Monte Carlo Investigations of Solvent Effects on the Decarboxylation Reaction of Neutral *N*-Carboxy-2-imidazolidinone in Aqueous Solution

Daqing Gao and Yuh-Kang Pan*,†

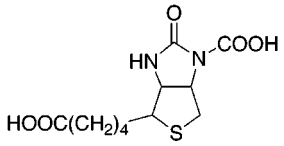
Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02467

Received March 12, 1999

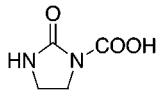
The solvent effects on the decarboxylation reaction of neutral *N*-carboxy-2-imidazolidinone in aqueous solution have been investigated by a combined quantum mechanical and molecular mechanical (QM/MM) Monte Carlo simulation method. In the present approach, the gas-phase intrinsic reaction coordinate of the reaction was 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 was determined via statistical perturbation theory using the combined QM/MM-AM1/TIP3P potential in Monte Carlo simulations. A two-stage mechanism consisting of the intramolecular proton transfer and N−C bond cleavage with the N−C bond breaking as the rate-limiting step was found. The computed free energy of activation in water is 21.0 ± 0.2 kcal/mol, in good agreement with the experimental value of 20.7 kcal/mol. Analyses of the structural and energetic nature of the differential solvation along the reaction coordinate 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. Over the past four decades, there have been extensive experimental studies in characterizing biotin enzymes and understanding the function of biotin in these enzymes.[2] In addition, a number of theoretical studies have been performed toward understanding biotin chemistry both in the gas phase and in aqueous solution at the atomic level,[3] including molecular orbital analyses of the sulfur atom participation in biotin-catalyzed carboxylations and a molecular dynamics simulation study of the binding of the high-affinity streptavidin−biotin complex.

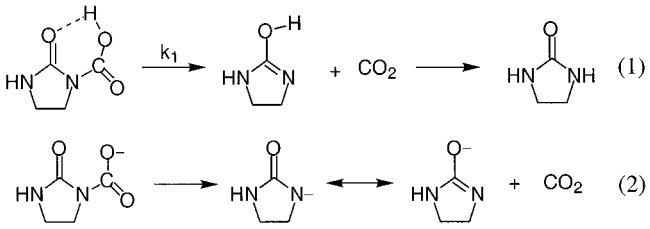


N(1')-carboxybiotin



1

Scheme 1



In a previous paper,[4] we reported the results for the decarboxylation reaction of the model compound *N*-carboxy-2-imidazolidinone *anion* in aqueous solution from a combined quantum mechanics and molecular mechanics (QM/MM) Monte Carlo simulation study. Our calculated free energy of activation, 22.7 ± 0.2 kcal/mol, is in excellent agreement with the experimental value of 23.2 kcal/mol obtained by Rahil et al. for the same model reaction.[5] Through these studies, insights on the energetics and factors that govern the dramatic solvent effects on the anionic process in aqueous solution have been obtained.

The decarboxylation of *N*-carboxy-2-imidazolidinone (**1**) in solution proceeds via two mechanisms, depending on the pH of the solution (Scheme 1).[5−8] Experimentally, at pH > 8, the rate of the reaction is independent of the pH,[5] consistent with a loss of carbon dioxide from the ionized carboxylate form (eq 2), which was modeled

† Tel: 617-552-3631. Fax: 617-552-2705. e-mail: yuhkang.pan@bc.edu.

(1) For a thorough review, see: (a) Knowles, J. R. *Annu. Rev. Biochem.* **1989**, *58*, 195. (b) *Annals of the New York Academy of Sciences*; Dakshinamurti, K., Bhagavan, H. N., Eds.; New York Academy of Sciences: New York, 1985; Vol. 447.

(2) (a) Lynen, F.; Knappe, J.; Lorch, E.; Jütting, G.; Ringelmann, E. *Angew. Chem.* **1959**, *71*, 481. (b) Kaziro, Y.; Hase, L. F.; Boyer, P. D.; Ochoa, S. *J. Biol. Chem.* **1962**, *237*, 1460. (c) Hegarty, A. F.; Pratt, R. F.; Guidici, T.; Bruice, T. C. *J. Am. Chem. Soc.* **1971**, *93*, 1428. (d) DeTitta, G. T.; Edmonds, J. W.; Stallings, W.; Donohue, J. *J. Am. Chem. Soc.* **1976**, *98*, 1920. (e) Stubbe, J.; Fish, S.; Abeles, R. H. *J. Biol. Chem.* **1980**, *255*, 236. (f) Fry, D. C.; Fox, T. L.; Lane, M. D.; Mildvan, A. S. *J. Am. Chem. Soc.* **1985**, *107*, 7659. (g) Berkessel, A.; Breslow, R. *Bioorg. Chem.* **1986**, *14*, 249. (h) Perrin, C. L.; Tammy, J. D. *J. Am. Chem. Soc.* **1987**, *109*, 5163. (i) Waldrop, G. L.; Rayment, I.; Holden, H. M. *Biochemistry* **1994**, *33*, 10249.

(3) (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*, 226.

(4) Gao, D.; Pan, Y.-K. *J. Org. Chem.* **1999**, *64*, 1159.

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

(6) Tipton, P. A.; Cleland, W. W. *J. Am. Chem. Soc.* **1988**, *110*, 5866.

(7) Attwood, P. V.; Tipton, P. A.; Cleland, W. W. *Biochemistry* **1986**, *25*, 8197.

(8) (a) Caplow, M. *J. Am. Chem. Soc.* **1965**, *87*, 5774. (b) Caplow, M.; Yager, M. *J. Am. Chem. Soc.* **1967**, *89*, 4513.

10.1021/jo9904563 CCC: $18.00
Published on Web 05/12/1999

previously.[4] Below pH 8, the observed rate increases with increased acidity, until a plateau is reached at pH 4, suggesting a reaction via a neutral transition state.[5] The acidity constant of **1** was determined to be $pK_a = 4.2$ by Rahil et al.[5]

In view of the importance of decarboxylation reactions in a variety of biological processes,[9,10] and to corroborate recent experimental studies of decarboxylation reactions of carboxybiotin and the model compounds,[5,6,8] it is essential to provide a computational assessment of solvent effects on decarboxylation reactions both in its anionic[4] and neutral form. In this paper, we report a combined QM/MM study of the decarboxylation reaction of the neutral *N*-carboxy-2-imidazolidinone in water. Valuable insights into the proton transfer and the N−C bond breaking processes, and the influence of hydration on the nature of the transition state, have been obtained. We have quantified and characterized all four possible unique conformers of acid **1** in the reactant state in aqueous solution. The calculated activation free energy of 21.0 ± 0.2 kcal/mol is in good agreement with the experimental value of 20.7 kcal/mol. In this case, a dramatic solvation process, which is markedly different from that of the decarboxylation of the corresponding anionic form of *N*-carboxy-2-imidazolidinone,[4] was observed along the reaction pathway. This study has also provided a clear picture of the nature of the transition state for decarboxylation reactions of neutral β-keto carboxylic acids in water, which has been highly controversial for over 60 years with regard to the proton transfer and polarity of the transition state.[10d,e,11] In this study, we adopt a "dual-level" computational approach. First, high-level ab initio calculations were carried out to dissect the intrinsic reactivity of the decarboxylation reaction in the gas phase. Solvent effects were 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. Transition structures (TS) and the intrinsic reaction coordinate (IRC)[12] path for the decarboxylation reaction of neutral *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.[13,14] A total of 69 structures were generated along the IRC, spanning the reaction coordinate, with the breaking $N-CO_2$ bond distance (R_{N-C}) varying from 1.39 to 2.90 Å. For convenience in the following discussion, we use this characteristic distance to denote the reaction coordinate, R_c, rather than the standard IRC values. Transition structures and stationary geometries were verified by vibrational frequency calculations with one and zero imaginary frequencies, respectively. Electron correlation effects were included through single-point energy 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).

