potentials[2] (QM/MM-MD) are much less expensive and have become very popular. Though a number of applications have been carried out at the *ab initio* and density functional levels, in most cases, particularly for enzymatic reactions, the definition of the quantum subsystem requires taking into account a large number of atoms. For such systems, QM/MM-MD simulations are often carried out at the semiempirical QM level.

In addition to these methods, effort has been devoted to the development of linear scaling algorithms allowing a full quantum mechanical treatment of very large systems.[3] The divide and conquer method[4] (D&C) belongs to this category and has the further advantage of being efficiently implemented on parallel computers. In principle, coupling with MD is also possible but computational limitations again require the use of semiempirical methods.

Indeed, semiempirical techniques such as AM1[5] and PM3[6] have been widely used to investigate chemical and biochemical reactions. Compared to *ab initio* approaches, they are much less time-consuming and therefore qualitative or semiquantitative information may be obtained at a reasonable computational cost. Unfortunately, semiempirical methods suffer from a number of problems in describing intermolecular interactions.[7−15] The water dimer case, in particular, has been discussed in detail. It is worth reviewing a few points. First, the global minimum in AM1 computations is a bifurcated complex and not a linear one.[7] Second, AM1 and PM3 potential energy curves for the water dimer contain spurious oscillations that are related to the core−core repulsion function.[16−19] Due to these oscillations, unphysical energy minima appear on the potential energy surface (PES). For example, there are two PM3 linear complexes,[20] one at $\sim d_{OO} = 2.8$ Å and another one at $\sim d_{OO} = 3.4$ Å, the latter being slightly less stable (see Figure 4 of ref 20 in which the geometry of the isolated water molecule was assumed for the monomers). The HH core−core functions are probably the main cause of these faults. They exhibit (in PM3) a stabilizing contribution of about 2 kcal/mol for an HH distance of 1.2 Å introducing important inconsistencies in the study of intermolecular interactions.[21]

As a result of these observations, one cannot expect to obtain reasonably good results for the structure of hydrated systems within traditional semiempirical Hamiltonians. Of course, this is a major limitation, because hydrogen bonding plays a

---

* To whom correspondence should be addressed.

[†] Université Henri Poincaré.

[‡] Universidad Autónoma del Estado de Morelos.

[§] The Pennsylvania State University.

[||] Harvard University.

10.1021/jp0459099 CCC: $30.25

**TABLE 1: PM3-PIF Parameters[20] in Atomic Units for the $g^{\mathrm{PIF}}(A, B)$ Function**

| A−B | $\alpha_{AB}$ | $\beta_{AB}$ | $\chi_{AB}$ | $\delta_{AB}$ | $\epsilon_{AB}$ |
|---|---|---|---|---|---|
| O−O | 12.45958 | 2.502734 | −266.2244 | 12857.53 | −89486.91 |
| O−H | 49.36554 | 2.276467 | −38.2875 | −145.05 | 651.82 |
| H−H | 0.51979 | 2.489980 | 50.1011 | −676.32 | 2711.77 |

fundamental role in most biochemical processes. Given the speed of semiempirical calculations and their ability to incorporate polarization and charge transfer effects, attempts to improve the current methods are extremely valuable.

Both AM1 and PM3 were derived from the MNDO[22] method that overestimates intermolecular repulsion and as a result is not able to predict hydrogen bonds. In terms of this property alone, the development of AM1 and PM3 resulted in a noticeable improvement.[23−26] The main modification was the introduction of Gaussian correction functions (GCFs) into the core−core repulsion terms. These GCFs were parametrized to force the adoption of the expected energy minima in hydrogen-bonded structures. However, since only a limited experimental data set was available for the parametrization, the GCFs introduced the observed artifacts into the potential energy surface.

Recently, Bernal-Uruchurtu et al.[20,27] have proposed to reparametrize the PM3 GCFs from high-level *ab initio* calculations to properly describe intermolecular interactions in the test case of the water dimer. Two approaches have been proposed. In the first one, original PM3 GCFs are kept to compute intramolecular interactions, whereas intermolecular interactions are described by a parametrizable interaction function (PIF).[20] This PIF has been designed to include short-range interactions through the use of an exponential term as well as long-range corrections to the semiempirical intermolecular energy (e.g., electrostatic, dispersion, and induction) through the use of terms in $R^{-6}$, $R^{-8}$, and $R^{-10}$ as follows:

$$\mathrm{PIF} = \sum_{\mathrm{A,B}}^{\mathrm{inter}} g^{\mathrm{PIF}}(\mathrm{A,B}) = \sum_{\mathrm{A,B}} \alpha_{\mathrm{A,B}} e^{-\beta_{\mathrm{AB}} R_{\mathrm{AB}}} + \frac{\chi_{\mathrm{AB}}}{{R_{\mathrm{AB}}}^{6}} + \frac{\delta_{\mathrm{AB}}}{{R_{\mathrm{AB}}}^{8}} + \frac{\epsilon_{\mathrm{AB}}}{{R_{\mathrm{AB}}}^{10}} \quad (1)$$

where A and B refer to hydrogen or oxygen atoms belonging to different water molecules, $R_{AB}$ is the distance between the two atoms, and $\alpha_{AB}$, $\beta_{AB}$, $\chi_{AB}$, $\delta_{AB}$, and $\epsilon_{AB}$ are the diatomic PIF parameters. Optimized values for these PIF parameters are reported Table 1.

In the second approach, Bernal-Uruchurtu et al.[27] have merged original PM3 GCFs with PIF into a unique core−core function valid to describe intramolecular interactions as well as intermolecular interactions. In this case, the GCFs have the form

$$g^{\mathrm{MAIS}}(\mathrm{A,B}) = \sum_{n=1}^{3} \alpha_{\mathrm{AB}}^{n} e^{-\beta_{\mathrm{AB}}(\gamma_{\mathrm{AB}} - R_{\mathrm{AB}})^2} \quad (2)$$

where $\alpha$, $\beta$, and $\gamma$ are diatomic parameters. Values for H−H, H−O, and O−O interactions have been optimized[27] to fit the *ab initio* potential energy surface for the water dimer (see Table 2). The modified PM3 method, termed PM3-MAIS (MAIS = method adapted for intermolecular studies), can also describe proton transfer in the water-hydronium system. In particular, a single-well PES is predicted, in good agreement with *ab initio* data, and in contrast to the standard PM3 Hamiltonian that predicts a double-well PES.[27] By extrapolating these results, one may expect PM3-PIF and PM3-MAIS to be suitable approaches for describing hydration phenomena in the condensed phase though further tests are clearly necessary.

