A QM/MM Monte Carlo Simulation Study of Solvent Effects on the Decarboxylation Reaction of *N*-Carboxy-2-imidazolidinone Anion in Aqueous Solution

Daqing Gao and Yuh-Kang Pan*

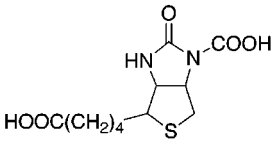
Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467

Received July 30, 1998

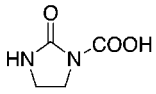
The solvent effects on the decarboxylation reaction of *N*-carboxy-2-imidazolidinone anion in aqueous solution have been investigated by a combined quantum mechanical and molecular mechanical (QM/MM) Monte Carlo simulation method. In the present approach, the gas-phase intrinsic reaction coordinate of the reaction is first obtained by ab initio molecular orbital calculations at the RHF/6-31+G(d) level. Then, the potential of mean force for the decarboxylation reaction is determined via statistical perturbation theory using the combined QM/MM-AM1/TIP3P potential in Monte Carlo simulations. The computed free energy of activation in water is 22.7 ± 0.2 kcal/mol, in very good agreement with the experimental value of 23.2 kcal/mol. Detailed insights into the structural and energetic nature of the differential solvation of the reactant and the transition state are presented.

Introduction

The coenzyme biotin (vitamin H) is an effective acceptor and donor of carbon dioxide in biosynthetic carboxylation pathways.[1] A carboxyl group is transferred through carboxylation and decarboxylation of biotin on N-1, a reaction catalyzed by biotin-dependent enzymes.[2] The biotin molecule itself has been a challenging target of synthetic organic chemistry,[3] as well as several theoretical studies.[4]



N(1')-carboxybiotin



N-carboxy-2-imidazolidinone

Recent experimental studies of solvent effects on the decarboxylation reaction of the model compound *N*-carboxy-2-imidazolidinone revealed that the reaction rate is retarded in going from acetonitrile to aqueous solution.[5] The origin of the rate retardation has been attributed to a more favorable stabilization of the ground state and destabilization of the transition state in polar protic water than in the less polar aprotic acetonitrile.

The decarboxylation of *N*-carboxy-2-imidazolidinone in solution proceeds via two mechanisms as a function of the pH (Scheme 1).[5] Experimentally, at pH > 8, the rate of the reaction is independent of the pH, consistent with a loss of carbon dioxide from the ionized carboxylate form. Below pH 8, the observed rate increases with increased acidity, requiring a reaction via a neutral transition state. The acidity constant has been determined to be pK_a = 4.2 by Rahil et al.[5] Studies of the solvent effects on the reaction rate were conducted at pH = 10.2 by these researchers, and a rate acceleration of 517 times in acetonitrile over that in aqueous solution has been observed. This remarkable solvent effect is comparable to the decarboxylation of 3-carboxybenzisoxazole, which has been extensively investigated by Kemp and coworkers.[6] The latter reaction has been the subject of several recent computational[7] and experimental investigations,[8] including a correlation of the catalytic properties of an antibody in decarboxylations to solvent polarities.

In view of the importance of decarboxylation reactions in a variety of biological processes,[9] and to corroborate recent experimental studies of decarboxylation reactions,[2f,5] computational investigations of the mechanisms of decarboxylation reactions are warranted. In this study, we employ a combined quantum mechanical and molecular mechanical (QM/MM) approach to model the decarboxy-

* Corresponding author. Tel: 617-552-3631. Fax: 617-552-2705. E-mail: yuhkang.pan@bc.edu.

(1) For a review, see: Knowles, J. R. *Annu. Rev. Biochem.* **1989**, *58*, 195.

(2) (a) Lynen, F.; Knappe, J.; Lorch, E.; Jütting, G.; Ringelmann, E. *Angew. Chem.* **1959**, *71*, 481. (b) Caplow, M.; Yager, M. *J. Am. Chem. Soc.* **1967**, *89*, 4513. (c) Guchhait, R. B.; Polakis, S. E.; Dimroth, P.; Stoll, E.; Moss, J.; Lane, M. D. *J. Biol. Chem.* **1974**, *249*, 6633. (d) Berkessel, A.; Breslow, R. *Bioorg. Chem.* **1986**, *14*, 249. (e) Attwood, P. V.; Tipton, P. A.; Cleland, W. W. *Biochemistry* **1986**, *25*, 8197. (f) Tipton, P. A.; Cleland, W. W. *J. Am. Chem. Soc.* **1988**, *110*, 5866.

(3) De Clercq, P. J. *Chem. Rev.* **1997**, *97*, 1755.

(4) (a) Thatcher, G. R. J.; Poirier, R.; Kluger, R. *J. Am. Chem. Soc.* **1986**, *108*, 2699. (b) DeTitta, G. T.; Blessing, R. H.; Moss, G. R.; King, H. F.; Sukumaran, D. K.; Roskwitalski, R. L. *J. Am. Chem. Soc.* **1994**, *116*, 6485. (c) Li, N; Maluendes, S.; Blessing, R. H.; Dupuis, M.; Moss, G. R.; DeTitta, G. T. *J. Am. Chem. Soc.* **1994**, *116*, 6494. (d) Grant, A. S. *J. Mol. Struct. (THEOCHEM)* **1998**, *422*, 79. (e) Miyamoto, S.; Kollman, P. A. *Proteins* **1993**, *16*, 2626.

(5) Rahil, J.; You, S.; Kluger, R. *J. Am. Chem. Soc.* **1996**, *118*, 12495.

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

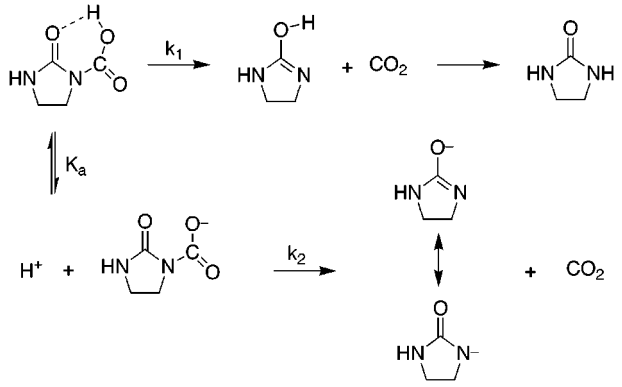
(7) (a) Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 8600. (b) Zipse, H.; Apaydin, G.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 8608.

(8) (a) Lewis, C.; Kreamer, T.; Robinson, S.; Hilvert, D. *Science* **1991**, *253*, 1019. (b) Lewis, C.; Paneth, P.; O'Leary, M. H.; Hilvert, D. *J. Am. Chem. Soc.* **1993**, *115*, 1410. (c) Grate, J. W.; McGill, R. A.; Hilvert, D. *J. Am. Chem. Soc.* **1993**, *115*, 8577. (d) Tarasow, T. M.; Lewis, C.; Hilvert, D. *J. Am. Chem. Soc.* **1994**, *116* , 7959.

(9) (a) Lee, J. K.; Houk, K. *Science* **1997**, *276*, 942. (b) Huang, C.-L.; Wu, C.-C.; Lien, M.-H. *J. Phys. Chem. A* **1997**, *101*, 7867. (c) Bach, R. D.; Canepa, C. *J. Org. Chem.* **1996**, *61*, 6346. (d) Bach, R. D.; Canepa, C. *J. Am. Chem. Soc.* **1997**, *119*, 11725. (e) Miller, B. G.; Traut, T. W.; Wolfenden, R. *J. Am. Chem. Soc.* **1998**, *120*, 2666.