Gas-phase enthalpy changes at 298 K are determined by eq 3:

$$\Delta H^{298} = \Delta E_e^0 + \Delta E_{vib}^0 + \Delta(\Delta E_{vib})^{298} + \Delta E_t^{298} + \Delta E_r^{298} + \Delta PV \quad (3)$$

where ΔE_e^0 is the electronic energy change, E_{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 and rotational energies and the work term, which are all computed classically using standard methods.[15] To compute the enthalpy change ΔH^{298}, harmonic frequencies were scaled by a factor of 0.90 to correct for the overestimation of vibrational frequencies at the HF level.[16] Scaled frequencies below 500 cm^{-1} were treated as classical rotations in calculating their vibrational contributions.

Solvent effects were also examined with the Onsager reaction field method and Tomasi's polarizable continuum model at the HF/6-31+G(d) and B3LYP/6-31+G(d) levels, respectively, using the gas-phase HF/6-31+G(d) optimized structures.[13]

QM/MM Potential. To investigate the solvent effect on the decarboxylation reactions in water, a combined quantum mechanical and molecular mechanical potential is used to describe solute−solvent interactions.[17−22] In the present approach, the solute molecule is treated by the semiempirical AM1 model,[23] while the solvent is approximated by the three-point, TIP3P model for water.[24] Thus, the effective Hamilto-

(9) (a) Kemp, D. S.; Cox, D. D.; Paul, K. G. *J. Am. Chem. Soc.* **1975**, *97*, 7312. (b) Tarasow, T. M.; Lewis, C.; Hilvert, D. *J. Am. Chem. Soc.* **1994**, *116*, 7959. (c) Smiley, J. A.; Benkovic, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 3877. (d) Sun, S.; Smith, G. S.; O'Leary, M. H.; Schowen, R. L. *J. Am. Chem. Soc.* **1997**, *119*, 1507. (e) Miller, B. G.; Traut, T. W.; Wolfenden, R. *J. Am. Chem. Soc.* **1998**, *120*, 2666.

(10) (a) Zipse, H.; Apaydin, G.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 8608. (b) Gao, J. *J. Am. Chem. Soc.* **1995**, *117*, 8600. (c) Huang, C.-L.; Wu, C.-C.; Lien, M.-H. *J. Phys. Chem. A* **1997**, *101*, 7867. (d) Bach, R. D.; Canepa, C. *J. Org. Chem.* **1996**, *61*, 6346. (e) Bach, R. D.; Canepa, C. *J. Am. Chem. Soc.* **1997**, *119*, 11725. (f) Lee, J. K.; Houk, K. *Science* **1997**, *276*, 942.

(11) (a) Westheimer, F. H. *Tetrahedron* **1995**, *51*, 3. (b) Logue, M. W.; Pollack, R. M.; Vitullo, V. P. *J. Am. Chem. Soc.* **1975**, *97*, 6868. (c) Straub, T. S.; Bender, M. L. *J. Am. Chem. Soc.* **1972**, *94*, 8881. (d) Brower, K. R.; Gay, B.; Konkol, T. L. *J. Am. Chem. Soc.* **1966**, *88*, 1681. (e) Bigley, D. B.; Thurman, J. C. *J. Chem. Soc. B* **1968**, 436. (f) Swain, C. G.; Bader, R. F. W.; Esteve, R. M., Jr.; Griffin, R. N. *J. Am. Chem. Soc.* **1961**, *83*, 1951. (g) Steinberger, R.; Westheimer, F. H. *J. Am. Chem. Soc.* **1951**, *73*, 429. (h) Westheimer, F. H.; Jones, W. A. *J. Am. Chem. Soc.* **1941**, *63*, 3283. (i) Pedersen, K. J. *J. Am. Chem. Soc.* **1938**, *60*, 595. (j) Pedersen. K. J. *J. Phys. Chem.* **1934**, *38*, 559.

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

(13) (a) 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. (b) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. L.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision E.2; Gaussian, Inc.: Pittsburgh, PA, 1995.

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

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

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

(17) (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.

(18) (a) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (b) Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 3448. (c) Tapia, O.; Colonna, F.; Angyan, J. G. *J. Chem. Phys. Phys.-Chim. Biol.* **1990**, *87*, 875. (d) Thole, B. T.; van Duijnen, P. Th. *Chem. Phys.* **1982**, *71*, 211.

(19) (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.

nian of the hybrid system is given by

$$\hat{H}_{\text{eff}} = \hat{H}^0 + \hat{H}_{\text{qm/mm}} + \hat{H}_{\text{mm}} \quad (4)$$

where $\hat{H}^0$ is the Hamiltonian for the solute molecule in the gas phase, $\hat{H}_{\text{mm}}$ is the molecular mechanics interaction energy for the solvent, and $\hat{H}_{\text{qm/mm}}$ is the solute−solvent interaction Hamiltonian. The combined AM1/MM approach has been extensively used and shown to yield good solvation results for a variety of systems,[19−21] including the decarboxylation reaction of the anionic form of *N*-carboxy-2-imidazolidinone in aqueous solution.[4] The technical details of the combined QM/MM potential have been summarized in several articles; interested readers are referred to refs 17−22 for additional details of the method.

Monte Carlo Simulations. Statistical mechanical Monte Carlo calculations have been carried out in the isothermal−isobaric (NPT) ensemble at 25 °C and 1 atm using the combined AM1/TIP3P potential. In these simulations, the N−CO_2 bond that breaks along the reaction coordinate was oriented to coincide with the *Z*-axis of a rectangular box with periodical boundary conditions[25] (approximately 20 × 20 × 30 Å^3). A total of 391 TIP3P water molecules are included in the fluid simulations. Twenty-seven out of the 69 structures generated by the ab initio IRC path calculations, plus an additional 7 structures that extended the reaction coordinate to R_c = 4.3 Å from the last IRC strucutre at R_c = 2.9 Å, were selected. Differences in hydration free energy were calculated via statistical perturbation theory using the Zwanzig equation with double-wide sampling.[26,27] In eq 5, R_i denotes a structure

$$\Delta G(R_c) = -k_B T \ln\langle \exp\{-[E(R_j) - E(R_i)]/k_B T\}\rangle_{R_i} \quad (5)$$

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 5 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 5 does not account for nonequilibrium solvation effects. Approaches that allow consideration of the activation barrier in the solute−solvent coordinate space have been used by Warshel and co-workers.[28] It would be of interest to investigate nonequilibrium solvation effects on decarboxylation reactions in future studies.

The intermolecular interactions were truncated between spherical cutoff distances of 8.5 and 9.0 Å for water−water interaction based on the oxygen separations. Solute−solvent interactions were feathered to zero between 9.0 and 9.5 Å based on solute atom−water oxygen distances. Standard Metropolis sampling was used.[29] 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.[30] 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 33 simulations were performed to span the entire reaction coordinate (1.39−4.3 Å). Each simulation consisted of at least 8 × 10^5 configurations of equilibration, followed by an additional 1.5 × 10^6 configurations of averaging. Uncertainties were computed from fluctuations in separate averages over batches of 1 × 10^5 configurations.

Tautomerization Free Energy Calculations. To examine the relative stabilities of different tautomers in aqueous solution, free energy perturbation calculations were performed to obtain the hydration free energy differences in aqueous solution according to eq 6:

$$\Delta\Delta G_{\text{hyd}}(A \rightarrow B) = \Delta G^{\text{el}}(A) - \Delta G^{\text{el}}(B) + \Delta G^{\text{vdw}}(A \rightarrow B) \quad (6)$$

where $\Delta G^{\text{el}}(A)$ and $\Delta G^{\text{el}}(B)$ are the free energy changes from electrostatic decoupling for molecule A and molecule B in which the electrostatic term of the QM/MM interaction Hamiltonian is annihilated[21] and $\Delta G^{\text{vdw}}(A \rightarrow B)$ is the free energy change of the van der Waals components between A and B. We have calculated the free energy changes of **3** to **2**, **4** to **2**, and **11** to **10** in water (Figure 1). For each molecule, 11 simulations were executed in the electrostatic decoupling step which consisted of 10^6 configurations of equilibration and 10^6 configurations of data collection, while five simulations were carried out to interconvert the "van der Waals tautomers". van der Waals parameters used in this work can be found in ref 21a.