**TABLE 2: PM3-MAIS Parameters[27] in Atomic Units for the $g^{\mathrm{MAIS}}(\mathrm{A,B})$ Function**

| A−B | $N$ | $\alpha$ | $\beta$ | $\gamma$ |
|---|---|---|---|---|
| O−O | 1 | −6.797745 | 0.326159 | 1.21137 |
| | 2 | 6.912401 | 0.320157 | 1.20945 |
| | 3 | 0.074600 | 1.268150 | 2.57953 |
| O−H | 1 | −0.071576 | 0.192055 | 1.44174 |
| | 2 | 0.015255 | 1.284460 | 2.19381 |
| | 3 | 0.029575 | 0.345024 | 2.58602 |
| H−H | 1 | 0.000362 | 0.385490 | 6.23090 |
| | 2 | 0.009138 | 0.227530 | 2.21168 |
| | 3 | 0.007175 | 3.013020 | 2.22572 |

The purpose of this work is to examine the ability of the PM3-PIF and PM3-MAIS parametrizations to describe the structure and basic properties of liquid water and to compare it with the standard AM1 and PM3 methods. This represents a significant test for the theoretical model since, as pointed out by Head-Gordon and Hura in their recent review devoted to the structure of water,[28] “simulation is a venue for telling us what is not water-like or what physics is missing”. This information is of crucial importance when searching for improved semiempirical methods.

We therefore report the first molecular dynamics simulations of liquid water using a full quantum mechanical description of the system at the semiempirical level. Quantum mechanical forces are computed “on the fly” using either standard SCF calculations or the D&C method, that permits the study of large system sizes (up to 216 water molecules here) in contrast to previously reported *ab initio* MD simulations.

## Methodology

We have performed molecular dynamics simulations of liquid water assuming a cubic box with periodic boundary conditions. Two different system sizes have been considered: 64 water molecules in a cubic box with a box length of 12.516 Å, and 216 molecules in a cubic box of length 18.774 Å. In all cases, the entire system was described using quantum mechanics with no geometric constraints imposed on the water monomers. At each simulation step, a semiempirical SCF calculation for the molecules in the box was carried out using conventional techniques (full Fock matrix diagonalization) or the D&C approach. In both cases, the minimum image convention is used to build the total Fock matrix. Systems of 64 water molecules were studied with conventional SCF calculations. However, to test the influence of the method used to compute the electron density, conventional SCF or D&C, two different simulations have been carried out in the case of PM3. For the 216-molecules system, only D&C calculations have been done. The molecular dynamics program ROAR[29] was coupled with a linear scaling semiempirical code (DivCon)[30] implemented in a parallel MPI version.[31] All simulations were run using the NVT ensemble at 300 K making use of a Nosé-Hoover thermostat[32] that had been implemented in ROAR.[33] For the smallest system, MD simulations were done using four different semiempirical Hamiltonians, namely AM1, PM3, PM3-PIF, and PM3-MAIS, whereas for the largest system, only PM3-MAIS was used. D&C calculations were performed using core regions of one water molecule, an inner buffer of 5.0 Å, and an outer buffer of 1.0 Å. This represents the best compromise between computational speed and accuracy.[34] Using a 0.2 fs time step and after a first equilibration phase, 10 ps data have been collected for deriving average properties.

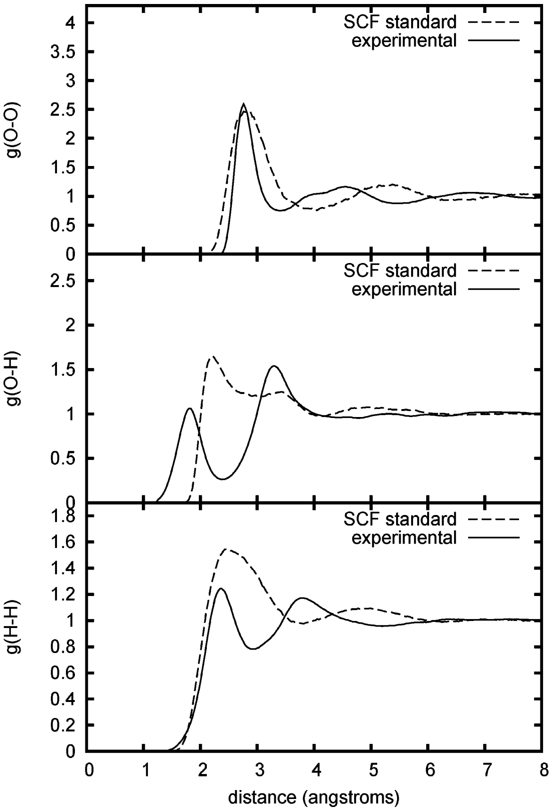

**Figure 1.** Radial distribution functions for liquid water obtained in the MD simulation with the AM1 method and experimental results.[35]

Computations were carried out on an AMD Athlon processor (1.3 Ghz, 512 Mb RAM) running under Linux. Benchmark calculations have shown average CPU time per MD step for the smaller system (64 molecules) was 7.4 s when using conventional SCF calculation, whereas D&C calculations led to higher CPU time (28.4 s per MD step) for the same system. The greater cost of D&C calculations compared to conventional SCF calculations is mainly due to the time required to perform each SCF iteration (1.1 s per SCF iteration vs 4.6 s per SCF iteration respectively for conventional and D&C calculations with an average of 6 SCF iteration per MD step in each case). In contrast, for the larger system (216 molecules), conventional SCF becomes too expensive. One conventional SCF iteration takes about 52.7 s (with an average of 9 SCF iterations per MD step), whereas one D&C SCF iteration takes only 15.9 s (with an average of 6 SCF iterations per MD step). The average CPU time per MD step for the large system was therefore 481 and 101 s respectively for the conventional and the D&C calculations.

## Results

We present the results obtained for radial distribution functions, average water molecule geometry, heat of vaporization, net atomic charges, charge transfer, and dipole moments in the liquid phase. Unless otherwise stated, property averages correspond to configurational plus time averages using the 64-molecules system.

