Hybrid ab Initio QM/MM Simulation of *N*-Methylacetamide in Aqueous Solution

Jiali Gao* and Marek Freindorf†

Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260

Received: December 19, 1996; In Final Form: February 21, 1997[⊗]

Combined quantum mechanical and molecular mechanical (QM/MM) simulations of *N*-methylacetamide in aqueous solution have been carried out to investigate the charge polarization of the solute and to explore the feasibility of hybrid QM/MM calculations using ab initio methods. In the present study, the ab initio Hartree−Fock theory along with the 3-21G basis set was used in the quantum mechanical calculations. Statistical mechanical Monte Carlo approach was then applied in molecular mechanical simulations, employing the empirical TIP3P model for water. Comparisons with results obtained from the hybrid semiempirical Austin model 1 (AM1)/TIP3P and Jorgensen's OPLS (optimized potential for liquid simulations) potential were made, and a good accord among the three methods has been obtained. The solute charge polarization was analyzed through population analyses and determination of polarization energies. We found that the polarization effects contribute 10−15% to the total solute−solvent interaction energy for *N*-methylacetamide in water.

Introduction

An accurate description of intermolecular interactions is an important task in computational chemistry. Traditionally, empirical potential functions have been used for studies of condensed phase and biopolymer systems such that the energy and forces on the atoms can be rapidly evaluated. Although empirical potentials can yield reasonable results for the equilibrium thermodynamic properties, there are several well-known shortcomings. Empirical methods are inappropriate for the study of chemical processes involving bond forming and breaking,[1] and the parametrization process is often time-consuming and difficult. Further, there is a lack of general procedures for the treatment of many-body polarization effects.[2] Recently, combined quantum mechanical (QM) and molecular mechanical (MM) methods have become popular because they have the potential for studying chemical processes in large molecular systems and in condensed phases.[3−8] In these calculations, the solute molecule is treated quantum mechanically, whereas the solvent molecules are approximated by empirical or MM force fields. Since the electronic structure of the solute molecule is determined for each microscopic state sampled during the computer simulation, it is no longer necessary to parametrize empirical potential functions for solute−solvent interactions. In addition, the method can be generally utilized to study chemical reactions in solution and in enzymes.

In this article, we examine the solvent effect on *N*-methylacetamide (NMA) in aqueous solution using a hybrid ab initio QM/MM potential in Monte Carlo simulations. *N*-Methylacetamide is the simplest molecular model of peptide linkage in proteins and has been the subject of extensive experimental[9−24] and theoretical investigations.[25−43] In this report, our emphasis is on the examination of the possibility of statistical Monte Carlo and molecular dynamics simulations of solution systems with the use of an ab initio hybrid QM/MM method. To this end, most studies employing hybrid QM/MM potentials have been limited to semiempirical approaches in the quantum mechanical part.[3−8] The main advantage of using a semiempirical method is its computational efficiency. However, semiempirical QM methods, including the empirical valence bond approach, are restricted by their parametrized nature. Thus, it is desirable to use ab initio molecular orbital or density functional theory to construct hybrid QM/MM Hamiltonians, since the accuracy of ab initio calculations can be systematically improved, either by using large basis sets or by including correlation effects.[44] Here, we are interested in a combined ab initio Hartree−Fock and MM potential using the split valence 3-21G basis set.[45] The choice for this basis set has taken into account the need for both computational accuracy and efficiency. Numerous studies have demonstrated that calculations using the 3-21G basis sets can yield reasonable results for molecular geometries, energies, conformational analyses, and reaction paths.[44] A larger basis set and the use of correlated methods certainly would be attractive; however, it will significantly increase the demands for computational resources, including CPU time, disk space, and memory requirement, which would prevent its application to large molecular systems. Although hybrid QM/MM methods have been used previously in energy minimization and single-point energy calculations, fluid simulations on the basis of a hybrid ab initio QM/MM potential have not yet been reported. In a recent study, we have demonstrated that the hybrid ab initio HF/3-21G (AI-3) and MM method, which is denoted as the AI-3/MM potential, can yield excellent energetic and structural results for bimolecular complexes in comparison with full ab initio HF/6-31G* calculations.[46] The work is now extended to fluid simulations.

In what follows, we first present a brief outline of the hybrid QM/MM method. Then, the approach is illustrated by fluid simulations of *N*-methylacetamide in water with comparison to results from Monte Carlo calculations using Jorgensen's OPLS and the hybrid AM1/MM potential.

Computational Details

In the present study, the solution system is divided into (1) a quantum mechanical region, consisting of the solute molecule *N*-methylacetamide, which is described by the HF/3-21G level of theory,[45] and (2) a molecular mechanical region, containing the solvent molecules, which are approximated by the three-site TIP3P model for water.[46,47] The total effective Hamiltonian of the system is given by

$$\hat{H}_{\mathrm{eff}} = \hat{H}^{\circ} + \hat{H}_{\mathrm{qm/mm}} + \hat{H}_{\mathrm{mm}} \qquad (1)$$

† Permanent address: Department of Theoretical Chemistry, Jagellonian University, Karasia 3, 30-060 Krakow, Poland.

⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1997.

