Ab Initio Studies on the Catalytic Mechanism of Aspartic Proteinases: Nucleophilic versus General Acid/General Base Mechanism

Hwangseo Park, Junghun Suh, and Sangyoub Lee*

Contribution from the Department of Chemistry and Center for Molecular Catalysis, Seoul National University, Seoul 151-742, South Korea

Received August 9, 1999. Revised Manuscript Received February 14, 2000

Abstract: *On the basis of ab initio calculations at the MP2/6-31G**//RHF/6-31G** level, we study the model reactions for the catalytic action of aspartatic proteinases. We elucidate the mechanistic features of two competing catalytic mechanisms by determining the reaction paths. In contrast to the previous theoretical studies which neglected the electron correlation, the concerted and the stepwise pathways are predicted to be almost equally favored in the general acid/general base mechanism. On the other hand, we find that a concerted reaction pathway is preferred to the stepwise one in the nucleophilic mechanism. We also find that both nucleophilic and general-acid/general-base mechanisms may be operative in a peptide hydrolysis by aspartic proteinases. For the model reaction under consideration, the former is energetically more favored if one considers only the potential energy profile along the intrinsic reaction coordinate, as has been done in previous theoretical studies. However, when the entropic contribution due to nuclear motions as well as the zero-point vibrational energies is included, the latter is predicted to be the preferred one by 1.9 kcal/mol. The covalent intermediate, which is the end-point minimum energy complex on the concerted nucleophilic pathway, is found to be a unstable one that is 18.1 kcal/mol higher in free energy than the incipient model enzyme−substrate complex. This fact is in accordance with the previous experimental implications that decomposition of the covalent intermediate is much faster than its formation. The present work provides a theoretical support for the persistent argument that the possibility of the nucleophilic mechanism cannot be excluded in the catalytic action of aspartic proteinases although the required experimental detection of the covalent intermediate has been unsuccessful so far. It is demonstrated that for all reaction pathways under consideration, the protonation of the nitrogen atom belonging to the peptidic bond is an essential step in crossing the activation barrier for the rupture of a peptide substrate. The relevance of the mechanistic features observed for the model reactions to an enzymatic reaction is discussed.*

Introduction

The aspartic proteinases (Apases) are a group of proteolytic enzymes containing two aspartyl (Asp) residues in their active sites. Among them are pepsin, penicillopepsin, renin, and HIV-1 protease that has been widely studied due to the pharmaceutical interest. X-ray crystallographic data for the various APases reveal that their active site regions are very similar in structure: two Asp residues lie close to each other at the center of a deep and extended active site cleft.[1] It is generally agreed that these two Asp residues play an essential role in the catalysis of peptide hydrolysis.[2] The protonation state of the Asp dyad has been considered as an important aspect in understanding the hydrolytic mechanism of Apases. In this regard, much effort has been devoted to the determination of pK_a values of the two catalytic Asp residues.[3]

Despite the detailed knowledge of the structure, the elucidation of a catalytic mechanism has been the subject of persisting controversy. Although the detailed mechanism by which APases cleave proteins is still uncertain, much attention has been paid to the two main possibilities for the detailed roles of the two Asp residues in the catalysis. The first is a nucleophilic attack of the enzyme on the aminocarbonyl carbon of the peptide bond and the expulsion of the amine component, yielding an anhydride intermediate (nucleophilic mechanism). The second alternative involves the general-base assisted attack of catalytic water molecule upon the carbonyl, followed by prototropic shifts and a direct elimination to give the product acid and amine (general-acid/general-base mechanism).

A distinctive mark of the nucleophilic mechanism is the formation of intermediates that contain a covalent bond formed by the transfer of acyl group from a peptide substrate to an Asp residue of enzyme. Despite a great deal of effort, experimental detection of this kind of covalent intermediates for a variety of peptide substrates has been unsuccessful.[4] On the other hand, the formation of a noncovalent enzyme-bound amide hydrate intermediate was confirmed for some peptide

* To whom correspondence should be addressed. Telephone: +82-2-875-4887. Fax: +82-2-874-3704. E-mail: sangyoub@snu.ac.kr.

(1) (a) Subramanian, E.; Swan, I. D. A.; Liu, M.; Davies, D. R.; Jenkins, J. A.; Tickle, I. J.; Blundell, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 556. (b) Hsu, I.-N.; Delbaere, L. T. J.; James, M. N. G.; Hofmann, T. *Nature* **1977**, *266*, 140.

(2) (a) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider J.; Kent, S. B. H. *Science* **1989**, *245*, 616. (b) Clement, G. E. *Prog. Bioorg. Chem.* **1973**, *2*, 177.

(3) (a) Trylska, J.; Antosiewicz, J.; Geller, M.; Hodge, C. N.; Klabe, R. M.; Head, M. S.; Gilson, M. K. *Protein Sci.* **1999**, *8*, 180 and references therein. (b) Hyland, L. J.; Tomaszek, T. A., Jr.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454.

10.1021/ja992849p CCC: $19.00
Published on Web 04/05/2000

substrates.[3b,5] This provides unmistakable support for the general-acid/general-base (GA/GB) mechanism. Nevertheless, the nucleophilic route has not been ruled out conclusively since it is possible that the covalent intermediates may not accumulate to an experimentally detectable amount if their formation is the rate-limiting step and any subsequent steps that lead to their removal occur much faster. Furthermore, there has been some supporting evidence for a nucleophilic role of active site Asp residue in the hydrolysis of some model peptide substrates. The first example was the observation of transpeptidation reactions for a variety of substrates in which an anhydride intermediate and an amino moiety would be involved.[6] From the kinetic studies on the hydrolysis of a geometrically distorted amide with dicarboxylic acids capable of forming cyclic anhydrides, Somajaji and Brown provided persuasive evidence for the formation of anhydride intermediates in the course of the reaction.[4c]

Although the earlier theoretical studies of Apases were focused mostly on the determination of stable conformations of active sites with different protonation states,[7] some ab initio results have already been reported for the catalytic reaction paths associated with the GA/GB mechanism of HIV-1 protease action, by using formamide molecule and the formate/formic acid pair as models for substrate and enzyme, respectively. Lee and co-workers determined the reaction path for the concerted GA/GB mechanism at the RHF/6-31G* level of theory,[8] and found that the forward hydrolysis barrier is about 22 kcal/mol with a reverse barrier of about 34 kcal/mol. More recently, Venturini and co-workers determined the stationary-state structures that appear on the reaction path of the stepwise GA/GB mechanism at the MP2/6-31+G* level of theory.[9] They also showed that in the GA/GB mechanism the stepwise pathway is energetically favored over the concerted one by about a 10 kcal/mol difference in activation energy at the RHF/6-31G* level, in accordance with the experimental proposal by Meek and co-workers.[3b,5]