**Radial Distribution Functions.** An important test for the different simulation protocols is the quality of the prediction of the structural parameters for liquid water obtained from X-ray or neutron diffraction experiments.[35] In Figures 1 and 2, we plot the radial distribution functions (RDFs) obtained using AM1 and PM3, respectively. In the case of PM3 (Figure 2), we also compare the RDFs obtained with standard-SCF and D&C calculations. As seen, both approaches lead to comparable results, thus confirming that the D&C approach can be used with confidence for these systems.

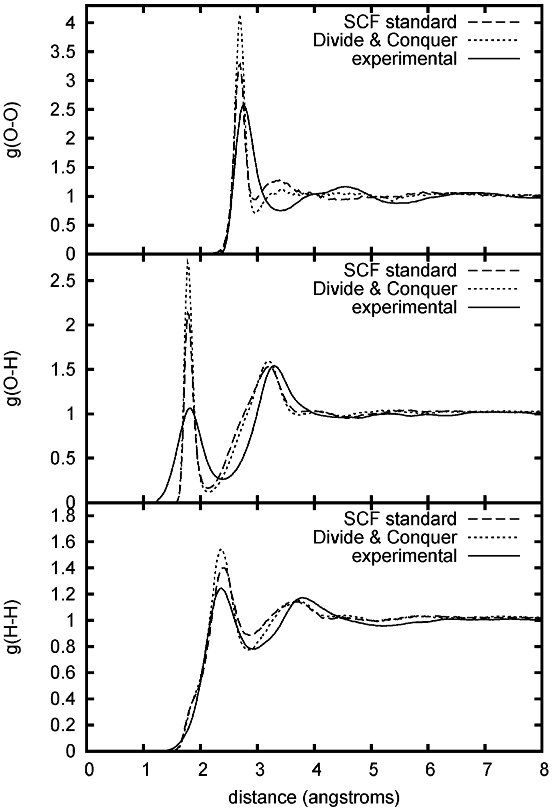

**Figure 2.** Radial distribution functions for liquid water obtained in the MD simulation with the PM3 method and experimental results.[35] Results using the standard SCF method and the D&C approach are included for comparison.

Comparing the AM1 and PM3 RDFs with the experimental curves clearly indicate an anomalous coordination pattern even in the first solvation shell. For example, considering the O−O functions, AM1 predicts a very broad first peak. Integration of the coordination number up to 4.0 Å (the position of the minimum separating the first and second peaks) one obtains $N = 9.8$. If one integrates up to 3.4 Å (the position of the minimum in the experimental curve), then $N = 6.7$. Both values are too large when compared with experiment ($N = 4.6$). In the case of PM3, we found a very sharp peak located at 2.7 Å displaying an overestimated peak height. One also notes the presence of an unphysical smaller peak at about 3.4 Å. The estimated coordination number for the first peak is $N = 2.7$ and for the second one, $N = 5.8$. If the integration is made up to 3.4 Å, one gets $N = 5.0$.

Similarly, the AM1 and PM3 results for the O−H RDF, the fingerprint of the H-bond network in the liquid, were not well reproduced. The function predicted by AM1 exhibits a first peak that is too broad, is shifted toward higher values of interatomic distances, and is not clearly separated from the second peak. In the case of PM3, the O−H RDF displays two peaks that are clearly separated but their positions are slightly shifted to smaller distances and they display an anomalous height ratio. Results for the H−H RDF are also inaccurate, especially in the case of AM1.

Clearly the structure of liquid water is poorly described with standard AM1 and PM3. Overall, the bad results obtained for the RDFs with these methods are the consequence of their inadequacy to reproduce the potential energy surface for the

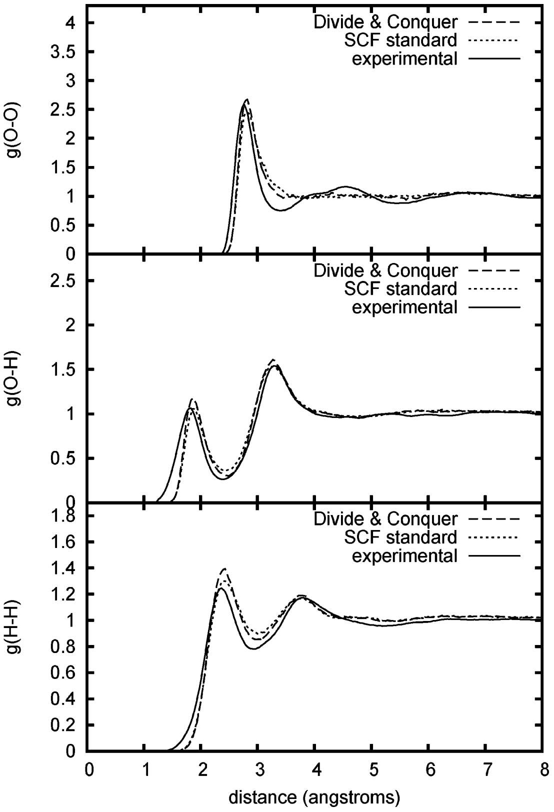

**Figure 3.** Radial distribution functions for liquid water obtained in the MD simulation with the PM3-PIF method and experimental results.[35] Results using the standard SCF method and the D&C approach are included for comparison.

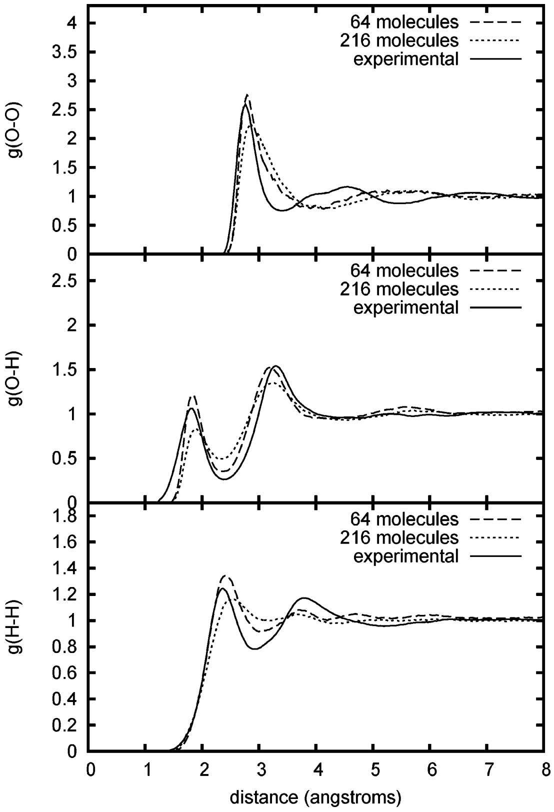

**Figure 4.** Radial distribution functions for liquid water obtained in the MD simulation with the PM3-MAIS method and experimental results.[35] Simulations were carried out for a box containing 64 or 216 water molecules.

water dimer, as discussed above. Artifacts such as oscillations in the water−water PES or HH stabilizing interactions due to GCFs are certainly the major problems.