S1089-5639(97)00041-8 CCC: $14.00

TABLE 1: Lennard-Jones Parameters for the OPLS, AM1/MM, and HF 3-21G/MM Potentials

	OPLS[30]			AM1/MM[6a]		HF 3-21G/MM[46]	
atom	*q* (*e*)	σ (Å)	ϵ (kcal/mol)	σ (Å)	ϵ (kcal/mol)	σ (Å)	ϵ (kcal/mol)
O	−0.53	2.96	0.210	2.95	0.20	3.60	0.15
N	−0.55	3.25	0.170	2.80	0.15	3.90	0.20
C	0.58	3.75	0.105	3.50	0.08	3.80	0.08
C_C	0.0	3.91	0.160	3.50	0.08	3.80	0.08
C_N	0.20	3.80	0.170	3.50	0.08	3.80	0.08
H_C				2.00	0.01	2.60	0.008
H_N	0.30	0.00	0.000	0.80	0.10	1.30	0.10
				TIP3P[47]			
O	−0.8340	3.1506	0.1521	3.1506	0.1521	3.1506	0.1521
H	0.4170	0.0	0.0	0.0	0.0	0.0	0.0

where $\hat{H}^\circ$ is the Hamiltonian for the solute molecule in the gas phase, $\hat{H}_{mm}$ is the molecular mechanical energy for the solvent, and $\hat{H}_{qm/mm}$ is the solute−solvent interaction Hamiltonian. Explicit descriptions of these Hamiltonians have been given in previous publications.[3−8] The total potential energy of the hybrid QM/MM system is determined by the expectation value of the molecular wave function, Φ, over the effective Hamiltonian, $\hat{H}_{eff}$:

$$E_{tot} = \langle\Phi|\hat{H}_{eff}|\Phi\rangle = E_{qm} + E_{qm/mm} + E_{mm} \qquad (2)$$

where Φ is the wave function of the solute molecule in solution and E_{mm} is the interaction energy of the solvent molecules, determined classically. In eq 2, E_{qm} and $E_{qm/mm}$ are evaluated by ab initio HF-SCF methods. These two terms represent respectively the energy of the solute molecule in solution and the solute−solvent interaction energy. $E_{qm/mm}$ includes two components, which are determined according to

$$E_{qm/mm} = \langle\Phi|\hat{H}^{el}_{qm/mm}|\Phi\rangle + \sum_{i}^{solute}\sum_{s}^{water} 4\epsilon_{is}\left[\left(\frac{\sigma_{is}}{R_{is}}\right)^{12} - \left(\frac{\sigma_{is}}{R_{is}}\right)^{6}\right] \qquad (3)$$

Here, $\hat{H}^{el}_{qm/mm}$ is the electronic part of the solute−solvent interaction Hamiltonian. The Lennard-Jones term in eq 3 accounts for short-range electron repulsions and dispersion interactions between the solute and solvent molecules. It is necessary because of the partition of a molecular system into inhomogeneous QM and MM regions.[5,6] We note that there are two empirical parameters, σ_{is} and ϵ_{is}, for each pair of interacting atoms. These parameters are determined from atomic parameters using standard combining rules such that $\sigma_{is} = (\sigma_i\sigma_s)^{1/2}$ and $\epsilon_{is} = (\epsilon_i\epsilon_s)^{1/2}$, where σ_i and ϵ_i are parameters for the "QM" solute atoms and σ_s and ϵ_s are parameters for the MM interaction sites. The atomic Lennard-Jones parameters for solvent water molecules are taken directly from the TIP3P model,[47] whereas the parameters for the solute molecule have been determined previously for the hybrid HF AI-3/TIP3P potential.[46] In the present calculation, the Lennard-Jones σ parameter for nitrogen has been slightly modified to obtain a better description of hydrogen-bonding interactions between NMA and water. All parameters are listed in Table 1.

We are particularly interested in the solute charge polarization induced by interactions with the solvent. As in previous studies,[6] the polarization effect is determined by the change in the solute wave function as illustrated in Figure 1. The

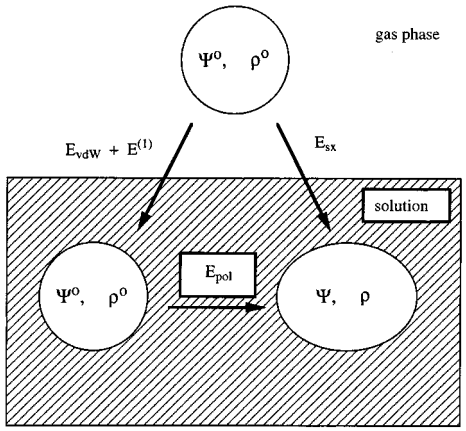

Figure 1. Schematic representation for the calculation of polarization energies. See refs 6 and 7.

associated polarization energy can be decomposed into two components:

$$E_{pol} = E_{dist} + E_{stab} \qquad (4)$$

where E_{dist} is the energy penalty necessary for polarizing the solute wave function (eq 5) and E_{stab} is the stabilization energy that is gained as a result of the solute charge polarization (eq 6).

$$E_{dist} = \langle\Phi|\hat{H}^\circ|\Phi\rangle - \langle\Phi^\circ|\hat{H}^\circ|\Phi^\circ\rangle \qquad (5)$$

$$E_{stab} = \langle\Phi|\hat{H}^{el}_{qm/mm}|\Phi\rangle - \langle\Phi^\circ|\hat{H}^{el}_{qm/mm}|\Phi^\circ\rangle \qquad (6)$$

Here Φ and Φ° are respectively the wave functions of the QM solute molecule in aqueous solution and in the gas phase. The total solute−solvent interaction energy can be written as follows:

$$E_{sx} = E_{vdW} + E^{(1)} + E_{pol} \qquad (7)$$

where $E^{(1)}$ is the solute−solvent interaction energy with the "permanent" or gas phase solute charge distribution (Φ°).[6] We note that the effect of the solvent polarization in response to the charge redistribution of the solute is not considered here because a significant increase in computational time will be required and would have been not practical for hybrid ab initio QM/MM simulations. However, polarizable solvent models have been implemented in semiempirical calculations. A detailed description can be found in refs 8a and 8j.

