An Automated Procedure for Simulating Chemical Reactions in Solution. Application to the Decarboxylation of 3-Carboxybenzisoxazole in Water

Jiali Gao

Contribution from the Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14214

Received July 5, 1994[⊗]

Abstract: A combined Monte Carlo quantum mechanical and molecular mechanical (QM/MM) simulation method is described for the investigation of the solvent effects on chemical reactions. In the present approach, ab initio molecular orbital calculations are first used to locate the transition state, from which the reaction path is determined by using Gaussian 90. Then, free energy changes between adjacent structures generated along this intrinsic reaction path are evaluated via statistical perturbation theory using the combined QM/MM-AM1/TIP3P potential. Since empirical parametrization of the reaction system is not needed in these calculations, the method presented here is essentially an automated procedure for simulating reactions in solution, which may be conveniently used by organic chemists. We have employed the procedure to examine the decarboxylation of 3-carboxybenzisoxazole in aqueous solution. The predicted free energy of activation is 26.1 ± 0.3 kcal/mol, in excellent agreement with the experimental value of 26.3 kcal/mol. Analyses of the contributing factors in solute-solvent interaction suggest that the aqueous solvent effect is primarily due to the difference in the intrinsic (in vacuo) charge distributions between the reactant and transition state. Solvent polarization contributes significantly to the solute-solvent interaction; however, the nature of the electronic polarization of the reactant and the transition state is markedly different.

Introduction

Computer simulations can provide valuable insights into the structure and dynamics of chemical reactions in solution.[1] Significant progress has been made over the past few years. In these studies, the computational procedure typically involves three major steps:[2] (1) determination of the minimum energy path (MEP) for the reaction in the gas phase, (2) development of intermolecular potential functions for solute-solvent and solvent-solvent interactions, and (3) computation of the potential of mean force (pmf) and trajectories of the reaction in solution by Monte Carlo or molecular dynamics simulations. Although the procedure appears to be straightforward, a major difficulty is the parametrization of empirical potential functions that can properly describe intermolecular interactions in solution along the entire reaction path. The problem is further complicated by specific consideration of the solvent polarization effect,[3] which is often treated as an average effect,[4] and by the lack of a general approach to chemical processes involving bond forming and breaking.[1c] Consequently, computer simulation of chemical reactions in solution has been limited to only a few well-defined model systems.[5-7]

To overcome these difficulties, a combined quantum mechanics and molecular mechanics (QM/MM) approach has been proposed.[8-10] In this method, part of the system, for example, the reactant molecules in a solution or the active site region of an enzyme, is treated explicitly by quantum mechanical methods during the fluid simulation. At the same time, the "environmental" solvent molecules are approximated by empirical models. Thus, it is no longer necessary to parametrize potential functions for the solute molecules in such a combined QM/MM method. Clearly, ab initio molecular orbital (MO) theory would be ideal for the electronic structure calculations because well-established methods are available.[11] However, these calculations are too demanding in computing resources to be of practical use for large systems of biological interest. Consequently, attention has been directed to semiempirical molecular orbital methods, particularly those developed by Dewar et al. and by Stewart.[12,13] Our recent studies demon-

[⊗] Abstract published in *Advance ACS Abstracts*, March 15, 1995.

(1) (a) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, 1991. (b) Karplus, M.; Petsko, G. A. *Nature* **1990**, *347*, 631. (c) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987. (d) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992.

(2) Jorgensen, W. L. *Adv. Chem. Phys.* **1988**, *70*, 469.

(3) (a) Caldwell, J.; Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 9144. (b) van Belle, D.; Couplet, I.; Prevost, M.; Wodak, S. J. *J. Mol. Biol.* **1987**, *198*, 721. (c) Barnes, P.; Finney, J. L.; Nicholas, J. D.; Quinn, J. E. *Nature* **1979**, *282*, 459. (d) Warshel, A.; LeVitt, M. *J. Mol. Biol.*, **1976**, *103*, 227.

(4) (a) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657. (b) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. (c) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187. (d) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2978.

(5) For recent applications, see: (a) Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 10966. (b) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 7430. (c) Zheng, Y.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 2733.

(6) Gao, J. *J. Am. Chem. Soc.* **1991**, *113*, 7796.

(7) Kim, H. J.; Hynes, J. T. *J. Am. Chem. Soc.* **1992**, *114*, 10508.

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

(9) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199.

(10) For a recent review, see: (a) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; 1995, Vol. 7, in press. Methods and applications of combined quantum mechanical and molecular mechanical potentials: (b) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537.

(11) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(12) (a) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902. (b) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

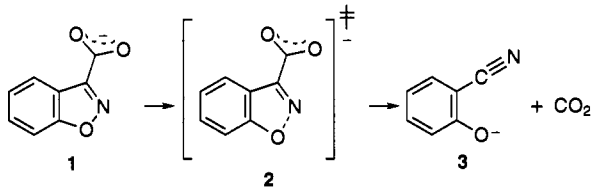
(13) (a) Stewart, J. J. P. *J. Computer-Aided Mol. Design* **1990**, *4*, 1. (b) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209, 221.

0002-7863/95/1517-8600$09.00/0 © 1995 American Chemical Society

strated that the combined AM1/TIP3P potential can yield excellent results for the free energies of hydration of organic compounds and can provide valuable insights into the solvent polarization effect, which is not available using the empirical potential function.[14] A shortcoming in the semiempirical MO calculations is the difficulty in systematically improving the structural and energetic results for certain reactions in the gas phase.

In this paper, we report the results from a computer simulation of the decarboxylation reaction of 3-carboxybenzisoxazole in water following a convenient procedure with the combined QM/MM-AM1/TIP3P potential.[8,10] In this method, ab initio molecular orbital calculations are first used to locate the transition state, from which the reaction path is determined by using GAUSSIAN 90.[15] Then, the structures generated along this intrinsic reaction coordinate (IRC) are used in Monte Carlo simulations, treating them quantum mechanically by the semiempirical AM1 theory.[8-10,12] Adjacent frames are perturbed to yield the change in free energy of solvation $\Delta\Delta G_{sol}$ via statistical perturbation theory.[16] This in essence is similar to the method employed previously by Jorgensen and co-workers.[2,5] However, the major difference and advantages are in the second, most crucial step in their approach, which is eliminated here because solute–solvent interactions are evaluated by quantum mechanical calculations during the fluid simulation.[8-10,14] Therefore, the present method is essentially an automated procedure for simulation of chemical reactions in solution.

We chose to study the decarboxylation of 3-carboxybenzisoxazole to carbon dioxide and salicylonitrile (**1** → **3**) as an application of this method because large solvent effects on the reaction rate have been observed.[17,18] In fact, in their elegant and thorough investigation of the decarboxylation reactions of 3-carboxybenzisoxazole and its derivatives in various soltents, Kemp and co-workers discovered that the rate can increase by 10^8-fold on going from aqueous solution to hexamethylphosphoramide (HMPA).[17a]



1 **2** **3**

From the kinetic data, these authors attributed the dramatic rate change to the balancing effect of the ground-state destabilization and the transition-state stabilization due to delocalization of the negative charges.[17] Furthermore, these authors concluded that there is little change in the transition state structure in different solvents in spite of the dramatic solvent-promoted rate acceleration.[17b] This has been confirmed by recent measurement of the carbon isotope effects for the decarboxylation in aqueous and dioxane solutions as well as in a catalytic antibody.[19] The mechanism of the reaction is also of considerable biological importance, for example, in the decarboxylation of pyruvic acid to acetyl coenzyme A through thiamine pyrophosphate cofactor in the citric acid cycle.[20] The present study, along with the work reported by Zipse and Houk,[21] provides the theoretical insight into the solvent effects on the decarboxylation reaction of 3-carboxybenzisoxazole in aqueous solution at the atomic level.

