J|A|C|S
ARTICLES
Published on Web 02/27/2003

Modeling the Reaction Mechanisms of the Imide Formation in an *N*-(*o*-Carboxybenzoyl)-L-amino Acid

Zhijian Wu, Fuqiang Ban, and Russell J. Boyd*

Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4J3 Canada

Received May 16, 2002; E-mail: russell.boyd@dal.ca

Abstract: Reaction mechanisms of the imide formation in an *N*-(*o*-carboxybenzoyl)-L-amino acid have been studied using density functional theory. Our results suggest that the reaction route initiated by protonation at the oxygen of the carboxyl group of the amino acid is favored, while those initiated by deprotonation at the oxygen of the carboxyl group of phthalic acid and at the amidic nitrogen are minor pathways. During the dehydration process, water functions as a catalyst. These conclusions are in good agreement with the experimental facts that at highly acidic conditions (hydrogen ion concentration $H_0 < -1$), imide formation is the most favorable pathway, whereas in the pH range 0−5, cyclization to the imide is not the dominant reaction. Our calculations also show that the carboxyl group of the amino acid is involved in the catalytic reaction in both the favored and minor pathways and that solvent effects have little influence on the reaction barriers.

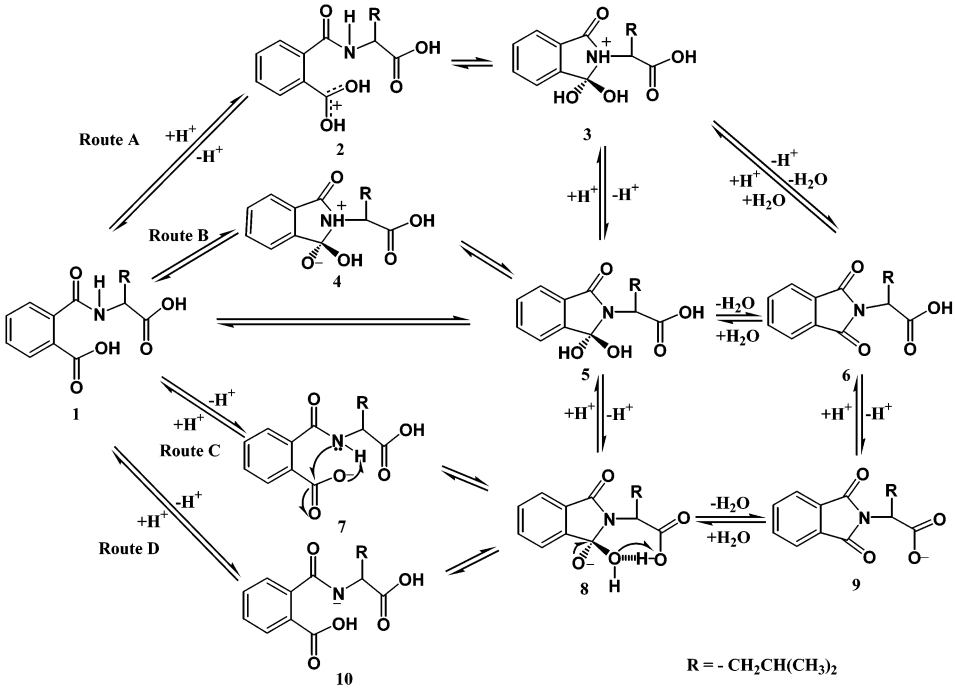
Introduction

The mechanisms by which enzymatic catalysis proceeds are important and very perplexing.[1,2] Carboxyl (acid) proteinases, conventionally described as aspartic proteinases,[2] have a very long history in enzymology. The distinct characteristic of aspartic proteinases is their pH-dependent catalytic action with an optimal pH range 1.9−4.0. The hydrolysis of a peptide bond is catalyzed by two carboxyl groups of the aspartic residue. In fact, as models for hydrolytic enzymes, intramolecular systems containing carboxyl groups have been extensively investigated for many decades and continue to be a subject of great interest.[3−26] Previous studies have examined *N*-(2-aminophenyl)-phthalamic acid,[6] phthalamic acid[7−9] *N*-methylphthalamic acid and *N*-acetylphthalamic acid,[9] phthalanilic acid,[10] alkyl hydrogen dialkylmaleates,[11] maleic acid,[12−14] and *N*-(*o*-carboxybenzoyl)-L-phenylalanine, *N*-(*o*-carboxybenzoyl)-L-leucine, and *N*-(*o*-carboxybenzoyl)-L-valine.[15,16] Amide hydrolysis (C−N bond cleavage) by an enzyme can be mimicked by intramolecular catalysis utilizing the carboxyl groups of the model systems. In addition to amide hydrolysis reactions, cyclization to the imide (formation of a C−N bond) has also been observed.[9,15−17] Thus, cyclization to the imide[9,15−17] and amide hydrolysis[6−16] are the two main types of reactions of interest that occur in the model systems. In particular, molecules containing two carboxyl groups, such as *N*-(*o*-carboxybenzoyl)-L-phenylalanine, *N*-(*o*-carboxybenzoyl)-L-leucine, and *N*-(*o*-carboxybenzoyl)-L-valine,[15,16] have been used to probe the catalytic activity of the neighboring carboxyl group of an amino acid that is involved in enzymatic reactions. Possible reaction mechanisms for cyclization to the imide and amide hydrolysis of *N*-(*o*-carboxybenzoyl)-L-leucine have also been presented experimentally by Nome and co-workers.[16] In this report, we focus on the imide formation. The reaction mechanism for amide hydrolysis will be presented in a separate contribution.[27] For cyclization to the imide, the three proposed routes (A, B, and C in Scheme 1) correspond to different pH ranges.[16] Cyclization depends strongly on the acidity of the solution. In route A with highly acidic conditions (hydrogen ion concentration $H_0 < -1$), there