In the present work, we investigate the mechanistic features and energetics of the nucleophilic mechanism of Apases-catalyzed peptide hydrolysis by examining the intrinsic reaction coordinates (IRC) of both concerted and stepwise pathways for a model reaction. This model system consists of the formamide and acetate/acetic acid pair, acting for a peptide substrate and active sites of Apases, respectively. Although various protonation states have been proposed for the free enzymes and the enzyme−inhibitor complexes,[3a,10] there has been experimental evidence that the two catalytic Asp residues of HIV-1 PR exist in opposite states of protonation when the enzyme is complexed with various substrates.[3b] As a model for the monoprotonated aspartyl dyad, we use the acetate/acetic acid pair. This kind of model approach has also been adopted in other theoretical studies.[8,9] The reason for taking the acetate/acetic acid pair rather than the formate/formic acid pair is to suppress the presence of a spurious C−H···O hydrogen bond[8] by removing the hydrogen attached to a sp^2 carbon. To compare the energetics of the competing catalytic mechanisms at the same level of theory, we also determine the minimum energy paths of concerted and stepwise GA/GB mechanisms by adding a catalytic water molecule to the above model system for the nucleophilic mechanism. Comparison of the present results for the concerted and stepwise GA/GB mechanisms with those obtained in refs 8 and 9, respectively, would also allow us to estimate the effect of a structural variation in active sites on the catalytic behavior of Apases, at least in a qualitative manner.

In addition to the potential energy profile, we also evaluate the variation of entropy along the catalytic paths. The entropic factor has not been considered in the previous theoretical studies of Apases,[8,9] although it is known to be an important factor in determining the reactivity of weakly bonded complexes due to the presence of low-frequency vibrations inherent to the intermolecular coordinates.[11] The free energies of activation for the four reaction paths under consideration are then compared to determine which mechanism is preferable for the catalytic action of APases. It will be shown that the entropy factor can play a significant role in determining the catalytic pathway of Apases. By examining the geometrical features of the stationary-state structures on each reaction pathway, we also indicate the basic elements of the catalytic actions that are inherent in the four competitive catalytic mechanisms of peptidic bond cleavage by APases.

Computational Details

All geometries corresponding to minima and transition states on the four reaction pathways are fully optimized at the RHF/6-31G** level of theory with the latest version of GAMESS code.[12] These geometry optimizations are performed with the aid of analytically determined gradients and search algorithms of the quasi-Newton−Raphson procedure.[13] The nature of each SCF stationary point is determined by the number of imaginary frequencies that are evaluated by diagonalizing the analytical matrix of energy second derivatives (Hessian). Each transition state structure possesses a single negative eigenvalue of the Hessian matrix, and the corresponding imaginary vibrational frequency is related to the time scale of the motion that carries the reaction system over the energy barrier. The intrinsic reaction coordinate (IRC) connecting a transition state to the neighboring stable structures is determined at the RHF/6-31G** level using the Gonzalez−Schlegel second-order (GS2) method.[14]

To get a better prediction of energetics, post-HF level calculations including the effect of electron correlation are performed at the RHF/6-31G** optimized geometries. These single-point calculations are carried out with the 6-31G** basis sets using the Møller−Plesset

(4) (a) Dunn, B. M.; Fink, A. L. *Biochemistry* **1984**, *23*, 5241. (b) Hoffmann, T.; Fink, A. L. *Biochemistry* **1984**, *23*, 5247. (c) Somajaji, V.; Keillor, J.; Brown, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 2625 and references therein.

(5) (a) Hyland, L. J.; Tomaszek, T. A., Jr.; Roberts, G. D.; Carr, S. A.; Magaad, V. W.; Bryan, H. L.; Fakhoury, S. A.; Moore, M. L.; Minnich, M. D.; Culp, J. S.; DesJarlais, R. L.; Meek, T. D. *Biochemistry* **1991**, *30*, 8441. (b) Rodriguez, E. J.; Angeles, T. S.; Meek, T. D. *Biochemistry* **1993**, *32*, 12380.

(6) (a) Takahashi, M.; Wang, T. T.; Hofmann, T. *Bicochem. Biophys. Res. Commun.* **1974**, *57*, 39. (b) Wang, T. T.; Hofmann, T. *Biochem. J.* **1976**, *153*, 691.

(7) (a) Beveridge, A. J.; Heywood, G. C. *Biochemistry* **1993**, *32*, 3325. (b) Beveridge, A. J.; Heywood, G. C. *J. Mol. Struct.* **1994**, *306*, 235.

(8) Lee, H.; Darden, T. A.; Pedersen, L. G. *J. Am. Chem. Soc.* **1996**, *118*, 3946.

(9) Venturini, A.; Lopez-Ortiz, F.; Alvarez, J. M.; Gonzalez, J. *J. Am. Chem. Soc.* **1998**, *120*, 1110.

(10) (a) Clement, G. E.; Snyder, S. L.; Price, H.; Cartmell, R. *J. Am. Chem. Soc.* **1968,** *90*, 5603. (b) Cornish-Bowden, A. J.; Knowles, J. R. *Biochem. J.* **1969**, *113*, 353. (c) Hunkapiller, M. W.; Richards, J. H. *Biochemistry* **1972**, *11*, 2829. (d) James, M. N. G.; Sielecki, A. *J. Mol. Biol.* **1983**, *163*, 299. (e) Pearl, L. H.; Blundell, T. L. *FEBS Lett.* **1984**, *174*, 96. (f) Green, D. W.; Ayknet, S.; Gierse, J. K.; Zupec, M. E. *Biochemistry* **1990**, *29*, 3126. (g) Wang, Y.-X.; Freedberg, D. I.; Yamazaki, T.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Kiso, Y.; Torchia, D. A. *Biochemistry* **1996**, *35*, 9945. (h) Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. *Nature Struct. Biol.* **1996**, *3*, 946.