The RDFs obtained with the PM3-PIF parametrization are shown in Figure 3. Compared with both standard semiempirical Hamiltonians, they exhibit a much better agreement with experiment, particularly for the OH and HH RDFs. The position and height of the two main peaks were predicted with reasonably good precision. Some problems remain for the OO RDF, however. First, although the first maximum is located at the correct position, 2.8 Å, the width of this peak is larger than expected and there is not a clear second shell. Though one must recall that many classical force-fields fail to describe the OO second shell, in the case of semiempirical methods, the reason is probably related to underestimated nonadditive interactions, as shown in the study of water clusters.[20] We shall come back to this point later. The coordination number obtained by integrating the OO RDF out to 3.4 Å is a little too large: $N = 5.1$. If the integration is made up to the position of the first minimum (3.9 Å), one gets $N = 7.9$. As expected, this last number is too large because of the width of the first OO peak, but it is slightly less than the AM1 first shell coordination number computed above. Overall, for the OH and HH RDFs, PM3-PIF represents a very good improvement over AM1 and PM3, and, concerning the OO RDF, it represents a slight improvement over AM1 without an unphysical behavior as with PM3.

The RDFs obtained with the PM3-MAIS parametrization are shown in Figure 4. They do not show large differences with PM3-PIF results. For the OO RDF, although the first peak is located at 2.8 Å like in the case of PM3-PIF, its width is slightly larger. As confirmed by the calculation of the coordination number: it is evaluated at 5.4 at the position of the first experimental minimum (3.4 Å). and at 8.4 at the position of the first minimum ($r = 3.9$ Å), A second water shell in the PM3-MAIS OO RDF can also be identified, albeit being located further away from the experimental one (its maximum is lying around 5.5 Å instead of 4.6 Å for the experimental curve). Regarding OH and HH RDFs, PM3-MAIS curves present very similar shapes as PM3-PIF ones.

In the PM3-MAIS case, we have also checked that the influence of increasing the number of water molecules in the box, from 64 to 216, does not modify the observed results. These results are also shown in Figure 4. The D&C method is used in this case. The RDFs predicted for the large and small systems have slight differences that might be due in part to insufficient simulation time. The coordination number obtained by integrating the OO RDF up to 3.4 Å is 5.4, i.e. the same value that was obtained for the 64-molecules system. Hence, for the purposes of our work, we consider that the structural properties are reasonably well described with the small system.

Owing to some uncertainties in the analysis of diffraction experiments, it is valuable to compare simulation results with the total intermolecular pair correlation function.[36] In Figure 5, we compare the $G(r)$ obtained using the four semiempirical approaches with the experimental function. As shown, only PM3-PIF and PM3-MAIS exhibit reasonable agreement, PM3-PIF being slightly closer to experimental results.

**Structure of the Water Molecule.** Since no geometrical constraints were imposed in our simulation, the study of structural changes induced in the single water molecule by the hydrogen-bond network was feasible.

The predicted changes of the geometry of the water molecule in the liquid are summarized in Table 3. Properties in the liquid

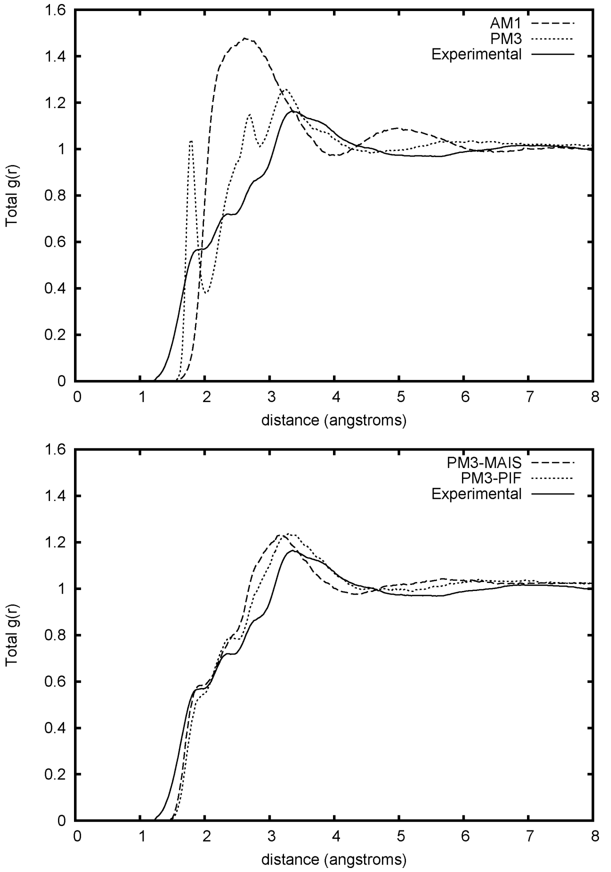

**Figure 5.** Comparison of the total pair correlation function for water at ambient conditions. We use the expression proposed in ref 36 for heavy water $G(r) = 0.489g_{DD}(r) + 0.421g_{OD}(r) + 0.09g_{OO}(r)$ using RDFs for light water (as noted in ref 36, light and heavy water are often assumed to have the same structure).