10.1021/jo981523g CCC: $18.00
Published on Web 01/26/1999

Scheme 1



lation reaction of *N*-carboxy-2-imidazolidinone in aqueous solution. Because the experiment of the decarboxylation reaction was conducted at pH = 10.2, the present study will focus on the decarboxylation reaction of the anion corresponding to the experimental conditions at high pH in aqueous solution. These results shed new light on the energetics and factors that govern the dramatic solvent effects on this model carboxybiotin compound in water at an atomic level. Here, high level ab initio calculations are first carried out to dissect the intrinsic reactivity of the decarboxylation reaction in the gas phase. Solvent effects are then determined by Monte Carlo free energy perturbation (FEP) simulations, making use of a combined QM/MM potential by treating the reactant molecule quantum-mechanically. In the following, computational details are first summarized, followed by results and discussion.

Computational Details

Ab Initio Calculations. The transition structure (TS) and the intrinsic reaction coordinate (IRC)[10] path for the decarboxylation reaction of the conjugate base of *N*-carboxy-2-imidazolidinone in the gas phase were determined by ab initio molecular orbital methods at the restricted Hartree−Fock (RHF) level with the 6-31+G(d) basis set. The basis functions include a set of s and p diffuse functions, and a set of polarization functions on all non-hydrogen atoms.[11,12] A total of 37 structures were generated along the IRC, spanning the reaction coordinate R_c (R_c is the breaking N−CO_2 bond distance) from 1.48 Å to 2.59 Å. The transition structure and stationary geometries are verified by vibrational frequency calculations with one and zero imaginary frequencies, respectively. Electron correlation effects are included through single-point calculations using Möller−Plesset perturbation theory up to the second order. Thus, the notation for the best energy calculations is MP2/6-31+G(d)//RHF/6-31+G(d). Enthalpy changes at 298 K are determined by eq 1:

$$\Delta H^{298} = \Delta E_e^0 + \Delta E_{vib}^0 + \Delta(\Delta E_{vib})^{298} + \Delta E_t^{298} + \Delta E_r^{298} + \Delta PV \quad (1)$$

where ΔE_e^0 is the electronic energy change, ΔE_{vib}^0 is the change in zero-point energy, and $\Delta(\Delta E_{vib})^{298}$ is the change in vibrational energy difference from 0 to 298 K. The last three terms are changes in translational, rotational energies, and the work term, which are all computed classically using standard methods.[13] To compute the enthalpy change ΔH^{298}, harmonic frequencies are scaled by a factor of 0.90 to correct for the overestimation of vibrational frequencies at the HF level.[14] Scaled frequencies below 500 cm^{-1} are treated as classical rotations in calculating their vibrational contributions.

QM/MM Potential. To investigate the solvent effects on the decarboxylation reactions in water, a combined quantum mechanical and molecular mechanical potential is used to describe the solute−solvent interactions.[15−19] The key idea in this approach is the partition of a condensed phase system into a QM and MM region. Thus, the solute molecule in the QM region is treated quantum mechanically, and the solvent MM region is approximated by a classical force field. Combined QM/MM methods have gained great attention in the past five years and have been applied to a variety of organic and biological systems, including studies of organic and enzymatic reactions, ion solvation, conformational equilibria, and kinetic isotope effects. The computational details have been described in several papers.[15−19] Thus, only the most relevant points are summarized here. The total effective Hamiltonian of the system is given by

$$\hat{H}_{eff} = \hat{H}^0 + \hat{H}_{qm/mm} + \hat{H}_{mm} \quad (2)$$

where $\hat{H}^0$ is the Hamiltonian for the solute molecule in the gas phase, $\hat{H}_{mm}$ is the molecular mechanics interaction Hamiltonian for the solvent, and $\hat{H}_{qm/mm}$ is the solute−solvent interaction Hamiltonian. The total energy of the system is expressed as the expectation value of the molecular wave function Φ as follows:

$$E_{tot} = \langle\Phi|\hat{H}_{eff}|\Phi\rangle = E_{qm} + E_{qm/mm} + E_{mm} \quad (3)$$

where Φ is the wave function of the solute molecule in solution, and E_{mm} is the interaction energy of the solvent

(10) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523.

(11) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B.G.; Wong, M. W.; Foresman, J. B.; Robb, M. A.; Head-Gordon, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. L.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92/DFT*, Revision G3; Gaussian, Inc.: Pittsburgh, PA, 1993.

(12) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

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

(14) Grev, R. S.; Janssen, C. L.; Schaefer, H. F., III. *J. Chem. Phys.* **1991**, *95*, 5128.

(15) (a) Warshel, A; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. (b) Warshel, A.; Lappicirella, A. *J. Am. Chem. Soc.* **1981**, *103*, 4664. (c) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199. (d) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523.

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

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

(18) (a) Thompson, M. A.; Schenter, G. K. *J. Phys. Chem.* **1995**, *99*, 6374. (b) Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1995**, *99*, 384. (c) Liu, H.; Shi, Y. *J. Comput. Chem.* **1994**, *15*, 1331. (d) Stanton, R. V.; Hartsough, D. S.; Merz, K. M., Jr. *J. Phys. Chem.* **1993**, *97*, 11868. (e) Wei, D.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *224*, 291. (f) Liu, H.; Muller-Plathe, F.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 1702. (g) Chatfield, D. C.; Brooks, B. R. *J. Am. Chem. Soc.* **1995**, *117*, 5561. (h) Barnes, J. A.; Williams, I. H. *Chem. Commun.* **1996**, 193. (i) Bakowis, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580.

(19) (a) Gao, J.; Xia, X. *Science* **1992**, *258*, 631. (b) Gao, J. In *Reviews in Computational Chemistry*; VCH: New York, 1995; Vol. 7. (c) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298.

which is determined classically. E_{qm} and $E_{qm/mm}$ are, respectively, the energy of the solute in solution, and the solute−solvent interaction energy. $E_{qm/mm}$ includes both electrostatic and van der Waals components:

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

where $\hat{H}^{el}_{qm/mm}$ is the electronic part of the solute−solvent interaction Hamiltonian. In the van der Waals term (eq 4), standard geometric combining rules are used to determine the Lennard-Jones parameters. Therefore, $\sigma_{is} = (\sigma_i\sigma_s)^{1/2}$ and $\epsilon_{is} = (\epsilon_i\epsilon_s)^{1/2}$, where σ_i and ϵ_i are parameters for the solute atoms, and their values are taken from ref 19a. σ_s and ϵ_s are parameters for the solvent, taken from Jorgensen's three point charge TIP3P water model.[20]

In studying solute−solvent interactions, it is of particular interest to examine the polarization effect on the solute molecule that is induced by the solvent. The polarization energy is determined by eq 5:

$$E_{pol} = E_{dist} + E_{stab} \quad (5)$$

where E_{dist} is the solute distortion energy which is the energy penalty necessary for polarizing the solute wave function, and E_{stab} is the stabilization energy gained due to the solute charge polarization.

$$E_{dist} = \langle\Phi|\hat{H}^0|\Phi\rangle - \langle\Phi^0|\hat{H}^0|\Phi^0\rangle \quad (6)$$

$$E_{stab} = \langle\Phi|\hat{H}^{el}_{qm/mm}|\Phi\rangle - \langle\Phi^0|\hat{H}^{el}_{qm/mm}|\Phi^0\rangle \quad (7)$$

With this definition of the solute polarization energy, the total solute−solvent interaction energy can be written as

$$E_{sx} = E_{vdW} + E^{(1)} + E_{pol} \quad (8)$$

where $E^{(1)}$ is the solute−solvent interaction energy with a gas-phase solute wave function (Φ^0), and E_{vdW} is the van der Waals energy, arising from the Lennard-Jones terms in eq 4. In this study, all of these energy terms are obtained to provide additional insights on specific solute−solvent interactions.