(11) (a) Pross, A. *Theoretical and Physical Principles of Organic Reactivity*; Wiely: New York, 1995; pp 130−136. (b) Hobza, P.; Zahradnik, R. *Top. Curr. Chem.* **1980**, *93*, 53.

(12) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.

(13) (a) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385. (b) Helgaker, T. *Chem. Phys. Lett.* **1991**, *182*, 305. (c) Bell, S.; Crighton, J. S. *J. Chem. Phys.* **1984**, *80*, 2464.

(14) Gonzalez, C.; Schlegel, B. H. *J. Chem. Phys.* **1989**, *90*, 2154.

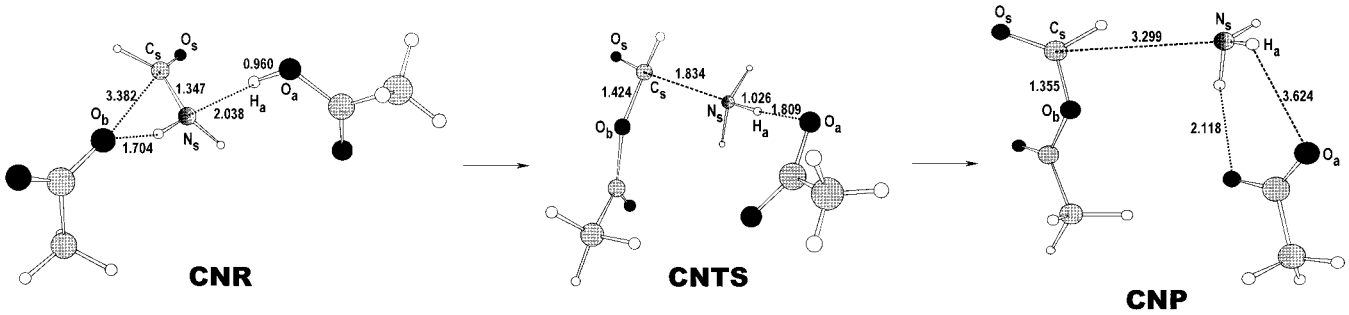

Figure 1. A pictorial description of the reaction pathway of the concerted nucleophilic mechanism for a peptide hydrolysis by Apases. Some selected interatomic distances are given (in Å).

second-order perturbation theory (MP2).[15] The electronic energies computed in this way are used to calculate the relative free energies (ΔG) that are given by

$$\Delta G = \Delta E_{\text{elec}} + \Delta H' - T\Delta S$$

The reference state of zero energy is arbitrarily taken as that resembling the isolated enzyme and substrate. Hence the reference system for the calculation of nucleophilic mechanisms consists of the acetate/acetic acid pair and an isolated formamide, while that for the GA/GB mechanism consists of an isolated formamide and the complex of the acetate/acetic acid pair and the water molecule. $\Delta H'$ denotes the enthalpy change due to thermal motions of the nuclei including the zero-point vibrational energies. ΔS denotes the entropy change. While the electronic energies are evaluated from the MP2/6-31G** level calculations, the vibrational frequencies used to estimate $\Delta H'$ and ΔS are obtained at the RHF/6-31G** level.

The entropies of transition states and minima are evaluated by summing the translational, rotational, and vibrational contributions ($S = S_{\text{trans}} + S_{\text{rot}} + S_{\text{vib}}$), each of which is in turn obtained from the corresponding molecular partition function. In the calculation of the molecular partition functions, rigid-rotor and harmonic-oscillator approximations are used to factorize the contributions from rotation and vibration. The single imaginary frequency mode of the transition state is excluded in the calculation of the vibrational partition function, while all normal modes are considered for the structures of energy minima.

All input structures for obtaining optimized reactant states are prepared from a molecular mechanics energy-minimization calculation using the AMBER force field under the constraints that the hydrogen belonging to the carboxylic acid group should be directed to nitrogen (concerted pathways) or oxygen (stepwise pathways) of the amide group. These constraints are based on the enzyme−substrate complex structures proposed experimentally by X-ray crystallography[16] and by a kinetic study.[5b]

Results and Discussion

Concerted Nucleophilic Mechanism. Figure 1 displays snapshots of the calculated IRC: the reactant structure (**CNR**), the transition state structure (**CNTS**), and the product structure (**CNP**). It is apparent that the carboxylate/carboxylic acid pair hydrolyzes the peptide substrate in a cooperative manner: the carboxylate group acts as a nucleophile, while the carboxylic acid group participates in the reaction as a general acid which protonates the nitrogen (N_s) atom of the substrate peptidic bond. As a consequence of this proton transfer, the N_s atom undergoes a pyramidalization[17] that causes the weakening of the peptidic C−N bond and the increase of the electrophilicity of the aminocarbonyl carbon (C_s). Thus the C−O bond formation between the C_s atom and carboxylate oxygen (O_b) and the breakage of the C_s-N_s bond occur simultaneously through the S_N2-type reaction, facilitated by the proton (H_a) transfer from the carboxylic acid group to the N_s atom.

The reaction pathway starts with the formation of an enzyme−substrate complex (**CNR**) in which the N_s atom is hydrogen bonded to the undissociated carboxylic acid group at a distance of 2.038 Å. A stronger hydrogen bond is observed between the negatively charged oxygen (O_b) of the carboxylate group and N−H hydrogen of the formamide that are separated by 1.704 Å. The $O_b\cdots H-N_s$ and $N_s\cdots H_a-O_a$ hydrogen bonds make major contributions to the stabilization of **CNR**.

The reaction proceeds from **CNR** with the approach of the O_b atom toward the C_s atom along a line perpendicular to the amidic plane and the gradual shift of the carboxylic proton (H_a) to the N_s atom. This protonation-induced pyramidalization at N_s increases the p-character in N_s hybridization, and weakens the C_s-N_s bond. Hence, the C_s atom becomes more susceptible to a nucleophilic attack, promoting the formation of the O_b-C_s bond.