All calculations were performed using the MCQUB/BOSS program provided by Professor Gao at SUNY, Buffalo.[31]

Results and Discussion

(a) Ab Initio Structures and Relative Energies in Water. The reactant molecule **1** may adopt four local conformations, **2**, **3**, **4**, **5**, with the hydroxy group being either in the syn or in the anti position as shown in Figure 1. The calculated energetic results of different tautomers in the gas phase and in aqueous solution are listed in Table 1. The anti conformer **2** is stabilized by an intramolecular hydrogen bond to the carbonyl oxygen and is predicted to be 4.8, 7.2, and 18.1 kcal/mol at the HF/6-31+G(d) level and 5.3, 6.8, and 16.8 kcal/mol at the MP2/6-31+G(d)//HF/6-31+G(d) level lower in energy than conformers **3**, **4**, and **5**, respectively. It can be seen that electron correlation does not play an important role for the relative energies of these tautomers. In combination with the RHF/6-31+G* vibrational frequencies and the HF/6-31+G* and MP2/6-31+G* energies, the free energy differences between **2** and **3** are predicted to be 4.1 and 4.6 kcal/mol, respectively. Free energies of structures **4** and **5** are also 6.3 and 15.5 kcal/mol higher than those of **2** at the highest level. Thus, **2** is the lowest energy conformer in the gas phase in all ab initio calculations

(20) (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. (j) Ho, L. L.; MacKerell, A. D., Jr.; Bash, P. A. *J. Phys. Chem.* **1996**, *100*, 4466. (k) Gao, J. *J. Comput. Chem.* **1997**, *18*, 1061 (l) Alhambra, C.; Wu, L.; Zhang, Z.-Y.; Gao, J. *J. Am. Chem. Soc.* **1998**, *120*, 3858. (m) Antonczak, S.; Monard, G.; Ruiz-López, M. F.; Rivail, J.-L. *J. Am. Chem. Soc.* **1998**, *120*, 8825.

(21) (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.

(22) Kaminski, G. A.; Jorgensen, W. L. *J. Phys. Chem. B* **1998**, *102*, 1787.

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

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

(25) (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.

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

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

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

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

(30) (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.

(31) (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.

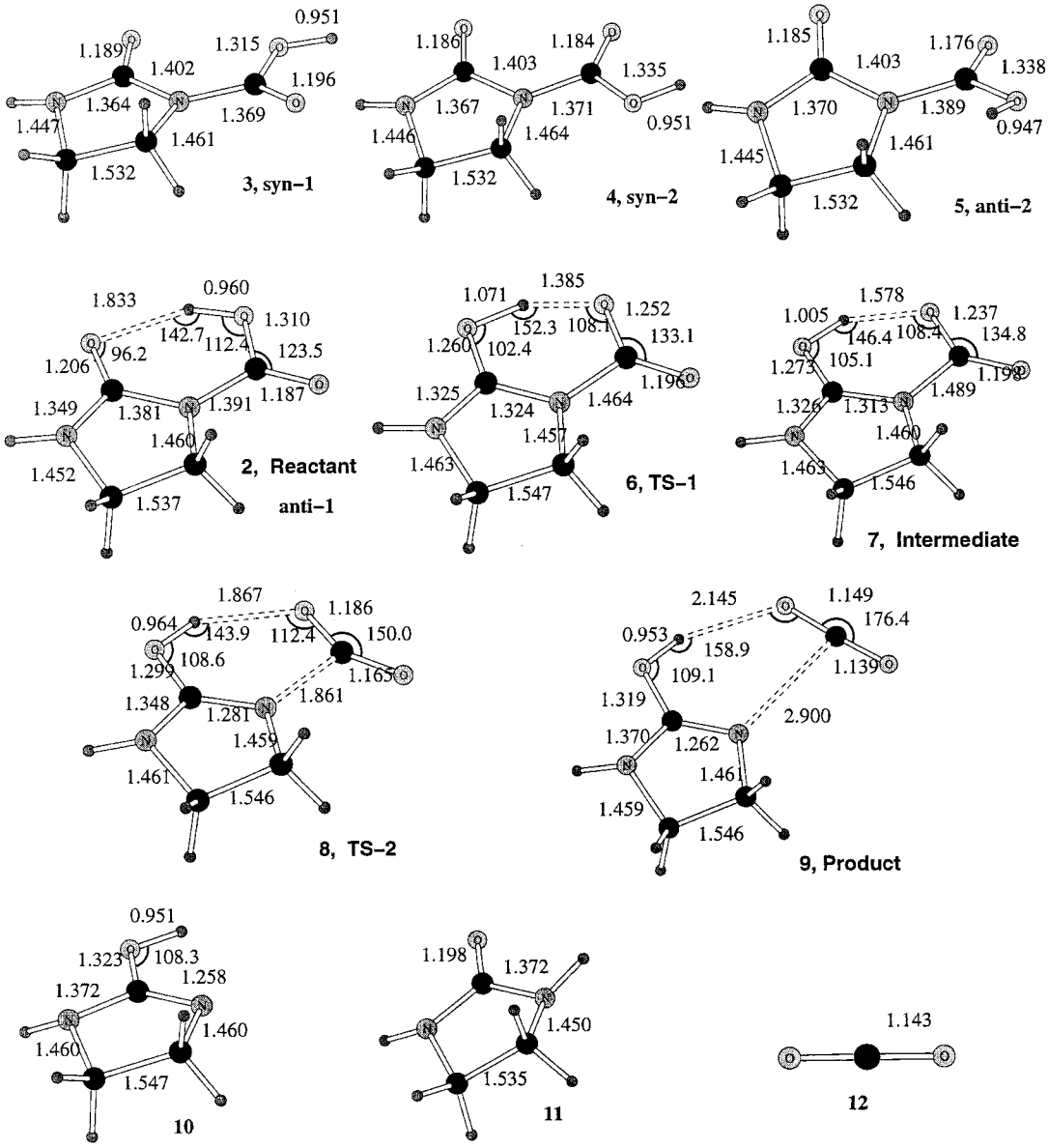


Figure 1. RHF/6-31+G(d) optimized structures. Bond lengths are in angstroms; angles are in degrees.

Table 1. Energy (kcal/mol) and Entropy (cal/mol·K) Changes for Tautomerizations

	2 → 3	**2 → 4**	**2 → 5**	**10 → 11**
		Gas Phase		
ΔE^0(AM1)	−0.2	0.3	6.6	−18.2
ΔE^0(HF/6-31+G(d))	4.8	7.2	18.1	−21.0
ΔH^{298}(HF/6-31+G(d))	4.5	7.1	17.5	−20.2
ΔS^{298}(HF/6-31+G(d))	1.2	1.3	2.3	0.0
ΔG^{298}(HF/6-31+G(d))	4.1	6.7	16.9	−20.2
ΔE^0(MP2/6-31+G(d))	5.3	6.8	16.8	−18.8
ΔH^{298}(MP2/6-31+G(d))	5.0	6.7	16.2	−18.3
ΔG^{298}(MP2/6-31+G(d))	4.6	6.3	15.5	−18.3
		Aqueous Solution		
ΔE_{aq}(HF/6-31+G(d), SCRF)	8.6	13.0	23.9	−23.9
ΔG_{aq}(HF/6-31+G(d), PCM)	3.1	4.8	12.4	−24.1
ΔG_{aq}(B3LYP/6-31+G(d),PCM)	4.5	6.1	13.5	−21.5
$\Delta\Delta G_{hyd}$(QM/MM)	−1.5 ± 0.3	−2.3 ± 0.3		−3.6 ± 0.2
ΔG_{aq}(QM/MM)	3.1 ± 0.3	3.9 ± 0.3		−14.7 ± 0.2

except that the AM1 Hamiltonian predicts that **3** is slightly preferred over **2** by 0.2 kcal/mol in electronic energies. For simple carboxylic acids, such as formic and acetic acid, the syn conformer is typically 5−6 kcal/mol more stable than the anti form in the gas phase,[32] although the energy difference is attenuated in solution thanks to the favorable solvation effect for the anti conformation.[33] The preference for the anti conformation

(32) (a) Li, Y.; Houk, K. N. *J. Am. Chem. Soc.* **1989**, *111*, 4505. (b) Wang, X.; Houk, K. N. *J. Am. Chem. Soc.* **1988**, *110*, 1870. (c) Wiberg, K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1987**, *109*, 5935. (d) Wiberg. K. B.; Laidig, K. E. *J. Am. Chem. Soc.* **1988**, *110*, 1872.

in **1** emphasizes the important role of intramolecular hydrogen bonding in the gas phase, which is more than enough to offset the energy difference between the syn and anti conformers of the carboxylic acid.

