

Hybrid Quantum and Molecular Mechanical Simulations: An Alternative Avenue to Solvent Effects in Organic Chemistry

JIALI GAO

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

Received November 13, 1995

Computer simulation has drastically changed chemists' perception and understanding of chemical reactions and interactions in solution.¹ Methods for modeling solute–solvent interactions may be broadly divided into two categories, according to the representation of the solvent. The earliest approach, which is still being widely used today, is based on a continuum treatment of the medium characterized by a bulk dielectric constant.² Continuum models can provide valuable insight into the solvent effect and are computationally efficient, but they lack specific information on intermolecular interactions.² The second approach is based on Monte Carlo (MC) and molecular dynamics (MD) simulation techniques, in which one or more solute molecules are placed in a periodic unit cell consisting of a few hundred or perhaps a few thousand solvent molecules.³ Thermodynamic properties are averaged over geometries and conformations sampled randomly in the MC simulation, or over a time-dependent trajectory in the MD calculation. These computational procedures are also used in modeling biological macromolecules,^{4,5} providing valuable information in areas such as rational drug design. In all simulation studies, the key element is the accuracy of the potential energy functions used to describe intermolecular interactions.

A rigorous quantum mechanical treatment of the entire fluid system would be ideal in determining the potential energy surface, but is too time-consuming to be practical because of the size and complexity of the system.⁵ Instead, molecular mechanics (MM) force fields, which treat atoms and molecules as van der Waals spheres, are traditionally used.^{3,4} Great efforts have been made to develop reliable potential functions. Now, a number of MM force fields are available for organic and biological systems.⁶ MM calculations can provide valuable information on molecular structure and thermodynamic properties in solution such as the free energy of solvation.⁴ Yet, the reorganization of electronic structures in making and breaking chemical bonds requires quantum mechanical treatment. Therefore, a logical alternative approach is to combine quantum mechanical and molecular mechanical (QM/MM) methods in MC and MD simulations.^{7–9} In such a combined or hybrid QM/MM potential, the solute molecule (or the reactant) is treated quantum-me-

chanically, whereas the rest of the system is approximated by MM force fields. An attractive feature of the method is that it synthesizes the accuracy/generality offered by QM calculations and the computational efficiency of MM representations. The procedure may be systematically improved by increasing either the level of the QM treatment of the solute or the sophistication of the MM force field.¹⁰ In this Account, the focus will be our work on organic systems in solution. References 7–9 contain additional technical details and reviews that reflect other research in this area.

Combined Quantum Mechanical and Molecular Mechanical Methods

Combined QM/MM potentials were first used in classical trajectory calculations in the gas phase by treating the π electron system quantum-mechanically on a σ framework represented by molecular mechanics.¹¹ Subsequently, in a seminal paper, Warshel and Levitt laid out the basic algorithm of the combined QM/MM method for biological systems.^{7a} The simulation techniques presented here are based on the procedure introduced by Field et al.⁸ Early calculations were primarily performed with energy minimizations.¹² MD and MC simulations were only practical

(1) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; VCH: New York, 1990.

(2) For recent reviews, see: (a) Tomasi, J. *Chem. Rev.* **1994**, *94*, 2027. (b) Cramer, C. J.; Truhlar, D. G. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1995; Vol. 6, pp 1–72.

(3) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: London, 1987.

(4) (a) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University: Cambridge, 1987. (b) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. *Adv. Chem. Phys.* **1988**, *71*, 1.

(5) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and in Solutions*; John Wiley & Sons: New York, 1991.

(6) (a) Burkert, U.; Allinger, N. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982. (b) Hagler, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319. (c) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187. (d) Jorgensen, W. L.; Tirado-Rives, J. *Am. Chem. Soc.* **1988**, *110*, 1657. (e) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(7) (a) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. (b) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523. (c) Topia, O.; Colonna, F.; Angyan, J. G. *J. Chim. Phys. Phys.-Chim. Biol.* **1990**, *87*, 875.

(8) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.

(9) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1995; Vol. 7, pp 119–185.

(10) Gao, J.; Furlani, T. R. *IEEE Comput. Sci. Eng.* **1995**, Fall issue, 24.

(11) (a) Wang, I. S. Y.; Karplus, M. *J. Am. Chem. Soc.* **1973**, *95*, 8160. (b) Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1972**, *94*, 5612.

(12) Weiner, S. J.; Leibel, G. L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 649.

Jiali Gao was born in Heilongjiang Province, China, in 1962. He received his B.S. from Beijing University in 1982, which marked the first collegial graduation after the "Cultural Revolution". He came to the U.S. through William von E. Doering's Chemical Graduate Program, and obtained his Ph.D. from Purdue University in 1987. Following a postdoctoral fellowship at Harvard, he joined the Department of Chemistry at the State University of New York at Buffalo in 1990, where he is presently an Associate Professor of Chemistry. He dedicates this Account to Professor William von E. Doering for his generosity and contributions to organic chemistry.