The present paper is organized as follows. A brief review of the theory and its implementation in the Monte Carlo simulation is given in the methodology section. This is followed by computational details. Results and discussion are presented next, followed by conclusions.

Methodology

(a) The Combined QM/MM Monte Carlo Simulation Method. Following the procedure described previously,[8-10] the reactant molecules (solutes) are represented by nuclei and electrons using the Hartree–Fock wave function Φ, which is written as a single Slater determinant of all occupied molecular orbitals (MO).[11] The MO's are linear combinations of an atomic orbital basis set, $\{\psi_i\}$. The surrounding solvent molecules from the MM region, whose atoms are approximated as interaction sites, and the interaction energies are evaluated by empirical (molecular mechanics) potential functions.[4] The total effective Hamiltonian of the system within the Born–Oppenheimer approximation is given by:

$$\hat{H}_{eff} = \hat{H}^o + \hat{H}_{qm/mm} + \hat{H}_{mm} \tag{1}$$

where $\hat{H}^o$ is the Hamiltonian for the reactant, $\hat{H}_{mm}$ is the solvent MM energy, and $\hat{H}_{qm/mm}$ is the solute–solvent (QM/MM) interaction Hamiltonian, which consists of electrostatic, $\hat{H}^{el}_{qm/mm}$, and van der Waals, E^{vdW}_{sx}, terms (eq 2).

$$\hat{H}_{qm/mm} = \hat{H}^{el}_{qm/mm} + E^{vdW}_{sx} \tag{2}$$

The van der Waals term (in Lennard-Jones form) in eq 2 contains the only adjustable parameters (σ_m and ϵ_m) in the present approach whose values have been determined previously.[14,22]

The total potential energy in the combined QM/MM force field is computed using eq 3,

$$E_{tot} = \langle\Phi|\hat{H}_{eff}|\Phi\rangle = E_{qm} + E_{qm/mm} + E^{vdW}_{sx} + E_{mm} \tag{3}$$

where Φ is the wave function of the solute molecule in solution. In eq 3, E_{qm} and $E_{qm/mm}$, which involve electronic positions, are determined through MO Hartree–Fock self-consistent-field (HF-SCF) calculations.[11] The quantum mechanical energies may be expressed in terms of the solute energy in the gas phase (E_{gas}) and the solute–solvent electrostatic interaction energy (E^{el}_{sx}):

$$E_{qm} + E_{qm/mm} = E_{qm}(gas) + E^{el}_{sx} \tag{4}$$

The second term in eq 4 can be further decomposed into contributions due to the solute "permanent" charge distribution (that in the gas phase), $E^{(1)}$, and solvent polarization, E_{pol}, where $E^{(1)} = \langle\Phi^o|\hat{H}^{el}_{qm/mm}|\Phi^o\rangle$, and Φ^o is the wave function of the solute in vacuo.[14]

The idea of using a combined QM/MM treatment of a molecular system is not new and has been extensively exploited.[8-10,23-31] Parallel to the explicit treatment of the solvent, continuum representations have

(14) Gao, J.; Xia, X. *Science* **1992**, *258*, 631.

(15) 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, PA, 1990.

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

(17) (a) Kemp, D. S.; Paul, K. G. *J. Am. Chem. Soc.* **1975**, *97*, 7305. (b) Kemp, D. S.; Cox, D. D.; Paul, K. G. *J. Am. Chem. Soc.* **1975**, *97*, 7312. (c) Casey, M. L.; Kemp, D. S.; Paul, K. G. *J. Org. Chem.* **1973**, *38*, 2294. (d) Kemp, D. S.; Paul, K. G. *J. Am. Chem. Soc.* **1970**, *92*, 2553.

(18) (a) Bunton, C. A.; Minch, M.; Sepulveda, W. *J. Phys. Chem.* **1971**, *75*, 2707. (b) Shirai, M.; Smid, J. *J. Am. Chem. Soc.* **1980**, *102*, 2863. (c) Kunitake, T.; Shinkai, S.; Klotz, I. M. *J. Org. Chem.* **1977**, *42*, 306. (d) Straub, T. S.; Bender, M. L. *J. Am. Chem. Soc.* **1972**, *102*, 7877.

(19) (a) Lewis, C.; Paneth, P.; O'Leary, M. H.; Hilvert, D. *J. Am. Chem. Soc.* **1993**, *115*, 1410. (b) Grate, J. W.; McGill, R. A.; Hilvert, D. *J. Am. Chem. Soc.* **1993**, *115*, 8577. (c) Lewis, C. T.; Kramer, T.; Robinson, S.; Hilvert, D. *Science* **1991**, *253*, 1019.

(20) Breslow, R. *J. Am. Chem. Soc.* **1958**, *80*, 3719.

(21) Zipse, H.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 8608–8617.

(22) Floris, F.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 616.

(23) Bernal, J. D.; Fowler, F. D. *J. Chem. Phys.* **1933**, *1*, 515.

(24) Warshel, A.; Russel, S. *Q. Rev. Biophys.* **1984**, *17*, 283.

(25) Warshel, A. *Computer Modeling of Chemical Reactions in Proteins and Solutions*; Wiley: New York, 1991.

been introduced into MO calculations.[32,33] Good agreement between computed and experimental solvation energies has been obtained.[32-34] However, in some cases, consideration of the specific solute−solvent interactions is needed.[35] Interested readers are directed to ref 32 for technical developments and recent applications in solution chemistry.

(b) Implementation. We have developed a computer program MCQUB (Monte Carlo QM/MM at the University of Buffalo), which treats solute−solvent interactions with the combined QM/MM potential.[36] QM energies are evaluated with the MOPAC program by Stewart.[37-39] Statistical mechanical Monte Carlo simulations were carried out with the BOSS program developed by Jorgensen,[40] in which the QM/MM potential has been incorporated. The implementation takes advantage of the facilities available in BOSS and MOPAC programs such that reaction profiles in solution can easily be computed.

In general, the QM energy for each new configuration requires a full HF-SCF calculation, although this is not necessary in practice because QM/MM interactions are truncated to save computer time by using a spherical cutoff R_{cut}. Consequently, solvent moves that are outside of the cutoff radius from the QM atoms will not affect the solute wave function and the QM energies, but the total energy of the system will of course be affected due to the MM term. Further, the density matrix for the QM molecule is only significantly influenced by the solvent molecules within a relatively small range,[41] which would require complete HF-SCF calculations. In the present simulation, a full HF-SCF calculation is iterated whenever there is a solute move, a solvent move that is within 5.5 Å of any QM atom, or 20 successive distant solvent moves. Note that this 5.5 Å limit is only used to determine the *frequency* of the SCF calculations; the QM energies are still computed using the standard R_{cut}. On average, one fifth of the total configurations sampled in the Monte Carlo simulation involve full HF-SCF calculations in this approach.