Statistical mechanical Monte Carlo simulations have been carried out for a cubic box containing 260 water molecules plus one NMA with the MCQUB/BOSS program.[48] The QM energies (HF/3-21G) are determined using the GAMESS program,[49] which has been modified to include the solute−solvent interaction terms, while the BOSS program presently handles input and output for the simulation. The isothermal−isobaric (NPT) ensemble at 25 °C and 1 atm is used with periodic boundary conditions, along with a spherical cutoff distance of 9 Å for the evaluation of solute−solvent and solvent−solvent interaction energies. The Monte Carlo simulations consist of at least 10^6 configurations of equilibration, which are followed by an additional 1.5×10^6 configurations of data averaging. All computations are performed on IBM RS/6000 Model 390 computers in our laboratory, which took roughly 10 days.

For comparison, these calculations are repeated with the use of the hybrid semiempirical AM1/TIP3P method[6] and Jorgensen's OPLS potential.[28,30] In the hybrid AM1/TIP3P calculation, the Lennard-Jones parameters for the solute molecule are taken

TABLE 2: Geometries and Interaction Energies (kcal/mol) for the Complexes of *N*-Methylacetamide with Water

	OPLS			AM1/MM			HF 3-21G/MM			6-31G*[a]		
	R_{OH}	ϕ	ΔE	R_{OH}	ϕ	ΔE	R_{OH}	ϕ	ΔE	R_{OH}	ϕ	ΔE
I	1.78	138	−6.5	1.71	133	−6.8	2.0	119	−6.6	1.98	115	−6.9
II	1.78	141	−7.2	1.70	144	−7.1	1.99	143	−6.6	1.98	134	−7.3
III	1.94	175	−5.9	1.77	176	−4.3	2.1	168	−5.6	2.12	177	−5.4

[a] Taken from ref 30 for *N*-methylacetamide−water complexes.

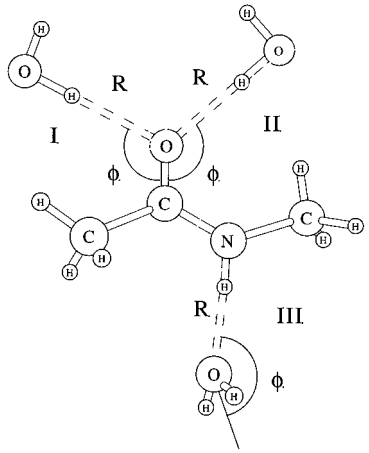

Figure 2. Structures considered for the complexes of *N*-methylacetamide with water.

from previous publications,[6] while Jorgensen's OPLS parameters are those from ref 30. To compare the results of these calculations, which employ different levels of approximation, all simulations have been performed utilizing the same conditions as in the hybrid ab initio AI-3/TIP3P calculation. The hybrid AM1/TIP3P simulation took about 10 h, while the use of the OPLS potential reduced the computational time to only 1.5 h.

Results and Discussion

Bimolecular Complexes. To assess the validity of the hybrid QM/MM potential for fluid simulations, hydrogen-bonding complexes of *N*-methylacetamide with a water molecule are first examined. In these calculations, NMA is treated quantum mechanically using the ab initio HF/3-21G method, whereas the water molecule is represented by the TIP3P model. Intermolecular geometry optimizations are executed using a program developed in our laboratory, with the NMA structure held fixed at the corresponding levels of theory. The experimental geometry was used for water. Thus, only intermolecular hydrogen bond distances and angles are optimized. The results are summarized in Table 2, and the structural arrangements are given in Figure 2. In addition, listed in Table 2 are results obtained using the OPLS and the hybrid AM1/TIP3P potential, along with full ab initio HF/6-31G* calculations. The data for the molecular mechanical and full ab initio calculations are taken from previous studies.[30]

The agreement in hydrogen-bonding energy between the hybrid AI-3/TIP3P model and the HF/6-31G* calculations is reasonable for both hydrogen bond acceptor (**I** and **II**) and donor (**III**) complexes. For structure **I**, the AI-3/TIP3P potential yields an interaction energy of −6.6 kcal/mol, comparing with values of −6.5, −6.8, and −6.9 from the OPLS, AM1/TIP3P, and HF/6-31G* calculations, respectively. The AI-3/TIP3P potential underestimates the hydrogen-bonding energy for structure **II** by 0.6 kcal/mol in comparison with the HF/6-31G* results. For the donor complex, **III**, the hybrid AI-3/TIP3P interaction energy is in good accord with the HF/6-31G* data, whereas the OPLS potential overestimates the interaction energy by 0.5 kcal/mol, and the hybrid AM1/TIP3P model yields weaker binding by about 1 kcal/mol. Note that the original OPLS potential predicts an interaction energy that is 1.2 kcal/mol more attractive than the ab initio data for **III**.[29]

TABLE 3: Computed Total Interaction Energies of NMA in Water and Energy Components (kcal/mol) at 25 °C and 1 atm[a]

energy	HF 3-21G/MM	AM1/MM
E_{pol}	−2.8 ± 0.1	−4.2 ± 0.2
E^1	−18.4 ± 0.5	−15.7 ± 0.3
E_{dist}	2.8 ± 0.1	4.0 ± 0.2
E_{vdW}	−6.2 ± 0.2	−6.7 ± 0.1
E_{stab}	−5.6 ± 0.2	−8.2 ± 0.3
E_{sx}	−27.4 ± 0.5	−26.5 ± 0.6

[a] E^1, which is defined as $\langle\Phi^\circ|\hat{H}^{el}_{qm/mm}|\Phi^\circ\rangle$, represents the electrostatic interaction energy of the solute with solvent, with the solute gas phase charge distribution (Φ°).