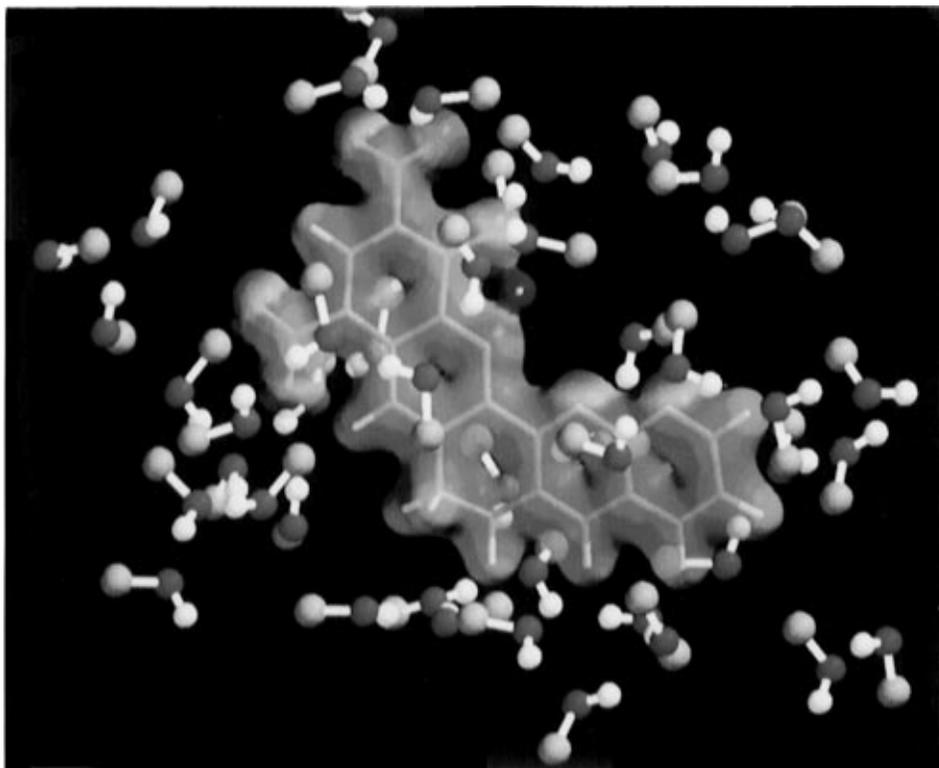


Figure 1. Schematic representation of the partition of an organic system in methanol, a computational model for the proton transfer reaction in creatinine. The electron density of the solute (creatinine) is colored by the electrostatic potential (blue is positive, and red is negative) to emphasize that the solute is treated quantum-mechanically. The solvent methanol molecules, in which the methyl group is represented by a single van der Waals sphere, are treated classically and are shown in a ball-and-stick model.

recently thanks to advances both in computer technology and in the development of QM algorithms.^{5,13–15}

Figure 1 illustrates the treatment of a solution system by the hybrid QM/MM method. The molecular system (either a dilute solution or an enzyme) is partitioned into a quantum mechanical region and a molecular mechanical region. For the QM region, electronic structure of the solute molecules is represented by a restricted Hartree–Fock (HF) wave function, Φ , which is written as a Slater determinant of all doubly occupied molecular orbitals (MO). The solvent atoms in the MM region are approximated by empirical force fields. Bond lengths and angles of the solvent molecules are often fixed to save computer time, although torsional motions are generally included in the MC simulation. This method can be extended to open-shell systems and excited state calculations, as illustrated later. The effective Hamiltonian of the system is given by eq 1,

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{QM}}^{\text{p}} + \hat{H}_{\text{QM/MM}}^{\text{el}} + \hat{H}_{\text{QM/MM}}^{\text{vdW}} + \hat{H}_{\text{MM}} \quad (1)$$

where $\hat{H}_{\text{qm}}^{\text{p}}$ is the Hamiltonian for the QM solute molecule, \hat{H}_{mm} is the solvent–solvent interaction energy, and $\hat{H}_{\text{qm/mm}}^{\text{el}}$ and $\hat{H}_{\text{qm/mm}}^{\text{vdW}}$ are the QM/MM electrostatic (eq 2) and van der Waals (vdW; eq 3) interaction Hamiltonians. In eqs 2 and 3, q_s and Z_m

$$\hat{H}_{\text{QM/MM}}^{\text{el}} = - \sum_{s=1}^S \sum_{i=1}^N \frac{eq_s}{r_{is}} + \sum_{s=1}^S \sum_{m=1}^M \frac{q_s Z_m}{R_{ms}} \quad (2)$$

$$\hat{H}_{\text{QM/MM}}^{\text{vdW}} = \sum_{s=1}^S \sum_{m=1}^M 4\epsilon_{ms} \left[\left(\frac{\sigma_{ms}}{R_{ms}} \right)^{12} - \left(\frac{\sigma_{ms}}{R_{ms}} \right)^6 \right] \quad (3)$$

are, respectively, charges on the solvent and solute nuclear centers, N is the total number of electrons in the solute, S and M are the total number of solvent and solute atoms, respectively, and r and R are distances of the solute electrons and nuclei from the solvent atoms.

Since electronic structures of the solvent molecules are not explicitly represented, the QM/MM vdW term is required in eq 1 to account for the electronic repulsion and dispersion interactions between the QM and MM regions.⁸ The Lennard-Jones parameters σ_{ms} and ϵ_{ms} in eq 3 are obtained by the use of geometric means of atomic parameters for the MM (σ_s and ϵ_s) and QM (σ_m and ϵ_m) regions, such that $\sigma_{ms} = (\sigma_m \sigma_s)^{1/2}$ and $\epsilon_{ms} = (\epsilon_m \epsilon_s)^{1/2}$. σ_s and ϵ_s are taken directly from the MM force field,⁶ whereas σ_m and ϵ_m are new empirical parameters for QM atoms in the hybrid QM/MM potential. The latter must be optimized for the particular QM and MM combination in order to yield good results for solute–solvent interactions.¹⁶ These parameters depend on the atomic number and hybridization of the QM atom, and are fully transferable from one system to another. They are determined by fitting results from hybrid QM/MM calculations to full ab initio or experimental data for hydrogen-bonded complexes.^{14,16}

The total energy of the fluid system at an instantaneous configuration during an MC or MD simulation is evaluated by the expectation value of the wave

(13) (a) Bash, P. A.; Field, M. J.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092. (b) Hwang, J.-K.; King, G.; Creighton, S.; Warshel, A. *J. Am. Chem. Soc.* **1988**, *110*, 5297.

(14) (a) Gao, J.; Xia, X. *Science* **1992**, *258*, 631. (b) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537.

(15) (a) Thompson, M. A. *J. Phys. Chem.* **1995**, *99*, 4794. (b) Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374. (c) Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 384. (d) Liu, H.; Shi, Y. *J. Comput. Chem.* **1994**, *15*, 1311.

(16) Gao, J. In *Modeling the Hydrogen Bond*; Smith, D. A., Ed.; ACS Symposium Series 569; American Chemical Society: Washington, DC, 1994; pp 8–20.

function and is given as follows:

$$E_{\text{tot}} = \langle \Phi | \hat{H}_{\text{eff}} | \Phi \rangle = E_{\text{QM}} + E_{\text{QM/MM}}^{\text{el}} + E_{\text{QM/MM}}^{\text{vdW}} + E_{\text{MM}} \quad (4)$$

Technically, the computational procedure involves modification of the one-electron integral term in standard Hartree–Fock calculations of gas-phase molecules.^{8,9,14} However, what is significant is that the hybrid QM/MM approach provides a practical bridge to extend theoretical organic chemistry beyond the gas-phase realm to quantum mechanical investigations in solution.