**TABLE 3: Predicted Geometrical Changes (in Å and Degrees) for the Water Molecule in Going from Gas to Liquid Phase**[a]

| method | | AM1 | PM3 | PM3-PIF | PM3-MAIS | exp[b] |
|---|---|---|---|---|---|---|
| $R$ (O−H) | gas | 0.961 | 0.951 | 0.951[c] | 0.955 | 0.9572 |
| | liquid | 0.965 | 0.962 | 0.966 | 0.971 | 0.970 |
| ∠ (HOH) | gas | 103.53 | 107.69 | 107.69[c] | 106.84 | 104.52 |
| | liquid | 103.71 | 108.08 | 106.80 | 106.02 | 106 |

[a] Gas-phase values correspond to optimized geometries. Liquid values correspond to configurational and time averages obtained for the 64-molecules system during 10 ps. [b] Reference 37. [c] Values identical to PM3 in agreement with PIF definition.

correspond to simulation averages. Gas-phase results correspond to the optimized geometry for the isolated molecule. We can see that the computed O−H bond length increases in the liquid (+0.004 Å, +0.012 Å, +0.015 Å, and +0.016 Å for AM1, PM3, PM3-PIF, and PM3-MAIS, respectively), PM3, PM3-PIF, and PM3-MAIS changes being closer to the experimentally observed value[37] than AM1. It is interesting to note that PM3-PIF and PM3-MAIS values are not far from that of CP-MD simulations[38] that predicted an O−H increase of 0.019 Å, and from QM/MM calculations[39] that predicted an O−H increase up to 0.008 Å. Changes in the bond angle were rather small in all cases though one may note that PM3-PIF and PM3-MAIS respectively predict a 0.9° and 0.8° decrease, in contrast to experiment, but in agreement with QM/MM calculations from Alhambra et al.[51] which found a relative decrease in the HOH bond angle, ranging from 1.3 ± 0.5° for hybrid HF/6-31G*/TIP3P Monte Carlo simulations to 0.5 ± 3° for DFT/TIP3P molecular dynamics simulations.

**Heat of Vaporization.** The heat of vaporization has been computed for all molecular dynamics using the following formula:[40]

$$\Delta H_{\text{vap}} = -E_{\text{int}} + RT \tag{3}$$

where $E_{\text{int}}$ is the mean intermolecular interaction energy calculated from

$$E_{\text{int}} = \frac{\langle E_n \rangle - n{*}E_1}{n} \tag{4}$$

where $\langle E_n \rangle$ is the mean potential energy of $n$ interacting water molecules averaged during a molecular dynamics simulation and $E_1$ is the water molecule gas phase energy. The heat of vaporization experimental value is 10.50 kcal/mol.[41] The calculated values are respectively 7.67 ± 0.56 kcal/mol for AM1, 4.68 ± 0.84 kcal/mol for PM3, 9.02 ± 0.69 kcal/mol for PM3-PIF, and 5.17 ± 0.74 kcal/mol for PM3-MAIS in the case of 64 water molecules and 4.66 ± 0.49 kcal/mol for PM3-MAIS in the case of 216 water molecules. All computed semiempirical heat of vaporizations range below the experimental value. The rather large difference between the PM3-PIF and PM3-MAIS values can mainly be explained by the use of different $g_{AB}$ core−core functions. Indeed, when one simply optimizes the water dimer, the root-mean-square deviation between the PM3-PIF and the PM3-MAIS structures is only 0.009 Å, whereas the interaction energy can be respectively evaluated at −5.11 kcal/mol for PM3-PIF and −4.71 kcal/mol for PM3-MAIS. Giving the nearly identical geometries between the two dimers, the electronic structures of both systems can be considered equal. Therefore, the 0.4 kcal/mol difference in stabilization energy between PM3-PIF and PM3-MAIS can only be explained by the use of different core−core functions associated with their respective parameters. This small and additive difference in the case of the water dimer structure can be extended to liquid water: for each water−water interaction, the PM3-PIF method predicts a slightly higher stabilization energy compared to PM3-MAIS, thus giving a rather large difference in heat of vaporization.

**Charge Transfer and Polarization Effects.** As recently pointed out by Finney,[42] the possibility of a significant charge transfer in liquid water has been disregarded by most works devoted to liquid water simulation. This analysis becomes possible in the present treatment since we calculate the wave function of the entire water sample in the simulation box. To discuss the charge distribution and charge transfer, we have computed net atomic charges using the CM1 method.[43] Note that a recent study by van der Vaart and Merz shows that AM1 and PM3 have an excellent agreement with high level *ab initio* calculations on charge-transfer effects in hydrogen bonded clusters.[44] In particular, these authors have shown (i) charge transfer does occur in hydrogen bonded clusters, and while its amount is small (e.g., between 0.01 and 0.05 electron per hydrogen bond), it is not an artifact of the computation and it does also appear in the absence of basis set superposition error; (ii) AM1 and PM3, using CM1 or CM2 charge descriptors, show an excellent agreement with high level MP2 *ab initio* computations, therefore showing the utility of revealing charge-transfer effects through the use of simpler semiempirical charge models and methods.

The net charge carried by a water molecule in the liquid is analyzed in Figure 6, which shows the corresponding distribution during the simulation. All of the methods predict a distribution showing a maximum probability at zero charge, as expected.

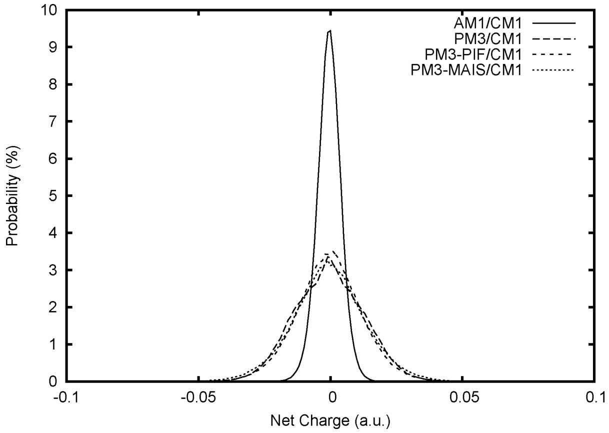

**Figure 6.** Predicted distribution of net charge on water molecules (CM1 values) for the four studied semiempirical methods.