Solvent effects on the conformational equilibrium between the syn and anti forms in **1** were examined using the Onsager reaction field method (SCRF)[34] and Tomasi's polarizable continuum model (PCM)[35] that have been implemented in Gaussian 94. A dielectric constant of 78.3 was used in these calculations. The polarizable continuum solvation model predicted only a small shift in conformation energy differences between **2** and **3**, changing from 4.8 kcal/mol in the gas phase to 3.1 kcal/mol at the HF/6-31+G(d) level in a continuum dielectric medium (water), while the Onsager SCRF model further exaggerates the energy differences in comparison with the gas-phase results.

Our extensive QM/MM Monte Carlo simulation calculations predict that the syn conformer (**3**) is better solvated than **2** by -1.5 ± 0.3 kcal/mol and that **4** is better solvated by -2.3 ± 0.3 kcal/mol. In combination with the gas-phase free energy differences ΔG^{298} determined at the MP2/6-31+G* level, our best estimate of the free energy difference, ΔG_{aq}(QM/MM) is 3.1 ± 0.3 kcal/mol between **2** and **3** and 3.9 ± 0.3 kcal/mol between **2** and **4** in aqueous solution. These results turn out to be in excellent agreement with the predictions from the continuum models (PCM).

These results confirm that the anti conformation **2** with the presence of an intramolecular hydrogen bond is preferred in aqueous solution. This finding is consistent with the experimental interpretation that the decarboxylation reactions of both neutral carboxybiotin and *N*-carboxy-2-imidazolidinone proceed through an intramolecular proton transfer to the ureido oxygen, leading to the enol form (**10**) of the product (Scheme 1).[6−8] In the experimental study of the mechanism of decarboxylation of carboxybiotin, Cleland et al. obtained a D_2O solvent isotope effect of 1.6 and a linear proton inventory, which shows that only one proton is in flight in the transition state.[6]

An alternative mechanism is a solvent-assisted process, in which the carboxyl proton is transferred to a solvent water molecule and a proton migrates to the ureido oxygen from the bulk solvent. In this case, the syn conformation of the carboxylic acid would be preferred for this stepwise proton-transfer process. However, the present ab initio molecular orbital calculations indicate that the syn conformation is less stable both in the gas phase and in solution. It appears that the stepwise, general acid−general base catalysis mechanism will not be as competitive as the intramolecular proton transfer. Consequently, our simulation studies will be focused on the sequence depicted in Scheme 1.

At the HF/6-31+G(d) level, the decarboxylation reaction of *N*-carboxy-2-imidazolidinone is found to proceed via a stepwise process with the initial proton transfer followed by the departure of carbon dioxide (**2** → **6** → **7** → **8** → **9**, Figure 1). Vibrational frequency calculations confirmed that both the proton transfer (**TS-1**) and decarboxylation (**TS-2**) saddle points are true transition states, and the IRC reaction-path-following calculations from these two transition states lead to the corresponding reactant (**2**), to a product complex (**9**), and to a common intermediate structure (**7**, **INT**). The intermediate is only 0.3 kcal/mol lower in energy than **TS-1** (**6**), whereas inclusion of electron correlation at fixed RHF/6-31+G(d) geometries increases its energy to 1.3 kcal/mol and free energy to 2.5 kcal/mol above **TS-1** at the MP2/6-31+G(d) level. Thus, it is not clear if this intermediate structure will survive geometry optimizations at a higher level of theory with inclusion of electron correlations. However, as will become clear *a posteriori* from our simulation results, the proton transfer and the decarboxylation processes in the overall reaction indeed follow two separate reaction paths. Figure 2 illustrates the two critical geometry variations along the proton-transfer coordinate and the decarboxylation coordinate. As the proton transfer takes place, the N−CO_2 distance is only increased by 0.1 Å from 1.391 Å in reactant **2** to 1.489 Å in intermediate **7**. On the other hand, the O−H distance varied only by 0.05 Å as the N−C bond is elongated and cleaved to a final distance of 2.900 Å in the product complex **9**. To reflect these structural features in the overall decarboxylation reaction of **2**, we decided to use (C═)O−H and N−CO_2 distances separately to depict the free energy change as a function of the reaction coordinate with the intermediate structure **7** as the switching point for these two distances (Figure 3).

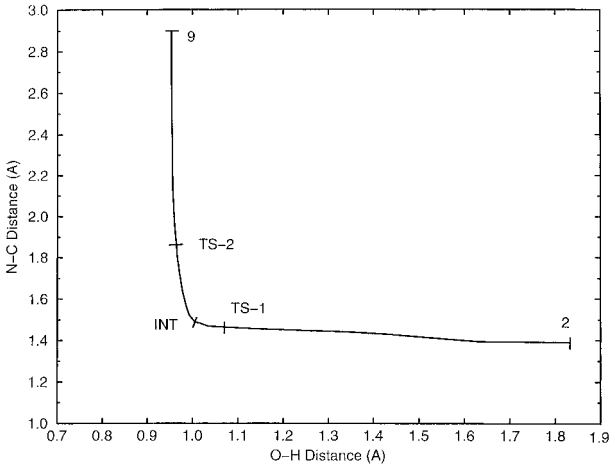


Figure 2. Comparison of the N−C bond distances vs the ureido oxygen (C═)O−H distances in the series **2** to **9**.

Gas-phase energetic results from ab initio calculations are summarized in Table 2. The activation energy for the first proton-transfer step is 19.4 kcal/mol at the HF/6-31+G(d) level. The activation energy for the second, decarboxylation step is 5.8 kcal/mol from intermediate **7**, giving rise to an overall activation energy of 24.9 kcal/mol (**TS-2** relative to reactant **2**). The AM1 Hamiltonian gives an activation energy $\Delta E^{\neq}$ of 29.3 kcal/mol with the HF/6-31+G(d) geometries. Inclusion of correlation effects at the MP2/6-31+G(d)//HF/6-31+G(d) level reduces the proton transfer and the overall barrier height to 14.0 and 19.0 kcal/mol, respectively. In combination with the RHF/6-31+G(d) vibrational frequencies and the MP2/6-31+G(d) electronic energies, the computed *overall* activation enthalpy, $\Delta H^{\neq}$, activation entropy, $\Delta S^{\neq}$, and Gibbs activation free energy, $\Delta G^{\neq}$ for the reaction at 298 K are

(33) (a) Gao, J.; Pavelites, J. J. *J. Am. Chem. Soc.* **1992**, *114*, 1912. (b) Tadayoni, B. M.; Rebek, J. Jr. *Bioog. Med. Chem. Lett.* **1991**, *1*, 13.

(34) (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.

(35) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem Phys.* **1981**, *55*, 117. (b) Wiberg, K. B.; Castejon, H.; Keith, T. A. *J. Comput. Chem.* **1996**, *17*, 185.