Of course, it is desirable to use ab initio molecular orbital methods or density functional theory to describe the solute electronic structures in the combined QM/MM potential. However, this is still too time-consuming to be practical for systems of the present size. Consequently, as in most previous studies, we have used the semiempirical AM1 model[21] to represent the *N*-carboxy-2-imidazolidinone molecule. The combined AM1/MM approach has been shown to yield good solvation results for a variety of systems, including a similar decarboxylation reaction of 3-carboxybenzisoxazole in water.[7]

Monte Carlo Simulations. Statistical mechanical Monte Carlo calculations were carried out in the isothermal−isobaric (NPT) ensemble at 25 °C and 1 atm using the combined AM1/TIP3P potential. In the simulation, the N−CO_2 bond that breaks along the reaction coordinate is oriented to coincide with the *Z*-axis of a rectangular box with periodical boundary conditions[22] (approximately 20 × 20 × 30 Å^3) containing 391 TIP3P water molecules. For convenience in the following discussion, the reaction coordinate (R_c) is represented by the N−CO_2 distance, rather than the IRC atomic units. Fifteen structures from ab initio IRC calculations, plus an additional 13 structures that extended the reaction coordinate to R_c = 5.2 Å from the last IRC structure at R_c = 2.6 Å, were selected. Differences in hydration free energy are calculated via statistical perturbation theory using the Zwanzig equation with double-wide sampling.[23,24]

$$\Delta G(R_c) = -k_B T \ln\langle\exp\{-[E(R_j) - E(R_i)]/k_B T\}\rangle_{R_i} \quad (9)$$

where R_i is a structure along the IRC reaction path, k_B is the Boltzmann constant, T is the temperature, and $E(R_j)$ and $E(R_i)$ are the total potential energies for state R_j and R_i. The brackets $\langle\cdots\rangle_{R_i}$ in eq 9 indicate ensemble averaging over the potential energy surface $E(R_i)$ for the reference structure R_i. It should be noted that geometrical mapping of eq 9 does not account for the nonequilibrium solvation effects which seems to be minor in the present system. Other approaches that allow one to evaluate the activation barrier in the solute−solvent coordinate space were recently developed.[25]

The intermolecular interactions are truncated between spherical cutoff distances of 8.5 and 9.0 Å for water−water interaction based on the oxygen separations. Solute−solvent interactions are feathered to zero between 9.0 and 9.5 Å based on solute atom−water oxygen distances. Standard Metropolis sampling was used.[26] To facilitate the statistics near the solute, the Owicki−Scheraga preferential sampling technique with $1/(r^2 + C)$ weight, where C = 150 Å^2, was also adopted.[27] The ranges for attempted translations and rigid-body rotations of the solute and solvent molecules were adjusted to yield ca. 45% acceptance rates for new configurations. Volume changes were restricted to within ±230 Å^3 at every 2375 configurations. A total of 27 simulations were performed to span the entire reaction coordinate (1.48−5.2 Å). Each simulation consisted of 5−8 × 10^5 configurations of equilibration, followed by additional 1.5 × 10^6 configurations of averaging. The uncertainties were computed from the fluctuations in separate averages over batches of 1 × 10^5 configurations. All calculations were performed using the MCQUB/BOSS program provided by Professor Gao at SUNY, Buffalo.[28]

Results and Discussion

(a) Ab Initio Structures and Energies. The fully optimized geometries of the reactant, transition state and

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

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

(22) (a) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991. (b) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids;* Clarendon Press: Oxford, 1987.

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

(24) (a) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050. (b) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.

(25) Warshel, A.; Muller, R. P. *J. Phys. Chem.* **1995**, *99*, 17516.

(26) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

(27) (a) Owicki, J. C.; Scheraga, H. A. *Chem. Phys. Lett.* **1977**, *47*, 600. (b) Owicki, J. C. *ACS Symp. Ser.* **1978**, *86*, 159. (c) Jorgensen, W. L.; Bigot, B.; Chandrasekhar, J. *J. Am. Chem. Soc.* **1982**, *104*, 4584.

(28) (a) Gao, J. *MCQUB*, SUNY at Buffalo, 1997. (b) Stewart, J. J. P. *MOPAC, Version 5. QCPE* **1986**, *6*, 455, No. 391. (c) Jorgensen, W. L. *BOSS, Version 2.9*; Yale University: New Haven, CT, 1990.

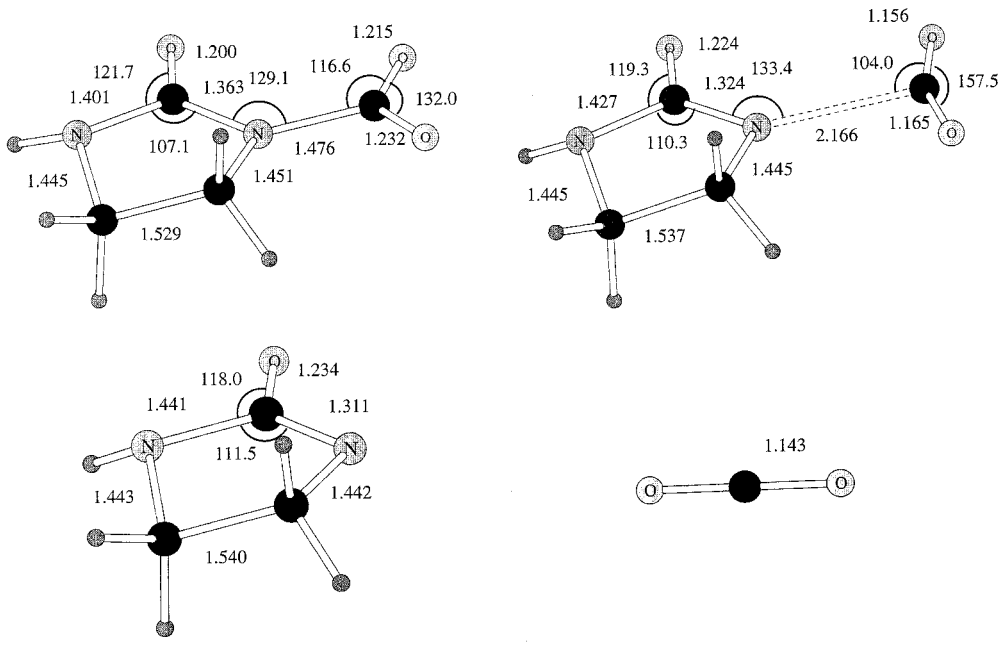


Figure 1. RHF/6-31+G(d) optimized geometries for the reactant, transition state, and products; bond lengths are in angstroms, angles are in degrees.