At the transition state (**CNTS** in Figure 1), the carboxylic proton H_a in the $N_s\cdots H_a\cdots O_a$ hydrogen bond is almost fully shifted to the N_s side; the N−H distance decreases from 2.038 Å in **CNR** to 1.026 Å in **CNTS**. In terms of the O_b-C_s bond formation also, we have a late transition state; the bond distance decreases considerably from 3.382 to 1.424 Å. On the other hand, the C_s-N_s bond cleavage lags behind. The bond is stretched from 1.347 Å in **CNR** to 1.834 Å in **CNTS**, while the separation between the C_s and N_s atoms is 3.299 Å in the product state, **CNP**. The single imaginary frequency associated with **CNTS** is dominated by the N_s-H_a bond forming motion as well as the C_s-N_s bond breaking motion, implying that the proton transfer to N_s is an essential step to cross the activation barrier for the cleavage of an amidic bond. A similar mechanistic feature has also been observed in the intramolecular hydrolysis of the amidic bond in *N*-methylmaleamic acids,[18] as well as in the hydrolysis reaction between water and formamide molecules.[19]

The minimum energy structure located at the end of the reaction path, **CNP** in Figure 1, corresponds to a complex of reaction products and an enzyme. In this complex the C_s-N_s bond is fully broken while the N_s-H_a bond formation is complete, leading to the formation of the ammonia molecule. An N−H···O hydrogen bond between the carboxylate group

(15) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(16) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schnerider, J.; Kent, S. B.; Wlodawer, A. *Biochemistry* **1991**, *30*, 1600.

(17) (a) Silva, A. M.; Cachau, R. E.; Sham, H. L.; Erickson, J. W. *J. Mol. Biol.* **1996**, *255*, 321. (b) The two optimized $H-N_s-C_s$ angles of the isolated formamide molecule are 121.6° and 119.1°. These are reduced to 116.7° and 116.4°, respectively, in the model enzyme−substrate complex (**CNR** in Figure 1).

(18) Park, H.; Suh, J.; Lee, S. *J. Mol. Struct.* (*THEOCHEM*) **1999**, *490*, 47.

(19) (a) Krug, J. P.; Popelier, P. L. A.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 7604. (b) Antonczak, S.; Ruiz-Lopez, M. F.; Rivail, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 3912.

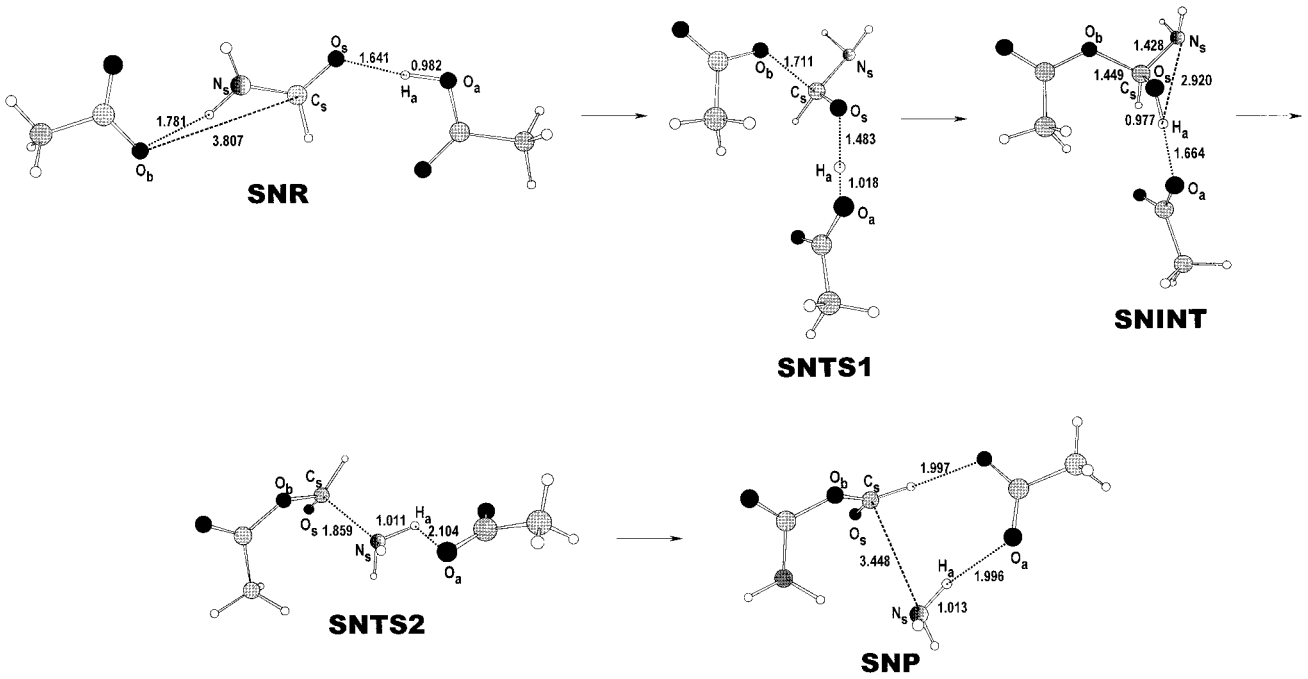

Figure 2. A pictorial description of the reaction pathway of the stepwise nucleophilic mechanism for a peptide hydrolysis by Apases. Some selected interatomic distances are given (in Å).

and ammonia molecule is formed with a relatively larger O···H distance, 2.118 Å. The complex **CNP** involves an anhydride intermediate formed between the enzyme and a substrate by a covalent bond. Such a covalent intermediate has been sought by experimentalists for a long time as supporting evidence for the nucleophilic mechanism in Apases-catalyzed peptide hydrolysis.

The MP2/6-31G**//RHF/6-31G** level of theory predicts that the free energy of **CNP** is higher than that of **CNR** by 18.1 kcal/mol at 298.15 K and 1 atm. This relative instability of **CNP** is likely due to the lack of a strong hydrogen bond interaction as in **CNR**, and could be one of the reasons that the anhydride intermediate has not been detected, for its decomposition would be much faster than its formation.[4]

Stepwise Nucleophilic Mechanism. Figure 2 displays the stationary-state structures located along the reaction path. The first minimum energy structure, **SNR**, corresponds to an enzyme−substrate complex. In contrast to the complex **CNR**, the target for the electrophilic attack of the carboxylic acid group on the amidic bond is the oxygen (O_s) rather than the nitrogen (N_s), resulting in the formation of a stronger $O_s\cdots H_a-O_a$ hydrogen bond (with the $O_s\cdots H_a$ distance of 1.641 Å) than the $N_s\cdots H_a-O_a$ one in **CNR**. Another distinctive feature is that the formamide molecule remains planar in **SNR**; the two optimized $H-N_s-C_s$ angles are 124.7° and 120.5°. The complex is further stabilized through the hydrogen bond interaction between the negatively charged oxygen O_b of the carboxylate group and a hydrogen attached to the N_s atom at a distance of 1.781 Å.