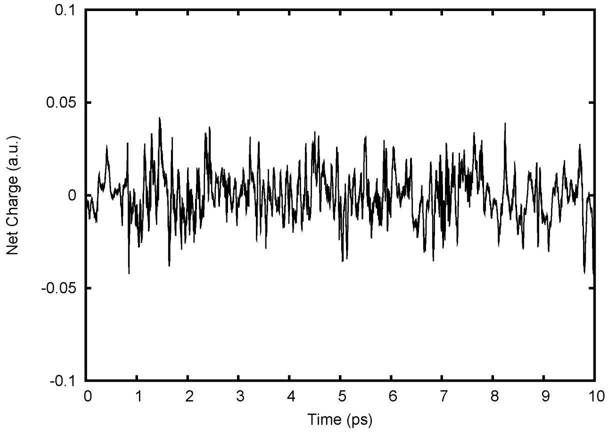

**Figure 7.** Time fluctuations of the net atomic charge of a particular water molecule (MD PM3-MAIS results with the 64-molecules system).

The AM1 distribution is very sharp, with charge-transfer values varying between −0.01 and 0.01 (all charge values in this paper are given in atomic units). PM3, PM3-PIF, and PM3-MAIS curves are broader and predict charge transfer in the range −0.05 and +0.05. The results of van der Vaart and Merz for the water dimer indicate that for a given geometry, AM1 and PM3 charge transfers are comparable. Larger differences obtained in the liquid phase, can be ascribed to structural differences in the predicted hydrogen-bonds. Time fluctuations of the net charge in a water molecule are illustrated in Figure 7. As shown, these fluctuations are rather fast since at the picosecond time scale the molecular charge may reach its maximum (positive) and minimum (negative) values several times.

The AM1 net atomic charge on the oxygen and hydrogen atoms changes from −0.706 and +0.353 in gas phase to −0.756 and +0.378 in the liquid (average values), respectively. The corresponding PM3 values change from −0.712 and +0.356 in gas phase to −0.786 and +0.393 in the liquid, all values identical in the case of PM3-PIF. Similarly, PM3-MAIS calculations in the liquid were close to the standard PM3 calculations: −0.787 and +0.394. Further analysis shows that charge polarization is basically due to electronic effects or, in other words, the geometry changes in the liquid represent a very small contribution to the overall polarization of the water molecule. This point can be illustrated by computing the CM1 atomic charges for an isolated molecule with the average geometry of the liquid. In the case of PM3, for instance, the corresponding charges were −0.717 and +0.358, which represents a small polarization contribution compared to the total effect (roughly 10%).

**TABLE 4: Modification of the Water Molecule Dipole Moment in Going from Gas to Liquid Phase**[a]

| | AM1 | PM3 | PM3-PIF | PM3-MAIS | exp |
|---|---|---|---|---|---|
| | | | Gas Phase | | |
| ⟨$\mu$⟩ | 2.02 | 1.92 | 1.92[d] | 1.95 | 1.854[b] |
| CM1 | 2.017 | 1.923 | 1.923[d] | 1.947 | |
| CM2 | 2.032 | 1.899 | 1.899[d] | 1.922 | |
| Mulliken | 1.094 | 0.966 | 0.966[d] | 0.974 | |
| | | | Liquid Water | | |
| ⟨$\mu$⟩ | | | | | 2.6[c]<br>(+0.75)[c] |
| CM1 | 2.169<br>(0.152) | 2.136<br>(0.213) | 2.174<br>(0.251) | 2.209<br>(0.262) | |
| CM2 | 2.163<br>(0.131) | 2.056<br>(0.157) | 2.100<br>(0.201) | 2.130<br>(0.208) | |
| Mulliken | 1.343<br>(0.249) | 1.258<br>(0.292) | 1.321<br>(0.355) | 1.276<br>(0.302) | |

[a] Gas-phase values are calculated using the optimized geometry and either the dipole moment operator or the net atomic charges. Liquid phase values are configurational and time averages calculated with net atomic charges for the 64-molecules system. Induced moments are in parenthesis. All values in Debyes. [b] Reference 48. [c] Reference 43. [d] Values identical to PM3 in agreement with PIF definition.

Polarization may also be analyzed in terms of the induced dipole moment. The value of the dipole moment in liquid water is a fundamental question that has attracted much interest in theoretical simulation studies. In the present work, computation of the dipole moment in the liquid is not straightforward due to the delocalization of the wave function over the entire system. Moreover, since charge transfer is allowed and molecules are not necessarily neutral, obtaining the instantaneous dipole moment is not straightforward. The same problem arises in CP-MD simulations, and several approaches have been employed. The most elaborated one has been proposed by Silvestrelli and Parrinello,[38] who use the maximally localized-Wannier function scheme in order to analyze the electron distribution. Here, we have adopted a simpler method consisting of the evaluation of the dipole moment of each molecule using its CM1 charges, that both correctly reproduce the gas phase dipole moment, as shown below, and has been successfully used to describe solvation properties of molecules in solution.[45,46] For comparison, dipole moments obtained using Mulliken and CM2 charges[47] have also been computed. Since only a fraction of water molecules are not neutral and the absolute charge transfer is not very high, this approximation is acceptable. For consistency, the dipole moment is always evaluated from the center of mass of the molecule.

Table 4 summarizes the results. In the gas phase, the dipole moment is obtained using the corresponding operator, but for the sake of comparison with calculations in the liquid, we also include in the table the results obtained from net atomic charges. The average distribution for the dipole moment obtained in the MD simulations using the four different semiempirical methods is plotted in Figure 8.

Let us first consider the gas phase values. When the dipole moment is computed rigorously, calculations lead to a slightly overestimated quantity, especially with AM1. When CM1 and CM2 net atomic charges are used instead, the results do not change much, particularly with the CM1 method. In contrast, Mulliken charges lead to substantially underestimated gas-phase dipole moments.