(c) The Reaction Path. The combined QM/MM potential has the advantage of exploring the potential energy *surface* rather than a profile for a chemical reaction in solution. Indeed, using the combined QM/MM-AM1/TIP3P model, we have recently constructed a two-dimensional free energy surface for the type II S_N2 reaction of $NH_3 + CH_3Cl \rightarrow CH_3NH_3^+ + Cl^-$ in aqueous solution to investigate the structural change of the transition state due to solvation.[42] Nonetheless, most often organic chemists are interested in the free energy profile along a predefined reaction path to discuss the reaction mechanism.

In the present study, we choose to use the intrinsic reaction coordinate (IRC) determined with GAUSSIAN 90 to approximate the reaction path and energy for the decarboxylation reaction of 3-carboxybenzisoxazole in the gas phase.[43,44] The change in free energy of hydration along this IRC is then evaluated via statistical perturbation theory in the combined QM/MM Monte Carlo simulation by perturbing the system of adjacent structures generated in the reaction path by using GAUSSIAN 90. The potential of mean force (pmf) in aqueous solution is obtained as a sum of the gas phase and solvation free energies.

It should be mentioned that we have used alternative procedures to construct the potential of mean force for chemical transformations in solution with the combined QM/MM-AM1/TIP3P model.[29,30] In a study of the relative basicity of the carboxylate syn and anti lone pairs in water, the pmf for the transformation of acetic acid from its syn to anti conformation was determined based on the AM1 gas-phase energy along the AM1 minimum energy path in Monte Carlo simulations.[29] Here, the AM1 results are in excellent agreement with experimental and high-level ab initio data. In the case of the cis/trans isomerization of the peptide bond in dimethylformamide, energetic results from AM1 calculations were found to be inadequate; however, the AM1 geometries along the MEP are in good accord with those predicted at the ab initio 6-31G(d) level.[30] Therefore, the AM1 energy was scaled to reproduce the activation free energy of isomerization determined at the MP4-(SDTQ)/6-31G(d) level, while the AM1 geometries were adopted in the evaluation of solvation free energies using the combined QM/MM potential.

Computational Details

The transition state (TS) and IRC path for the decarboxylation reaction of 3-carboxybenzisoxazole in the gas phase are determined by using GAUSSIAN 90.[15] All computations including vibrational frequencies are carried out at the 3-21G level.[11] A total of 39 structures are generated along the IRC, spanning the reaction coordinate r_c (for the convenience of later discussion, r_c is represented as the distance of the breaking $C_3-C_{O_2}$ bond) from 1.56 to 3.30 Å. Vibrational frequency calculations confirm that the TS and stationary structures have one and zero imaginary frequencies, respectively. For structures along the reaction path, vibrational frequencies are computed by using a reaction path Hamiltonian to project out translational and rotational degrees of freedom as well as one corresponding to the reaction path before diagonalization of the force constant matrix.[44] In computing the thermodynamic properties, the frequencies are scaled by a factor of 0.9 for evaluation of the zero-point energies, their thermal corrections, and the vibrational entropies.[11] Frequencies below 550 cm^{-1} are treated as classical rotations in the zero-point energy calculation, while the imaginary frequency is ignored in all computations.

(26) (a) Weiner, S. J.; Seibel, G. L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 649. (b) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (c) Rullmann, J. A. C.; Bellido, M. N.; van Duijnen, P. T. *J. Mol. Biol.* **1989**, *206*, 101. (d) Waszkowycz, B.; Hillier, I. H.; Gensmantel, N.; Payling, D. W. *J. Chem. Soc., Perkin Trans. 2* **1991**, 225. (e) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826.

(27) Bash, P. A.; Field, M. J.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8192.

(28) (a) Hwang, J.-K.; King, G.; Greighton, S.; Warshel, A. *J. Am. Chem. Soc.* **1988**, *110*, 5297. (b) Yadav, A.; Jackson, R. M.; Holbrook, J. J.; Warshel, A. *J. Am. Chem. Soc.* **1991**, *113*, 4800.

(29) Gao, J.; Pavelites, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 1912.

(30) Gao, J. *J. Am. Chem. Soc.* **1993**, *115*, 2930.

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

(32) For recent thorough reviews of this subject, see: (a) Cramer, C. J.; Truhlar, D. G. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; 1994, Vol. 6, pp 1−72. (b) Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Design* **1992**, *6*, 629. (c) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(33) (a) Tapia, O.; Goscinski, O. *Mol. Phys.* **1975**, *29*, 1653. (b) Karelson, M. M.; Zerner, M. *J. Phys. Chem.* **1992**, *96*, 6950. (c) Ford, G. P.; Wang, B. *J. Am. Chem. Soc.* **1992**, *114*, 10563. (d) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776. (e) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (f) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239. (g) Negre, M.; Orozco, M.; Luque, F. J. *Chem. Phys. Lett.* **1992**, *196*, 27. (h) Pascual-Ahuir, J. L.; Silla, D. P.; Tomasi, J. *J. Comput. Chem.* **1987**, *8*, 778. (i) Rivail, J. L.; Terryn, B. *J. Chim. Phys. Chim. Biol.* **1982**, *79*, 2. (j) Rinaldi, D.; Rivail, J. L.; Rguini, N. *J. Comput. Chem.* **1992**, *13*, 675. (k) Dillet, V.; Rinaldi, D.; Angyan, J. G. *Chem. Phys. Lett.* **1993**, *202*, 18. (l) Dillet, V.; Rinaldi, D.; Rivail, J.-L. *J. Phys. Chem.* **1994**, *98*, 5034. (m) Young, P.; Green, D. V. S.; Hillier, I. H.; Burton, N. A. *Mol. Phys.* **1993**, *80*, 503. (n) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305. (o) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213. (p) De Vries, A. H.; van Duijnen, P. Th.; Juffer, A. H.; Rullmann, J. A. C.; Dijkman, J. P.; Merenga, H.; Thole, B. T. *J. Comput. Chem.* **1995**, *16*, 37.

(34) (a) Chipot, C.; Gorb, L. G.; Rivail, J.-L. *J. Phys. Chem.* **1994**, *98*, 1601. (b) Hall, R. J.; Burton, N. A.; Hillier, I. H.; Young, P. E. *Chem. Phys. Lett.* **1994**, *220*, 129. (c) Sola, M.; Lledos, A.; Duran, M.; Bertran, J.; Abboud, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 2873. (d) Basilevsky, M. V.; Chudinov, G. E.; Napolov, D. V. *J. Phys. Chem.* **1993**, *97*, 3270. (e) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1992**, *114*, 8794.

(35) Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; VCH: Weinheim, Germany, 1990.

(36) Gao, J. *MCQUB (Monte Carlo QM/MM at the University at Buffalo)*; SUNY at Buffalo: Buffalo, NY, 1994.

(37) Stewart, J. J. P. *MOPAC, Version 5. QCPE* **1986**, *6*, 455, No. 391.

(38) Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.

(39) Dewar, M. J. S.; Thiel, W. *Theor. Chim. Acta* **1977**, *46*, 89.

(40) Jorgensen, W. L. *BOSS, Version 2.9*; Yale University: New Haven, CT, 1990.

(41) Gao, J. Unpublished results.

(42) (a) Gao, J.; Xia, X. *J. Am. Chem. Soc.* **1993**, *115*, 9667. (b) Xia, X. Ph.D. Thesis; SUNY: Buffalo, NY, 1993.

(43) (a) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523. (b) Fukui, K. *Acc. Chem. Res.* **1981**, *14*, 363.