(1) Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790.
(2) Page, M. I.; Williams, A. *Enzyme Mechanisms*; Royal Society of Chemistry: London, 1987; p 229.
(3) Hartwell, E.; Hodgson, D. R. W.; Kirby, A. J. *J. Am. Chem. Soc.* **2000**, *122*, 9326.
(4) Bowden, K.; Brownhill, A. *J. Chem. Soc., Perkin Trans. 2* **1997**, 219.
(5) Barber, S. E.; Dean, K. E. S.; Kirby, A. J. *Can. J. Chem.* **1999**, *77*, 792.
(6) Perry, C. J. *J. Chem. Soc., Perkin Trans. 2* **1997**, 977.
(7) Granados, A. M.; de Rossi, R. H. *J. Org. Chem.* **2001**, *66*, 1548.
(8) Bender, M. L.; Chow, Y.-L.; Chloupek, F. *J. Am. Chem. Soc.* **1958**, *80*, 5380.
(9) Brown, J.; Su, S. C.; Shafer, J. A. *J. Am. Chem. Soc.* **1966**, *88*, 4468.
(10) Hawkins, M. D. *J. Chem. Soc., Perkin Trans. 2* **1976**, 642.
(11) Aldersley, M. F.; Kirby, A. J.; Lancaster, P. W. *J. Chem. Soc., Perkin Trans. 2* **1974**, 1504.
(12) Aldersley, M. F.; Kirby, A. J.; Lancaster, P. W.; McDonald, R. S.; Smith, C. R. *J. Chem. Soc., Perkin Trans. 2* **1974**, 1487.
(13) Kirby, A. J.; McDonald, R. S.; Smith, C. R. *J. Chem. Soc., Perkin Trans. 2* **1974**, 1495.
(14) Kirby, A. J.; Lancaster, P. W. *J. Chem. Soc., Perkin Trans. 2* **1972**, 1206.
(15) Onofrio, A. B.; Joussef, A. C.; Nome, F. *Synth. Commun.* **1999**, *29*, 3039.
(16) Onofrio, A. B.; Gesser, J. C.; Joussef, A. C.; Nome, F. *J. Chem. Soc., Perkin Trans. 2* **2001**, 1863.
(17) Fife; T. H.; Chauffe, L. *J. Org. Chem.* **2000**, *65*, 3579.
(18) Yunes, S. F.; Gesser, J. Chaimovich, C. H.; Nome, F. *J. Phys. Org. Chem.* **1997**, *10*, 461.
(19) Barros, T. C.; Yunes, S.; Menegon, G.; Nome, F.; Chaimovich, H.; Politi, M. J.; Dias, L. G.; Cuccovia, I. M. *J. Chem. Soc., Perkin Trans. 2* **2001**, 2342.
(20) Park, H. Suh J.; Lee, S. *THEOCHEM* **1999**, *490*, 47.
(21) Hori, K.; Kamimura, A.; Ando, K.; Mizumura, M.; Ihara, Y. *Tetrahedron* **1997**, *53*, 4317.
(22) Antonczak, S.; Ruiz-López, M. F.; Rivail, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 3912.
(23) Antonczak, S.; Ruiz-López, M. F.; Rivail, J. L. *J. Mol. Model.* **1997**, *3*, 434.
(24) Bakowies, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 5712.
(25) Guo, J.-X.; Ho, J.-J. *J. Phys. Chem. A* **1999**, *103*, 6433.
(26) Naundorf, H.; Worth, G. A.; Meyer, H.-D.; Kühn, O. *J. Phys. Chem. A* **2002**, *106*, 719.
(27) Wu, Z. J.; Ban, F.; Boyd, R. J. Manuscript submitted

10.1021/ja020700z CCC: $25.00

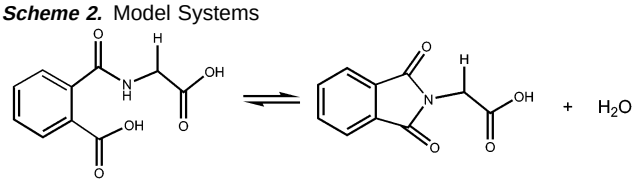
Scheme 1. Summary of the Experimentally Proposed Schematic Routes for Imide Formation



are two possible mechanisms, that is, **1→2→3→6** and **1→2→3→5→6**. The intermediates are cationic for mechanism **1→2→3→6** and a mixture of cationic and neutral species for mechanism **1→2→3→5→6**. In the pH range 0−5, cyclization is not the dominant reaction. In route B (pH range 0−2), the two possible mechanisms are **1→4→5→6** and **1→5→6**, and the intermediates are zwitterionic or neutral. In route C (pH range 2−5), the mechanism is **1→7→8→9→6** and the intermediates are anionic. In addition to the pathways presented by Nome and co-workers,[16] a previous study on *N*-methylphthalamic acid[9] proposed pathways for the cyclization to the imide which have been adapted for *N*-(*o*-carboxybenzoyl)-L-leucine as shown in Scheme 1. The relationships between pH ranges and pathways are not clearly assigned. These two additional pathways (**1→10→8→9→6** and **1→10→8→5→6**) of route D are initiated by deprotonation at the amidic nitrogen of **10**. Possible routes proposed on the basis of experiments[9,16] for cyclization to the imide are summarized in Scheme 1. The species in routes C and D are all anionic but with different deprotonation sites.