If one considers the induced dipole moment in the liquid, one finds that the results depend substantially on the semiem-

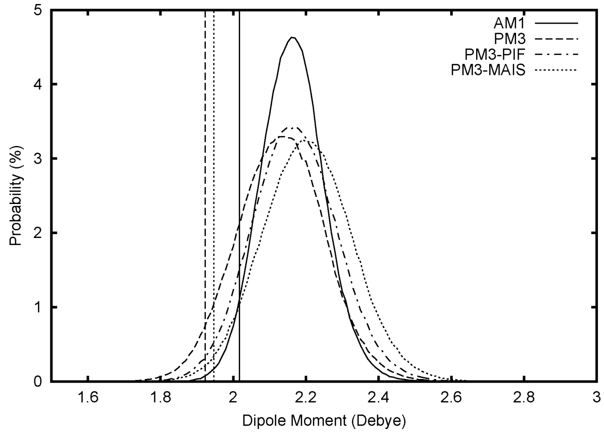

**Figure 8.** Predicted distribution for the dipole moment of the water molecule in the liquid phase. Gas-phase dipole moments are indicated with vertical lines. Values are calculated using CM1 net atomic charges.

pirical method and on the approach used to compute the atomic charges. PM3-PIF and PM3-MAIS predict the largest polarization effect. Nevertheless, all computed induced moments are too small when compared with experimental data. For instance, using CM1 charges, PM3-PIF and PM3-MAIS respectively lead to an induced dipole moment of 0.25 or 0.26 D, which is a little larger than that obtained with either AM1 (0.15D) or PM3 (0.21D) but much lower than the experimental estimate of 0.75D based on data for the Ih phase of ice.[48] It is also lower than theoretical predictions from CP-MD simulations (1.08D)[38] or DFT/MM MD simulations[49] (0.82 D) (see ref 44 and references therein for other theoretical estimations). However, our calculation is consistent with previous computations made at the AM1 level using either a continuum model (0.17D) or QM/MM simulations (0.30D),[50] or a molecular-orbital derived polarization potential (0.43D).[51]

Another rough estimation of the dipole moment of the water molecule in the liquid may be done using its average geometry and the average net atomic charges on oxygen and hydrogen atoms. By doing so, one gets $\mu = 2.169$, 2.136, 2.174, and 2.209 D using CM1 charges obtained with AM1, PM3, PM3-PIF, and PM3-MAIS, respectively. These values are practically the same as the averages obtained for the instantaneous dipole moment given in Table 4, confirming the limited influence of charge-transfer effects.

Clearly, the semiempirical methods tend to underestimate the polarization effect in the liquid. Similarly, many-body effects in water clusters have been shown to be too low at the PM3 level.[20] Water molecule polarizability is indeed substantially underestimated at the AM1 and PM3 levels (a perturbational calculations leads to 3.4 au with both methods, while the experimental value is 9.91.[52]). This may be related to the NDDO approximation and/or to the use of a minimal basis set. Several approaches have been proposed[53] to correct semiempirical molecular polarizabilities a posteriori but one cannot expect to improve the description of polarization effects in extended systems by a simple re-parametrization of core−core terms (like in PM3-MAIS) or even of electronic terms. Reexamination of the fundamental semiempirical approximations seems to be necessary.

## Conclusions

In this work, we have carried out the first molecular dynamics simulations of liquid water fully described by a semiempirical quantum mechanical Hamiltonian. For large simulation boxes, we used the D&C approach instead of the standard SCF procedure. A detailed description of the computational framework will be reported elsewhere. Here, we have focused on an evaluation study to test the ability of several semiempirical methods to describe the properties of this system. We have shown that PM3-PIF and PM3-MAIS, in contrast with the standard semiempirical models AM1 and PM3, appear to be a good choice, despite a few shortcomings. In particular, PM3-PIF results in an improvement of the radial distribution functions over AM1 and PM3 together with a good representation of intermolecular energies. In fact, neither AM1 nor PM3 correctly predict the structure of the first coordination shell. These methods suffer from important defects in the water dimer potential energy surface that are magnified in the simulation of liquid structure.

Polarization effects were shown to be underestimated by all methods and this seems to be due to an intrinsic limitation of the NDDO approximation and to the use of a minimal basis set. The use of a full quantum mechanical potential in MD simulations allows for the analysis of charge transfer effects, not easily accessible with classical force fields. Up to now, there are few reliable studies on the extension of these effects to condensed phases. Our results for liquid water show a small but nonnegligible effect.

Although our approach has limitations, we believe that the results found in this work encourage the use of the PM3-PIF and PM3-MAIS parameters to deal with water−water interactions in semiempirical studies of chemical and biochemical problems in aqueous media.

**Acknowledgment.** Part of this work was supported with an ECOS Nord − ANUIES Grant MP98-01 and CONACYT 38326-E Grant. K.M.M. thanks the NSF via Grant MCB-0211639 for financial support.

## References and Notes

(1) (a) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471. (b) Lippert, J.; Hutter, G.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124.

(2) For a review, see for instance: (a) Gao J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: Weinheim, Germany, 1996; Vol. 7, 119−185. (b) Amara P.; Field M. In *Encyclopedia of Computational Chemistry*; P.v.R. Schleyer, P. v. R., Ed.; Wiley & Sons: New York, 1998; Vol. 1, pp 431−437. (c) Ruiz-López, M. F.; Rivail, J. L. In *Encyclopedia of Computational Chemistry*; P.v.R. Schleyer, P. v. R., Ed.; Wiley & Sons: New York, 1998; Vol. 1, pp 437−448. (d) Monard, G.; Merz, K. M. Jr. *Acc. Chem. Res.* **1999**, *32*, 904.

(3) (a) Goedecker, S.; Scuseria, G. E., *Comput. Sci. Eng.* **2003**, *5*, 14. (b) Goedecker, S.; *Rev. Mod. Phys.* **1999**, *71*, 1085.

(4) (a) Yang, W.; Lee, T.-S., *J. Chem. Phys.* **1995**, *103*, 5674. (b) Dixon, S. L.; Merz, K. M., Jr., *J. Chem. Phys.* **1996**, *104*, 6643. (c) Dixon, S. L.; Merz, K. M. Jr., *J. Chem. Phys.* **1997**, *107*, 879.

(5) Dewar, M. J. S.; Zoebish, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(6) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.

(7) Ventura, O. N.; Coitiño E. L.; Lledós A.; Bertrán J. *J. Mol. Struct., Theochem* **1989**, *187*, 55−68.

(8) Coitiño, E. L.; Irving, K.; Rama, J.; Iglesias, A.; Paulino, M.; Ventura, O. N. *J. Mol. Struct., Theochem* **1990**, *210*, 405.

(9) Khalil, M.; Woods, R. J.; Weaver, D. F.; Smith, V. H., Jr. *J. Comput. Chem.* **1991**, *12*, 584−593.

(10) Coitiño, E. L.; Ventura, O. N.; Sosa, R. M. *J. Mol. Struct., Theochem* **1992**, *254*, 315.