(44) (a) Miller, W. H.; Handy, N. C.; Adams, J. E. *J. Chem. Phys.* **1980**, *72*, 99. (b) Isaacson, A. D.; Truhlar, D. G. *J. Chem. Phys.* **1982**, *76*, 1380. (c) Page, M.; Doubleday, C.; McIver, J. W., Jr. *J. Chem. Phys.* **1990**, *93*, 5634.

(45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

Table 1. Thermodynamic Results for the Decarboxylation of 3-Carboxybenzisoxazole in the Gas Phase at 298 K[a]

	reactant → transition state	reactant → product
ΔE^{o}(3-21G)	19.3	−34.5
ΔE_v^{o}	−2.0	−1.9
$\Delta\Delta E_v^{298}$	0.1	0.1
ΔH^{298}	17.4	−34.6
ΔS	6.4	40.4
ΔG^{298}	15.4	−46.6
ΔE^{o}(MP2/6-31+G(d))[b]	18.5	−27.3
ΔG^{298}(MP2/6-31+G(d))	14.6	−39.7

[a] Energies and entropies are given in kcal/mol and cal/(mol·K) [eu]. [b] Reference 21.

Statistical mechanical Monte Carlo calculations are carried out in the isothermal−isobaric (NPT) ensemble at 25 °C and 1 atm using a combined AM1/TIP3P potential in which the solute (3-carboxybenzisoxazole) is treated quantum mechanically.[36,40] A rectangular box containing 390 water molecules (ca. 20 × 20 × 30 Å^3) is used in the fluid simulations. Seventeen structures, roughly with an interval of 0.1 Å apart, are selected from the ab initio IRC calculations, and the differences in hydration free energy are evaluated via statistical perturbation theory. In these simulations, the $C_3-C_{O_2}$ bond that is breaking along the reaction coordinate is oriented to coincide with the z-axis of the water box. The standard Metropolis sampling procedure is adopted along with the Owicki−Scheraga preferential sampling technique using $1/(r^2 + c)$ weighting, where c = 150 Å^2, to facilitate the statistics near the solute molecule.[46] At least 5 × 10^5 configurations are taken in the equilibration stage, and an additional 1.5 × 10^6 configurations are collected for averaging. The intermolecular interactions are feathered to zero between spherical cutoff distances (R_{cut}) of 8.5 to 9.0 Å for water−water interactions based upon the oxygen separations. Solute−solvent interactions are truncated between 9.0 and 9.5 Å based on solute atom−water oxygen distances. New configurations are generated by randomly selecting a molecule, translating it in all three Cartesian directions, and rotating it along a randomly chosen axis. An acceptance rate of about 40% is maintained by using ranges of ±0.11 Å and 10° for molecular motions. Volume changes are restricted to within ±190 Å^3 at every 2375 configurations.

Results and Discussion

(a) The Gas-Phase Reaction Profile. The decarboxylation reaction in the absence of medium effects was examined with ab initio calculations at the 3-21G level. In all calculations, a planar geometry with C_s symmetry was used. The transition state was first located and used as the starting point to carry out the reaction path following calculations. Vibrational frequencies obtained with the 3-21G basis set confirmed the assignment of the TS and stationary points and were employed to determine the thermochemical properties. The relative energies of the reactant, transition state, and products are listed in Table 1 and provide a good calibration of the theoretical model. As expected previously,[17] the reaction occurs via a concerted mechanism with the decarboxylation and ring-opening processes taking place simultaneously. The TS has an energy of 19.3 kcal/mol above the ground state, and the energy of reaction is −34.3 kcal/mol at the 3-21G level. Zipse and Houk[21] have thoroughly investigated the basis set and correlation effects on the computed energies.[21] They found that a large basis set such as 6-31+G(d) increases the barrier height by about 7 kcal/mol, while inclusion of the electron-correlation effect at the MP2/6-31+G(d) level lowers the energy by about 8 kcal/mol. These two opposing factors almost cancel out, leading to a good

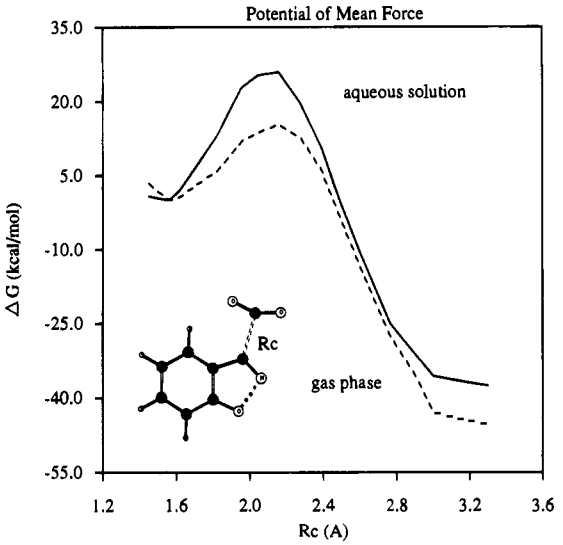


Figure 1. Calculated free energy profile in the gas phase (dashed curve) and the potential of mean force in water (solid curve) for the decarboxylation reaction of 3-carboxybenzisoxazole. The reaction coordinate is represented for convenience as the $C_3-C_{O_2}$ distance along the 3-21G IRC.

accord between the 3-21G and the MP2/6-31+G(d) energies. The energy obtained at the MP2/6-31G(d)//6-31G(d) level is 19.1 kcal/mol.[21] The good performance of the 3-21G basis set, due to fortuitous cancellation of errors, suggests that it may be able to provide a uniformly good description of the reaction profile (Figure 1), and it is used throughout this study.

The thermodynamic quantities are computed using the 3-21G vibrational frequencies scaled by a factor of 0.9. The resulting activation enthalpy, $\Delta H^{\ddagger}$, and free energy, $\Delta G^{\ddagger}$, are 17.4 and 15.4 kcal/mol, respectively. Gas-phase experimental data are not available for comparison, although $\Delta H^{\ddagger}$ values of 23−32 kcal/mol in various solvents ranging from HMPA to water have been reported by Kemp and Paul.[17a] In view of the large solvent effect involved in this reaction, the theoretical results appear to be reasonable. The computed entropy of activation $\Delta S^{\ddagger}$ (6.4 cal/(mol·K)) is much smaller than the experimental value in solution (18−20 ± 4 cal/(mol·K)).[17a] The large difference might be due to the solvent effect. In the case of 3-carboxy-4-hydroxybenzisoxazole, where the reaction rate is not affected by change of solvents due to the presence of an intramolecular hydrogen bond, the observed $\Delta S^{\ddagger}$ is 4 cal/(mol·K), in close agreement with the present study.[17b] Finally, the calculated enthalpy of reaction for the conversion **1** → **3** is −34.6 kcal/mol.