Due to the size of the molecule and the complexity of the reaction pathways, only a few theoretical computations[18−26] have been reported for reaction mechanisms of intramolecular systems containing carboxyl groups. Most of the theoretical studies have concentrated on the pathway of amide hydrolysis.[18−26] The previous theoretical studies have been carried out mainly at the Hartree−Fock (HF) level, with single-point energies calculated using second-order Møller−Plesset perturbation theory (MP2). Examples include PM3[18] and HF/6-31G(d)[19] studies of 1,8-naphthalic acid, a HF/6-31G study of *N*-methlmaleamic acids,[20] HF/6-31G and HF/6-31+G studies of *N*-methyl acetamide, acetanilide, and *N*-acetyl imidazole,[21] HF/3-21G,[22] BLYP/6-31G(d,p),[23] and combined quantum chemical and statistical mechanical studies of formamide,[24] and a MP2/6-31+G(d,p) study of N-substituted formamides.[25] In addition, a multiconfiguration time-dependent Hartree dynamics study on phthalic acid monomethylester has been published.[26] The development of density functional theory (DFT) with modern exchange-correlation functionals has paved the way to study equilibrium structures and reaction pathways for intramolecular catalysis in fairly large molecules. These DFT methods have been tested extensively against other accurate correlation methods for the prediction of reliable geometries[28] and reaction barriers[29] and have been applied to large biological systems.[30,31] On the other hand, since experiment alone is usually not enough to deduce reaction mechanisms,[32] theoretical work is necessary to gain further insight into reaction mechanisms. In this work, we report the first theoretical study of the reaction mechanism of the imide formation shown in Scheme 1 using the model system shown in Scheme 2. Our model system with two carboxyl groups is based on previous experimental studies[15,16] and is designed to investigate intramolecular catalytic effects. The final product is produced by eliminating a water molecule through various tetrahedral intermediates. The smallest amino acid, glycine, is used in our model system (*N*-(*o*-carboxybenzoyl)glycine). Density functional theory is employed for the study of reaction mechanisms of the imide formation. Solvent effects have also been considered because they often play a crucial role in biological processes.

Scheme 2. Model Systems

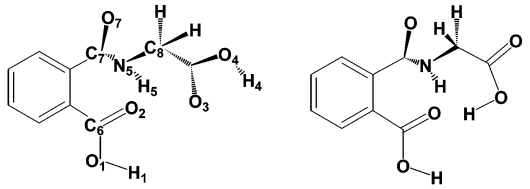


Computational Methods

All geometry optimizations were performed with the B3LYP hybrid density functional in conjunction with the 6-31G(d,p) basis set using

(28) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.
(29) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936.
(30) Ban, F.; Gauld, J. W.; Boyd, R. J. *J. Am. Chem. Soc.* **2001**, *123*, 7320.
(31) Himo, F.; Siegbahn, P. E. M. *J. Phys. Chem. B.* **2000**, *104*, 7502.
(32) Fersht, A. R.; *Enzyme Structure and Mechanism*; W. H. Freeman and Co: New York, 1985.

Scheme 3. Optimized Geometries of the Initial Reactant **1a** and **1b**



1a **1b**

the Gaussian 98 suite of programs.[33] The B3LYP functional is a combination of Becke's three-parameter hybrid exchange functional,[34,35] as implemented in Gaussian 98,[36] and the Lee−Yang−Parr correlation functional.[37] Harmonic vibrational frequencies and zero-point vibrational energies (ZPVEs) were obtained at the same level of theory. Relative energies were obtained by performing single-point calculations at the B3LYP level with the 6-311G(2df,p) basis set using the above optimized geometries and by including the zero-point vibrational energy i.e., B3LYP/6-311G(2df,p)//B3LYP/6-31G(d,p)+ZPVE. The entropy contributions to the free energies at 298.15 K were derived from B3LYP/6-31G(d,p) frequency calculations.

The solvent effect on the potential energy surface was investigated by single-point calculations at the B3LYP/6-311G(2df,p) level using the conductor-like polarized continuum solvent model (CPCM)[38] (denoted CPCM-B3LYP/6-311G(2df,p)) on the geometries obtained by using the Onsager model [39] with a dielectric constant of 78.39 at the B3LYP/6-31G(d,p) level (denoted Onsager-B3LYP/6-31G(d,p)). The zero-point energy corrections obtained from the Onsager-B3LYP/6-31G(d,p) frequency calculations are included in the calculations of the relative energies in the solvation of water, i.e., CPCM-B3LYP/6-311G-(2df,p)//Onsager-B3LYP/6-31G(d,p)+Onsager-ZPVE. Entropy contributions to the free energies of solvation at 298.15 K were derived from Onsager-B3LYP/6-31G(d,p) frequency calculations. All energies are in kJ mol^{-1}, and bond distances are in angstroms (Å).

Since the crystal structures are not available for *N*-(*o*-carboxybenzoyl)-L-leucine or the model compounds, the first challenge is to obtain a reasonable structure for reactant **1**. The structure of reactant **1** is very flexible and contains numerous conformers. Rather than carrying out a computationally expensive search of the conformational space of reactant **1**, we began with the final product (**6** in Scheme 1) since it has fewer conformers. We then went backward step-by-step to reach the reactant. An intrinsic reaction coordinate (IRC) calculation was used to locate the reactant and product from the transition state. Thus, reactant **1** from route A (**1a** in Scheme 3) and from routes C, D (**1b** in Scheme 3) is obtained. Reactant **1** from route B was not found (see Results and Discussion in next section for details). To reduce the computational effort, the semiempirical PM3 method was used for a preliminary conformational search, but only conformers from the B3LYP/6-31G-(d,p) calculations are reported in this work.

(33) Frisch, M. J.; Trucks, G. W.; Schlegel. H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(34) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
(35) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(36) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
(37) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(38) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.
(39) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.