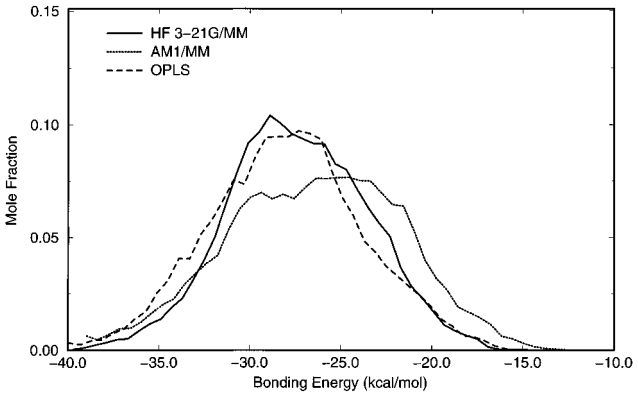

Figure 3. Computed distributions of total solute−solvent bonding energies (kcal/mol) for NMA in water. Units for the ordinates are mol % per kcal/mol.

The most significant improvement over the previous combined AM1/MM potential, as well as the empirical MM force field, is in the optimized geometrical variables. In the present AI-3/TIP3P optimization, hydrogen-bond distances are nearly in perfect agreement with the HF/6-31G* results. In contrast, because charge polarization is much smaller in the semiempirical AM1 model, the hydrogen bond distances are about 0.3 Å shorter than the corresponding HF/6-31G* values. A similar trend is necessary in empirical MM force fields, in which hydrogen bond distances are typically 0.2−0.4 Å too short comparing with HF/6-31G* results in order to compensate for the condensed phase polarization effects in fluid simulations.[50] Hydrogen bond angles predicted by the AI-3/TIP3P model are in accord with the ab initio results, though the angular flexibility typically is large for these complexes.

Energetics. Energetic results for NMA−water interactions in aqueous solution are listed in Table 3. The E_{sx} term gives the total solute−solvent interaction energy averaged over the trajectory during the Monte Carlo simulations. Overall, the agreement among the three potentials is remarkable in view of the different approximations that have been made in these models. Specifically, the total bonding energies of NMA in water are predicted to be −27.4 ± 0.5, −26.5 ± 0.6, and −28.0 ± 0.8 kcal/mol using the AI-3/TIP3P, AM1/TIP3P, and OPLS/TIP3P potentials, respectively. The energetic environment surrounding NMA in the solution is depicted by the bonding energy distribution functions in Figure 3. The NMA molecule experiences roughly an energy range of about 25 kcal/mol, between −40 and −15 kcal/mol, with an average reflected by the total solute−solvent interaction energy. However, it is

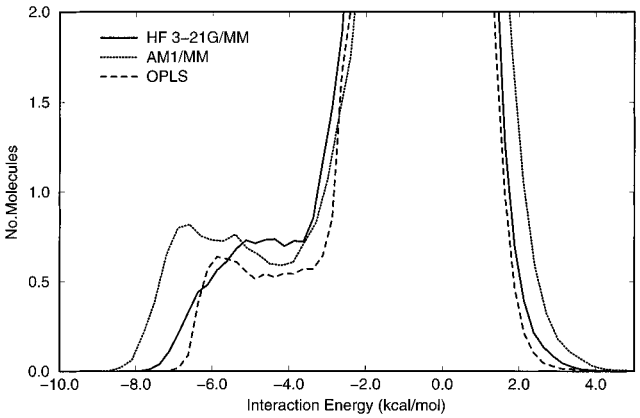

Figure 4. Computed distributions of individual solute−solvent interaction energies (kcal/mol) for NMA in water. Units for the ordinate are number of water molecules per kcal/mol.

apparent from Figure 3 that the interaction energy between NMA and water has somewhat wider distributions, perhaps due to the greater polarization effects predicted using the AM1/TIP3P model. Overall, the agreement with calculations using an optimized empirical potential for liquid simulations provides strong support to the validity of hybrid QM/MM methods for fluid simulations.

The energy pair distribution function contains important information on the hydrogen-bonding interactions between NMA and individual water molecules in the solution. Figure 4 compiles the enermy pair distribution functions obtained from the three separate simulations. The low-energy bands are from specific hydrogen-bonding interactions between the solute and solvent, while the central peak at interaction energy of close to zero represents interactions with distant water molecules in bulk. The lowest interaction energy appears at about −7 kcal/mol for simulations using the OPLS potential, which mirrors the interaction energies considered for the biomolecular complexes (Figure 1). On the other hand, for both hybrid potentials the lowest energy band starts at less than −8 kcal/mol, lower than the interaction energy for the bimolecular complexes. This is the consequence of the solvent polarization effect, which enhances the charge separation in NMA, leading to stronger interactions with water molecules in solution. This is a significant advantage of hybrid QM/MM potentials because the solute charge distributions are adjusted to the positions and charges of the surrounding environment.[6] However, in the empirical OPLS potential, atomic charges are fixed throughout the liquid simulation.[28−31] For the energy pair distribution function, the AI-3/TIP3P results seem to be in better agreement with the OPLS potential, while the hybrid AM1/TIP3P model yields stronger interactions with water.