Table 1. Thermodynamic Results for the Decarboxylation of *N*-Carboxy-2-imidazolidinone in the Gas Phase at 298 K[a]

	reactant → transition state	reactant → ion-dipole complex	reactant → products ($R_c = \infty$)
ΔE^0 (HF/6+31+G*)	17.1	16.2	23.9
ΔE_v^0	−1.3	−0.9	−2.4
$\Delta\Delta E_v^{298}$	−0.2	0.1	−1.4
ΔH^{298}	15.6	15.3	22.2
ΔS^{298}	4.9	12.5	28.1
ΔG^{298}	14.2	11.7	13.8
ΔE^0 (MP2/6+31+G*)	13.4	14.9	24.8
ΔH^{298} (MP2/6+31+G*)	12.0	14.1	23.0
ΔG^{298} (MP2/6+31+G*)	10.5	10.4	14.7
ΔH (AM1, mopac)	19.6		

[a] Energies and entropies are given in kcal/mol and cal/mol·K.

products are illustrated in Figure 1. It is found that the N_3 nitrogen has a nonplanar pyramidal structure in both the reactant and the TS. From reactant to the TS, the major structural changes are that the C_2-N_1 (IUPAC numbering) distance is shortened by 0.039 Å, while the C_2-O and C_2-N_3 distances are elongated by 0.024 Å and 0.026 Å, respectively. The N_1-CO_2 bond distance increases from 1.476 Å to 2.166 Å, and the carboxylate angle is widened by 25.5° in going from the reactant state to the TS.

Gas-phase energy results from ab initio calculations are summarized in Table 1. At the RHF/6-31+G(d)//RHF/6-31+G(d) level, the calculated activation energy is 17.1 kcal/mol, while inclusion of the electronic correlation effect at the MP2/6-31+G(d)//RHF/6-31+G(d) level reduces the activation energy to 13.4 kcal/mol. For comparison, the AM1 Hamiltonian yields an activation enthalpy $\Delta H^{\ddagger}$ (25 °C) = 19.6 kcal/mol with the RHF/6-31+G(d) structure, suggesting that the AM1 model gives a reasonable description of the present decarboxylation reaction. In combination with the RHF/6-31+G(d) vibrational frequencies and the RHF/6-31+G(d) electronic energies, the computed activation enthalpy, $\Delta H^{\ddagger}$, activation entropy, $\Delta S^{\ddagger}$, and Gibbs activation free energy, $\Delta G^{\ddagger}$, for the reaction at 298 K are 15.6 kcal/mol, 4.9 cal/mol·K, and 14.2 kcal/mol, respectively. When the MP2/6-31+G(d) energy differences are used, the computed $\Delta H^{\ddagger}$ and $\Delta G^{\ddagger}$ are 12.0 kcal/mol and 10.5 kcal/mol, respectively. Gas phase experimental activation energies are not available for the decarboxylation reaction of *N*-carboxy-2-imidazolidinone, although $\Delta H^{\ddagger}$ values of 19.0 and 20.8 kcal/mol at pH of 6.1 and 10.2 were obtained from the Eyring plots using the experimental data.[29] Comparison can be made to the barrier height for the decarboxylation of 3-carboxybenzisoxazole, which has a predicted $\Delta E^{\ddagger}$ of 19.1 kcal/mol at the MP2/6-31G(d)//6-31G(d) level.[7b] Note that the free energy of the product ion−dipole complex (R_c = 2.6 Å) is only 2.5 kcal/mol and 0.1 kcal/mol below the transition state at the HF/6-31+G(d) level and the MP2/6-31+G(d) level, respectively. The overall reaction ($R_c = \infty$) is endothermic by ca. 14 kcal/mol (ΔG) in the gas phase.

Mulliken population analysis from RHF/6-31+G(d) calculations shows that about 75% of the negative charge on the carboxylate group is transferred to the ring system in going from the reactant to the transition state. It can be expected that this will result in the loss of hydrogen-bonding interactions in the TS for this anion decarboxylation reaction in aqueous solution from the fluid simulations.

(b) The Potential of Mean Force in Aqueous Solution. The main goal of this study is to determine the solvent effects on the activation barrier, and the potential of mean force[30] for the decarboxylation of *N*-carboxy-2-imidazolidinone in water. This has been accomplished through a series of Monte Carlo free energy

(29) The activation parameters reported in ref 5 were in error. Using the reported rate constants, the Eyring plots yield the following activation parameters in water: $\Delta H^{\ddagger}$ = 20.8 kcal/mol, $\Delta S^{\ddagger}$ = −8.0 cal/mol·K, and $\Delta G^{\ddagger}$ = 23.2 kcal/mol at pH 10.2; $\Delta H^{\ddagger}$ = 19.0 kcal/mol, $\Delta S^{\ddagger}$ = −5.8 cal/mol·K, and $\Delta G^{\ddagger}$ = 20.7 kcal/mol at pH 6.1. Corrections: Rahil, J.; You, S.; Kluger, R. *J. Am. Chem. Soc.* **1998**, *120*, 2692.

(30) (a) Gao, J. *J. Am. Chem. Soc.* **1991**, *113*, 7796. (b) Peng, Z.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 2733. (c) Chang, N.-Y.; Lim, C. *J. Am. Chem. Soc.* **1998**, *120*, 2156.

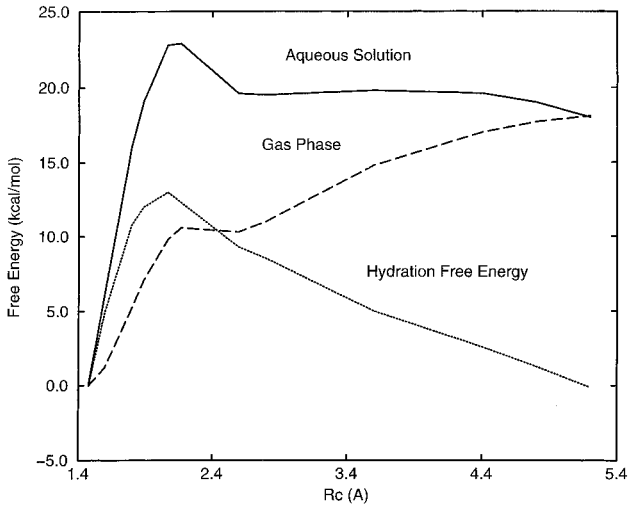


Figure 2. Computed free energy profile in the gas phase (dashed curve), hydration free energy (dotted curve), and the potential of mean force in aqueous solution (solid curve) for the decarboxylation reaction of *N*-carboxy-2-imidazolidinone. The reaction coordinate is represented as the N_1-C_{O2} distance along the 6-31+G(d) intrinsic reaction coordinate.

perturbation (FEP) simulations along the reaction path generated from the ab initio HF/6-31+G(d) calculations. Key findings are shown in Figure 2, which depicts the free energy profile for the decarboxylation reaction in the gas phase and in aqueous solution. The most dramatic feature in Figure 2 is the large increase in barrier height in going from the gas phase to aqueous solution. The reactant anion is predicted to be better solvated than the transition structure at $R_c = 2.2$ Å by 12.2 ± 0.2 kcal/mol. In conjunction with the gas phase activation energy (10.5 kcal/mol) obtained at the MP2/6-31+G(d)//6-31+G(d) level, the present QM/MM simulations yield an aqueous $\Delta G^{\ddagger}$ of 22.7 ± 0.2 kcal/mol, which is in excellent agreement with the experimental value of 23.2 kcal/mol at pH = 10.2.[29] The good agreement with experiment, together with an early study of the decarboxylation reaction of 3-carboxybenzisoxazole in water,[7] demonstrate that the use of a combined QM-AM1/MM potential can provide a reasonable description for decarboxylation reactions in solution.