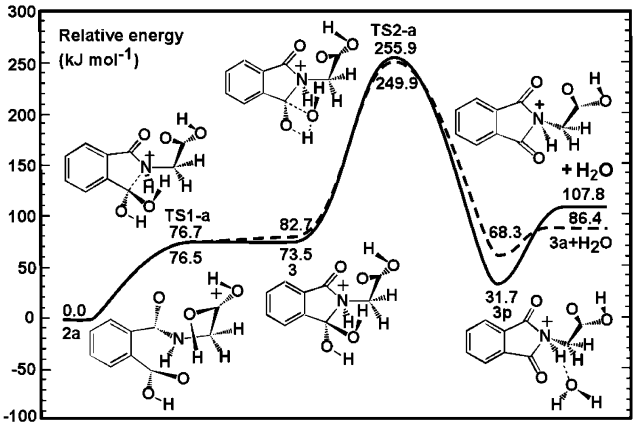


Figure 1. Schematic energy profile of the route **2a**→**3**→**3p**→**3a** + H_2O. Solid line represents the gas phase, the dashed line, the solution. Only the ZPVE correction is included.

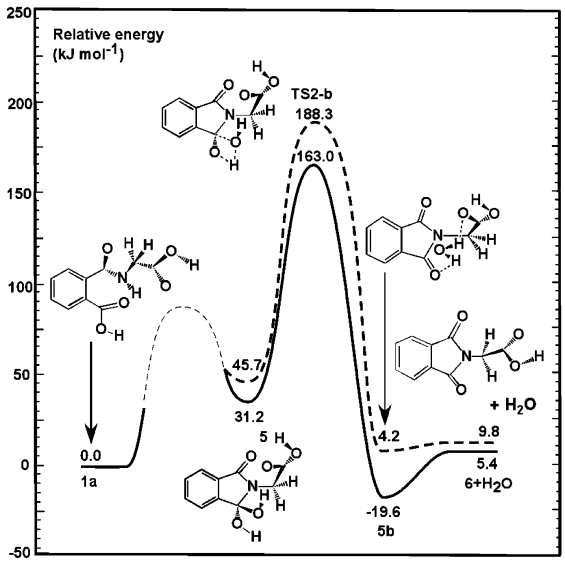


Figure 2. Schematic energy profile of the route **1a**→**5**→**5b**→**6**+H_2O. Solid line represents the gas phase, the dashed line, the solution. Only the ZPVE correction is included.

Results and Discussion

The computed energy profiles for the reaction routes are shown in Figures 1−10. Figures 1−5 include only the ZPVE correction, while the entropy contribution at 298.15 K is included in Figures 6−10. For each of the routes studied, two transition states have been found. **TS1** denotes the transition state of C−N bond formation, **TS2** the transition state of the direct dehydration step, and **TS3** the transition state of the dehydration through a water molecule. Our discussion will focus on the results of only the ZPVE corrected case since the entropy contribution at 298.15 K shows a similar energy profile.

Route A (1a→2→3→6 and 1a→2→3→5→6). In this route, there exist two pathways from **3** (Scheme 1), depending on the sequence of deprotonation and dehydration. If dehydration is followed by deprotonation, the route is **1a→2→3→6**, otherwise, it is **1a→2→3→5→6**. The dehydration process can proceed either directly or assisted by a water molecule. However, for the water-assisted dehydration, we only located the dehydration transition state **TS3-b** (in Figure 3) from **5** to form **6**, whereas for the charged species, the transition state of the water-assisted

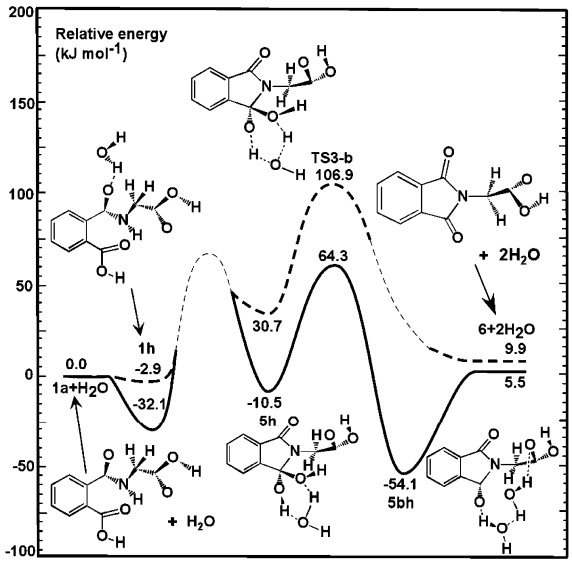


Figure 3. Schematic energy profile of the route **1a** + H_2O→**1h**→**5h**→**5bh**→**6** + $2H_2O$. Solid line represents the gas phase, the dashed, line the solution. Only the ZPVE correction is included.

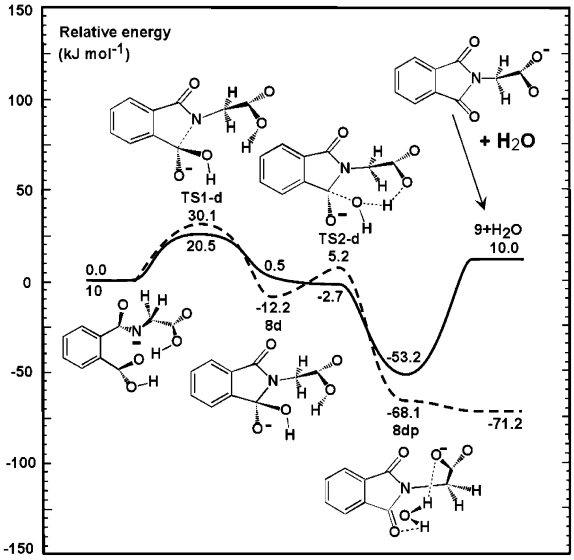


Figure 5. Schematic energy profile of the route **10**→**8d**→**8dp**→**9** + H_2O. Solid line represents the gas phase, the dashed line, the solution. Only the ZPVE correction is included.