Ab initio or density functional theory (DFT) methods would be desirable for use in combined QM/MM calculations because computational accuracy can be systematically improved.¹⁷ Indeed, both ab initio and DFT procedures have been implemented and used in QM/MM calculations.^{18,19} However, ab initio and DFT methods are too slow to be practical in MC and MD simulations of organic and biological systems because millions of electronic structure calculations are needed. For systems of organic and biological interest, computationally efficient QM methods must be employed at the present time in routine calculations. Consequently, the semiempirical Austin model 1 (AM1) theory developed by Dewar and co-workers and the closely related PM3 method by Stewart have been primarily utilized in QM/MM simulations.²⁰ Warshel has extensively used an empirical valence bond (EVB) approach, in which empirical parameters are calibrated against experimental data.^{5,7}

Bimolecular Interactions and Solvation Free Energies

Applying hybrid QM/MM methods, we have studied a variety of bimolecular complexes between water and organic compounds.^{16,19} In these investigations, the organic species are treated quantum-mechanically, and water is approximated by Jorgensen's three-point charge (TIP3P) model.^{21a} Figure 2 depicts an example of the structural arrangements for acetic acid and water complexes. These studies provided crucial validation of the hybrid QM/MM method for modeling intermolecular interactions. In addition, the Lennard-Jones parameters embedded in the hybrid QM/MM potential (eq 3) are optimized through these calculations.^{16,19}

Hydrogen-bonding energies predicted by the combined AM1/TIP3P model were found to be in accord with full ab initio 6-31G(d) and available experimental data. This is demonstrated by a root-mean-square (rms) deviation of 0.9 kcal/mol for 67 complexes.^{8,16} However, the semiempirical AM1 method tends to underestimate charge separations, particularly for the

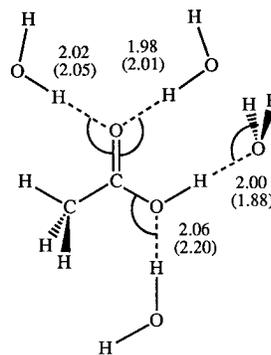


Figure 2. Illustration of bimolecular complexes for acetic acid with water. A variety of similar complexes have been studied using the combined QM/MM method.

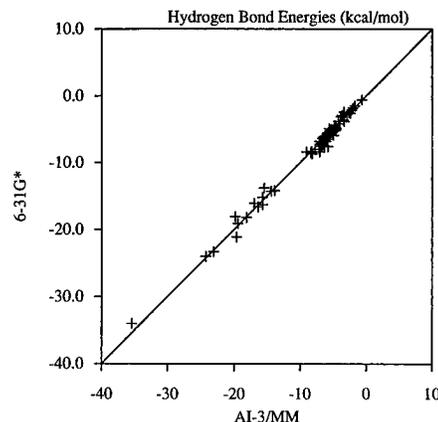


Figure 3. Comparison of interaction energies (kcal/mol) obtained from the hybrid ab initio 3-21G/TIP3P model and 6-31G(d) optimizations.

hydroxyl and amino groups.¹⁶ As a result, to maintain reasonable agreement with ab initio interaction energies, hydrogen bond distances predicted with the hybrid AM1/TIP3P potential are about 0.2–0.4 Å shorter than the corresponding HF/6-31G(d) values for some complexes.¹⁶ This certainly is a deficiency of the AM1/TIP3P model, and demands further investigations to improve its quality. It should be noted that empirical potentials such as Jorgensen's OPLS (optimized potential for liquid simulations) functions also yield shorter hydrogen bonds by approximately 0.2 Å for these complexes.^{21b}

We have recently incorporated the ab initio GAMESS package into our simulation program.^{22a} Using the 3-21G basis set,^{17a} we observed improvements in both energetics and structure over the semiempirical AM1/MM combination.¹⁹ Figure 3 shows a comparison of interaction energies obtained from the hybrid HF 3-21G/MM and full HF/6-31G(d) optimizations for over 80 complexes. Figure 3 shows good correlation with an rms deviation of 0.5 kcal/mol. Most significant is the improvement in the predicted hydrogen-bond distances; the average deviation from HF/6-31G(d) values is less than 0.1 Å.

The key thermodynamic quantity needed to characterize chemical equilibria in solution is the free

(17) (a) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986. (b) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: London, 1989.

(18) (a) Stanton, R. V.; Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1993**, *97*, 11868. (b) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050. (c) Wei, D.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *224*, 291.

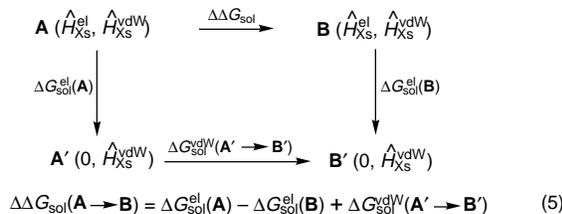
(19) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386.

(20) (a) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902. (b) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209, 221.

(21) (a) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926. (b) Jorgensen, W. L. *Chemtracts: Org. Chem.* **1991**, *4*, 91.

(22) (a) 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. *GAMESS*, version 11, 1994. (b) Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A. *Gaussian 90*; Gaussian Inc.: Pittsburgh, 1990.

energy difference.²³ A number of methods are available for computing the difference in free energy of solvation between two solutes.² The most widely used technique is the free energy perturbation (FEP) method,^{23,24} which involves “mutation” of one solute into another through a series of simulations that gradually change the geometrical and potential function parameters. However, it is difficult to annihilate or create electrons and nuclei in HF calculations employing a hybrid QM/MM potential. To circumvent this problem, we have developed an elaborate, two-step procedure for FEP calculations of molecules of different sizes (**A** and **B**).¹⁴ The method is depicted in the thermodynamic cycle below:



In the first step, the solvent electrostatic term in the QM/MM Hamiltonian (eq 1) is dwindled away, leaving only the vdW term for solute–solvent interactions. Then, the fictitious species **A'** and **B'** are interconverted using standard MM procedures.²³ Typically, each conversion involves five to ten steps to yield the free energy changes in eq 5.^{14,23} Note that Warshel has used an alternative route, in which MM potentials are first used to carry out fluid simulations. Then, the free energy difference between the QM and MM potential surfaces is determined on the basis of the MM trajectory.^{5,7}

The relative free energies of hydration for a series of organic compounds and ions were determined at 25 °C and 1 atm using the AM1/TIP3P potential in 1992.¹⁴ The results were found to be in accord with experimental data,²⁵ with errors similar to those in pure MM force field calculations. For example, the difference in free energy of hydration between methanol and ethane, a system widely used as a test case for free energy perturbation (FEP) calculations, was computed to be 6.2 kcal/mol, which may be compared with the experimental value of 6.9 kcal/mol.²⁵ These calculations were recently extended to determine the hydration free energies of the five nucleotide bases.²⁶ The results from our QM/MM Monte Carlo simulations are consistent with other theoretical predictions.²⁷ Still, several molecules exhibited larger deviations from experiment; the computed $\Delta\Delta G_{\text{hyd}}$ for *N*-methylacetamide (NMA) was -8.5 kcal/mol, while Wolfenden's equilibrium experiments yielded a value of -12 kcal/mol (relative to ethane).^{14,25} The disagreement warns that the performance of the hybrid AM1/TIP3P model must be verified for a given system.