The calculated geometries at the 3-21G level for the ground state (**1**) and the TS (**2**) are shown in Figure 2. The most notable structural changes between the reactant and the TS are elongation of both the $C_3-C_{O_2}$ and O−N bond distances from 1.56 and 1.49 Å to 2.16 and 1.84 Å, respectively. This is accompanied by a widening of the carboxylate angle from 131.5° to 151.1°, and the C_8-C_3-N angle from 112.1° to 117.8°. Zipse and Houk found that geometry optimization using a larger basis set at the 6-31G(d) level result in a transition state with a longer $C_3-C_{O_2}$ distance and a shorter separation between N and O in the isoxazole ring, though the 3-21G and 6-31G(d) structures for the reactant and products are in good agreement.[21] Zipse and Houk also optimized the TS structure with one water molecule hydrogen bonded to the carboxylate group. Although the $C_3-C_{O_2}$ distance is shortened and the N−O distance is lengthened, surprisingly, they found that the 3-21G and 6-31G(d) geometries are in reasonable agreement.[21] This suggests

(46) (a) Owicki, J. C.; Scheraga, H. A. *Chem. Phys. Lett.* **1977**, *47*, 600. (b) Owicki, J. C. *ACS Symp. Ser.* **1978**, *No. 86*, 159. (c) Jorgensen, W. L.; Bigot, B.; Chandrasekhar, J. *J. Am. Chem. Soc.* **1982**, *104*, 4584.

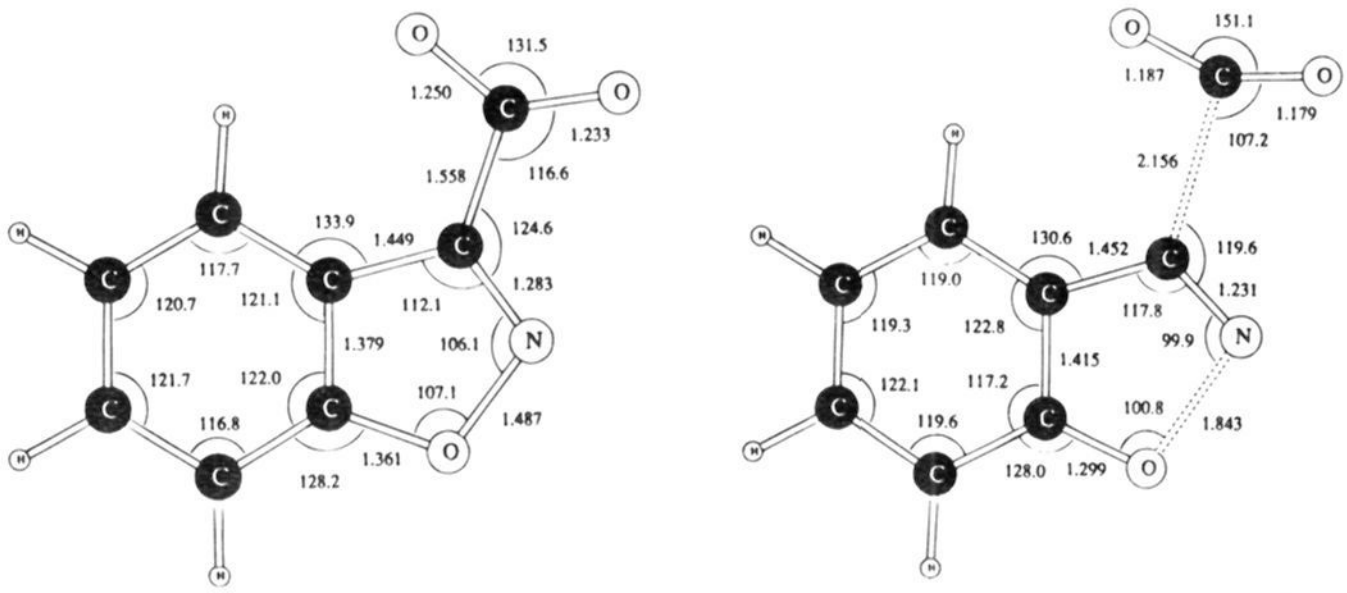


Figure 2. Optimized 3-21G geometries for the reactant and transition state; bond lengths in angstroms, angles in degrees.

that the 3-21G IRC path might be reasonable for use in Monte Carlo simulations.

(b) The Potential of Mean Force in Aqueous Solution. The principal goal of the present study is to calculate the potential of mean force (pmf) for the decarboxylation reaction and the aqueous solvent effect on $\Delta G^{\ddagger}$. Statistical perturbation theory along with a double-wide sampling technique was used in Monte Carlo simulations to compute the change in ΔG_{sol}.[40] This required a total of 16 simulations to span the entire r_c. Standard errors were estimated over averages of blocks of 10^5 configurations.

Key results are illustrated in Figure 1, which compares the aqueous-phase pmf (solid curve) with the gas-phase free-energy profile (dashed curve). The most dramatic feature is the large increase in activation free energy due to solvation. In water, the reactant is better solvated than the TS by 10.6 ± 0.3 kcal/mol. In conjunction with the gas-phase ab initio 3-21G results (Table 1), the calculation predicts an aqueous $\Delta G^{\ddagger}$ of 26.1 ± 0.3 kcal/mol, which is in quantitative agreement with the experimental value (26.3 ± 1.5 kcal/mol).[17a] If the MP2/6-31+G(d)//3-21G value was used, the computed $\Delta G^{\ddagger}$ would be 25.3 kcal/mol. The agreement emphasizes the validity of the present method.

The second significant result in Figure 1 is that the position of the transition state is not affected by the large solvent effect in aqueous solution. This is particularly surprising in view of the large increase in $\Delta G^{\ddagger}$ in water, and the sensitivity of the decarboxylation reactions to solvents in general.[17-19] It might be expected that strong hydrogen-bonding interactions between the carboxylate group and solvent in the ground state, which become much weaker at the transition state due to delocalization of charges, would yield an early or tighter TS than that in the gas phase. Indeed, Zipse and Houk demonstrate that inclusion of one or two water molecules hydrogen bound to the leaving carboxylate yields a TS with a shorter $C_3-C_{O_2}$ distance than that of **2** alone.[21] The computed ΔG_{sol} does show a maximum corresponding to $r_c = 2.04$ Å, or about 0.11 Å earlier than the location of the TS in the gas phase. This is in accord with findings of Zipse and Houk (a $R(C_3-C_{O_2})$ of 2.08 Å with one or two waters).[21] However, the change in ΔG_{sol} (−0.9 ± 0.1 kcal/mol) from $r_c = 2.04$ Å to the TS ($r_c = 2.16$ Å) is not sufficient to offset the increase in the intrinsic (in vacuo) free energy, leaving the position of the TS unchanged along the reaction coordinate in water. Interestingly, the simulation results are consistent with previous predictions based on experimental findings. Kemp and Paul, based upon the observation of a constancy in the Hammett coefficients and $\Delta S^{\ddagger}$ for this reaction, concluded that little change in transition state structures occurs in various solvents.[17b] More recently, Lewis et al., who measured the carbon kinetic isotope effect for the decarboxylation of 5-nitro-3-carboxybenzisoxazole in water, dioxane, and a catalytic antibody, found that the isotope effects are similar under all three conditions, suggesting again that the TS does not change significantly in different environments.[19] These observations suggest that use of the gas-phase IRC in aqueous simulations appears to be reasonable for this reaction.

Table 2. Computed Solute−Solvent Interaction Energies (kcal/mol) in Aqueous Solution at 25 °C and 1 atm

	reactant	transition state	product
E_{vdW}	−12.2 ± 0.2	−14.1 ± 0.2	−13.7 ± 0.2
$E_{elec}^{(1)}$	−123.5 ± 0.7	−95.1 ± 0.8	−105.1 ± 0.8
E_{pol}	−9.8 ± 0.2	−10.1 ± 0.2	−8.1 ± 0.2
E_{sx}	−145.5 ± 1.0	−119.3 ± 0.9	−127.0 ± 1.0
ΔE_{sx}	0.0	26.2 ± 1.3	18.5 ± 1.4
$\Delta\Delta G_{sol}$	0.0	10.6 ± 0.3[a]	7.8 ± 0.4[a]

[a] Determined from free energy simulations of the reaction profile shown in Figure 1.