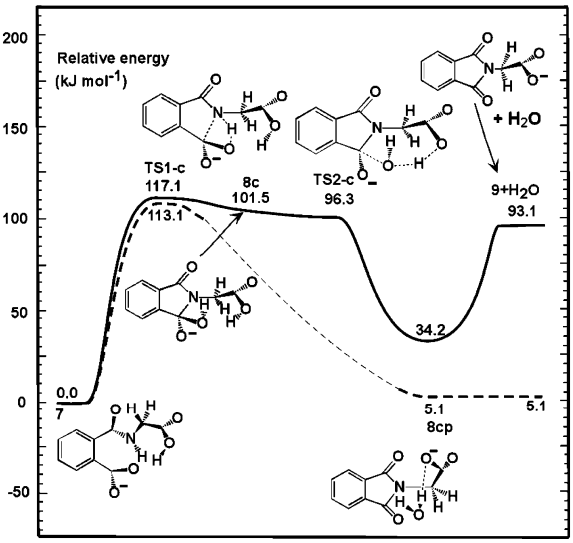


Figure 4. Schematic energy profile of the route **7**→**8c**→**8cp**→**9** + H_2O. Solid line represents the gas phase, the dashed line, the solution. Only the ZPVE correction is included.

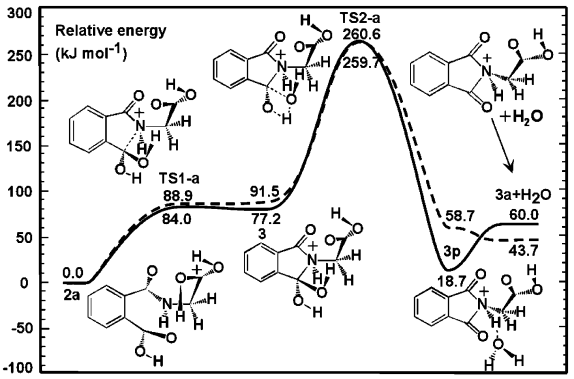


Figure 6. Schematic energy profile of the route **2a**→**3**→**3p**→**3a** + H_2O. Solid line represents the gas phase, the dashed line, the solution. The entropy contribution at 298.15 K and the ZPVE correction is included.

dehydration from **3** + H_2O to form **3p** + H_2O (see **3** and **3p** in Figure 1) was not located at the present level of theory.

In the conformational search of this route, we noted an interesting phenomenon in the IRC calculation to locate structure **2**: the hydrogen at O_2 (see Scheme 3 for labels of atoms) of the carboxyl group of phthalic acid is transferred to O_3 of the carboxyl group of the amino acid. This gives structure **2a** (Scheme 4), not structure **2** (Scheme 1) as proposed in the work of Nome and co-workers.[16] The energy difference between **2** and **2a**, and the transition state for the hydrogen transfer between O_2 and O_3 is given in Table 1. It can be seen that there is no energy barrier to form **2a** from **2**. The largest energy difference between **2** and **2a** occurs in solution at 298.15 K with a magnitude of 19.5 kJ mol^{-1}. The calculated proton affinities at O_3 of **2a** and O_2 of **2** are very close to each other (897.0 and 895.7 kJ mol^{-1}, respectively). On the other hand, there are significant differences in the geometries of **2** and **2a**. In the gas phase, the O_3-H bond distance in **2a** is 1.046 Å, while the O_2-H distance is 1.027 Å in **2**. The O_2-O_3 distances are 2.46 Å (**2a**), and 2.51 Å (**2**) in the gas phase and 2.52 Å in both **2a** and **2** in solution. It is known that if the interatomic distance between oxygens is within 2.5 ± 0.1 Å,[40] a low-barrier hydrogen bond (LBHB) will be formed. Therefore, the hydrogen moves easily between the two oxygens. Our calculated barriers and geometrical data suggest that an O_2-H-O_3 LBHB is formed in **2a** and **2**, which are in equilibrium (that slightly favors **2a**). The carboxyl group of the amino acid is involved in the imide formation via **2a** and **3** (Figure 1) under highly acidic conditions. In addition, the above calculation also suggests that since the intramolecular reactions are very much dependent on the relative positions of the reacting groups, the conformational analysis must be done carefully.

Conformer **1a** (Scheme 3) is easily protonated at the oxygen (O_3) of the amino acid due to its high proton affinity (897.0 kJ mol^{-1}). From Figure 1, one can see that in the gas phase the first transition state, **TS1-a**, between **2a** and **3**, is higher than **2a** by 76.7 kJ mol^{-1}, while **3** is only marginally lower than **TS1-a** by 3.2 kJ mol^{-1}. This means that **TS1-a** and intermediate **3** are nearly isoenergetic. In the dehydration process, however,

(40) Cleland, W. W. *Arch. Biochem. Biophys.* **2000**, *382*, 1.

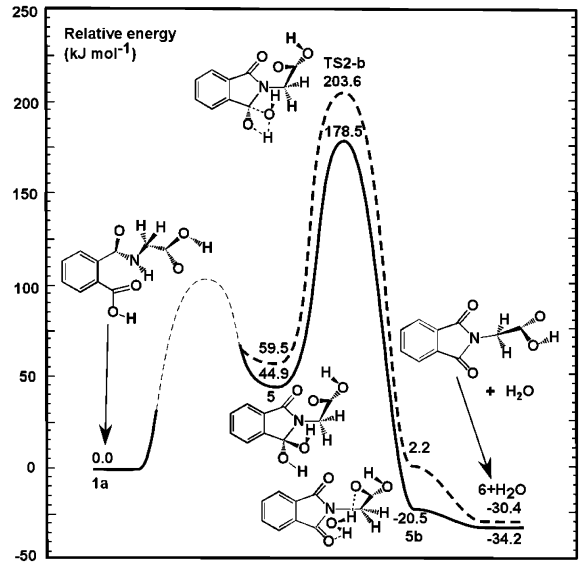


Figure 7. Schematic energy profile of the route **1a**→**5**→**5b**→**6** + H_2O. Solid line represents the gas phase, the dashed line, the solution. The entropy contribution at 298.15 K and the ZPVE correction is included.