Overall, a hybrid QM/MM potential making use of the semiempirical AM1 theory can provide reasonable

Table 1. Computed Polarization Energies and Components (kcal/mol)^a

compound	$E^{(1)}$	ΔE_{stab}	ΔE_{dist}	E_{pol}	$\Delta G_{\text{pol}}(\text{SCRf})$
H ₂ O	-13.7	-3.7	1.8	-1.9	-0.7
CH ₃ OH	-10.3	-3.4	1.7	-1.8	-0.5
CH ₃ NH ₂	-4.6	-0.9	0.4	-0.5	-0.3
CH ₃ OCH ₃	-5.3	-1.5	0.8	-0.8	-0.4
CH ₃ COCH ₃	-8.2	-3.1	1.5	-1.5	-1.2
CH ₃ CO ₂ H	-14.6	-3.9	1.9	-1.9	-1.1
NMA	-16.2	-7.9	3.9	-4.0	-2.0
adenine	-19.5	-7.6	3.8	-3.8	-4.5
cytosine	-34.8	-16.7	8.3	-8.5	-4.4
guanine	-37.2	-15.6	7.8	-7.8	-5.6
thymine	-21.2	-7.9	4.0	-4.0	-3.1
uracil	-23.8	-7.8	3.9	-3.9	-3.8

^a Energies are taken from simulations reported in refs 14 and 26. Free energies for the SCRf(AM1) models are from ref 27.

descriptions for intermolecular interactions in the gas phase as well as in solution, and the method may be applied to study important chemical processes in solution.

Solvent Polarization Effects

An important feature in the hybrid QM/MM calculation is that the solute wave function is determined both in the gas phase and in solution, allowing an *a priori* evaluation of the solvent polarization effects on the solute electronic structure. Using a perturbation approach,¹⁰ the solute electrostatic interaction energy in solution can be expressed by eq 6, where Φ and Φ^0

$$\Delta E_{\text{QM/MM}}^{\text{el}} = \langle \Phi | \hat{H}_{\text{QM}}^{\text{el}} + \hat{H}_{\text{QM/MM}}^{\text{el}} | \Phi \rangle - \langle \Phi^0 | \hat{H}_{\text{QM}}^{\text{el}} | \Phi^0 \rangle = \Delta E^{(1)} + \Delta E_{\text{pol}} \quad (6)$$

are solution and gas-phase wave functions for the solute, $\Delta E^{(1)}$ is the first-order perturbation ($\langle \Phi^0 | \hat{H}_{\text{QM/MM}}^{\text{el}} | \Phi^0 \rangle$), representing the interaction energy of the “unpolarized” solute with the solvent. All higher order perturbation terms consist of modifications to the wave function Φ^0 and give rise to solvent polarization energy.^{14a} ΔE_{pol} , which is always favorable (negative), can be further decomposed into two components: a solute–solvent stabilization term, ΔE_{stab} , and an energy penalty for distorting the solute wave function from its gas-phase equilibrium state into that in solution, ΔE_{dist} . According to classic linear response theory, $\Delta E_{\text{stab}} = -2\Delta E_{\text{dist}}$, which is reproduced by our numerical results listed in Table 1 for several representative molecules in aqueous solution.^{14a} In addition, the MC–QM/MM simulation data are compared with results from continuum solvation models (Table 1).²⁷ The agreement was remarkable, despite the differences in computational procedures. In all cases, the polarization effect is significant, contributing 10–20% to the electrostatic energy. This emphasizes the importance of including specific polarization terms in the MM force field for studies of protein–DNA interaction and molecular recognition. Indeed, polarization energies determined from combined QM/MM simulations provide valuable information needed to parametrize polarizable intermolecular potential functions (PIPF). A set of PIPF potentials for protein simulations is currently under development.²⁸

(28) Gao, J.; Habibollahzadeh, D.; Shao, L. *J. Phys. Chem.* **1995**, *99*, 16460.

(23) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.

(24) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

(25) (a) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016. (b) Wolfenden, R.; Anderson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849.

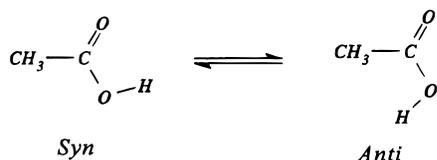
(26) Gao, J. *Biophys. Chem.* **1994**, *51*, 253.

(27) (a) Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1992**, *198*, 74. (b) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629. (c) Orozco, M.; Luque, F. J. *Biopolymers* **1993**, *33*, 1851. (d) Young, P. E.; Hillier, I. H. *Chem. Phys. Lett.* **1993**, *215*, 405. (e) Orozco, M.; Luque, F. J.; Habibollahzadeh, D.; Gao, J. *J. Chem. Phys.* **1995**, *102*, 6145.

Molecular electronic polarization is also revealed by the computed dipole moments in MC-QM/MM simulations.^{14a,27e} Most systems exhibit a significant dipolar increase in water. In particular, the dipole moments for the nucleotide bases are predicted to be 39–75% greater than their gas-phase values, a finding in excellent agreement with results by Cramer and Truhlar, and by Orozco et al., among others.²⁷ In addition, through the application of combined AM1/MM potential in Monte Carlo simulations, molecular electrostatic potentials (MEP) for organic compounds have been evaluated in water.²⁹ The *average* MEP was then utilized to derive atomic charges. In the past, this procedure has been widely used to develop partial charges for MM force fields; however, previous calculations have used gas-phase MEP-derived charges for fluid simulations,^{6e} neglecting the solvent polarization effect. Our approach takes into account the solvent mean field through MC simulations, thereby producing atomic charges that include solvent environmental effects.²⁹ Recently, Marrone et al. used a similar procedure to determine partial charges of 18-crown-6 and K⁺ in methanol, although only a single solvent structure was used without ensemble averaging.³⁰