Finally, the free energy of reaction is estimated from Figure 1 to be about −37 ± 1 kcal/mol in water. Although extending the simulations to a longer distance than 3.3 Å shown in Figure 1 for the separation between carbon dioxide and salicylonitrile could bring some change in the value of the free energy of reaction, it is not likely to be significantly more than a few kilocalories per mole since it is known that anion and neutral species bind weakly in aqueous solution.[47] The theoretical result may be compared with the *enthalpy* of reaction (about −30 kcal/mol) estimated by Kemp and Paul;[17] the agreement is good.

(c) Differential Hydration of the Transition State and Reactant. The origin of the differential hydration is characterized by further examination of the computed distribution functions and energetic results for the ground state, transition state, and products in separate simulations that involved averaging over 1.5×10^6 configurations. Of particular interest are the results given in Table 2, which lists the components of the solute−solvent interaction energies. Clearly, the reactant is significantly better solvated than the transition state; the total solute−solvent interaction energy, E_{sx}, is 26.2 kcal/mol more favorable for the reactant than the TS, which is primarily responsible for the increase in $\Delta G^{\ddagger}$. Of course, the interaction *energy* difference is much greater than the change in *free energy* (10.6 kcal/mol) from statistical perturbation calculations because strong solute−solvent interaction is also accompanied by larger solvent disruption.[48] Roux et al. and others have shown that

(47) Schneider, H.-J.; Schiestel, T.; Zimmermann, P. *J. Am. Chem. Soc.* **1992**, *114*, 7698.
(48) Jorgensen, W. L.; Gao, J. *J. Phys. Chem.* **1986**, *90*, 2174.

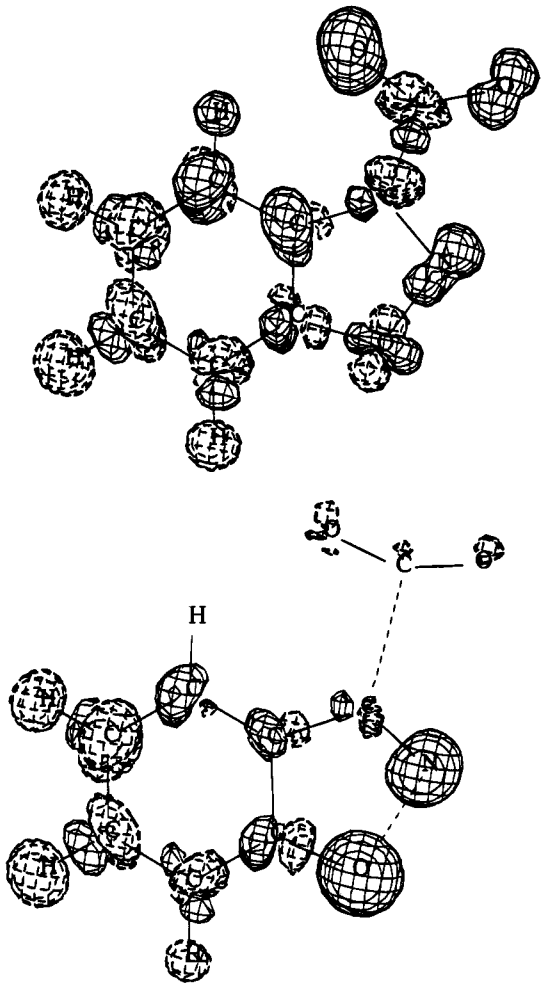


Figure 3. Electron density difference (EDD) plots for the reactant (top) and transition state (bottom). Dotted contours represent regions where electron density is depleted and solid curves indicate areas where there is a gain in electron density on transferring the solute from the gas phase into water.

the free energy of solvation for ionic systems is exactly one half of the total interaction energy based on an integral equation method.[24,49]

Table 2 also reveals that the solvation difference between the TS and reactant comes almost entirely from the $E^{(1)}$ term, a contribution due to the "permanent" charge distribution.[14] This is because the charge distribution is more delocalized in the TS than in the ground state. The polarization effect does contribute significantly to the total solute−solvent interaction energy for all species; however, the polarization energies are the same within statistical error for both the TS (**2**) and reactant (**1**) (−10.1 and −9.8 ± 0.2 kcal/mol, respectively). The smaller value (−8.1 ± 0.2 kcal/mol) estimated for the products is probably due to poor polarization of CO_2 in water. To investigate the nature of the solvent polarization effect on the transition state and reactant, electron density difference (EDD) plots are displayed in Figure 3 for two instantaneous configurations from Monte Carlo simulations of the two species.[50] The EDD plots are obtained by subtracting the gas-phase electron density from that of the solute in water determined with the AM1 wave function. In Figure 3, the plots for the TS and the ground state exhibit some remarkable similarity and surprising differences. In both cases, the major depletion in electron density is from the phenyl ring; however, electron densities are shifted to the carboxylate group in the ground state, and to the isoxazole oxygen and nitrogen in the transition state, respectively.

(49) Roux, B.; Yu, H. A.; Karplus, M. *J. Phys. Chem.* **1990**, *94*, 4683.
(50) Jorgensen, W. L.; Severance, D. *A Modified Version Psi/88*; Department of Chemistry, Yale University.

Table 3. Computed Partial Atomic Charges (e) Using the AM1 Wave Function

	reactant		transition state		product	
atom[a]	gas phase	aqueous	gas phase	aqueous	gas phase	aqueous
O	−0.183	−0.188	−0.398	−0.511	−0.472	−0.620
N	0.019	−0.021	0.037	−0.099	−0.242	−0.237
C_3	−0.135	−0.129	−0.315	−0.264	0.024	0.020
C_4	−0.023	−0.074	−0.031	−0.061	−0.006	−0.040
H_4	0.180	0.159	0.131	0.133	0.097	0.128
C_5	−0.205	−0.159	−0.274	−0.210	−0.319	−0.275
H_5	0.114	0.157	0.101	0.152	0.096	0.154
C_6	−0.104	−0.067	−0.074	−0.067	−0.058	−0.037
H_6	0.107	0.167	0.089	0.158	0.079	0.146
C_7	−0.174	−0.133	−0.261	−0.232	−0.303	−0.333
H_7	0.127	0.172	0.113	0.152	0.102	0.122
C_8	−0.133	−0.185	−0.176	−0.208	−0.303	−0.316
C_9	0.044	0.039	0.186	0.168	0.310	0.291
C	0.411	0.442	0.486	0.491	0.435	0.455
O_2	−0.549	−0.602	−0.313	−0.303	−0.221	−0.248
O_3	−0.497	−0.579	−0.303	−0.299	−0.218	−0.210

[a] Atoms are numbered following the IUPAC rules except for the carboxylate group where O_2 and O_3 are oxygens trans and cis to the nitrogen.