Interestingly, the polarization energy can be directly evaluated from hybrid QM/MM approaches since the solvent effect is directly incorporated into the HF-SCF computation. Following the analysis outlined in ref 6, the polarization energy for a solute in solution is decomposed into two terms (eq 3): (1) a solute electron distortion energy, E_{dist}, which yields a positive value for reorganizing the solute electron distribution in solution; and (2) a solute electrostatic stabilization energy, E_{stab}, which is a net gain in the interaction energy between the polarized solute and the bulk solvent over that of an unpolarized solute (Figure 1). Usually, the energy penalty for distorting or polarizing the solute charge density to create its electron distribution in solution is one-half of the gain in interaction energy. Numerical simulation results provide a good assessment of such a linear response relationship.[51] The total polarization energy of *N*-methylacetamide in water is computed to be -2.8 ± 0.1 kcal/mol at the HF 3-21G/TIP3P level (Table 3), while a value of -4.2 ± 0.2 kcal/mol is obtained using the hybrid AM1/TIP3P model. The present calculation indicates that the polarization energy estimated with the hybrid AM1/TIP3P potential is greater than that of the ab initio HF 3-21G model by 1.4 kcal/mol. This is reflected by the wider bonding energy distribution for the AM1/TIP3P model (Figure 3). The difference in the predicted polarization energy between the two hybrid methods may be a consequence of the shorter interaction distances for bimolecular complexes in the hybrid AM1/TIP3P potential, which would impose stronger polarization effects on the solute molecule. However, it is also possible that the difference is due to the use of a semiempirical QM method. Of interest is to note the linear response behavior of the polarization energies computed at both ab initio and semiempirical levels. For the AI-3/MM potential, the E_{dist} term is determined to be 2.8 kcal/mol, exactly one-half of the E_{stab} term in magnitude (−5.6 kcal/mol). For the AM1/TIP3P model, the relationship also roughly holds with E_{dist} and E_{stab} values of 4.0 and −8.2 kcal/mol, respectively. Overall, the polarization energy for NMA in water contributes about 10−15% to the total solute−solvent interaction energy.

TABLE 4: Computed Partial Atomic Charges (*e*) and Dipole Moments (D)

	HF 3-21G/MM		AM1/MM	
atom	gas phase	aqueous	gas phase	aqueous
N	−0.905	-0.904 ± 0.001	−0.392	-0.365 ± 0.002
C	0.846	0.852 ± 0.001	0.300	0.329 ± 0.001
O	−0.631	-0.718 ± 0.004	−0.370	-0.514 ± 0.003
H_N	0.349	0.394 ± 0.002	0.220	0.269 ± 0.002
CH_3(N)	0.320	0.319 ± 0.003	0.171	0.182 ± 0.003
CH_3(C)	0.020	0.058 ± 0.003	0.069	0.099 ± 0.003
μ (D)	3.85	5.15 ± 0.03	3.51	5.20 ± 0.04

Atomic Charges. To further explore the solute charge redistribution due to interactions with the solvent, Table 4 compares the computed atomic charges for NMA in the gas phase with those obtained in solution from the AI-3/TIP3P and AM1/TIP3P potentials. In determining the atomic charges, Mulliken population analysis has been used since the primary purpose is to illustrate the *change* in atomic charges.[52] We only note here that atomic charges can also be divided by fitting the QM electrostatic potentials averaged during the liquid simulation.[53] For clarity, only the total charge on the methyl groups are given in Table 4. Standard deviations for the computed charges are about 0.002 *e* on average. As expected, the largest charge variations occur on the carbonyl oxygen and amide hydrogen atoms, with net changes of −0.144 and 0.049 *e*, respectively, using the AM1/TIP3P model, and of −0.087 and 0.045 *e* using the AI-3/TIP3P potential. For comparison, continuum self-consistent reaction field (SCRF) calculations at the HF/6-31G* level revealed charge variations of −0.115 and 0.029 *e* for these two atoms, in good accord with our simulation results.[38] Charge variations on the amide methyl group, CH_3-(N), is minimal, for which both the AI-3/TIP3P and AM1/TIP3P model predict nearly zero charge transfer. However, hybrid QM/MM models indicate that the acetyl group, CH_3(C), has a decrease of about 0.03−0.04 *e* in electron density (or an increase in overall positive partial charges) for the AM1/TIP3P and AI-3/TIP3P potential, along with a loss of about 0.05−0.09 *e* from the amino group ($NHCH_3$). Concomitantly, this is accompanied by the large electron density increase at the carbonyl oxygen. Thus, there is a significant intragroup charge transfer in the solvation of *N*-methylacetamide, leading to increased conjugation effects at the amide bond. This result echos the finding that rotational barriers around the peptide bonds have great solvent effects and are increased by about 2−4 kcal/mol in

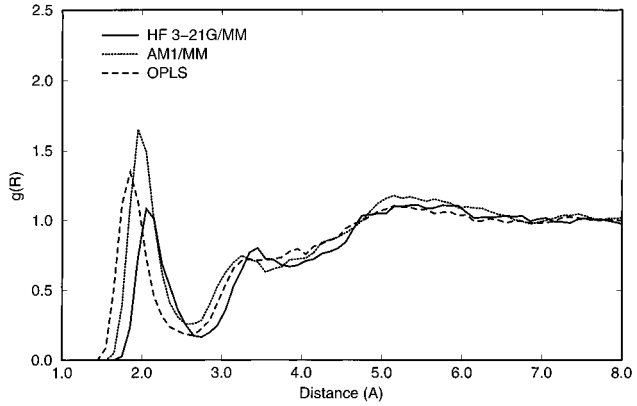

Figure 5. Carbonyl oxygen−water hydrogen radial distribution functions computed for *N*-methylacetamide in water.