The first reaction step involves the nucleophilic attack of the carboxylate oxygen O_b upon the C_s atom, accompanied by the proton (H_a) transfer from the carboxylic acid group to the carbonyl oxygen of the formamide, leading to the formation of the first transition state **SNTS1**. As a consequence of these geometrical changes the planarity of the peptidic bond is lifted. The two $H-N_s-C_s$ angles of the formamide molecule in **SNTS1** are reduced to 107.6° and 110.4°, respectively, and the bond between C_s and O_s lengthens with the increase of the p-character in C_s hybridization. Despite the 2.1 Å decrease in the O_b-C_s distance, the transition state is relatively early in terms of both C_s-O_s bond lengthening and O_s-H_a bond formation. The changes in these two bond lengths are 0.034 and 0.158 Å, which correspond to 25% and 24% advancements toward the formation of an intermediate, **SNINT**. The single imaginary frequency of **SNTS1** is dominated by the O_b-C_s and O_s-H_a bond forming motions, implying that the activation energy arises from distortion of the planar geometry of the amide link through the protonation of O_s and the simultaneous nucleophilic attack on C_s.

Passing through the first transition state, the reaction system falls into an intermediate minimum energy structure, **SNINT**, which is tetrahedral at the C_s atom. In this complex, the O_b-C_s bond is fully developed with the proton transfer from the carboxylic acid group to O_s atom being completed. The second reaction step is then initiated by lengthening of the C_s-N_s bond in **SNINT**, accompanied by a proton transfer from the O_s to the N_s atom. During this intramolecular proton transfer, the O_a atom of the carboxylate group in **SNINT** approaches the N_s atom to form a relatively weak O···H hydrogen bond with the ammonia that is formed in the second transition state, **SNTS2**. Although the C_s atom still remains tetrahedral in this transition structure, the C_s-N_s bond lengthens by 0.41 Å as compared to that in **SNINT**. A most significant feature associated with **SNTS2** is that the H_a atom attached to O_s in **SNINT** is fully transferred to N_s, indicating again that protonation of the amidic nitrogen is essential for the rupture of the amidic C−N bond.

From **SNTS2**, a further increase in the C_s-N_s distance leads to the final minimum energy structure, **SNP**, which is planar at the C_s atom due to the full breakage of the C_s-N_s bond. This complex is similar in structure to the **CNP** found in the concerted pathway except for the presence of a hydrogen bond between a carboxylate oxygen and the hydrogen atom attached to the C_s. As can be seen in Figure 2, the two hydrogens attached to C_s and N_s atoms are symmetrically shared by two oxygen atoms of the catalytic carboxylate group. Although the C−H···O hydrogen bond may be a spurious one due to the simplicity of the model, this novel-type hydrogen bond is now

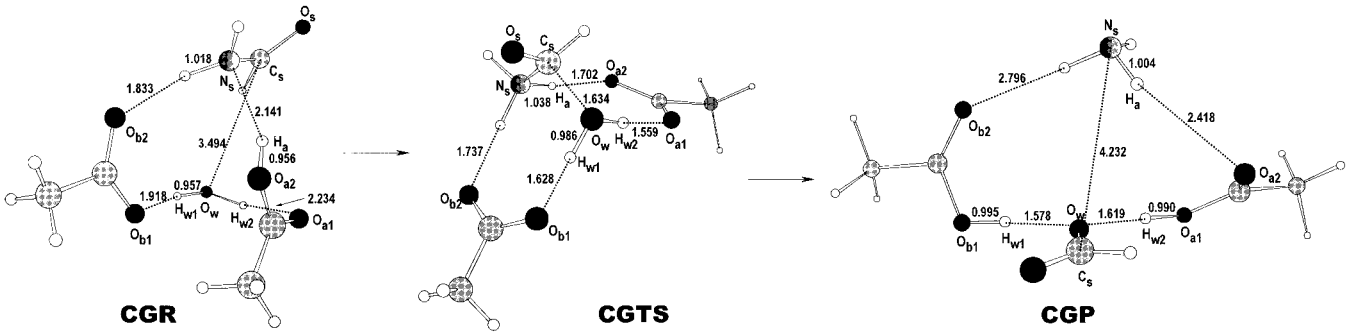

Figure 3. A pictorial description of the reaction pathway of the concerted GA/GB mechanism for a peptide hydrolysis by Apases. Some selected interatomic distances are given (in Å). The C_s-N_s distances in **CGR** and **CGTS** are 1.355 and 1.527 Å, respectively.

known to be an important kind of intra- or intermolecuar interaction in some chemical and biological systems.[20]

Concerted GA/GB Mechanism. A crucial factor that distinguishes the GA/GB mechanism from the nucleophilic counterpart is the participation of the catalytic water molecule as a nucleophile. In this case the carboxylate group of one Asp residue acts as a general base instead of a nucleophile, deprotonating the lytic water molecule and thereby increasing the nucleophilicity of the water oxygen. The minimum energy path for the concerted GA/GB mechanism is depicted in Figure 3.

The first minimum energy structure on the IRC (**CGR** in Figure 3) may correspond to an enzyme−substrate complex in which the hydrogen (H_a) belonging to the carboxylic acid group of an Asp residue is directed to the nitrogen atom (N_s) of the amide group, forming a weak hydrogen bond at a distance of 2.141 Å. In this structure, two hydrogens of the catalytic water molecule are hydrogen bonded to the two Asp residues: one (H_{w1}) with a negatively charged oxygen (O_{b1}) of the carboxylate group and the other (H_{w2}) with the carbonyl oxygen (O_{a1}) of the carboxylic acid group with distances of 1.918 and 2.234 Å, respectively. Similar to **CNR**, the strongest hydrogen bond is established between one of the hydrogens attached to the amide nitrogen and the second oxygen atom (O_{b2}) of the carboxylate group.