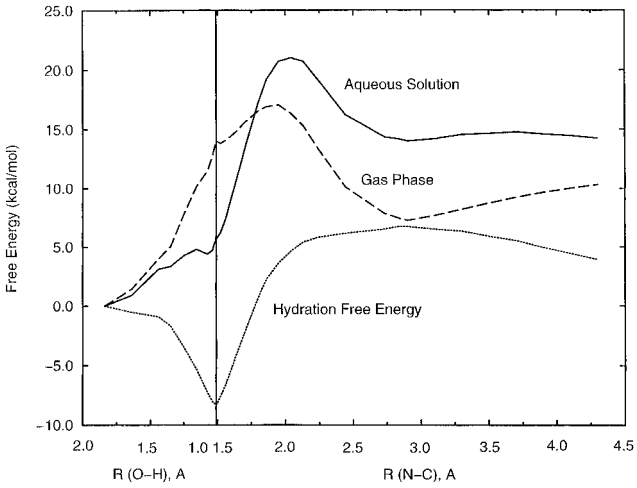


Figure 3. Computed free energy profile in the gas phase (dashed curve), hydration free energy (dotted curve), and the potential of mean force in aqueous solution (solid curve) for the decarboxylation reaction of neutral *N*-carboxy-2-imidazolidinone. The reaction coordinate is represented as the (C═)O−H distance and the N−C distance with the intermediate structure **7** as the switching point along the 6-31+G(d) intrinsic reaction coordinate.

17.4 kcal/mol, 1.5 cal/mol·K, and 16.9 kcal/mol, respectively. Experimental activation parameters for the decarboxylation reaction in the gas phase are not available for direct comparison, although $\Delta H^{\neq}$ values of 19.0 and 20.8 kcal/mol at pH of 6.1 and 10.2 in water have been obtained from the Eyring plots using Kluger's experimental rate constants.[36]

It should be noted that the calculated gas-phase activation free energy of 16.9 kcal/mol for the decarboxylation of neutral *N*-carboxy-2-imidazolidinone is 6.4 kcal/mol greater than the activation free energy (10.5 kcal/mol) for the decarboxylation of its conjugate base, the carboxylate anion.[4] However, the experimental decarboxylation free energy barrier for the neutral form is 2.5 kcal/mol lower than that for the anionic form in aqueous solution (20.7 vs 23.2 kcal/mol). This indicates that hydration has a dramatic effect on the decarboxylation reactions of *N*-carboxy-2-imidazolidinone that reverses the order of reactivity of the neutral and anionic forms of the model compounds for biotin (see below).

The tautomerization free energy of the isolated enol intermediate **10** to the final product **11** is −18.3 kcal/mol (Table 1), which gives an overall exothermic reaction, with a free energy of reaction of −10.2 kcal/mol for **2** → **11** + **12** (CO_2) in the gas phase.

Experiments[6,8,37] have established that the decarboxylation reaction of both neutral carboxybiotin and neutral *N*-carboxy-2-imidazolidinone proceeds through an intramolecular proton transfer to the ureido oxygen, rather than to the N1 nitrogen, leading to the enol form product. This is followed by a rapid tautomerization of the enol form to the keto form as CO_2 diffuses away. The present computer simulation will focus on the solvent effects on the activation free energy for this rate-limiting step, leading to the enol product of the decarboxylation reaction of neutral *N*-carboxy-2-imidazolidinone in aqueous solution.

(b) Reaction 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[38] for the decarboxylation of neutral *N*-carboxy-2-imidazolidinone in water. This has been accomplished through a series of Monte Carlo free energy perturbation (FEP) simulations along the reaction path generated from the ab initio HF/6-31+G(d) calculations. Key findings are shown in Figure 3, which depicts the free energy profiles for the decarboxylation reaction in the gas phase and in aqueous solution as a function of the proton-transfer coordinate, the (C═)O−H distance, and the decarboxylation coordinate, the N−C distance. To illustrate the progression of the overall reaction, the two geometrical variables are anchored (switched) at the structure corresponding to the gas-phase intermediate (**7**). Thus, the first half of Figure 3 corresponds to the free energy change for the proton-transfer process where the N−CO_2 bond distance has little variation. The second half of Figure 3 shows the free energy profile for the decarboxylation step where the (C═)O−H distance is essentially unchanged. It should be emphasized that although we have switched the representation of the reaction coordinate, there is no discontinuity in the reaction profile in Figure 3. This choice is only used for clarity in the presentation and discussion. Note that the gas-phase free energy profile was constructed at the MP2/6-31+G(d)//HF/6-31+G(d) level.

The most dramatic feature in Figure 3 is that there are clearly two stages for the decarboxylation reaction in aqueous solution, although, as in the gas phase, the true identity and the existence of the putative intermediate is questionable. There has been controversy at the experimental level as to whether the proton is fully or partially transferred in the transition state. In the early experimental studies of *N*-carboxy-2-imidazolidinone, Caplow failed to observe a solvent deuterium isotope effect, suggesting that the proton transfer is virtually complete in the activated complex.[8b] Cleland's recent experiment on the decarboxylation of carboxybiotin did show a D_2O solvent isotope effect of 1.6 and a rather low ^{13}C isotope effect of 1.012, suggesting an early transition state in which proton motion leads N−C bond breaking.[6] In contrast to the gas-phase reaction profile, which has a significant $\Delta G^{\neq}$ for the proton transfer step (11.5 kcal/mol), followed by a relatively small barrier for the decarboxylation from the intermediate, the reaction profile in aqueous solution is characterized by an initial flat region that is only about 5 kcal/mol above the reactant state. As the proton-transfer step is completed, the decarboxylation process experiences a sharp increase in activation free energy. Thus, the overall activation free energy is estimated to be 21.0 ± 0.2 kcal/mol, combining the gas-phase MP2/6-31+G(d) results and the combined QM/MM solvation free energies. This result is in excellent agreement with the experimental value of 20.7 kcal/mol for the decarboxylation reaction of *N*-carboxy-2-imidazolidinone at pH = 6.1 in water.[5,36] The accord between the theoretical and experimental results for this neutral

(36) 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: $\Delta H^{\neq}$ = 20.8 kcal/mol, $\Delta S^{\neq}$ = −8.0 cal/mol·K, and $\Delta G^{\neq}$ = 23.2 kcal/mol at pH 10.2; $\Delta H^{\neq}$ = 19.0 kcal/mol, $\Delta S^{\neq}$ = −5.8 cal/mol·K, and $\Delta G^{\neq}$ = 20.7 kcal/mol at pH 6.1. Corrections: Rahil, J.; You, S.; Kluger, R. *J. Am. Chem. Soc.* **1998**, *120*, 2692.

(37) Olah, G. A.; White, A. M. *J. Am. Chem. Soc.* **1968**, *90*, 6087.

(38) (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.

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

	2 → TS-1, 6	**2 → INT, 7**	**2 → TS-2, 8**	**2 → 9**	**2 → 10 + 12**	**2 → 11 + 12**
ΔE^0(HF/6-31+G(d))	19.4	19.1	24.9	13.9	18.7	−2.3
ΔE_v^0	−2.8	−1.2	−1.6	−1.4	−3.1	−2.5
$\Delta\Delta E_v^{298}$	−0.2	0.1	0.0	0.4	−1.0	−1.2
ΔH^{298}	16.5	18.0	23.3	12.9	16.3	−3.9
ΔS^{298}	−1.2	0.6	1.5	11.5	30.0	30.3
ΔG^{298}	16.9	17.8	22.9	9.5	7.3	−12.9
ΔE^0(MP2/6+31+G(d))	14.0	15.3	19.0	12.1	19.2	0.4
ΔH^{298}(MP2/6+31+G(d))	11.1	14.2	17.4	11.1	17.2	−1.1
ΔG^{298}(MP2/6+31+G(d))	11.5	14.0	16.9	7.7	8.2	−10.2

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

and previous anionic decarboxylations demonstrated the reliability of the present dual-level computational approach, suggesting that the simulation results may be further analyzed to shed light on the factors that govern the solvent effects on this decarboxylation reaction.