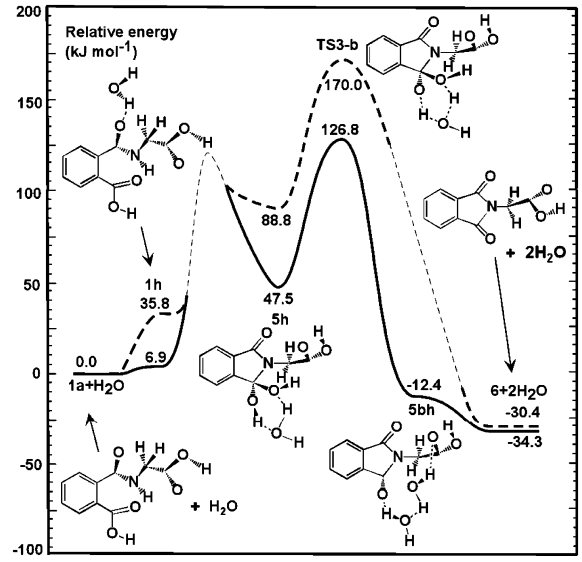


Figure 8. Schematic energy profile of the route **1a** + H_2O→**1h**→**5h**→**6** + H_2O. Solid line represents the gas phase, the dashed line, the solution. The entropy contribution at 298.15 K and the ZPVE correction is included.

a very high transition state (**TS2-a,** 182.4 kJ mol^{-1} from **3** and 255.9 kJ mol^{-1} from **2a**) is obtained. This indicates that dehydration from **3** to form cationic intermediate **3p** would be unfavorable. The complex **3p** and the separated system **3a** + H_2O have a relative energy of 31.7 kJ mol^{-1} and 107.8 kJ mol^{-1} compared with **2a**, respectively. The final product **6** can be derived by deprotonation of **3a** at an energy cost of 802.9 kJ mol^{-1}.

Similar trends are observed in solution as in the gas phase (Figure 1). The solvent lowers **TS1-a** by 0.2 kJ mol^{-1} and **TS2-a** by 6.0 kJ mol^{-1}. We also note that in solution, **3** is slightly higher in energy than **TS1-a** by 6.2 kJ mol^{-1} (after considering entropy contributions, the difference drops to 1.6 kJ mol^{-1}, Figure 5). This phenomenon may be explained by the geometrical changes of **3** and **TS1-a**. From our calculations, it can be seen that at **TS1** the C−N bond distance in the gas phase is longer than that in solution for all the routes studied (Table 2).

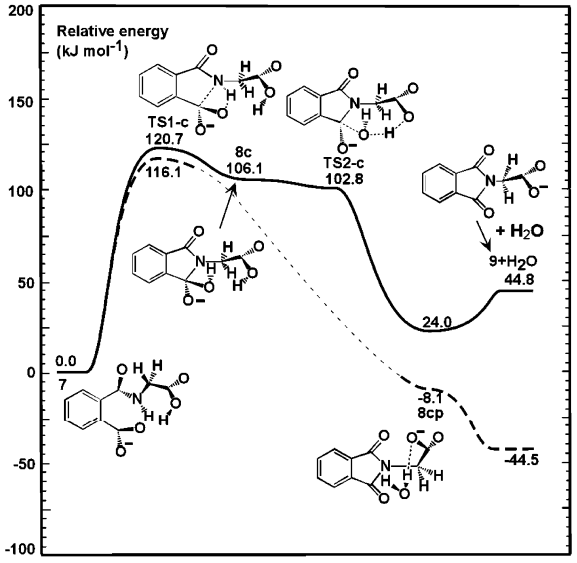


Figure 9. Schematic energy profile of the route **7**→**8c**→**8cp**→**9** + H_2O. Solid line represents the gas phase, the dashed line, the solution. The entropy contribution at 298.15 K and the ZPVE correction is included.

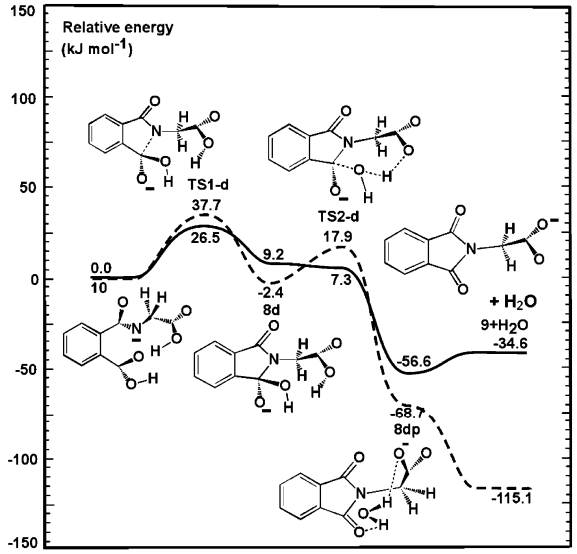
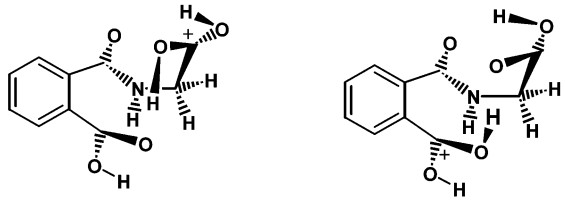


Figure 10. Schematic energy profile of the route **10**→**8d**→**8dp**→**9** + H_2O. Solid line represents the gas phase, the dashed line, the solution. The entropy contribution at 298.15 K and the ZPVE correction is included.