At structure **CGR**, the reaction is initiated by the approach of water oxygen (O_w) to the C_s atom, facilitated by proton transfer from the carboxylic acid group of one Asp residue to the amide nitrogen and partial deprotonation of the water molecule by the carboxylate group of the other Asp residue. When the O_w-C_s distance is reduced to 1.634 Å, the reaction system reaches the transition state (**CGTS** in Figure 3). In this transition-state structure a proton (H_a) is almost fully transferred to the N_s atom, rupturing the planarity of the amide linkage. In contrast to significant progress in O−C bond formation, the breaking of the C−N bond lags behind: while the O−C distance is contracted from 3.494 to 1.634 Å, the C−N bond is stretched from 1.355 to 1.526 Å. It is noteworthy that $O_w-H_{w2}\cdots O_{a1}$, $O_w-H_{w1}\cdots O_{b1}$, and $N_s-H\cdots O_{b2}$ hydrogen bonds get stronger in going from **CGR** to **CGTS**.

Further advancement in O_w-C_s bond formation, driven by transfers of two protons of the water molecule to the two Asp residues, leads to the end point minimum (**CGP** in Figure 3), which may correspond to reaction products complexed with the two catalytic Asp side chains. In this structurethe O_w-C_s bond is fully formed with breakage of the N_s-C_s bond, leading to the formation of the ammonia molecule and formate hydrogen bonded with two carboxylic acids.

Stepwise GA/GB Mechanism. Figure 4 displays the structures of the reaction system at the energy minima and transition states along the stepwise reaction pathway for the GA/GB mechanism. In the optimized structure of the model enzyme−substrate complex (**SGR** in Figure 4), the water molecule is hydrogen bonded to the negatively charged oxygen atom (O_{b1}) of the catalytic carboxylate group with a distance of 1.938 Å. A stronger hydrogen bond is formed between the carbonyl oxygen (O_s) of the formamide and the hydrogen (H_a) of the carboxylic acid group with a distance of 1.655 Å, which has a comparable bond strength as that found in **SNR**. This partial protonation of O_s again leaves the planarity of the amidic link intact.

From the complex **SGR**, the reaction proceeds as the oxygen of water (O_w) approaches the C_s atom along a line approximately perpendicular to the amidic plane. This nucleophilic attack is facilitated by two proton-transfer processes: H_a from the carboxylic acid group to O_s, and H_{w1} from the water molecule to O_{b1}. The former increases the susceptibility of C_s to a nucleophilic attack through the polarization of the C_s-O_s bond, while the latter enhances the nucleophilicity of O_w. During these changes a new hydrogen bond is established between the second hydrogen (H_{w2}) of the water molecule and the carbonyl oxygen (O_{a1}) of the carboxylic acid group.

When the C_s-O_w distance is reduced to 1.781 Å, the reacting system reaches the first transition state, **SGTS1**, in which the N_s atom has undergone a considerable pyramidalization, as monitored by two $H-N_s-C_s$ angles (see Figure 4). The proton (H_a) belonging to the carboxylic acid group is almost fully transferred to O_s, forming the $O_s-H_a\cdots O_{a2}$ hydrogen bond with the $H_a\cdots O_{a2}$ distance of 1.550 Å. In this short strong hydrogen bond the distance between the H-acceptor (O_{a2}) and the H-donor (O_s) is found to be 2.543 Å, which is comparable to the sum of van der Waals radii of the two oxygen atoms (2.55 Å).[21] This result provides evidence for the establishment of a low-barrier hydrogen bond (LBHB) during the catalytic action of Apases, which has also been invoked in various enzymatic reactions to explain the rate enhancement.[22] The involvement of the LBHB has already been appreciated in the catalytic mechanism of HIV-I protease.[5b] However, in the previous studies, including the ab initio study by Venturini and co-workers,[9] such LBHB was found to appear in the minimum energy structure corresponding to **SGR** rather than at the transition state. Since it is

(20) (a) Jeffrey, G. A. *An introduction to Hydrogen bonding*; Oxford University Press: Oxford, 1997, and references therein. (b) Desiraju, G. *Acc. Chem. Res.* **1991**, *24*, 290. (c) Song, J.-S.; Szalda, D. J.; Bullock, R. M. *J. Am. Chem. Soc.* **1996**, *118*, 11134. (d) Karle, I. L.; Ranganathan, D.; Haridas, V. *J. Am. Chem. Soc.* **1996**, *118*, 7128.

(21) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927.

(22) (a) Scheiner, S.; Kar, T. J. *Am. Chem. Soc.* **1995**, *117*, 6970. (b) Schwartz, B.; Drueckhammer, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 11902. (c) Shan, S.; Herschlag, D. H. *Science* **1996**, *272*, 97.