Second, it is interesting to note that the position of the transition state in aqueous solution is unchanged in comparison with that in the gas phase (Figure 2). Although this is surprising in view of the large solvent effect on the activation free energy, similar observations have been obtained for the decarboxylation of 3-carboxybenzisoxazole from Monte Carlo QM/MM simulations,[7] and from the interpretation of the experimental $\Delta S^{\ddagger}$ in various solvents,[6] as well as from carbon-13 kinetic isotope effects in solutions and in a catalytic antibody.[8] The computed difference in free energy of hydration along the reaction coordinate does show a shift with the maximum in $\Delta\Delta G_{hydr}(R_c)$ at 2.07 Å, which is about 0.1 Å earlier than the location of the transition state in the gas phase. However, the solvation effect is not enough to offset the increase in the intrinsic free energy (gas phase) change. Thus, the overall position of the transition structure in water remains unchanged. This *a posteriori* finding indicates that the use of gas phase IRC path for simulations in solution is reasonable for decarboxylation reactions.

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

	reactant	ts	product ($R_c = 5.2$ Å)
E_{vdW}	-8.3 ± 0.2	-8.8 ± 0.2	-10.2 ± 0.3
E_{vert}	-140.3 ± 1.0	-117.8 ± 1.0	-143.2 ± 1.3
E_{pol}	-8.5 ± 0.3	-5.8 ± 0.2	-7.3 ± 0.3
E_{stab}	-17.0 ± 0.6	-11.5 ± 0.3	-14.3 ± 0.6
E_{dist}	8.4 ± 0.3	5.7 ± 0.2	7.0 ± 0.3
E_{sx}	-157.2 ± 1.1	-132.5 ± 1.1	-160.7 ± 1.4
ΔE_{sx}	0.0	24.7 ± 1.6	-3.5 ± 1.8
$\Delta\Delta G_{sol}$	0.0	12.2 ± 0.2	-0.1 ± 0.4

Third, the free energy of the present reaction (up to $R_c = 5.2$ Å) in water, and in the gas phase, is predicted to be 17.4 kcal/mol and 17.5 kcal/mol, respectively. Note that the gas phase free energy increases by 7.1 kcal/mol when extending the calculations from the last IRC ion−dipole complex structure ($R_c = 2.6$ Å) to the extended structure at $R_c = 5.2$ Å; and the endothermicity of the initial decarboxylation step that is being modeled in the simulation study is accompanied by a proton transfer from water to the product imidazolidinone anion, and the formation of bicarbonate HCO_3^- from CO_2 and OH^- at pH 10.2. We expect that this shall make the overall decarboxylation reaction a facile exothermic process in solution in practice, although, it is noted that the calculated reaction energy (ΔE) of water with *N*-carboxy-2-imidazolidinone anion leading to bicarbonate and neutral 2-imidazolidinone is −7.5 kcal/mol in the gas phase, and 2.0 kcal/mol in aqueous solution with the Onsager self-consistent reaction field (SCRF) model[31] at the HF/6-31+G(d) level, respectively.

To gain additional insights into the origin of the differential hydration effect between the transition structure and the ground state reactant, additional simulations for the reactant, transition state, and product (at $R_c = 5.2$ Å) were carried out, which involved averages over 1.5×10^6 configurations in Monte Carlo calculations. Table 2 summarizes the results of the energy components that contribute to the total solute−solvent interactions for each species. Clearly, the solute−solvent interaction energy for the reactant structure is 24.7 kcal/mol lower than that for the transition structure in water. This is the primary reason for the predicted increase in barrier height ($\Delta G^{\ddagger}$) in water, although it should be noted that the interaction energies listed in Table 2 are much greater than the difference in free energy, which also includes the accompanying solvent reorganization energies and entropic effects. Note that the solvent reorganization energy can be captured in a more rigorous way by a different approach including mapping the solute−solvent coordinate space, rather than by using a solute geometrical mapping of eq 9.[25] For ionic systems, the free energy change is roughly one half of the difference in the interaction energies according to linear response theory.[32] This is consistent with the present numerical results from the combined QM/MM simulations.

The total solute−solvent interaction energy consists of van der Waals (E_{vdW}) and electrostatic terms (E_{elec}). The latter may be further divided into a vertical interaction

(31) (a) Onsager,.L. *J. Am. Chem. Soc.* **1936**, *58*, 1486. (b) Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776. (c) Wiberg, K. B.; Castejon, H.; Keith, T. A. *J. Comput. Chem.* **1996**, *17*, 185.

(32) (a) Tapia, O.; Silvl, B. *J. Phys. Chem.* **1980**, *84*, 2646. (b) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (c) Orozco, M.; Luque, F. J.; Habibollahzadeh, D.; Gao, J. *J. Chem. Phys.* **1995**, *103*, 9112.

Table 3. Computed Partial Atomic Charges (e) Using AM1 Wave Function[a]

	reactant		ts		product (R_c = 5.2 Å)	
atom	gas phase	aqueous	gas phase	aqueous	gas phase	aqueous
O	−0.381	−0.479	−0.488	−0.617	−0.527	−0.632
N_1	−0.399	−0.413	−0.538	−0.543	−0.489	−0.597
C_2	0.382	0.403	0.318	0.306	0.285	0.301
N_3	−0.390	−0.381	−0.390	−0.402	−0.392	−0.411
H_3	0.181	0.247	0.153	0.181	0.143	0.186
C_4	−0.066	−0.069	−0.085	−0.088	−0.094	−0.090
H_{41}	0.069	0.118	0.053	0.107	0.046	0.100
H_{42}	0.044	0.100	0.026	0.061	0.019	0.063
C_5	−0.139	−0.030	−0.036	−0.066	−0.045	−0.067
H_{51}	0.093	0.111	0.051	0.080	0.030	0.082
H_{52}	0.070	0.086	0.035	0.068	0.023	0.064
C	0.451	0.494	0.502	0.528	0.419	0.448
O_2	−0.487	−0.585	−0.276	−0.309	−0.209	−0.235
O_3	−0.553	−0.602	−0.324	−0.305	−0.210	−0.212