(c) Analysis of Solvation Effects on the Decarboxylation. The free energy of the present reaction (up to $R_c = 4.3$ Å) in water is predicted to be 14.3 kcal/mol. The endothermicity is overcome by the subsequent enolization of **10** to **11** when CO_2 diffuses away. The tautomerization free energy (ΔG_{aq}) of **10** to **11** in water is −21.5 kcal/mol at the B3LYP/6-31+G* level with the PCM model.[39] The free energy change (ΔG_{aq}) of **10** to **11** is predicted to be −14.7 ± 0.2 kcal/mol with our QM/MM simulation method (Table 1). This will make the overall neutral decarboxylation reaction (**2** → **11** + CO_2) a facile exothermic process in water.

Figure 3 reveals that the initial proton-transfer process is significantly *stabilized* in aqueous solution with a decrease in activation free energy $\Delta\Delta G_1^{\neq}$ of 8.3 kcal/mol in comparison to the gas-phase reaction. This, however, does not include the quantum mechanical tunneling effect, which might be significant for this process.[40] Nevertheless, the general features of the solvent effects are expected to be valid even when these effects are explicitly considered.[41] The dramatic solvent effect is accompanied by a concomitant increase in the molecular dipole moment in going from reactant **2** (6.9 D) to the intermediate species **7** (11.3 D) in aqueous solution. This is consistent with the development of a zwitterionic species as the proton migrates from the carboxylic group to the ureido oxygen (Scheme 1). It is interesting to note that there is also a differential solvation effect between the reactant state (**2**) and the **TS-1**/**INT** state in water as demonstrated by the computed *induced* dipole moments (μ_{ind}). For reactant **2**, the computed μ_{ind} is 1.7 D ($\mu_{ind} = \mu_{aq} - \mu_{gas}$; $\mu_{gas} = 5.2$ D), which may be compared with a change of $\mu_{ind} = 2.7$ D for the intermediate structure **7** (Table 3).

On the other hand, solvent effects greatly *increase* the activation free energy for the decarboxylation step, as found in other decarboxylation reactions including the anionic *N*-carboxy-2-imidazolidinone reaction in water.[4] The observed solvent effect is consistent with the fact that the partial charges (see below) in the zwitterionic species are being annihilated as the N−CO_2 bond length increases and that charges are more dispersed at the transition state structure, **TS-2**. This is echoed by a computed aqueous dipole moment of 6.77 ± 0.07 D for **TS-2**, which is about 4 D less than the intermediate structure **7**. Due to poor solvation, the location of the transition structure (**TS 2**) in aqueous solution is found to be shifted to a later stage ($R_{N-C} = 2.04$ Å) as compared with that in the gas phase ($R_{N-C} = 1.95$ Å). The process on going from the best hydrated intermediate structure **7** to the transition state in solution covers a loss of 13.0 kcal/mol in solvation free energy, more than enough to offset the gain in ΔG_{hydr} in going from reactant **2** to intermediate structure **7**. Consequently, the overall solvent effect is to increase the activation free energy of the entire reaction from 16.9 kcal/mol in the gas phase to a predicted 21.0 ± 0.2 kcal/mol in aqueous solution.

Table 3. Computed Solute−Solvent Interaction Energies (kcal/mol) and Dipole Moments (debye) in Water at 25 °C and 1 atm

	2, reactact	**6, TS-1**	**7, INT**	**8, TS-2**
E_{vdW}	−10.9 ± 0.2	−10.3 ± 0.2	−9.7 ± 0.3	−10.7 ± 0.2
E_{vert}	−18.8 ± 0.4	−31.6 ± 1.4	−35.5 ± 1.2	−14.9 ± 0.7
E_{pol}	−2.9 ± 0.1	−5.0 ± 0.4	−6.0 ± 0.4	−2.3 ± 0.2
E_{sx}	−32.6 ± 0.5	−46.9 ± 1.5	−51.2 ± 1.3	−27.9 ± 0.7
ΔE_{sx}	0.0	−14.3 ± 1.6	−18.6 ± 1.4	4.7 ± 0.9
$\Delta\Delta G_{hyd}$	0.0	−7.2 ± 0.1	−8.3 ± 0.1	2.3 ± 0.2
μ (in water)	6.9 ± 0.1	10.6 ± 0.1	11.3 ± 0.1	6.8 ± 0.1
μ (in vacuo)	5.2	8.2	8.6	5.3
μ_{ind}	1.7 ± 0.1	2.4 ± 0.1	2.7 ± 0.1	1.5 ± 0.1

It is interesting to further dissect the differences and similarities in solvent effects between the decarboxylation of the neutral and anionic forms of *N*-carboxy-2-imidazolidinone in water. For the reaction of *N*-carboxy-2-imidazolidinone ion, corresponding to high pH conditions (pH = 10.2), the solvent effect is relatively simple in that the transition structure is relatively poorly solvated, compared to the reactant state, because atomic charges are more delocalized in the transition state. Consequently, ground-state stabilization leads to a net increase in activation free energy by 12.2 kcal/mol over that (10.5 kcal/mol) of the gas-phase process. Combining with the gas-phase results at the MP2/6-31+G(d)//HF/6-31+G(d) level, the predicted $\Delta G^{\neq}$ for the ionic reaction 2 is 22.7 ± 0.2 kcal/mol,[4] in accord with the experimental estimate of 23.2 kcal/mol.[5,36] Similar findings in the size of solvation effect have been obtained for a different anionic decarboxylation reaction involving 3-carboxybenzisoxazole in water from Monte Carlo QM/MM simulations.[10a,b] However, two opposing factors contribute to the overall shape of the reaction profile for the decarboxylation reaction of the *neutral N*-carboxy-2-imidazolidinone species in water, which correspond to low pH conditions in the experiment (pH = 6.1).[5] Intrinsically, the proton

(39) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(40) Hwang, J. K.; Warshel, A. *J. Am. Chem. Soc.* **1996**, *118*, 11745.

(41) For a review of recent theoretical work on proton-transfer reactions in solution, see: Bertran, J. *Theor. Chem. Acc.* **1998**, *99*, 143 and references therein.

transfer (reaction 1) is not favored in the gas phase because of the low basicity of a carbonyl group relative to that of a carboxylate ion and of the creation of a zwitterion. On the other hand, decarboxylation is catalyzed by the initial proton transfer with a barrier of only 2.9 kcal/mol in free energy for the second reaction step (from the intermediate species). This value is considerably smaller than that for a typical anionic decarboxylation reaction, which is about 11 kcal/mol (e.g., 10.5 kcal/mol for reaction 2).[4,10] Nevertheless, the result of the unfavorable proton-transfer step is that the overall decarboxylation reaction of the neutral molecule has a greater barrier than that of the anionic compound (16.9 vs 10.5 kcal/mol). In aqueous solution, the proton-transfer process has a much more favorable solvent effect, whereas the decarboxylation step is retarded in a manner similar to that in other decarboxylations. As a result, the solvent effect on the overall reaction is relatively small (21.0 ± 0.2 kcal/mol), leading to a reversal in reactivity between the neutral and the anionic forms of *N*-carboxy-2-imidazolidinone in water. The prediction that the activation free energy for reaction 1 is 1.7 kcal/mol smaller than that of reaction 2 from our dual-level ab initio and combined QM/MM simulations is in good accord with the experimental result (2.5 kcal/mol). Our combined QM/MM simulation analysis indicates that this is the result of a dramatic reversal of the intrinsic reactivity of the two species due to solvent effects.