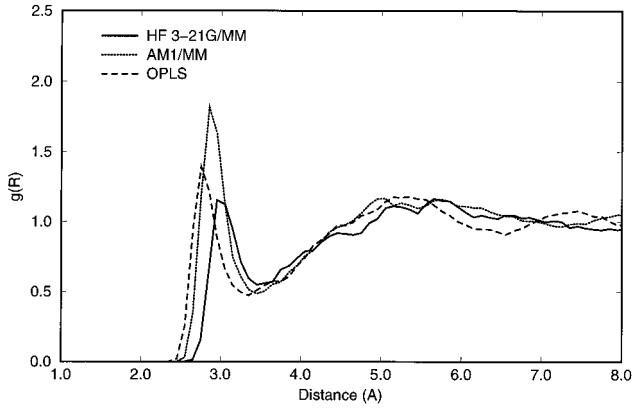

Figure 6. Carbonyl oxygen−water oxygen radial distribution functions computed for *N*-methylacetamide in water.

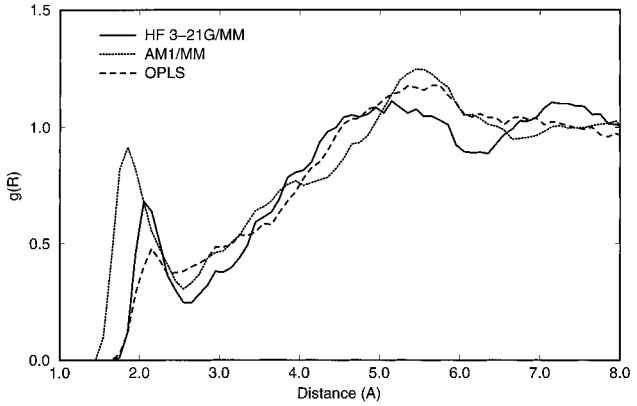

Figure 7. Amino hydrogen−water oxygen radial distribution functions computed for *N*-methylacetamide in water.

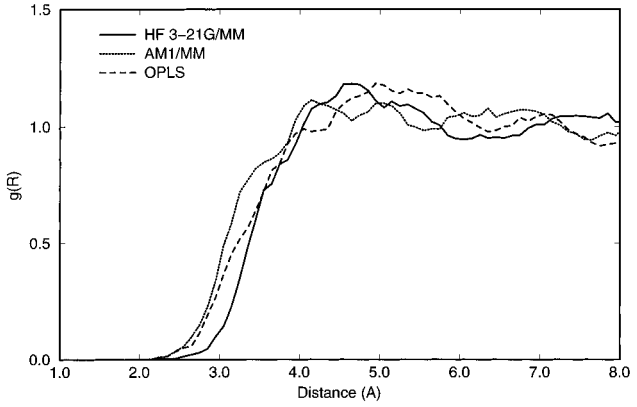

Figure 8. Nitrogen−water hydrogen radial distribution functions computed for *N*-methylacetamide in water.

aqueous solution relative to in organic solvents due to such an enhanced ground state delocalizations.[54]

A further indication of the polarization effect is the change in molecular dipole moment due to solvation (Table 4). The average dipole moment for NMA in water is 5.15 ± 0.03 D from hybrid AI-3/TIP3P simulations, which represents an induced dipole of 1.30 D over the gas phase value at the HF/3-21G level (3.85 D). The HF/3-21G dipole moment for NMA is in good accord with the experimental value of 3.78 D.[55] The semiempirical AM1/TIP3P model gives an induced dipole moment of 1.69 ± 0.04 D, which is somewhat greater than the AI3-TIP3P result. Comparison can be made to continuum SCRF HF/6-31G* calculations, which yielded dipole moments of 3.94 and 5.27 D in the gas phase and in water or an induced dipole of 1.33 D. This falls between our hybrid QM/MM simulation results. It is interesting to note that the dipole moment for NMA predicted with the OPLS partial charges is 3.85 D, which is much smaller than the value of about 5 D predicted here for NMA in water. However, surprisingly, interaction energies with water do not appear to be significantly different in these models.

Radial Distribution Functions. Specific hydrogen-bonding interactions are revealed in the radial distribution functions (rdf's) in Figures 5−8. In these figures, the first atom for an *xy* distribution, $g_{xy}(r)$, refers to an atom of the solute NMA, and the second atom is either the oxygen or the hydrogen of water. All radial distribution functions have been normalized to the bulk density of solvent atoms. The error range in these plots is estimated to be one-half of the bin size (0.05 Å) used in data collection.

The OO and OH distributions in Figures 5 and 6 have well-defined first peaks, indicating strong hydrogen-bonding interactions between the carbonyl oxygen with water hydrogen. The first peak for the OO rdf was predicted to occur at 3.00, 2.85, and 2.75 Å using the AI-3/TIP3P, AM1/TIP3P, and OPLS potentials, respectively. A similar trend is observed for the first peaks in the OH rdf's at 2.10, 1.95, and 1.85 Å from the three models. Integration to the minimum of each OH rdf gives nearest neighbors of 2.1, 2.5, and 2.1 for the AI-3/TIP3P, AM1/TIP3P, and OPLS calculations.