Conformational Equilibria in Solution

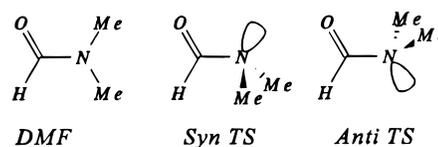
Computer simulation has a long history in studying solvent effects on conformational equilibria.³¹ This involves computation of free energy differences in solvation of various conformers using the FEP method. The first QM/MM application to conformational problems in solution was carried out in 1992 on the relative basicity of the *syn* and *anti* lone pairs of acetate ion in water.³² Recognizing that the basicity difference between the two lone pairs is related to the free energy difference of the *syn* and *anti* conformers of the conjugated acid, we computed the potential of mean force (pmf) for the hydroxy rotation in acetic acid in aqueous solution. The results showed a significant solvent stabilization of the *anti* structure by 4.8 kcal/mol,³² which largely offsets the gas-phase energy difference of 5.9 kcal/mol in favor of the *syn* conformer.³³ As a result, the p*K*_a difference between *syn*- and *anti*-acetic acid is reduced to less than 1 p*K*_a unit in water, in accord with Rebek and Zimmermann's experimental estimates,³⁴ and with Pranata's OPLS-type calculation.^{33c}



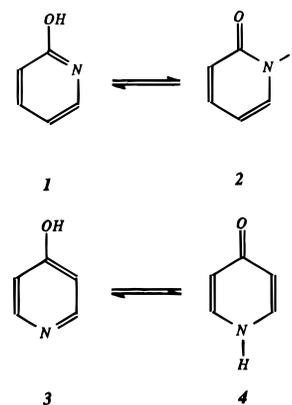
MC-QM/MM simulations were applied to the *cis*/*trans* isomerization about the tertiary peptide bond

- (29) Gao, J.; Luque, F. J.; Orozco, M. *J. Chem. Phys.* **1993**, *98*, 2975.
 (30) Marrone, T. J.; Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1994**, *98*, 1341.
 (31) See, for example, Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304.
 (32) (a) Gao, J.; Pavelites, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 1912.
 (b) Gao, J. *J. Phys. Chem.* **1992**, *96*, 6432.
 (33) (a) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1988**, *110*, 1872. (b) Li, Y.; Houk, K. N. *J. Am. Chem. Soc.* **1989**, *111*, 4505. (c) Pranata, J. *J. Comput. Chem.* **1993**, *14*, 685.
 (34) (a) Tadayoni, B. M.; Rebek, J., Jr. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 13. (b) Cramer, K. D.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1990**, *112*, 3680.

in dimethylformamide (DMF) in water, chloroform, and CCl₄.³⁵ An aqueous environment was found to increase the activation barrier of isomerization by ca. 1 kcal/mol, somewhat less than NMR experiments on analogous compounds and previous MC calculations.^{36,37} An interesting finding from this study, which has also emerged as a common feature in other processes, is that differential polarization between the ground and transition states plays a crucial role in the mechanism of solvation.^{9,35} For DMF, the planar ground state is more polarizable than the pyramidal transition state (TS), resulting in much greater polarization stabilization of the ground state over the TS. The polarization effect is accompanied by an increase in dipole moment by 1.53 D compared to those of the two TSs (0.4 and 1.0 D) in water.³⁵



In a more recent study, the MC-QM/MM method was used to study the tautomeric equilibria of 2- and 4-pyridone in water and chloroform.³⁸ In agreement with experimental findings, the gas-phase preference of the hydroxy tautomers is predicted to be dramatically reversed in water in favor of the keto forms by -4.6 ± 0.2 and -3.7 ± 0.2 kcal/mol for **2** and **4**, respectively. In the less polar solvent CHCl₃, on the other hand, **3** is still found to be the preferred form, but the two tautomers (**3** and **4**) exist in competitive amounts. Remarkably, electronic polarization is com-



puted to contribute 40–60% to the free energy changes, while the dipole moment of **4** is estimated to be 10.3 D in water, an increase of 4 D over the gas-phase value. The result compares well with Cramer and Truhlar's calculation (10.8 D) using a generalized Born solvation model.^{27b} Other studies employing combined QM/MM potentials include computation of cation- π interactions between tetramethylammonium ion and benzene in water,^{39a} conformational equilibria in dimethoxyethane,^{39b} host-guest interactions between

- (35) (a) Gao, J. *J. Am. Chem. Soc.* **1993**, *115*, 2930. (b) Gao, J. *Proc. Indian Acad. Sci., Chem. Sci.* **1994**, *106*, 507.
 (36) (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.
 (37) (a) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 7535. (b) Wiberg, K. B.; Rablen, P. R.; Keith, T. A. *J. Am. Chem. Soc.* **1995**, *117*, 4261.
 (38) Gao, J.; Shao, L. *J. Phys. Chem.* **1994**, *98*, 13772.

K⁺ and 18-crown-6,^{39c} and solvent effects on proton transfer potential surfaces.^{39d}

Chemical Reactions in Solution

Of great interest is simulation of chemical reactions involving changes in covalent bonds. The hybrid QM/MM method is impeccably suited for these studies.^{5,7-9} Our approach is similar to that used by Jorgensen, and involves determination of the free energy profile $\Delta G(\mathbf{R})$, or the potential of mean force (pmf), as a function of the reaction coordinate \mathbf{R} .⁴⁰ However, in hybrid QM/MM simulations, it is no longer necessary to specifically parametrize empirical potential functions for a new reaction.^{40b} Two components contribute to $\Delta G(\mathbf{R})$: the free energy change in the gas phase, $\Delta G_{\text{gas}}(\mathbf{R})$, and the free energy of solvation, $\Delta G_{\text{sol}}(\mathbf{R})$. The gas-phase properties can be obtained by high-level ab initio calculations using programs such as Gaussian 90 or GAMESS,²² which yield structural and energetic results along with the reaction path. MC-QM/MM simulations are used to produce changes in free energy of solvation via the Zwanzig perturbation formula:^{23,24}

$$\Delta G_{\text{sol}}(\mathbf{R} + \Delta \mathbf{R}) - \Delta G_{\text{sol}}(\mathbf{R}) = -kT \ln \langle e^{-[E(\mathbf{R} + \Delta \mathbf{R}) - E(\mathbf{R})]/kT} \rangle_{E(\mathbf{R})} \quad (7)$$

where $\Delta \mathbf{R}$ represents a small change in the reaction coordinate, typically about 0.1–0.15 Å for bond distances and about 5° for bond angles. The brackets $\langle \dots \rangle_{E(\mathbf{R})}$ indicate an ensemble average over the potential surface $E(\mathbf{R})$. The reaction path in solution can also be constructed using hybrid QM/MM potentials by evaluating the multidimensional free energy surface.^{41a}