It should be noted that although Figure 3 only represents the result of a single, typical configuration out of the 1.5 million sampled in the simulation, the illustrations are consistent with the charge distributions averaged over all the liquid configurations. Partial atomic charges for these species in the gas phase and in aqueous solution have thus been computed from the Mulliken population analyses.[11] Standard deviations for the computed atomic charges are about 0.003 e on average. Although the values should only be regarded as a provocative indication of the charge distributions, the qualitative trends are still valuable. Table 3 recorded the average atomic charges for the reactant, TS, and product molecules in water along with those in the gas phase from AM1 calculations. The numerical results in Table 3 mirror the qualitative drawings in Figure 3. In the ground state, the carboxylate group in water has a total charge of −0.739 e, which is 0.104 e more negative than the gas-phase value due to hydration. The changes for the isoxazole atoms are negligible. For the TS structure in water, there is virtually no change in atomic charge on the CO_2 group (−0.111 e in water vs −0.130 e in the gas phase), but the net gains in electron density on the isoxazole oxygen (0.113 e) and nitrogen (0.136 e) are significant, indicating a shift of the hydrogen bonding site from the carboxylate group in the ground state to the isoxazole oxygen in the transition state. Now that the quantitative polarization energies are predicted to be similar in hydration of the TS and the ground state (Table 2), the origins of the polarization effects are quite different from the two species (Figure 3).

The tremendous solvent-promoted rate acceleration for the decarboxylation of 3-carboxybenzisoxazoles has been attributed to the stabilization of the charge-delocalized TS in organic solvent and the ground state destabilization owing to loss of hydrogen bonding between the carboxylate and water.[17] However, the balance of the two factors is not clear. The present results demonstrate that the polarization effect contributes equally to the hydration of the TS and reactant, suggesting that stabilization of the carboxylate ion in the ground state is the major factor that inhibits the reaction in aqueous solution. To investigate whether the polarization effect can stabilize the TS

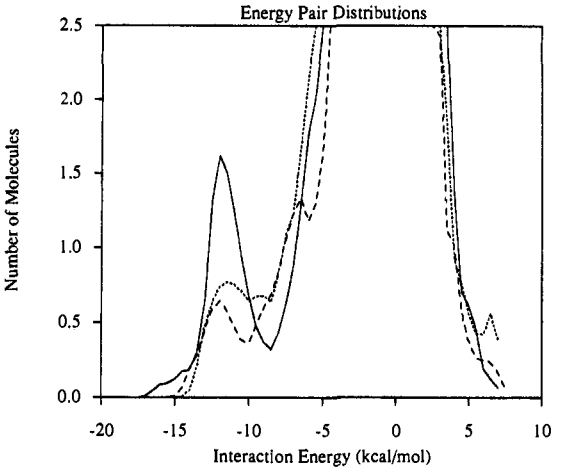


Figure 4. Computed solute−solvent energy pair distributions for the reactant, transition state, and product ($r_c = 3.3$ Å). The ordinate gives the number of water molecules bound by the solute with the energy shown on the abscissa. Units for the ordinate are molecules per kilocalories per mole. Solid curves are for the reactant, dashed curves are for the transition state, and dotted curves are for the product. This convention is used throughout.

in aprotic solvents, simulations using a non-aqueous solvent and the combined QM/MM potential are required.

Some additional insights on the solute−solvent interactions are provided by the energy pair distributions shown in Figure 4. The sharp first peak for 3-carboxybenzisoxazole (solid curve) is assigned to the hydrogen-bonding interactions between the carboxylate ion and water, which is primarily responsible for the most favorable solute−solvent interaction energy listed in Table 2. This peak is replaced by a small band of similar interaction energy for the transition state, which is accompanied by a second peak at an interaction energy of about −6 kcal/mol (dashed curve). In view of the charge patterns noted above, the first peak may be assigned to hydrogen-bonding interaction between water and O_1, and the second peak is due to the residual hydrogen bonds to the CO_2 oxygens. Finally, the distribution for the product shows a single hydrogen-bonding peak from the salicylonitrile oxygen (dotted curve). Note that the relative heights of the hydrogen-bonding peaks, which all occur with about the same interaction energy, are consistent with the computed total solute−solvent interaction energies (Table 2).

Specific hydrogen-bonding interactions are revealed in the radial distribution functions (rdf's) computed for the reactant, TS, and products (Figures 5−8). In these figures the first atom for an xy distribution, $g_{xy}(r)$, refers to an atom of the solute, and the second atom is either the oxygen or hydrogen of water. For the solute O_2 and O_3 are carboxylate oxygens trans and cis to the nitrogen atom, respectively, and O and N are the oxygen and nitrogen atoms in the isoxazole ring. All rdf's have been normalized to the bulk density for atom y. The error range in these plots is about one half of the bin size (0.05 Å) for data collection.

The carboxylate oxygen−water oxygen and hydrogen rdf's are shown in Figures 5 and 6. The most striking features are the first peaks in these plots, which reflect hydrogen bonding of water to carboxylate (CO_2) oxygens. For the reactant hydrogen bonding is apparent, but it is diminished in magnitude for the transition state and even smaller for the product. Integration of the O_2−H and O_3−H first peaks to the minima at 2.60 Å yields 3.4 hydrogens each for the reactant, 2.0 for the transition state, and about 1.0 for the product, which are the same as the number of oxygens from integration of the O−O peaks out to 3.4 Å. Thus, the carboxylate or CO_2 group forms

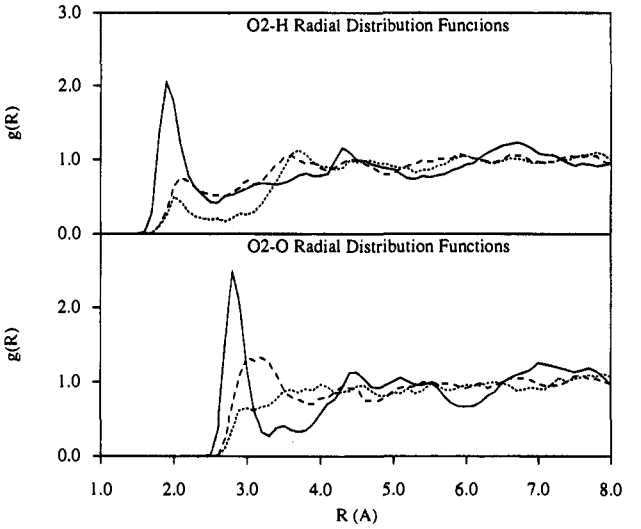


Figure 5. Computed O_2−O and O_2−H radial distribution functions for the reactant, transition state, and product. Distances are in angstroms throughout.

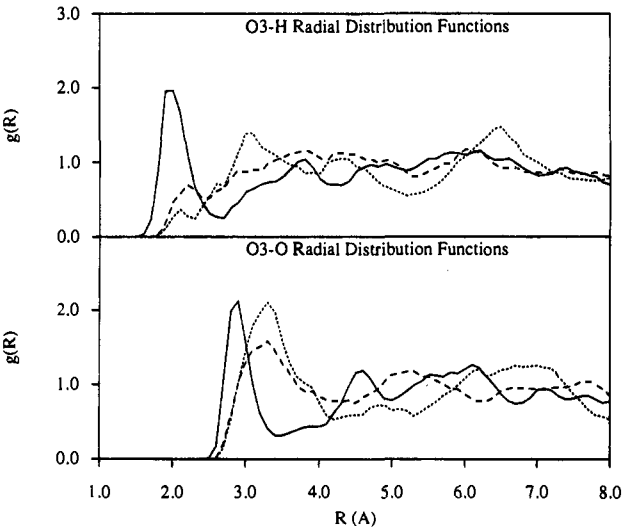


Figure 6. Computed O_3−O and O_3−H rdf's.