Scheme 4. Optimized Geometries of **2** and **2a**



2a **2**

For route A, the decrease of the C−N bond distance (0.033 Å) from the gas phase to solution is around 10 times larger than that in route C (**TS1-c**, 0.002 Å) and D (**TS1-d**, 0.004 Å). For **3**, the C−N bond distance is elongated by 0.012 Å in solution compared with the gas phase in route A. In route D (conformer **8d**), however, the opposite situation is found, i.e., the bond distance in the gas phase is longer than that in solution. Therefore, the overall effects of shortening the C−N bond

Table 1. Relative Energies (in kJ mol^{-1}) of **2**, Transition State of Hydrogen Transfer between O_2 and O_3 (**TS**) Compared with **2a** in the Gas Phase and Solution at 0 and 298.15 K, Respectively

	0 K			298.15 K		
	2	**TS**	**2a**	**2**	**TS**	**2a**
gas	0.0	2.1	0.1	0.0	−0.7	−1.4
solution	0.0	−5.9	−13.2	0.0	−13.4	−19.5

Table 2. C_6-N_5 Bond Distances (Å) at the TS1 Transition State and Its Product in Both the Gas Phase and in Solution

	TS			product		
route	structure	gas	solution	structure	gas	solution
A	**TS1-a**	1.981	1.948	**3**	1.653	1.665
C	**TS1-c**	2.031	2.029	**8c**	1.567	NA[a]
D	**TS1-d**	1.993	1.989	**8d**	1.544	1.533

[a] Not available.

distance in **TS1-a** and elongating C−N in **3** in solution compared with the gas phase may explain the higher energy of **3** than **TS1-a**, although they can be regarded as being isoenergetic.

As mentioned above, from **3**, the reaction can also proceed through another route. First, **3** is deprotonated to give neutral intermediate **5** (Scheme 1). The calculated energy for deprotonation from **3** to form **5** is 864.8kJ mol^{-1}. For the direct dehydration process, the dehydration of **5** will lead to neutral complex **5b** with an energy cost of only 131.8 kJ mol^{-1} (**TS2-b**, Figure 2) in the gas phase and 142.6 kJ mol^{-1} in solution compared with **5**. Relative to the initial reactant **1a**, **TS2-b** is 163.0 kJ mol^{-1} higher in the gas phase and 188.3 kJ mol^{-1} higher in solution. The transition state between **1** and **5** was not located in this study (for details see the discussion on route B below). Complex **5b** and the separated system **6** + H_2O lie lower by −50.8 kJ mol^{-1} and −25.8 kJ mol^{-1} than **5**, respectively. However, if dehydration proceeds through a water molecule, i.e., water reacts as a catalyst, the reaction barrier drops significantly (Figure 3). The dehydration barrier of **TS3-b** is only 74.8 kJ mol^{-1} in the gas phase and 76.2 kJ mol^{-1} in solution. Compared with **1h, TS3-b** is 96.4 kJ mol^{-1} higher in the gas phase and 109.8 kJ mol^{-1} higher in solution. In solution, **5bh** was not located in our calculations (indicated by the thin dashed line in Figures 3 and 8). The optimization process shows that the water molecules are completely separated and gives **6** + $2H_2O$.

The entropy corrected energy profile at 298.15 K (Figures 6 and 7) shows almost the same pattern. The energy barriers of **TS1-a** and **TS2-a** (Figure 6) are from 5.0 to 13.0 kJ mol^{-1} higher than the ZPVE-only corrected case in both phases. Relative to **5**, **TS2-b** and **TS3-b** (Figures 7 and 8) at 298.15 K are only about 5.0 kJ mol^{-1} higher than the ZPVE-only corrected results.

In brief, the reaction route **5→5b** (Figures 2 and 3), in particular assisted by a water molecule (Figure 3), can proceed more easily than the route **3→3a** (Figure 1). Therefore, this suggests that route **1→2→3→5→6** with a mixture of cationic and neutral intermediates (Scheme 1) is more favorable than **1→2→3→6** with only cationic intermediates.

Route B (1→4→5→6 and **1→5→6).** We were unable to locate zwitterionic intermediate **4** (Scheme 1) in both the gas phase and solution because no C−N bond is formed. Another possibility in this route is the direct formation of **5** from **1** through C−N bond formation and simultaneous proton transfer from nitrogen to the carboxyl group of phthalic acid (**1→5→6**). A transition state for this route was located at the PM3 and B3LYP/3-21G* levels but does not exist at the B3LYP/6-31G-(d,p) level. The above results suggest that route B is not a possible mechanism for the cyclization to the imide. Further study both theoretically and experimentally is needed to elucidate the reaction mechanisms for the pH range 0−2 with the lowest observed rate constants.

Route C (1b→7→8c→9→6). Since routes C and D have **8** in common (Scheme 1), we use **8c** to represent the conformer in route C, **8d** in route D. The deprotonation at oxygen (O_1) from **1b** (Scheme 3) costs 1336.8 kJ mol^{-1}, much higher (by 472.0 kJ mol^{-1}) than that needed from **3** to form **5** in route A. This suggests that route A (**1→2→3→5→6**) is more likely to occur than route C.

Transition state **TS1-c** between **7** and **8c** in Figure 4 shows that the formation of the C−N bond and proton transfer from N to the carboxylate anion occur simultaneously, in agreement with Nome and co-workers' expectation.[16] In the gas phase, the first barrier of **TS1-c** is 117.1 kJ mol^{-1}. After cyclization, **8c** is 101.5 kJ mol^{-1} higher than **7**. In the dehydration, the energy of **TS2-c** lies lower than that of its reactant **8c** by 5.2 kJ mol^{-1}, and the reaction proceeds to give monoanionic complex **8cp** with a relative energy of 34.2 kJ mol^{-1}. The separated system **9** + H_2O has an energy 93.1 kJ mol^{-1} higher than that of **7**. The proton affinity at oxygen (O_3) of the carboxyl group of the amino acid in **9** is 1425.0 kJ mol^{-1}. The above calculation suggests that, due to the participation of the carboxyl group of the amino acid, the energy of **TS2-c** is lower than that of its reactant **8c**, and the reaction proceeds easily to give the final product **6**. This is in agreement with the experimental fact [16] that in the pH range 2−5 the second carboxyl group takes part in the formation of the imide. Theoretically, there exists another route, that is, **1→7→8→5→6** (Scheme 1); however, since it will need 74.8 kJ mol^{-1} to give **5bh** from **5h** (Figure 3) or 131.8 kJ mol^{-1} to give **5b** from **5** in the gas phase (Figure 2), this is certainly a minor pathway. Therefore, the route **1→7→8→9→6** (Scheme 1) is more favorable than **1→7→8→5→6**.