Nucleophilic Substitution Reactions. The first reaction studied was the S_N2 Menshutkin reaction of H₃N + CH₃Cl → CH₃NH₃⁺ + Cl⁻, which represents a great challenge to the theory because of the large solvent effect associated with the charge separation.⁴¹ To assess the structural and energetic changes, a two-dimensional free energy surface was constructed through a total of 87 MC-QM/MM simulations using the combined AM1/TIP3P potential (Figure 4). The computed ΔG^\ddagger (26.3 ± 0.3 kcal/mol) is consistent with the experimental activation energy of 23.5 kcal/mol for the analogous reaction of NH₃ + CH₃I in water.⁴² An interesting observation from this study is the large change of the TS structure due to solvent effects (Figure 4), a result that is anticipated according to the Hammond postulate.⁴³ The MC simulation indicates that the TS of the Menshutkin reaction occurs much earlier in water than in the gas phase, with an increase of 0.30 Å in the C–N distance and a concomitant decrease of 0.15 Å in the C–Cl bond length. The gas-phase and aqueous reaction profiles are

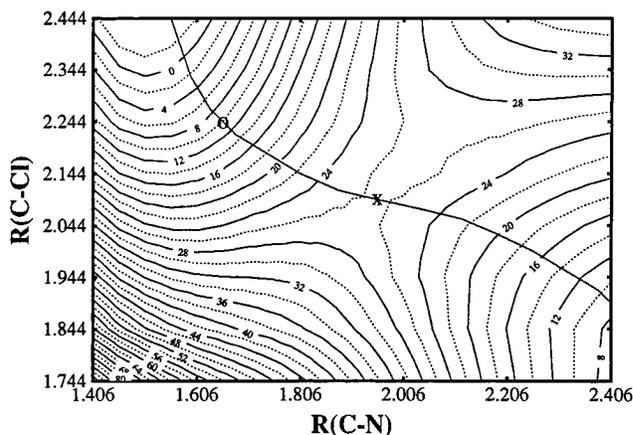


Figure 4. Computed free energy surface for the Menshutkin reaction of NH₃ + CH₃Cl → CH₃NH₃⁺ + Cl⁻ in water. Transition states in the gas phase and aqueous solution are marked by an O and an X, respectively. Energies are given in kilocalories per mole and distances in angstroms.

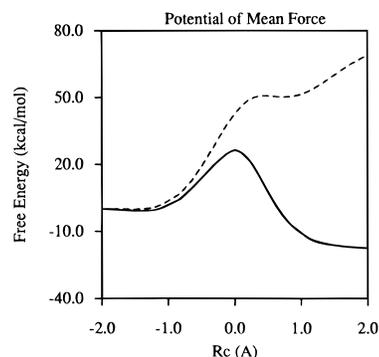


Figure 5. Potential of mean force for the Menshutkin reaction in water (solid curve) and in the gas phase (dashed curve). The reaction coordinate is indicated by the curve across the diagram in Figure 4.

compared in Figure 5, which demonstrates the striking solvent effect for this reaction.

Other QM/MM studies of nucleophilic substitution reactions include the S_N2 reaction of Cl⁻ + CH₃Cl in water with the use of combined AM1(MNDO)/TIP3P and EVB models,¹³ and the S_N1 ionization reaction (CH₃)₃CCl → (CH₃)₃C⁺ + Cl⁻ with the AM1/TIP3P potential.^{15c} Liu and Shi investigated the nucleophilic addition reaction between hydroxide ion and formaldehyde in water using a combined AM1/GROMOS method, and found changes in the reaction path in aqueous solution.^{15d}

Pericyclic Reactions. The next reaction to be investigated was the [3,3] sigmatropic Claisen rearrangement of allyl vinyl ether (AVE),⁴⁴ which was observed to have an intriguing aqueous rate acceleration.⁴⁵ MC-QM/MM calculations yield a TS stabilization ($\Delta\Delta G^\ddagger$) of -3.5 ± 0.2 kcal/mol, which translates to a 368-fold rate acceleration in water.⁴⁴ The results are in good accord with previous findings by Severance and Jorgensen from molecular mechanics MC simulations (-3.85 kcal/mol), and by Storer et al. employing their reaction specific solvation model, SM4-SRP (-2.3 kcal/mol).⁴⁶ Severance and Jorgensen attributed the solvent effect to increased hydrogen bonding at the transition state.^{46b} Consistent with this explanation, hybrid QM/MM simulations revealed additional information on the change of the charge distribution

(39) (a) Gao, J.; Chou, L. W.; Auerbach, A. *Biophys. J.* **1993**, *65*, 43. (b) Liu, H.; Muller-Plathe, F.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 1702. (c) Thompson, M. A.; Glendenning, E. D.; Feller, D. *J. Phys. Chem.* **1994**, *98*, 10465. (d) Gao, J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1993**, *27*, 491.

(40) (a) Jorgensen, W. L. *Adv. Chem. Phys.* **1988**, *70*, 469. (b) Gao, J.; Xia, X. In *Structure and Reactivity in Aqueous Solution: Characterization of Chemical Biological Systems*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium Series 568; American Chemical Society: Washington, DC, 1994; pp 212–228.

(41) (a) Gao, J.; Xia, X. *J. Am. Chem. Soc.* **1993**, *115*, 9667. (b) Gao, J. *J. Am. Chem. Soc.* **1991**, *113*, 7796. (c) Sola, M.; Lledos, A.; Duran, M.; Bertran, J.; Abboud, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 2873.

(42) Okamoto, K.; Fukui, S.; Nitta, I.; Shingu, H. *Bull. Chem. Soc. Jpn.* **1967**, *40*, 2354.

(43) Isaacs, N. S. *Physical Organic Chemistry*; Wiley: New York, 1987.

(44) Gao, J. *J. Am. Chem. Soc.* **1994**, *116*, 1563.

(45) White, W. N.; Wolfarth, E. F. *J. Org. Chem.* **1970**, *32*, 2196.