For the amide HO and NH rdf's (Figures 7 and 8), differences are shown among the three models. The first peak in the amide HO rdf occurs at 2.05 Å with the AI-3/TIP3P potential, and the hybrid AM1/TIP3P model gives a maximum peak at 1.85 Å. The OPLS potential has a somewhat larger value than the two hybrid QM/MM values at 2.15 Å, although it is in reasonable agreement with the AI-3/TIP3P potential. The short hydrogen bond distance predicted by the hybrid AM1/TIP3P is a documented shortcoming of the model, which is necessary in order to yield good interaction energies in comparison with ab initio results. The number of oxygen molecules hydrogen bonded to the amide hydrogen is estimated to be 0.7, 1.0, and 0.5 by integrating the HO rdf's for the AI-3/TIP3P, AM1/TIP3P, and OPLS potentials. The small coordination number to the amide hydrogen for the modified OPLS potential, which was designed to describe NMA−water interactions for both the cis and trans configurations,[30] is in reasonable accord with the AI-3/TIP3P potential. We note that the original OPLS potential, which overestimates the interaction energy between the amide hydrogen and water for NMA by 1.2 kcal/mol, predicted 0.9 water oxygens from the amide hydrogen.[29] The NH rdf's displayed in Figure 8 show little structural features for all three potential functions used. For the hybrid AM1/TIP3P and the OPLS potential, N−H distances are about 0.2−0.3 Å shorter

than using the hybrid AI-3/TIP3P model. Figure 8 indicates that, in all cases, solvent molecules do not participate in hydrogen-bonding interactions with the nitrogen atom perpendicular to the amide plane. All other radial distribution functions show little specific structural interest, and thus they are not displayed here.

Generally, we find that the radial distribution functions obtained from the hybrid AI-3/TIP3P calculations display greater interaction distances between the solute and solvent atoms. The hybrid AM1/TIP3P model as well as the OPLS potential gives relatively shorter distances for certain interaction pairs. Although the general features of the radial distribution functions for hydrogen-bonding pairs are similar in all calculations, the specific details in hydrogen-bonding distances can be different by as much as 0.2−0.3 Å. Thus, it is necessary to examine bimolecular complexes by comparison with high-level ab initio calculations when these potential functions are chosen for fluid simulations. The present study indicates that the hybrid HF 3-21G/MM approach yields the best structural results in comparison with full ab initio HF/6-31G* calculations in the gas phase and more consistent hydrogen bond distances from the computed $g(r)$'s in solution.

Conclusions

Hybrid QM/MM simulations of *N*-methylacetamide in aqueous solution has been carried out using the combined ab initio HF 3-21G/MM potential in Monte Carlo calculations. In the present study, the solute molecule is treated by the ab initio Hartree−Fock theory with the use of the 3-21G basis function, while the solvent molecules are represented by the three-point charge TIP3P model for water. The hybrid HF 3-21G/MM potential is found to yield good results for bimolecular complexes of *N*-methylacetamide with water in comparison with full ab initio HF/6-31G* results. The predicted hydrogen bond distances for NMA−water complexes are found to be in excellent agreement with the ab initio data, whereas both empirical OPLS and hybrid semiempirical AM1/MM methods give shorter hydrogen bond distances. The solvent effect on the solute electronic polarization has been analyzed. At the hybrid HF 3-21G/MM level, a polarization energy of −2.8 kcal/mol is obtained, which comprises 10% of the total solute−solvent interaction energy. The semiempirical method (AM1/TIP3P model) yields great polarization effects with a computed polarization energy of −4.2 kcal/mol or about 15% of the total solute−solvent interaction energy. The energy gained in the solute polarization is found to be twice the energy costs for polarizing the solute charge distribution, consistent with the linear response relationship. The induced dipole moment of NMA in water is predicted to be 1.30 ± 0.03 D over its gas phase value (3.85 D) using the ab initio 3-21G/MM potential and 1.69 ± 0.04 D with the semiempirical AM1/MM model. The charge polarization was echoed by an intragroup charge transfer of about 0.05−0.09 *e* from the amino unit to the carbonyl group across the peptide bond. Thus, the conjugation effect of the peptide bond is predicted to be increased in aqueous solution over that in the gas phase or nonpolar solvents. Comparisons of structural findings with results obtained using the parametrized empirical OPLS potential and the hybrid semiempirical AM1/TIP3P model indicate that the hybrid HF 3-21G/MM potential can provide an excellent description of solute−solvent interactions. The present study demonstrates the feasibility of ab initio hybrid QM/MM methods for fluid simulations.

Acknowledgment. Gratitude is expressed to the National Science Foundation, the Environmental Protection Agency, and the National Institutes of Health for support of this research. Part of the computations have been performed using the computing fascilities at the Center for Advanced Molecular Biology and Immunology of SUNY, Buffalo.

References and Notes

(1) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acid*; Cambridge University Press: Cambridge, 1987.

(2) (a) Barnes, P.; Finney, J. L.; Nicholas, J. D.; Quinn, J. E. *Nature* **1979**, *282*, 459. (b) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341. (c) Gao, J.; Pavelites, J. J.; Habibollazadeh, D. *J. Phys. Chem.* **1996**, *100*, 2689.

(3) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.

(4) (a) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (b) Tapia, O.; Colonna, F.; Angyan, J. G. *J. Chim. Phys. Phys.-Chim. Biol.* **1990**, *87*, 875. (c) Thole, B. T.; van Duijnen, P. Th. *Chem. Phys.* **1982**, *71*, 211.

(5) (a) Bash, P. A.; Field, M. J.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092. (b) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. (c) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826.

(6) (a) Gao, J.; Xia, X. *Science* **1992**, *258*, 631. (b) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537. (c) Gao, J.; Pavelites, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 1912. (c) Gao, J.; Li, N.; Freindor, M. *J. Am. Chem. Soc.* **1996**, *118*, 4912.

(7) (a) Gao, J. In *Reviews in Computational Chemistry*; VCH: New York, 1995; Vol. 7, pp 119−185. (b) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298.