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

energy (E_{vert}), which is the interaction energy between the solute and solvent with fixed gas-phase charge distribution for the solute, and a polarization energy (E_{pol}), which is due to the change in solute wave function caused by the solvent polarization effect. There is little variation in the van der Waals energy as the reaction proceeds from the reactant to products. Table 2 reveals that the difference in solvation energy between the reactant and the TS is largely from the E_{vert} term, which is due to the intrinsic (gas phase) charge distributions. In the transition state, the charge distribution is more delocalized, making the effective radius of the anion greater than the reactant, which has the charges primarily localized on the carboxylate group. Thus, according to the Born model of solvation, the effectively "larger" delocalized ionic transition state would be less well solvated than the reactant state.[33] There is a modest contribution from differential polarization effects between the reactant and the transition state. The combined QM/MM calculations suggest that the transition state is not as well polarized as the reactant, and the contribution from E_{pol} is 2.7 kcal/mol smaller than in the ground state. To further investigate this effect, the atomic charge distributions averaged over all the liquid configurations were analyzed. Partial atomic charges for all the species in the gas phase and in aqueous solution have been calculated from the Mulliken population analysis, which gives a reasonable qualitative description of atomic charge distributions. Standard deviations for the computed atomic charges are about 0.002 electrons on average. The atomic charges for the reactant, TS, and product molecules, from AM1 calculations, are given in Table 3. For the ground state in water, the carboxylate group has a total charge of −0.693 e, which is 0.104 e more negative than the gas-phase value due to hydration; no significant charge changes for other atoms are observed. For the TS structure, there is virtually no change in atomic charge on the CO_2 group due to solvation (−0.086 e in water vs −0.098 e in the gas phase), whereas there is a significant increase on the carbonyl oxygen in the imidazolidinone ring (−0.129 e). Thus, both the

(33) (a) Born, M. *Z. Physik*. **1920**, *1*, 45. (b) Warshel, A.; Russel, S. *Q. Rev. Biophys*. **1984**, *17*, 283. (c) Roux, B.; Yu, H. A.; Karplus, M. *J. Phys. Chem*. **1990**, *94*, 4683.

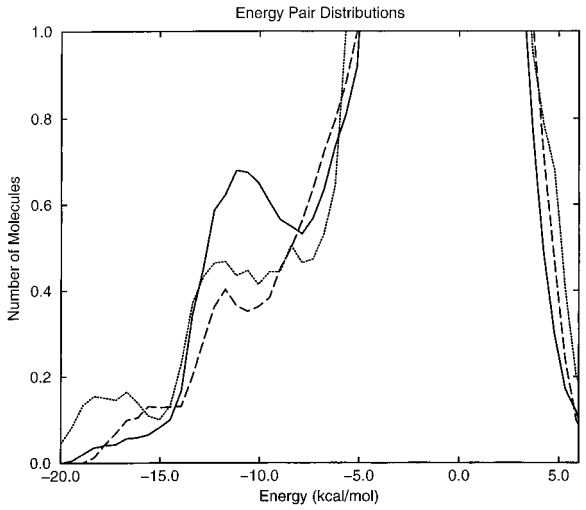


Figure 3. Computed solute−solvent energy pair distributions for the reactant, transition state, and product (R_c = 5.2 Å). The ordinate gives the number of water molecules coordinated with the solute with the interaction energy shown on the abscissa. Solid curves are for the reactant, dashed curves are for the transition state, and dotted curves are for the product. Units for the ordinate are no. of molecules per kcal/mol.

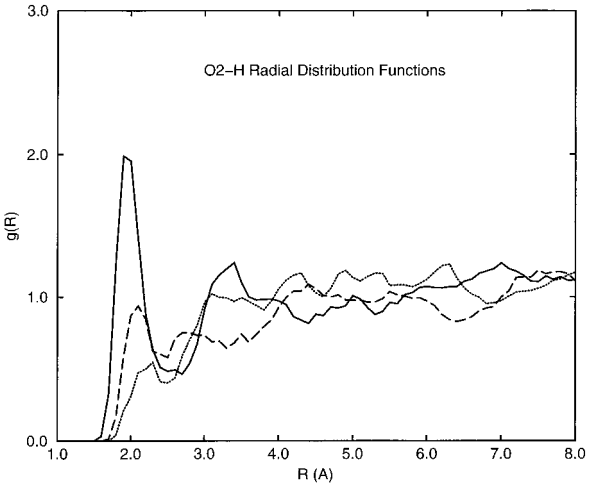


Figure 4. Computed O_2−H radial distribution functions for the reactant, transition state, and product.

polarization and the hydrogen bonds reshuffling (see below) between the reactant and the TS are the major factors that control the observed solvent effect on this decarboxylation reaction in aqueous solution.

In addition, Figure 3 records the energy pair distributions which show the spectrum of individual solute-water interactions in the liquids. The low-energy bands indicate the hydrogen-bonding interactions of solvent water with the solute molecule. Peaks at −10 to −15 kcal/mol are due to the hydrogen bonds between the carboxylate group and water. Peaks at −15 to −20 kcal/mol are attributed to the hydrogen bonds between the developing anion imidazolidinone N_1 and water. The energetic results here mirror the following radial distribution functions.

(c) Radial Distribution Functions. Radial distribution functions are shown in Figures 4−13. In these figures, the rdf $g_{xy}(r) + \langle N_y(r,r = dr)\rangle/(4\pi r^2\rho_y^0 dr)$ gives the probability of finding an atom y at a distance r from atom x. Peaks in rdf's are associated with solvation shells or specific neighbors and can be integrated to yield

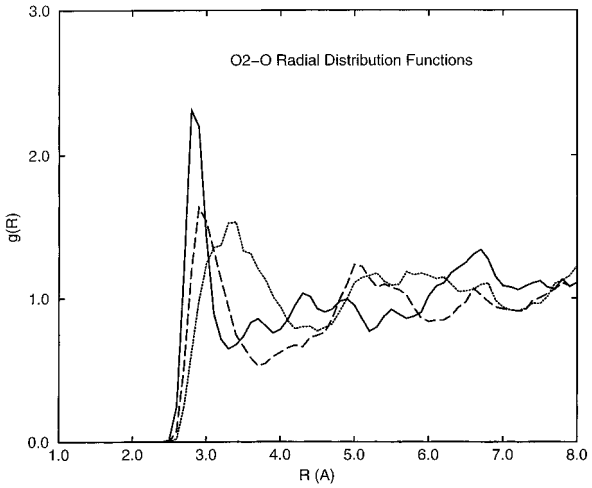


Figure 5. Computed O_2-O rdf's.

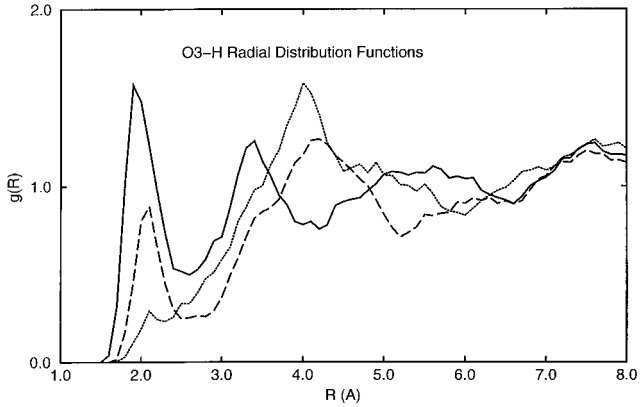


Figure 6. Computed O_3-H rdf's.