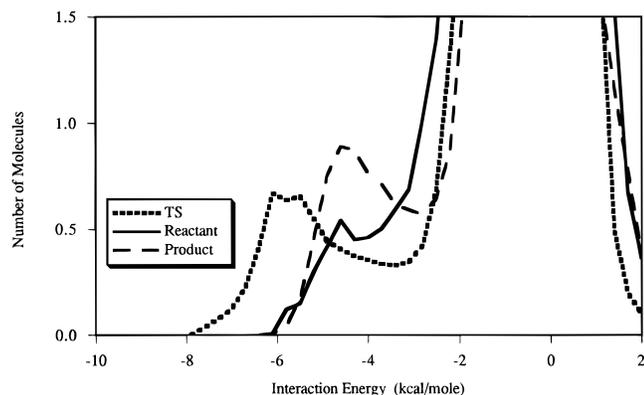
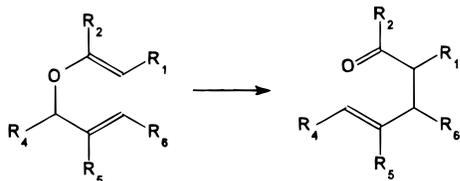


Figure 6. Pair energy distribution functions for the Diels–Alder reaction of cyclopentadiene and methyl vinyl ketone in water. Units for the ordinate are number of molecules per 1 kcal/mol.

during the reaction. The differential polarization effect between the TS and the reactant contributes 35% of the total $\Delta\Delta G^\ddagger$.^{44,46a} In addition, intrafragment polarization rather than inter-fragment charge separation between the allylic and oxallyl groups is found to be key to the computed solvent effects.⁴⁴

This work has been extended to substituted AVE systems.⁴⁷ Substitution of electron-withdrawing groups on the oxallyl unit and electron-donating substituents on the allylic fragment results in geometry changes



in transition structures, giving rise to a loose TS, enhanced ionic character, and more profound solvent effects.^{46c,47} Therefore, 2-cyano-6-methoxy-substituted allyl vinyl ether was predicted to have a synergistic effect with a solvent-induced rate increase of 6×10^5 times the rate in the gas phase, or 1700 times faster than the parent AVE in water.⁴⁷

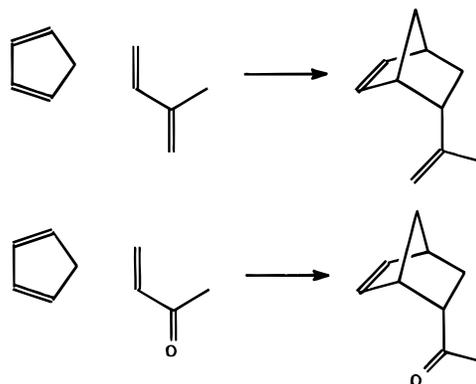
The Diels–Alder reactions of cyclopentadiene (CP) with methyl vinyl ether (MVE) and with isoprene were recently investigated using the MC–QM/MM method.^{48a} The computed aqueous stabilization of the TS is -3.5 ± 0.4 kcal/mol, in reasonable agreement with Breslow's experimental value (-3.85 kcal/mol),^{49a} and Jorgensen's MC prediction (-3.2 kcal/mol) using 6-31G-(d) CHELPG charges.^{48b} The enhanced hydrogen-bonding interaction at the TS predicted by Jorgensen is clearly revealed by the energy distribution functions shown in Figure 6. In the hydrophobic system (CP + isoprene), a greater solvent effect (-4.6 ± 0.3 kcal/

(46) (a) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1992**, *114*, 8794. (b) Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 10966. (c) Storer, J. W.; Giesen, D. J.; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G.; Liotard, D. A. In *Structure and Reactivity in Aqueous Solution*; Cramer, C. J., Truhlar, D. G., Eds.; ACS Symposium Series 568; American Chemical Society: Washington, DC, 1994; pp 24–49.

(47) Sehgal, A.; Shao, L.; Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 11337.

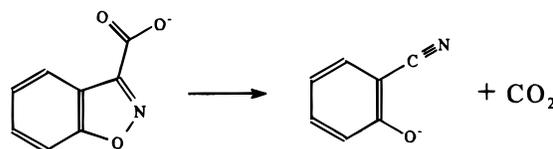
(48) (a) Furlani, T. R.; Gao, J. *J. Org. Chem.*, submitted for publication. (b) Jorgensen, W. L.; Blake, J. F.; Lim, D.; Severance, D. L. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1727.

(49) (a) Rideout, D. C.; Breslow, R. *J. Am. Chem. Soc.* **1980**, *102*, 7816. (b) Breslow, R.; Zhu, Z. *J. Am. Chem. Soc.* **1995**, *117*, 9923. (c) Engberts, J. B. F. N. *Pure Appl. Chem.* **1995**, *67*, 823. (d) Sangwan, N. K.; Schneider, H.-J. *J. Chem. Soc., Perkin Trans. 2* **1989**, 1223.



mol) is observed, and these results appear to be consistent with recent experiments on the dimerization of CP in water.^{49b} These findings indicate that hydrophobic effects play an important role in the rate enhancement of Diels–Alder reactions.

Other Reactions. The MC–QM/MM method was used to determine the pmf for the decarboxylation reaction of 3-carboxybenzisoxazole in water,⁵⁰ a reaction extensively studied by Kemp and co-workers for its gigantic solvent effects.⁵¹ The gas-phase reaction



path is determined at the HF/3-21G level, while the solvent effect is evaluated with the hybrid AM1/TIP3P model. An activation free energy ΔG^\ddagger of 26.1 ± 0.3 kcal/mol in water was computed, in agreement with Kemp's experimental value of 26.4 ± 1.5 kcal/mol.⁵¹ Solvent effects in other organic solutions were investigated by Houk and co-workers using OPLS-type potentials in MC simulations.⁵² The computed kinetic isotope effects are in agreement with experiment.^{51,52}

Making use of the gas-phase and aqueous wave functions obtained from MC–QM/MM simulations, we have calculated the change in electron density in the hydration process for the reactant and TS structures ($\Delta\rho = \rho_{\text{sol}} - \rho_{\text{gas}}$).^{14,50} The charge density analysis, which would not have been possible in molecular mechanics calculations, revealed different origins in the electronic polarization for the ground and transition states.⁵⁰ In the ground state, charge polarization occurs on the carboxylate group, whereas it takes place on the isoxazole oxygen and nitrogen at the TS.

Simulating Electronic Excited States and Photochemical Processes

An area that is virtually unexplored by computer simulations is photochemical reactions and the solvation of electronically excited states.⁵³ Yet, many chemical and biochemical events take place in the excited states. Following the pioneering work of Warshel,⁵⁴ progress has been made in our laboratory

(50) Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 8600.