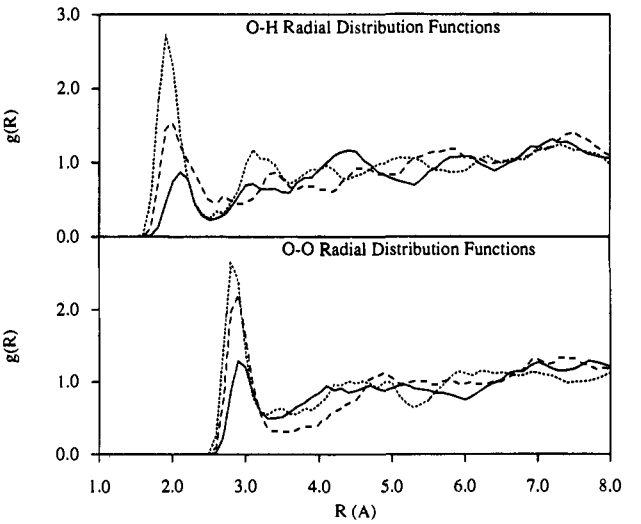


Figure 7. Computed O−O and O−H rdf's.

on average a total of 7 hydrogen bonds in the ground state, a value in exact accord with the prediction for carboxylate ions in water.[48] It should be mentioned that for the TS and product, not only does the strength of hydrogen bonding interactions

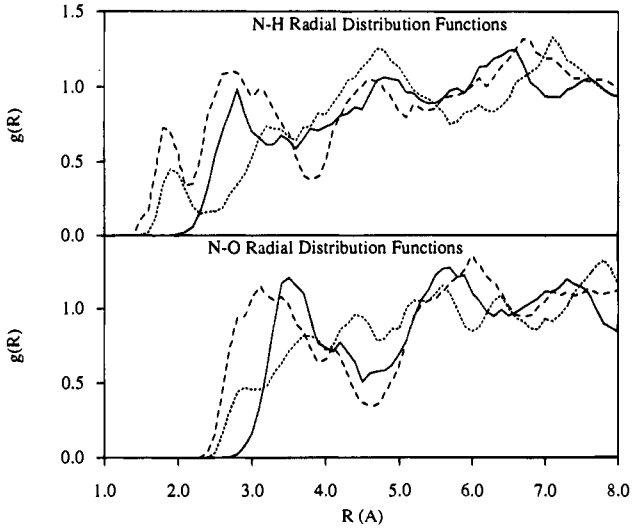


Figure 8. Computed N−O and N−H rdf's.

decrease as revealed in the energy pair distributions (Figure 4) but also the number of hydrogen bonds diminishes to about 4 and 1, respectively.

However, the disappearance of hydrogen bonding at the carboxylate is accompanied by the occurrence of a sharp first peak in the O−O and O−H rdf's (Figure 7). The trends show a progressive shortening of the positions of these peaks as the reaction proceeds. Integration of the O−H (O−O) rdf's to their first minima at 2.5 (3.3) Å reveals 1.5 (2.1), 3.0 (3.1), and 3.5 (3.6) hydrogens (oxygens) for the reactant, TS, and product, respectively. In addition, the N−H and N−O rdf's (Figure 8) exhibits a small hydrogen bonding peak for the TS and product, which does not exist in the rdf for the reactant. The integrals of the hydrogen-bonding peaks in the N−H rdf's contain 0.9 and 0.7 hydrogen for the TS and product. Interestingly, the number of hydrogen bonds to the nitrogen is greater for the TS structure than for the product, which is consistent with the prediction of a larger solvent polarization effect on the TS than the product.[17b]

Concluding Remarks

A Monte Carlo simulation method using the combined AM1/TIP3P potential is used to evaluate the solvation free energy along the reaction path of the decarboxylation reaction of 3-carboxybenzisoxazole in aqueous solution. At the 3-21G//3-21G level, the transformation occurs via a concerted mechanism with a gas-phase $\Delta G^{\ddagger}$ of 15.4 kcal/mol at 298 K. From Monte Carlo simulations, the barrier height in water is predicted to be 26.1 ± 0.3 kcal/mol, in excellent agreement with experiment. Consistent with the experimental findings, the solvent effect has little effect on the location of the transition state in spite of the dramatic increase in $\Delta G^{\ddagger}$ in aqueous solution. Analyses of the contributing factors in solute−solvent interaction suggest that the aqueous solvent effect is primarily due to the difference in the intrinsic (in vacuo) charge distributions for the reactant and transition state. Although solvent polarization contributes significantly to the solute−solvent interactions, there is little difference between the reactant and transition state. Significally, the origin of the electronic polarization of the reactant and transition state is found to be markedly different. Analyses of electron density distributions indicate that electronic polarization is predominant on the carboxylate group for the reactant, but it is shifted to the isoxazole oxygen and nitrogen atoms in the TS.

Grate et al. suggest that efforts to improve antibody's catalytic efficiency should focus on reducing hydrogen bond donation.[19b] However, our calculation demonstrates that hydrogen bonding to the phenoxide oxygen will induce a stronger polarization in the transition state, consequently leading to stabilization of the transition state. The low catalytic rate of immunoglobin is perhaps not entirely due to the possible presence of hydrogen bonds in the binding site. It is also possible and more likely because there is a lack of hydrogen bonds that can stabilize the transition state. If the "active site" of the catalytic antibody is tailored in such a fashion as to contain a basic residue (lone pair of electrons or an anion) to interact with the CO_2 group, and an acidic residue capable of donating hydrogen bonds near the phenoxide oxygen, it will reduce the electronic polarization of the carboxylate group and enhance the polarization of the cyanopenoxide. Thus, the design shall bring about both ground state destabilization (reduced polarization of the carboxylate which is predominant in the ground state) and transition state stabilization (enhanced polarization of the phenoxide). Clearly, introducing hydrogen-bonding interactions at specific sites is just as important as reducing such interactions at other positions for this decarboxylation reaction.

The present study illustrated the method of computer simulation of chemical reactions in solution using a combined QM/MM potential. Computer simulations of chemical reactions in solution began in the 1980's, which are exemplified by the calculation of a model S_N2 reaction involving exchange of chloride and methyl chloride in water.[1a,51] In that study, an empirical potential function was used to describe the solute−solvent interactions, which involved laborious parametrizations and is difficult to generalize to other complex reactions in solution. The method employed here[8-10] provides an alternative for evaluating intermolecular interactions. Since the solute−solvent interaction is treated by quantum mechanical calculations in the fluid simulation, it is no longer necessary to parametrize empirical potential functions. Furthermore, electronic structure and polarization can be explicitly evaluated, providing valuable insight into the nature of molecular interactions in solution that are not possible using the empirical method.[14] Although the AM1 theory used in the present study is an approximate, semiempirical method, it can be applied to molecules of biological interest due to its computational efficiency, and the capacity of the method can easily be extended along with the ever increasing computer speed.

Acknowledgment. Gratitude is expressed to the National Science Foundation, the Pittsburgh Supercomputing Center, and the National Institutes of Health for support of this research. I thank Dr. Xinfu Xia for computational assistance.

JA9421314

(51) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 6867; **1985**, *107*, 154.