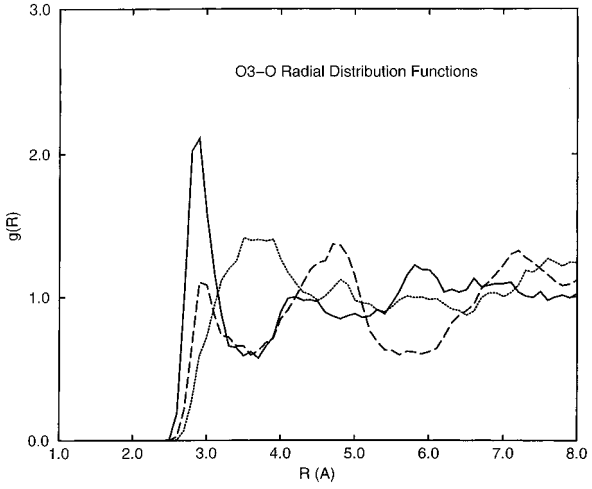


Figure 7. Computed O_3-O rdf's.

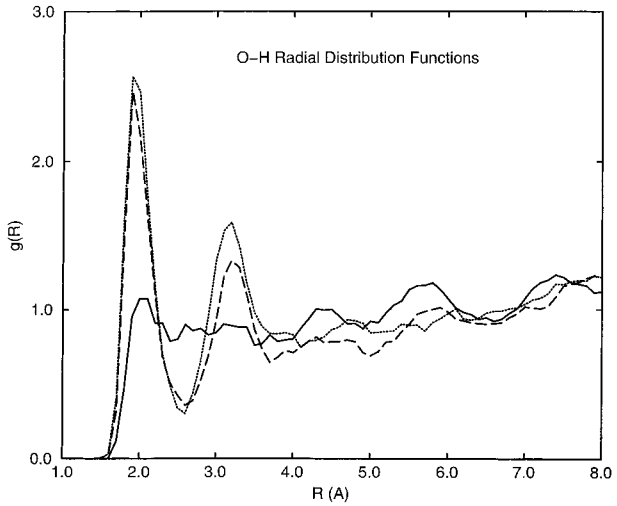


Figure 8. Computed O−H rdf's.

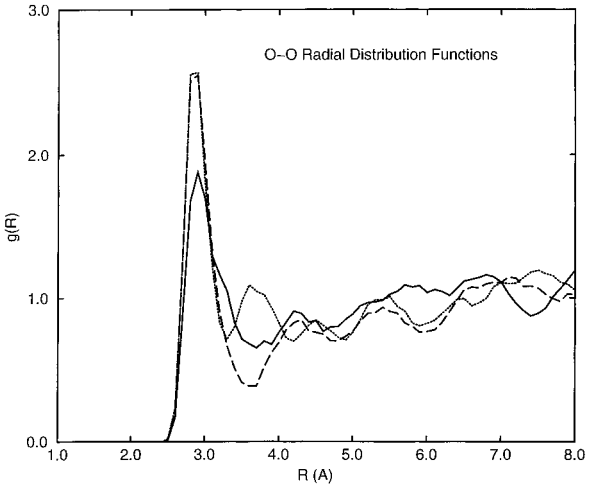


Figure 9. Computed O−O rdf's.

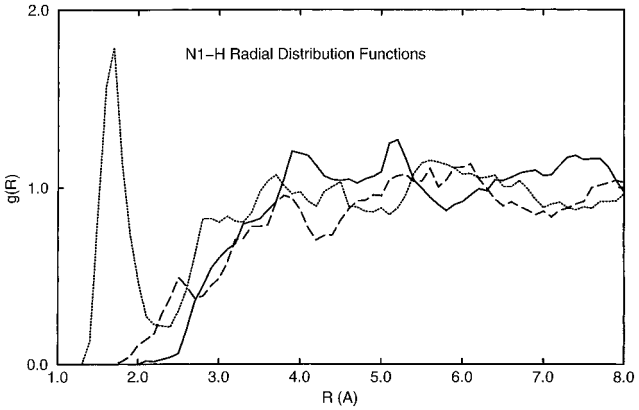


Figure 10. Computed N_1-H rdf's.

coordination numbers. All rdf's have been normalized to the bulk density for atom y. O_2 and O_3 are carboxylate oxygens cis and trans to the carbonyl oxygen, respectively. Figures 4−7 show the carboxylate oxygen−water hydrogen and oxygen rdf's. The first peaks reflect hydrogen bonding of water to the carboxylate oxygens. For the reactant, the hydrogen-bonding interactions are clearly reflected by the sharp peaks in the rdfs. On the other hand, the hydrogen-bonding peaks are significantly reduced for the TS and product, indicating that the hydrogen bond strength is weakened. Integration of the O_2-H and O_3-H first peaks to the minima at 2.6 Å yields 3.8 and 3.2 hydrogen bonds for the reactant, 2.0 and 1.5 hydrogen bonds for the TS, and 1.2 and 0.4 hydrogen bonds for the product. Thus, the total numbers of nearest neighbors around the CO_2 unit are 7.0, 3.5, and 1.6 for the reactant, TS, and product, respectively.

The decrease in hydrogen bonds to the CO_2 group is accompanied by an increase in hydrogen bonds at the

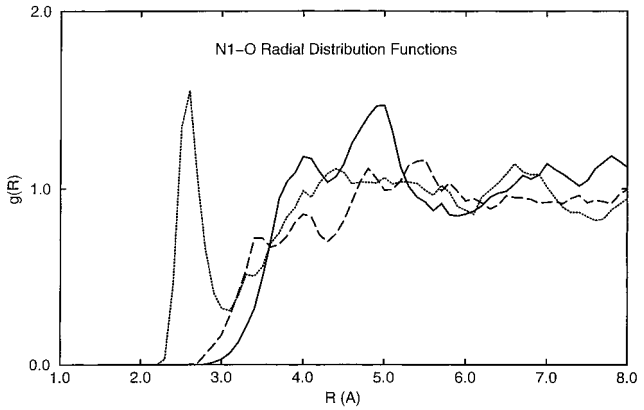


Figure 11. Computed N_1−O rdf's.

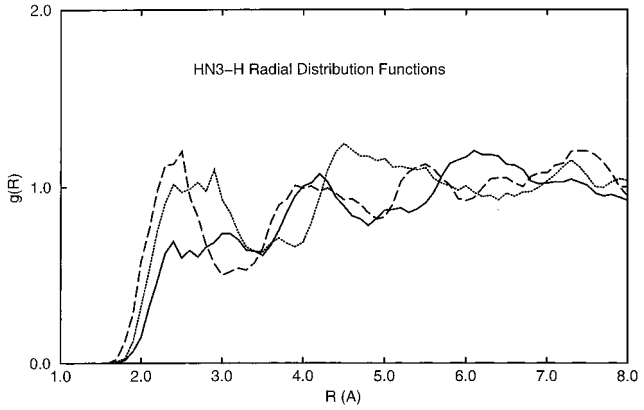


Figure 12. Computed H_{N3}−H rdf's.