(51) Kemp, D. S.; Paul, K. G. *J. Am. Chem. Soc.* **1975**, *97*, 7305.

(52) Zipse, H.; Apaydin, G.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 8608.

(53) (a) Blair, J. T.; Krogh-Jespersen, K.; Levy, R. M. *J. Am. Chem. Soc.* **1989**, *111*, 6948. (b) DeBolt, S. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 7515.

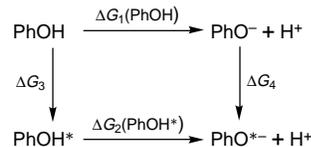
(54) Luzhkov, V.; Warshel, A. *J. Am. Chem. Soc.* **1991**, *113*, 4491.

and others.^{15b,55} We recently developed a combined QM configuration interaction and MM (QMCI/MM) method for simulating solvent effects on molecules in the excited states.⁵⁵ In this method, a CI wave function is used to describe the QM solute molecule, and to calculate ground and excited energies of the solute in solution.

The procedure was applied to examine the solvatochromic shifts of the $n \rightarrow \pi^*$ transition of acetone in water and organic solvents, including CH_3OH , CH_3CN , CHCl_3 , and CCl_4 .⁵⁵ Solvent spectral shifts have been widely used by chemists to probe solvent polarity and solute-solvent interactions.¹ Theoretical investigations of the solvent effect on electronic spectra were pioneered by McRae and Bayliss, and the early computational techniques have been described in several reviews.⁵⁶ Our approach differs from most previous studies in that the solvent molecules are explicitly represented in quantum mechanical calculations. Specifically, a 5-orbital/6-electron CI wave function that includes 100 configurations was used in MC-QM/MM simulations.⁵⁵ These computations were performed on the ground-state potential surface; thereby the excitation energy was obtained for the solute molecule surrounded by the ground-state solvent cage. Experimentally, the $S_0 \rightarrow S_1$ excitation was observed to be blue-shifted by 1560 and 570 cm^{-1} in water and methanol, and red-shifted by 490 cm^{-1} in CCl_4 , relative to that in the gas phase.^{56a} For comparison, MC-QMCI/MM calculations yielded blue shifts of 1694, 763, and 5 cm^{-1} , respectively, in these three solvents.⁵⁵ The large blue shifts of the $n \rightarrow \pi^*$ transition in polar solvents are due to hydrogen-bonding interactions that stabilize the ground state more than the excited state owing to a decrease in molecular dipole moments upon photoexcitation. The red shift in nonpolar solvents results from a change in dispersion interactions. Such interactions, however, were not included in our initial calculation, and represent an area for future investigation. Recently, Rosch and Zerner described a procedure that can treat the dispersion effects in these calculations.⁵⁷ Further, Thompson and Schenter have used a hybrid QM/MM model with MM polarization terms to study the excited states in the photosynthetic reaction center.^{15b} Studies of the spectral shifts in retinal and protonated Schiff bases in solution and in bacteriorhodopsin are currently in progress in our laboratory.

In another study, the MC-QMCI/MM method was used to determine the ground- and excited-state $\text{p}K_a$

difference of phenol in water.⁵⁸ In the thermodynamic cycle below, the ground- and excited-state $\text{p}K_a$ difference for phenol, or the free energy difference of the two horizontal equations, is related to the difference between the two vertical transformations (eq 8). High-



$$2.303RT\Delta\text{p}K_a = \Delta G_1(\text{PhOH}) - \Delta G_2(\text{PhOH}^*) = \Delta G_3 - \Delta G_4 \quad (8)$$

level QM calculations can be used to determine the geometry and energy difference between the ground- and excited-state species in the gas phase, while the solvation free energies are determined by FEP calculations. Our MC-QM/MM simulations yield a predicted $\Delta\text{p}K_a(\text{PhOH} \rightarrow \text{PhOH}^*)$ of $-8.6 \text{ p}K_a$ units,⁵⁸ on the basis of gas-phase excitation energies determined at the CASPT2 level of theory,⁵⁹ and solvation free energies from a series of Monte Carlo QMCI simulations that involved 100 configurations. This gives rise to a $\text{p}K_a$ of 1.4 for the singlet excited PhOH^* . For comparison, the experimental estimate of the acidity change is about $-6 \text{ p}K_a$ units ($\text{p}K_a(\text{PhOH}) = 10.0$, $\text{p}K_a(\text{PhOH}^*) = 3.6-4.1$) using the Förster cycle.⁶⁰ In view of the large experimental uncertainties in determining the excited-state $\text{p}K_a$, theoretical calculations provide an alternative approach to the investigation of solvation of ground- and excited-state molecules.

Concluding Remarks

The ability to model organic reactions in solution has greatly increased in the past decade. Previous studies of chemical reactions have primarily utilized molecular mechanics force fields, which involve laborious and difficult parametrization. The methods and examples presented here, which combine quantum mechanics and molecular mechanics in statistical mechanics simulations, provide an alternative avenue to study organic processes in solution. Hybrid QM/MM simulations offer the opportunity to obtain valuable insights into the solvent effects on solute electronic structure and molecular charge polarization. These properties would be otherwise difficult to compute using empirical MM force fields. Although semiempirical methods have so far been used in most applications, the procedure can easily be extended to the more systematic *ab initio* method, and new computational methodologies and applications will undoubtedly continue to emerge.

The author gratefully acknowledges the efforts of many co-workers listed in the references for their assistance and the financial support of the National Science Foundation, the National Institutes of Health, and the Environmental Protection Agency.

AR950140R

(55) Gao, J. *J. Am. Chem. Soc.* **1994**, *116*, 9324.

(56) (a) Bayliss, N. S.; McRae, E. G. *J. Phys. Chem.* **1954**, *58*, 1002. (b) Liptay, W. In *Modern Quantum Chemistry*; Sinanoglu, O., Ed.; Academic Press: New York, 1966; Part II, pp 173-198. (c) Amos, A. T.; Burrows, B. L. *Adv. Quantum Chem.* **1973**, *7*, 289.

(57) Rosch, N.; Zerner, M. C. *J. Phys. Chem.* **1994**, *98*, 5817.

(58) Gao, J.; Li, N.; Freindorf, M. *J. Am. Chem. Soc.*, in press.

(59) Widmark, P.-O.; Malmqvist, P.-A.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.

(60) (a) Förster, T. *Z. Electrochem.* **1950**, *54*, 42. (b) Ireland, J. F.; Wyatt, P. A. H. *Prog. Phys. Org. Chem.* **1976**, *12*, 131.