To gain additional insights into the origin of the differential hydration effects along the reaction pathway, additional simulations for reactant **2**, **TS-1**, intermediate **7**, and **TS-2**, were carried out, which involved equilibrations over 0.8×10^6 configurations and averages over 1.5×10^6 configurations in Monte Carlo calculations. Table 3 summarizes the results of the energy components that contribute to the total solute−solvent interactions for each species. 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 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.[21] There is little variation in the van der Waals energy as the reaction proceeds from reactant **2** to **TS-2**. The solute−solvent interaction energy for the intermediate structure **7** is 18.6 kcal/mol lower than that for reactant **2.** This makes **7** better hydrated by −8.3 kcal/mol in solvation free energy than **2**. This is the primary reason for the easy proton transfer in water with much lower activation free energy as compared to the gas-phase value. Note that the solution energy also includes solvent reorganization contributions. However, to accurately capture the solvent reorganization energy, a more rigorous approach, including mapping the solute−solvent coordinate space, may be used.[28] There is a substantial contribution from differential polarization effects between reactant **2** and the intermediate structure **7**. The combined QM/MM calculations suggest that the intermediate structure **7** is more polarized than reactant **2** by −3.1 kcal/mol.

This implies a charge redistribution between the reactant **2** and the intermediate structure **7**, which will lead to differential hydrogen-bonding interactions. This is evident from the Mulliken partial charges averaged over all the liquid configurations given in Table 4. Partial atomic charges from Mulliken population analysis for all the species in the gas phase and in aqueous solution, which gives a reasonable qualitative description of atomic charge distributions, have been obtained from AM1 calculations. Standard deviations for the computed atomic charges are about 0.002 electron on average. During the whole proton-transfer process in water, the transferred proton carries nearly the same amount of positive charge of 0.3 e. However, for the carboxylate group (CO_2) in water, the total charge of the carboxylate group in **7** is −0.537 e, which represents a gain of −0.309 e in comparison with the total charge of the carboxylate group in reactant **2** (−0.228 e). The significant charge development of the carboxylate group in **7** indicates that intermediate structure **7** is zwitterionic in nature. The charge development of the carboxylate group lies mainly in the O2 oxygen which is cis to the ureido oxygen (−0.556 e in **7** vs −0.319 e in **2**). It is interesting to note that there are only subtle differences in atomic charges on the carboxylate group between **2** and **TS-2** in water. The charges are dispersed again on going from the zwitterionic **7** to **TS-2**, even though the proton is fully transferred in **TS-2**. The Mulliken population analysis presented here suggests that intermediate **7** will be better solvated in water due to the significant charge separation in the zwitterionic structure and there will be no substantial differential solvation effects on **2** and **TS-2**. This is seen in the following energy and radial distribution functions.

Table 4. Computed Partial Atomic Charges (e) Using AM1 Wave Function[a]

	2, reactant		**7, INT**		**8, TS-2**	
atom	gas phase	aqueous	gas phase	aqueous	gas phase	aqueous
O	−0.381	−0.436	−0.248	−0.241	−0.238	−0.248
N_1	−0.415	−0.402	−0.360	−0.391	−0.400	−0.426
C_2	0.413	0.427	0.370	0.379	0.288	0.301
N_3	−0.374	−0.356	−0.339	−0.293	−0.332	−0.322
H_3	0.253	0.283	0.267	0.318	0.243	0.282
C_4	−0.034	−0.036	−0.041	−0.038	−0.059	−0.052
H_{41}	0.108	0.141	0.116	0.147	0.114	0.127
H_{42}	0.086	0.097	0.098	0.115	0.083	0.101
C_5	−0.019	−0.022	−0.037	−0.038	−0.061	−0.062
H_{51}	0.118	0.110	0.131	0.124	0.117	0.112
H_{52}	0.104	0.121	0.117	0.119	0.107	0.105
C	0.452	0.483	0.460	0.497	0.497	0.534
O_2	−0.280	−0.319	−0.497	−0.556	−0.378	−0.408
O_3	−0.317	−0.392	−0.381	−0.478	−0.285	−0.339
H	0.285	0.302	0.346	0.335	0.304	0.294

[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.

(d) Distribution Functions. Figure 4 gives the energy pair distributions which show the spectrum of individual solute−water interactions such as hydrogen-bonding and dipolar interactions in the liquids. Note that a shoulder occurs near −7 kcal/mol for intermediate **7**, indicating stronger hydrogen-bonding interactions. These are not apparent for **2** and **TS-2**. The lowest interaction energy for **7** appears at −10.5 kcal/mol, which is about 4 kcal/mol more attractive than that for **2** and **TS-2**.

The solute−solvent bonding energy distributions for reactant **2**, intermediate **7**, and **TS-2** are compared in Figure 5. Intermediate **7** experiences a broad energy environment covering a range of 40 kcal/mol between −76 and −36 kcal/mol. Reactant **2** and **TS-2** have similar energy environments with the distributions of **2** left-

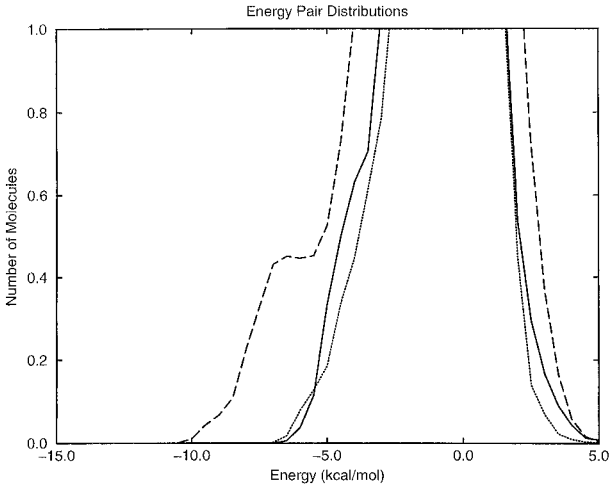


Figure 4. Computed solute−solvent energy pair distributions for reactant **2** (solid curve), INT **7** (dashed curve), and **TS-2** (dotted curve). The ordinate gives the number of water molecules coordinated with the solute with the interaction energy shown on the abscissa. Unit for the ordinate is number of molecules per kcal/mol.

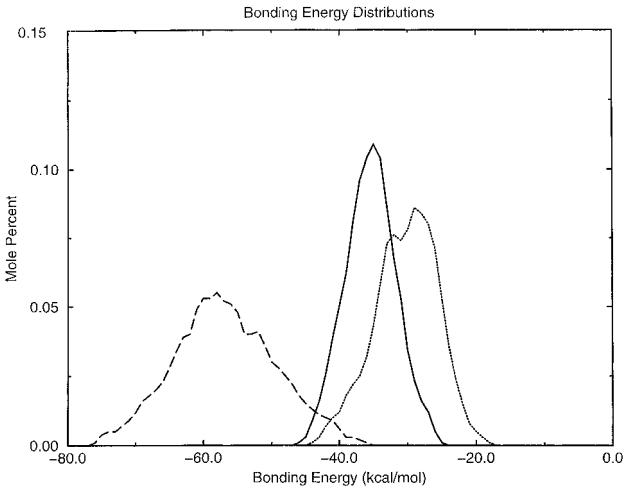


Figure 5. Computed distributions of the total solute−solvent bonding energies for **2**, INT **7**, and **TS-2**. Unit for the ordinate is mol % per kcal/mol.

shifted by 5 kcal/mol relative tothat of **TS-2**. Note that the average of the distributions is the total solute−solvent bonding energy term of E_{sx} listed in Table 3. The more attractive solute−solvent interactions for **7** are apparently observed.