(11) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 3892.

(12) Rusinska-Roszak, D.; Lozynski, M.; Mack H. G.; Korn, M. *J. Mol. Struct., Theochem* **1995**, *342*, 33−41.

(13) Cativiela, C.; Dillet, V.; García, J. I.; Mayoral, J. A.; Ruiz-López, M. F.; Salvatella, L. *J. Mol. Struct., Theochem.* **1995**, *331*, 37.

(14) Dos Santos, H. F.; De Almeida, W. B. *J. Mol. Struct., Theochem.* **1995**, *335*, 129.

(15) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219.

(16) Bu$\beta$, V.; Messinger, J.; Heuser, N. *Q. C. P. E. Bull.* **1991**, *11*, 5.

(17) Csonka, G. I. *J. Comput. Chem.* **1993**, *14*, 895.

(18) Csonka, G. I.; Éliás, K.; Csizmadia, I. G. *J. Comput. Chem.* **1997**, *18*, 330.

(19) Csonka, G. I.; Ángyán, J. *J. Mol. Struct., Theochem* **1997**, *393*, 31.

(20) Bernal-Uruchurtu, M. I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F. *J. Comput. Chem.* **2000**, *21*, 572.

(21) Harb, W.; Bernal-Uruchurtu, M. I.; Ruiz- López, M. F. *Theor. Chem. Acc.*, in press.

(22) (a) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899. (b) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4907.

(23) Dannenberg, J. J. *J. Phys. Chem.* **1988**, *92*, 6869.

(24) Jurema, M. W.; Shields, G. C. *J. Comput. Chem.* **1993**, *14*, 89.

(25) Kallies, B.; Mitzner, R. *J. Mol. Mod.* **1995**, *1*, 68.

(26) Dannenberg, J. J. *J. Mol. Struct, Theochem* **1997**, *401*, 279.

(27) Bernal-Uruchurtu, M. I.; Ruiz-López, M. F. *Chem. Phys. Lett.* **2000**, *330*, 118.

(28) Head-Gordon T.; Hura, G. *Chem. Rev.* **2002**, *102*, 2651.

(29) Cheng, A.; Stanton, R. S.; Vincent, J. J.; van der Vaart, A.; Damodaran, K. V.; Dixon, S. L.; Hartsough, D. S.; Mori, M.; Best, S. A.; Monard, G.; Garcia, M.; Van Zant, L. C.; Merz, K. M., Jr. *ROAR* 2.0; The Pennsylvania State University: University Park, PA, 1999.

(30) Dixon, S. L.; van der Vaart, A.; Gogonea, V.; Vincent, J. J.; Brothers, E. N.; Suárez, D.; Westerhoff, L. M.; Merz, K. M., Jr. *DivCon99*; The Pennyslvania State University: University Park, PA, 1999.

(31) Vincent, J. J.; Merz, K. M., Jr. *Theor. Chem. Acc.* **1998**, *99*, 220.

(32) Nosé, S. *Mol. Phys.* **1984**, *52*, 255; Hoover, W. G. *Phys. Rev.* **1985**, *31* A, 1695. Nosé, S. *Mol. Phys.* **1986**, *57*, 187.

(33) Cheng, A.; Merz, K. M. Jr. *J. Phys. Chem.* **1996**, *100*, 1927.

(34) van der Vaart, A.; Suarez, D.; Merz, K. M., Jr. *J. Chem. Phys.* **2000**, *113*, 10512.

(35) (a) Soper A. K.; Bruni, F.; Ricci M. A. *J. Chem. Phys.* **1997**, *106*, 247. (b) Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9140.

(36) Guillot, B. *J. Mol. Liquids* **2002**, *111*, 219.

(37) Experimental data for the water molecule structure are from Benedict, W. S.; Gailar, N.; Plyler, E. K. *J. Chem. Phys.* **1973**, *24*, 2254 (gas phase) and from Ichikawa, K.; Kameda, Y.; Yamaguchi, T.; Wakita, H.; Misawa, M. *Mol. Phys.* **1991**, *73*, 79 (liquid phase).

(38) Silvestrelli, P. L.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 3572.

(39) Alhambra, C.; Byun, K.; Gao, J. In *Combined Quantum Mechanical and Molecular Mechanical Methods*; Gao, J., Thompson, M. A., Eds.; ACS Symposium Series 712; American Chemical Society: Washington, DC, 1998; p 35.

(40) Fox, T.; Kollman, P. A. *J. Phys. Chem. B* **1998**, *102*, 8070.

(41) (a) Giauque, W. F.; Stout, J. W. *J. Am. Chem. Soc.* **1936**, *58*, 1144. (b) Giguere, P. A.; Carmichael, J. L. *J. Chem. Eng. Data* **1962**, *7*, 526.

(42) Finney J. L. *J. Mol. Liq.* **2001**, *90*, 3003.

(43) Storer J. W.; Giesen D. J.; Cramer C. J.; Truhlar D. G. *J. Comput. Aided Mol. Des.* **1995**, *9*, 87.

(44) van der Vaart, A.; Merz, K. M. Jr. *J. Chem. Phys.* **2002**, *116*, 7380.

(45) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385.

(46) 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.

(47) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820.

(48) Coulson, C. A.; Eisenberg, D. *Proc. R. Soc. London, Ser. A* **1966**, *290*, 4546.

(49) Chalmet S.; Ruiz-López, M. F. *J. Chem. Phys.* **2001**, *115*, 5220.

(50) Gao, J.; Luque, F. J.; Orozco, M. *J. Chem. Phys.* **1993**, *98*, 2975.

(51) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346.

(52) Murphy, W. F. *J. Chem. Phys.* **1977**, *67*, 5877.

(53) Dewar, M. J. S.; Stewart, J. J. P. *Chem. Phys.* **1984**, *111*, 416. Schürer, G.; Gedeck, P.; Gottschalk, M.; Clark, T. *Int. J. Quantum Chem.* **1999**, *75*, 17. Kagawa, H.; Ichimura, A.; Kamka, N. A.; Mori, K. *J. Mol. Struct, Theochem* **2001**, *546*, 127.

(54) Clough, S. A.; Beers, Y.; Klein, G. P.; Rothman, L. S. *J. Chem. Phys.* **1973**, *59*, 2254.