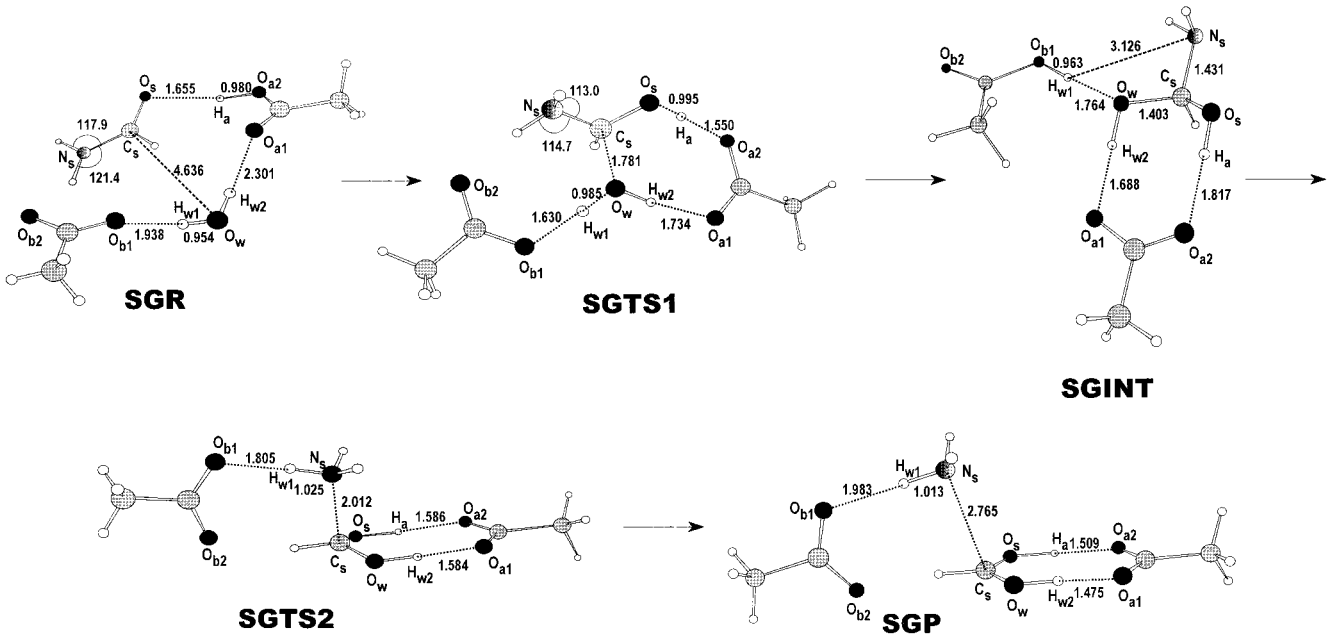

Figure 4. A pictorial description of the reaction pathway of the stepwise GA/GB mechanism for a peptide hydrolysis by Apases. Some selected interatomic distances and angles are given (in Å and deg), respectively.

the stabilization of the transition state rather than the reactant structure that is more relevant to the rate enhancement, we believe that the finding in this study should be more appropriate to the enzymatic catalysis.

A further approach of O_w to C_s at the first transition state carries the reaction system into the second energy minimum (**SGINT** in Figure 4), where the O_w-C_s bond is fully formed with the H_{w1} atom being completely transferred to the catalytic carboxylate group. As already pointed out by Venturini and coworkers,[9] the complex **SGINT** may correspond to the enzyme-bound noncovalent amide hydrate intermediate that has been characterized by experiment and molecular modeling as supports for the GA/GB mechanism. In this complex the distance between N_s and the hydrogen (H_{w1}) belonging to the carboxylic group is 3.126 Å, which is in a good agreement with the previous ab initio result (3.065 Å) using the formic acid/formate pair as a model for enzyme,[9] as well as that obtained by a molecular dynamics study (3.1 Å).[5a]

The next reaction step involves an intermolecular proton transfer from the carboxylic acid group to the amine group (the approach of H_{w1} toward N_s) and simultaneous elongation of the C_s-N_s bond, leading to formation of the second transition state, **SGTS2**. During these changes the $O_{b1}-H_{w1}\cdots O_w$ hydrogen bond in **SGINT** is broken and replaced by a new weaker $O_{b1}\cdots H_{w1}-N_s$ hydrogen bond; the $O_{b1}\cdots H_{w1}$ distance in the latter is longer than the $O_w\cdots H_{w1}$ distance in the former by 0.041 Å (see Figure 4). On the other hand, the two hydrogen bonds between hydroxyl groups attached to C_s and two carboxylate oxygens O_{a1} and O_{a2} get stronger in going from **SGINT** to **SGTS2**, revealing a local C_{2v} symmetry. In these two symmetrical hydrogen bonds, the distances between H-donor and H-acceptor fall within 2.59 Å, supporting again the involvement of the LBHB in the catalytic mechanism of Apases. The transition state **SGTS2** is late in terms of both C_s-N_s bond breakage and N_s-H_{w1} bond formation; the two bond lengths are 2.012 and 1.025 Å, respectively. Thus the protonation of the N_s atom is again necessary to cross the activation barrier for the second reaction step.

The reaction system then transforms to the final minimum energy structure, **SGP**, which corresponds to the reaction products complexed with the catalytic residues. In this structure the C_s-N_s bond is fully broken, leading to the formation of the ammonia molecule and causing the C_s atom to be planar.

Energetics and Relevance to the Enzymatic Catalysis. Earlier experimental findings indicate that the hydrogen-bonding pattern between the carboxylic acid group of a catalytic Asp residue and a substrate plays an essential role in determining the catalytic pathway: in the stepwise pathway the target for electrophilic attack of an Asp residue is the oxygen atom of the substrate amide group, while that in the concerted pathway is amide nitrogen. All four reactant state structures considered in this work are consistent with such geometrical features found in the enzyme−substrate complex: the reactant state structures that involve N-protonation of the formamide molecule (**CNR** and **CGR**) undergo concerted reactions, while those in which the formamide reveals O-protonation (**SNR** and **SGR**) lead to the stepwise pathway.

Previous theoretical studies of Apases (specifically the HIV-protease)[8,9] were based on the calculation of the potential energy profile along the IRC. However, it is well-known that entropic effects arising from the thermal nuclear motions transverse to the IRC, including the zero-point vibrational motions, are just as important in determining the reactivity of weakly bonded complexes.[11] We therefore compute the various thermodynamic potentials at 25 °C and 1 atm by using the usual statistical mechanical expressions as adopted in the GAMESS code. Vibrational frequencies of the complexes required for the calculations are evaluated at the RHF/6-31G** level of theory. Table 1 displays the relative contributions of various energetic factors to the free energies of the complexes.

We note that in all cases the variation of the entropic term along the reaction path is much smaller in magnitude than the electronic energy variation. That is, the major contribution to the free energy of activation ($\Delta G^{\ddagger}$) comes from the activation enthalpy ($\Delta H^{\ddagger}$) rather than from the activation entropy term ($-T\Delta S^{\ddagger}$). Nonetheless, it is essential to include the entropic contribution. When the electronic energies are considered alone, the concerted nucleophilic path is the most preferred one, but

Table 1. Relative Contributions of Various Energetic Factors to the Free Energies of the Stationary States along the Reaction Pathways[a]

energy minima and transition states	ΔE_{elec}	$\Delta H'$	$-T\Delta S$	ΔG
concerted nucleophilic				
CNR	−7.1 (−6.2)	4.9	7.8	5.6
CNTS	13.0 (24.8)	4.5	13.9	31.4
CNP	10.8 (12.6)	4.1	8.8	23.7
stepwise nucleophilic				
SNR	−18.9 (−16.5)	5.1	9.9	−3.9
SNTS1	2.8 (12.3)	4.1	14.5	21.4
SNINT	−3.0 (2.8)	6.1	14.4	17.5
SNTS2	15.8 (26.8)	4.5	12.4	32.7
SNP	8.2 (10.2)	3.6	10.3	22.2
concerted GA/GB				
CGR	−12.5 (−9.7)	7.2	12.1	6.8
CGTS	8.1 (27.3)	6.5	16.1	30.7
CGP	−16.0 (−12.8)	5.6	10.9	0.5
stepwise GA/GB				
SGR	−16.3 (−14.4)	4.8	10.1	−1.4
SGTS1	−0.9 (10.2)	4.4	15.3	18.8
SGINT	−8.6 (−4.3)	6.6	13.6	11.6
SGTS2	6.2 (14.3)	4.0	12.9	23.1
SGP	5.3 (9.9)	2.6	12.6	20.5