Specific hydrogen-bonding interactions are revealed in the radial distibution functions (not shown here). Peaks in rdf's are associated with solvation shells or specific neighbors and can be integrated to yield coordination numbers. Hydrogen bond numbers with solvent water for the ureido oxygen and the carboxylate oxygens cis and trans to the ureido oxygen are 1.0, 1.2, and 2.0 for reactant **2**, 1.0, 1.7, and 2.4 for intermediate **7**, and 0.9, 1.1, and 1.5 for **TS-2**, respectively. The total average hydrogen bonds formed by solute oxygens are 4.2 for reactant **2**, 5.1 for the intermediate **7**, and 4.0 for the **TS-2** structure. Thus, it is likely that intermediate **7** is better solvated than reactant **2** through forming one more hydrogen bond with solvent water, while interactions with solvent water molecules between the **TS-2** structure and reactant **2** are attenuated by the strength rather

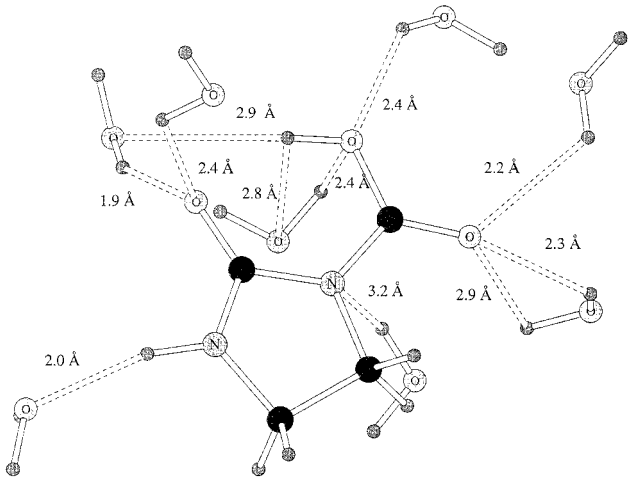


Figure 6. Illustration of one configuration from the simulations of reactant **2** in water. All hydrogen bonds are indicated.

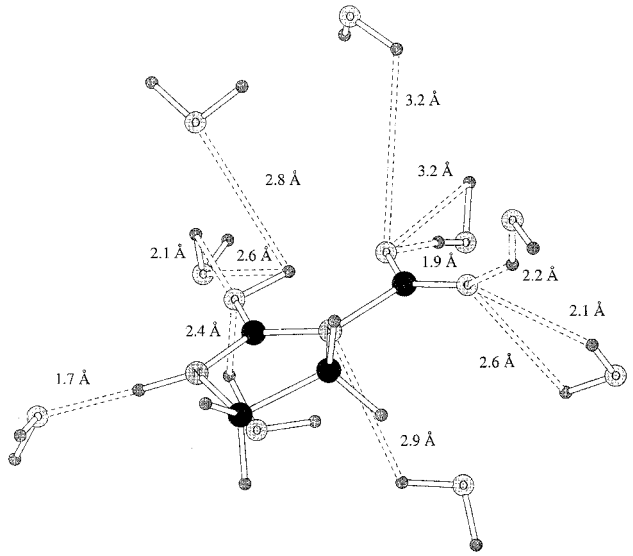


Figure 7. Illustration of one configuration from the simulations of INT **7** in water.

than by the number of hydrogen bonds. This feature is reflected by displaying the configurations obtained from Monte Carlo simulations.

The plots of the last configuration from the Monte Carlo simulations of reactant **2**, intermediate **7**, and **TS-2** are shown in Figures 6−8. For clarity, only water molecules which are hydrogen bonded to the solute are shown with the hydrogen bond O···H distances labeled in the plots. It can be seen that there are one strong and two weak hydrogen-bonding interactions for the carboxyl oxygen in **7**, while two typical hydrogen bonds are observed in reactant **2**. In most cases, the hydrogen bond lengths in **7** are shorter than those in **2** and **TS-2**. The hydrogen bond numbers are conserved between **2** and **TS-2**. However, the average hydrogen bond distances are lengthened from **2** to **TS-2**, suggesting weaker hydration of **TS-2**. Thus, although the total solvation effect on the two critical points of **2** and **TS-2** on the reaction surface is not large, the whole reaction does experience a dramatic solvent effect. Better solvation for the proton-transfer process is observed due to greater hydrogen-bonding interactions and stronger polarization effects when the reaction proceeds on going from **2** to **7**. The

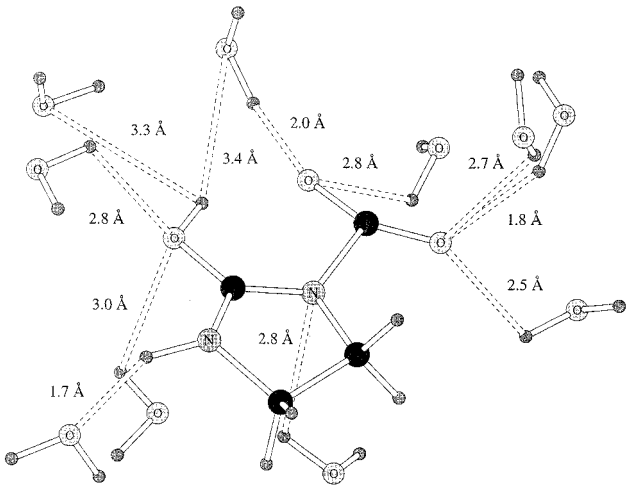


Figure 8. Illustration of one configuration from the simulations of **TS-2** in water.

favorable solvation for the decarboxylation step is lost quickly along the reaction pathway on going from **7** to **TS-2** (a structure very close to the true TS in solution), leading to a modest unfavorable solvation effect for the total reaction.

Concluding Remarks

We have shown a "thrilling movie" of the decarboxylation reaction of the neutral *N*-carboxy-2-imidazolidinone in aqueous solution along the reaction coordinate using Monte Carlo simulations. A two-step mechanism with the proton-transfer ahead of the rate-limiting N−C bond breaking was found. The gas-phase free energy of activation is 16.9 kcal/mol, while it rises to 21.0 ± 0.2 kcal/mol in water, which is in excellent agreement with the experimental value of 20.7 kcal/mol.[5,36] The proton-transfer step was found to be much faster in water than in the gas phase.

The present study illustrates the first complete reaction profile for a *neutral* intramolecular hydrogen transfer and decarboxylation reaction in aqueous solution. Valuable structural and energetic information for this reaction have been obtained. Implications of this work also give clear, detailed insights into the nature of the transition states of neutral decarboxylation reactions of other β-keto carboxylic acids. There has been a long history of experimental studies of neutral decarboxylation reactions of β-keto carboxylic acids in solution.[11] We have presented a detailed picture regarding the mechanism and sequence of the proton transfer and decarboxylation that has been controversial from experimental investigations.[11] Transition states with both full and partial proton transfers have been proposed on the basis of different experimental observations. The confusion comes mainly from the speculation on the relation between the polarity of the transition state and the extent of the proton transfer. In this work, it is clearly shown that the proton is fully transferred to the carbonyl oxygen in the rate-limiting transition state but the polarity of the TS is nearly the same as that of the reactant. No significant solvation effect for the overall reaction was obtained. This feature seems to be general for other β-keto acids, since recent ab initio reaction field (SCRF) calculations on the reactants and transition states of several β-keto acids such as formylacetic acid, malonic acid, and acetoacetic acid showed that the polarities of the reactants and the transition states are similar and the activation energies are less affected by different solvents.[10c,d,e] Indeed, it has been experimentally observed that the rate of the decarboxylation reactions of neutral β-keto acids has little to do with the dielectric constant of the medium.[11]

In contrast to neutral decarboxylation reactions, there is a pronounced solvation effect on anionic decarboxylations in aqueous solution.[4] The reactant is better solvated by 12 kcal/mol than the transition state for the anionic decarboxylation reaction of *N*-carboxy-2-imidazolidinone in aqueous solution. This work makes it understandable that, in going from the aqueous solution to a lower dielectric medium, with dismissed hydrogen bonding capability the anionic reaction will be accelerated due to desolvation of the reactant, while there is only a modest overall solvent effect for the rate of the neutral decarboxylation reaction. Thus, the present computer simulation study of the biotin model compound *N*-carboxy-2-imidazolidinone in aqueous solution will provide an interesting reference for future computational and experimental studies of the enzymatic decarboxylation reactions of carboxybiotin at physiological pH and temperature,[42] when the structure of a biotin-dependent decarboxylase and the structural details of the active site are accurately determined to atomic resolution.

Acknowledgment. NATO (CRG941209) is thanked for support.

JO9904563

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