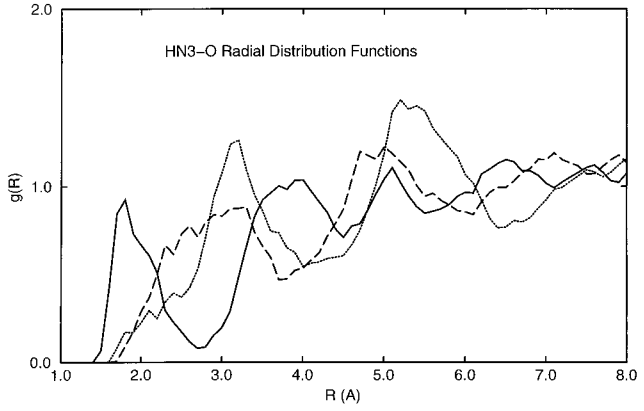


Figure 13. Computed H_{N3}−O rdf's.

imidazolidinone ring oxygen and N_1 nitrogen. Figures 8 and 9 show the ring carbonyl oxygen−water rdf's. Integration of the O−H first peaks to the minima at 2.5 Å yields 2.7, 3.9, and 4.1 hydrogen bonds for the reactant, TS, and product, respectively. It is interesting to note that in the crystal structure of biotin−avidin complex,[34] as well as from molecular dynamics simulation study of biotin,[4e] the ureido oxygen was found to form three hydrogen bonds. In addition, the N−H and N−O rdf's exhibit interesting features in hydrogen bonding patterns. The number of hydrogen bonds are invariant within computational error, by donating one hydrogen bond to water throughout the reaction (Figures 12, 13). However, the number of hydrogen bonds at the N_1 position changes from 0 for the reactant, to 1.1 for the TS, and to 2.0 for the product. In this case, N_1 acts as a hydrogen bond acceptor from the solvent as a negative partial charge is developed during the reaction. These features are reflected by the short N_1−H peaks in Figure 10. These observations are consistent with the energetic patterns shown in Figure 3. The loss of hydrogen bonds to the carboxylate group is reflected by the decrease in the peak at −10 to −15 kcal/mol. This is accompanied by the appearance of a more energetic peak at −15 to −20 kcal/mol, attributed to hydrogen bonding interactions between imidazolidinone anion and water. The stronger polarization effect for the reactant than the TS,

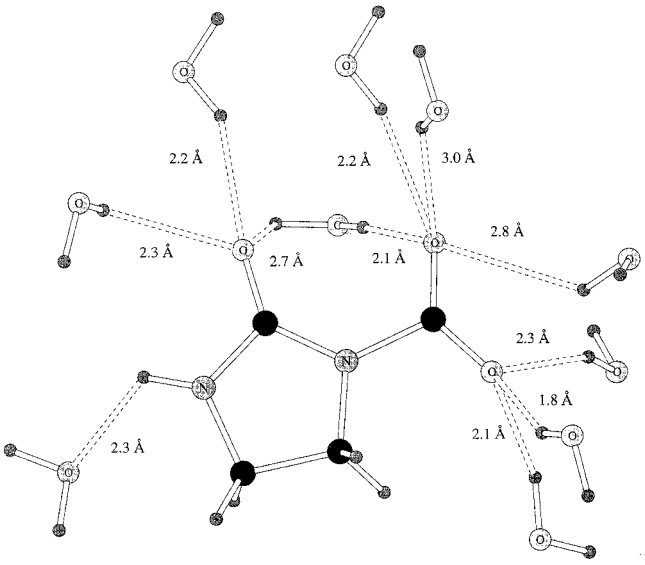


Figure 14. Illustration of one configuration from the simulation of reactant in water. All hydrogen bonds are indicated.

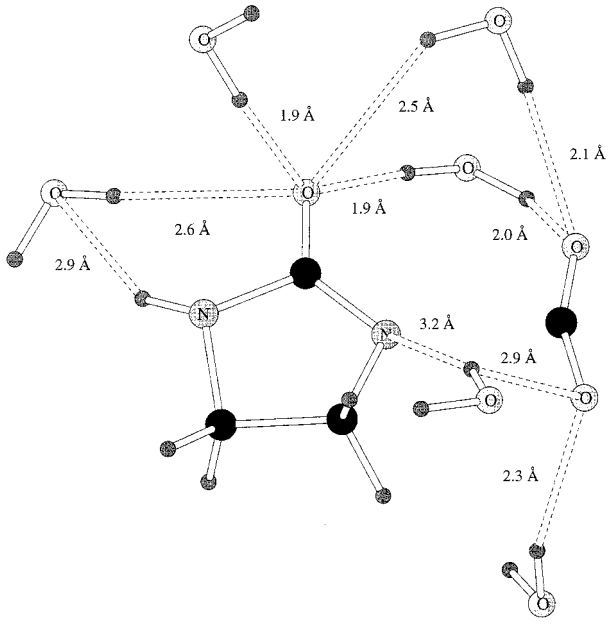


Figure 15. Illustration of one configuration from the simulation of transition state in water. All hydrogen bonds are indicated.

(34) (a) Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R. *Science* **1989**, *243*, 85. (b) Weber, P. C.; Wendoloski, J. J.; Pantoliano, M. W.; Salemme, F. R.; *J. Am. Chem. Soc.* **1992**, *114*, 3197. (c) Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M. *J. Mol. Biol.* **1993**, *231*, 698.

and the approximately one hydrogen bond loss in going from the reactant to the TS, results in a transition state destabilization effect which is the principal factor for the rate retardation of this decarboxylation reaction in water.

In closing, plots of the final configurations from simulations of reactant and TS are shown in Figures 14 and 15. For clarity, only the water molecules which are hydrogen bonded to the solute are shown with the hydrogen bond O···H distances labeled in the plots. It is interesting to see that two water molecules form four hydrogen bonds with the solute in a bifurcated fashion in the transition state due to the appropriate distance between the two acceptor oxygens, while there is only one such type of configuration for the reactant.

The present work has provided a detailed analysis of solvent effect on the title decarboxylation reaction in aqueous solution. The high activation energy in aqueous solution is a result of better solvation of the reactant by water. However, it should be noted that, in a recent paper,[35] Warshel and Florian challenged the notion that enzymatic catalysis originates from the ground state destabilization due to the nonpolar character of the enzyme active sites; instead, they proposed that the active sites of enzymes are in fact very polar and heterogeneous, and the origin of enzyme catalysis is due to a large electrostatic stabilization of the transition state, and the catalytic power of enzymes is stored in their folding energy in the form of the preorganized polar environment and not in the enzyme substrate interaction.

(35) Warshel, A.; Florian, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950.

Conclusions

The decarboxylation reaction of the *N*-carboxy-2-imidazolidinone anion in aqueous solution has been studied with the Monte Carlo simulation method using the combined AM1/TIP3P potential. In the gas phase, the calculated activation free energy $\Delta G^{\ddagger}$ is 14.2 kcal/mol at the RHF/6-31+G(d)//6-31+G(d) level, which is reduced to 10.5 kcal/mol if electron correlation effects are included at the MP2/6-31+G(d)//6-31+G(d) level. Combined with the MP2 gas-phase activation free energy and the solvation free energy difference between the ground state and the transition state obtained from Monte Carlo simulations, the barrier height in water is predicted to be 22.7 $\pm$ 0.2 kcal/mol, which is in excellent agreement with the experimental activation free energy of 23.2 kcal/mol. It was also found that solvation has little effect on the location of the transition state despite the dramatic increase in $\Delta G^{\ddagger}$ in aqueous solution. The origin of the increase in $\Delta G^{\ddagger}$ in aqueous solution was attributed to the intrinsic charge redistribution for the reactant and transition state, which leads to increased hydrogen-bonding interactions for the ground state, as well as more significant polarization effects for the reactant than the transition state. Understanding of chemical reactions in solution at the atomic level will help in the design of selective inhibitors or biomimetic catalysts such as artificial antibodies and enzymes.

Acknowledgment. This work is partially supported by NATO (CRG941209).

JO981523G