A similar trend is observed in solution as in the gas phase. However, the monoanionic intermediate **8c** and transition structure **TS2-c** of dehydration cannot be located in the present study (shown by the thin dashed line in Figure 4). The calculation shows that the intermediate structure **8c** goes directly to product **8cp** in the optimization. This suggests that dehydration is facile in solution. With inclusion of the entropy contribution at 298.15 K (Figure 9), the energy profiles rise slightly except for the separated system of **9** + H_2O compared with the ZPVE-only corrected case.

Route D (1b→10→8d→9→6). The deprotonation from amidic nitrogen (N_5) in **1b** (Scheme 3) needs 1416.1 kJ mol^{-1}, even higher (by 79.3 kJ mol^{-1}) than that needed from O_1 in route C. After **1b** is deprotonated from nitrogen, however, it can be seen that the energy barrier (**TS1-d**) (Figure 5) between **10** and **8d** is the lowest among all the reaction routes considered in this study. In the gas phase, it costs only 20.5 kJ mol^{-1} to form **8d** from **10**. **8d** dehydrates to form monoanionic complex **8dp**. The energy of **TS2-d** lies lower than **10** by 2.7 kJ mol^{-1}, and lower than its reactant **8d** by 3.2 kJ mol^{-1}. The product

complex **8dp** has an energy −53.2 kJ mol^{-1} lower than **10**. The separated system **9** + H_2O is 10.0 kJ mol^{-1} higher than **10**. As in route C, another route **1b→10→8→5→6** may also be possible, but it is less favorable.

In solution (Figure 5), the energy profile shows a pattern similar to that predicted in the gas phase. **TS1-d** and **TS2-d** are 9.6 and 7.9 kJ mol^{-1} higher than those in the gas phase compared with **10**, while the energies of **8d**, **8dp**, and the separated system **9** + H_2O are significantly lower. It can also be seen that in the solution, **TS2-d** lies higher than **8d** by 17.4 kJ mol^{-1}, while in the gas phase, **TS2-d** lies slightly lower than **8d** by 3.2 kJ mol^{-1}.

With inclusion of the entropy contribution (Figure 10), the energy barriers rise for **TS1-d** and **TS2-d** from 6.0 to 12.0 kJ mol^{-1} for both phases, while the energy drops significantly again for **9** + H_2O. In this route, it can be seen that the carboxyl group of the amino acid is also involved in the reaction. Routes C and D are the two competitive routes.

In brief, the above calculations suggest that solvent effects do not have much influence on the reaction barriers.

Conclusions

Reaction mechanisms of the imide formation in the model system *N*-(*o*-carboxybenzoyl)-L-amino acid have been investigated by use of the B3LYP density functional method. Studies on the amide hydrolysis will be presented in a separate paper.

Our calculations reveal that the route **1→2→3→5→6** (route A) in Scheme 1 is the most favorable reaction route. This corresponds to the highly acidic region studied experimentally. During the dehydration process, water functions as a catalyst. Route B (**1→4→5→6** and **1→5→6**) with pH range 0−2 is ruled out as a likely mechanism. Other routes, corresponding to high pH, are less favorable. These minor routes include **1→2→3→6**, **1→7→8→9→6**, and **1→10→8→9→6**. The most unlikely routes to occur are **1→7→8→5→6** and **1→10→8→5→6**. On the other hand, routes **1→7→8→9→6** and **1→10→8→9→6** are competitive routes. These conclusions are consistent with the experimental results[16] that at highly acidic conditions (hydrogen ion concentration $H_0 < -1$), imide formation is the most favorable pathway (route A, **1→2→3→5→6**), while in the pH range 2−5 (route C), cyclization is not the dominant reaction. Our calculations also show that the carboxyl group of the amino acid is involved in the reaction in both the highly acidic region (hydrogen ion concentration $H_0 < -1$) and the pH range 2−5.

Table 3. C_6-O_2 Bond Distances (Å) at the TS2 Transition State and Its Product in Both the Gas Phase and in Solution

	reactant			TS		
route	structure	gas	solution	structure	gas	solution
A[a]	**3**	1.357	1.355	**TS2-a**	1.612	1.603
A[a]	**5**	1.393	1.391	**TS2-b**	1.693	1.683
A[b]	**5h**	1.419	1.393	**TS3-b**	1.653	1.636
C[a]	**8c**	1.580	NA[c]	**TS2-c**	1.847	NA[c]
D[a]	**8d**	1.559	1.600	**TS2-d**	1.892	1.718

[a] Direct dehydration process. [b] Water assisted dehydration process. [c] Not available.

The C−N bond distance in the transition states (**TS1**) of all routes is shorter in solution than in the gas phase. The shortening is especially large for route A (Table 2). This trend is also observed for the C−O bond distance at the transition state (**TS2**) of dehydration (Table 3).

The energy profiles at 298.15 K show a pattern similar to those at 0 K for both phases. Solvent effects have little influence on the reaction barriers for the imide formation.

Acknowledgment. We gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada and the Killam Trusts for financial support. We also thank the reviewers for fruitful suggestions and comments.

Supporting Information Available: Archive entries of the B3LYP/6-31G(d,p) optimized geometries (Table S1) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA020700Z