(8) (a) Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374. (b) Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 384. (c) Liu, H.; Shi, Y. *J. Comput. Chem.* **1994**, *15*, 1311. (d) Stanton, R. V.; Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1993**, *97*, 11868. (e) Wei, D.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *224*, 291. (f) Liu, H.; Muller-Plathe, F.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 1702. (g) Liu, H.; Muller-Plathe, F.; van Gunsteren, W. F. *J. Mol. Biol.* **1996**, *286*, 454. (h) Chatfield, D. C.; Brooks, B. R. *J. Am. Chem. Soc.* **1995**, *117*, 5561. (i) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580. (j) Gao, J. *J. Comput. Chem.*, in press.

(9) Drakenberg, T.; Forsen, S. *J. Chem. Soc., Chem. Commun.* **1971**, 1404.

(10) Wolfenden, R. *Biochemistry* **1978**, *17*, 201.

(11) Spencer, N. J. *J. Phys. Chem.* **1981**, *85*, 1236.

(12) Pralat, K.; Jadzyn, J.; Balanicka, S. *J. Phys. Chem.* **1983**, *87*, 1385.

(13) Ataka, S.; Takeuchi, H.; Tasumi, M. *J. Mol. Struct.* **1984**, *113*, 147.

(14) Dudik, J. M.; Johnsonn, C. R.; Sanford, A. A. *J. Phys. Chem.* **1985**, *18*, 3805.

(15) Harhay, G. P.; Hudson, B. S. *J. Phys. Chem.* **1993**, *97*, 8158.

(16) Mayne, L. C.; Ziegler, L. D.; Hudson, B. *J. Phys. Chem.* **1985**, *89*, 3395.

(17) Rodrigo, M. M.; Tarazona, M. P.; Saiz, E. *J. Phys. Chem.* **1986**, *90*, 2236.

(18) Radzicka, A.; Pedersen, L.; Wolfenden, R. *Biochemistry* **1988**, *27*, 4538.

(19) Krimm, S.; Song, S.; Asher, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 4290.

(20) Song, S.; Asher, S. A.; Krimm, S.; Shaw, K. D. *J. Am. Chem. Soc.* **1991**, *113*, 1155.

(21) Wang, Y.; Purrello, R.; Savas, G.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6368.

(22) Wang, Y.; Purrello, R.; Savas, G.; Spiro, T. G. *J. Am. Chem. Soc.* **1991**, *113*, 6359.

(23) Li, Y.; Garrell, R. L.; Houk, K. N. *J. Am. Chem. Soc.* **1991**, *15*, 5895.

(24) Chen, X. G.; Schweitzer-Stenner, R.; Krimm, S.; Mirkin, N. G.; Asher, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 11141.

(25) Pullman, A. *Theor. Chim. Acta* **1974**, *33*, 87.

(26) Nitzsche, L. E.; Davidson, E. R. *J. Am. Chem. Soc.* **1978**, *23*, 7201.

(27) Radom, L.; Riggs, N. V. *Aust. J. Chem.* **1982**, *35*, 1071.

(28) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569.

(29) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 1489.

(30) Jorgensen, W. L.; Gao, J. *J. Am. Chem. Soc.* **1988**, *110*, 4212.

(31) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 3469.

(32) Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770.

(33) Yu, H. A.; Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 2425.

(34) Mirkin, N. G.; Krimm, S. *J. Am. Chem. Soc.* **1991**, *113*, 9742.

(35) Cieplak, P.; Kollman, P. *J. Comput. Chem.* **1991**, *12*, 1232.

(36) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc*. **1992**, *114*, 7535.
(37) Guo, H.; Karplus, M. *J. Phys. Chem*. **1992**, *96*, 7272.
(38) Luque, F. J.; Orozco, M. *J. Org. Chem*. **1993**, *58*, 6397.
(39) Guo, H.; Karplus, M. *J. Phys. Chem*. **1994**, *98*, 7104.
(40) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, III, W. A.; Hong, B. *J. Am. Chem. Soc*. **1994**, *116*, 11875.
(41) Morgantini, P. Y.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 6057.
(42) Ding, Y.; Bernardo, D. N.; Jespersen, K. K.; Levy, R. M. *J. Phys. Chem*. **1995**, *99*, 11575.
(43) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc*. **1996**, *118*, 672.
(44) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
(45) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc*. **1980**, *102*, 939.
(46) Freindorf, M.; Gao, J. *J. Comput. Chem*. **1996**, *17*, 386.
(47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926.
(48) (a) Gao, J. MCQUB Department of Chemistry, State University of New York at Buffalo, 1996. (b) Jorgensen, W. L. BOSS, Version 2.9, Department of Chemistry, Yale University, 1990.
(49) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. S. *J. Comput. Chem*. **1993**, *14*, 1347.
(50) Jorgensen, W. L. *Chemtracts: Org. Chem*. **1991**, *4*, 91.
(51) (a) Tapia, O.; Silvl, B. *J. Phys. Chem*. **1980**, *84*, 2646. (b) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys*. **1981**, *55*, 117.
(52) Mulliken, R. S. *J. Chem. Phys*. **1955**, *23*, 1833.
(53) (a) Gao, J.; Luque, F. J.; Orozco, M. *J. Chem. Phys*. **1993**, *98*, 2975. (b) Singh, U. C.; Kollman, P. A. *J. Comput. Chem*. **1984**, *5*, 129.
(54) (a) Drakenberg, T.; Dahlqvist, K.-I.; Forsen, S. *J. Phys. Chem*. **1972**, *76*, 2178. (b) Ross, B. D.; True, N. S. *J. Am. Chem. Soc*. **1984**, *106*, 2451. (c) Gao, J. *J. Am. Chem. Soc*. **1993**, *115*, 2930. (d) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc*. **1992**, *114*, 7535.
(55) Weast, R. C., Ed. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, 1985−1986; Vol. 66.