[a] The energy values for the complexes involved in the nucleophilic pathway are measured with respect to that of the system comprising the acetate/acetic acid pair and an isolated formamide. On the other hand, the reference system for the GA/GB mechanism consists of an isolated formamide and the complex of the acetate/acetic acid pair and water molecule. Numbers in parentheses are the energy calculated at the RHF/6-31G** level. All energy values are given in kcal/mol.

with the inclusion of the entropic factor both the concerted and stepwise GA/GB mechanisms turn out to be the better reaction pathways.

In contrast to the previous theoretical studies[8,9] in which the electron correlation effects are not taken into account, the activation energy of the concerted GA/GB mechanism is found to be about 1.9 kcal/mol lower than that of the stepwise counterpart at the MP2/6-31G**//RHF/6-31G** level of theory. Comparing the potential energies of stationary-state structures in the two reaction pathways evaluated at the MP2 and RHF levels, it follows immediately that the change in the preferred reaction pathway from stepwise GA/GB to concerted GA/GB can be attributed to the extraordinarily large stabilization of **CGTS** compared to **SGTS2** by the electron correlation effects. This result indicates that the inclusion of electron correlation is essential in evaluating the activation energy for a reaction pathway. A similar energetic feature was also observed for a Cope elimination reaction: due to the preferential stabilization of transition state by electron correlation effects, the activation energy determined at the MP2 level is about 17 kcal/mol lower than that obtained at the RHF level.[23] With the inclusion of entropic factors, however, the stepwise GA/GB pathway becomes almost equally favorable as the concerted one. The latter is predicted to be only slightly favored by a 0.6 kcal/mol difference in activation free energies. Thus, the present theoretical study indicates that both concerted and stepwise GA/GB mechanisms, which were proposed independently by Jaskolski et al.[16] and Meek et al.[5b] based on X-ray structure and kinetic studies, respectively, can be competitive in the catalytic action of APases.

From Table 1, we observe that the complex **SNR** is more stable than **CNR** by about 9.5 kcal/mol. By comparing the $N_s\cdots H_a$ distance in **CNR** (2.038 Å) to the $O_s\cdots H_a$ distance in **SNR** (1.641 Å), we may deduce that this energy difference stems from the difference in the extent of stabilization through the hydrogen bond; in general, oxygen would be a better H-acceptor than nitrogen. This is consistent with previous experimental findings that in formamide molecule oxygen is protonated in preference to nitrogen in an acidic medium.[24] Despite the stabilization of **SNR** (Figure 2), formation of the first transition state **SNTS1** requires 25.3 kcal/mol. In going from the intermediate **SNINT**, which has the energy of 21.4 kcal/mol above the **SNR**, to the second transition state **SNTS2**, the breaking of the amide bond poses a further barrier of 15.2 kcal/mol. Hence the overall free energy of activation $\Delta G^{\ddagger}$ amounts to 36.6 kcal/mol, as given by the energy difference between **SNR** and **SNTS2**. On the other hand, the activation barrier for the concerted nucleophilic mechanism, which is the energy difference between **CNR** and **CNTS**, is found to be 25.8 kcal/mol. On the basis of the previous experimental findings that the rate-determining step occurs after substrate binding and precedes the formation of any covalent intermediate,[4] our theoretical results then predict that the nucleophilic mechanism of Apases-catalyzed peptide hydrolysis would prefer the concerted pathway to the stepwise one.

As mentioned above, the free energy of activation $\Delta G^{\ddagger}$ for the concerted GA/GB mechanism, which is the free energy difference between **CGR** and **CGTS**, is 23.9 kcal/mol and lower than that of the nucleophilic mechanism by 1.9 kcal/mol. Here, the importance of the entropy factor in determining the favored catalytic mechanism can be appreciated by comparing the difference in activation energies with that in activation entropies for the two mechanisms. As can be seen in Table 1, the extent of increase in the entropic term ($-T\Delta S$) in going from **CGR** to **CGTS** is 2.1 kcal/mol smaller than that in going from **CNR** to **CGTS**. This relatively large difference in activation entropy leads to the preference of the concerted GA/GB mechanism by overbalancing the 0.5 kcal/mol difference in activation energy.

The small difference in $\Delta G^{\ddagger}$, however, indicates that the nucleophilic mechanism cannot be excluded conclusively in Apases-catalyzed hydrolysis of a peptide. The present theoretical result thus appeals to the previous experimental observations that the catalytic action of Apases may occur through the formation of the transient anhydride intermediate.[4] In this regard, it is also noteworthy that some Apase-inhibitor complexes have been found to lack the catalytic water molecule that can be a distinctive mark of the GA/GB mechanism.[25] Here, it should be remarked that the choice of catalytic pathway in the real enzymatic reaction may depend on the active-site environmental effects including the solvent effect that has not been considered in the present work.

Consistent with the previous experimental and theoretical studies, the hydrogen bond interactions between the two catalytic Asp residues and substrate, as can be seen in Figures 1−4, are found to play an essential role in stabilizing the transition states for all four reaction pathways under investigation. In this regard, we observe an interesting geometrical feature: the functional groups involved in the hydrogen bonds are pre-aligned in the incipient minimum energy structure so that the necessary hydrogen bonds can be readily established in the subsequent transition state. As can be seen in Figures 1−4, the hydrogen bonds $O_a\cdots H_a-N_s$ in **CNTS**, $O_{a2}\cdots H_a-N_s$ in **CGTS**, and $O_{a2}\cdots H_a-O_s$ in **SGTS1** are formed, respectively, from $O_a-H_a\cdots N_s$ in **CNR**, $O_{a2}-H_a\cdots N_s$ in **CGR**, and $O_{a2}-H_a\cdots O_s$ in **SGR** as the proton is transferred from the donor to the acceptor

(23) Komaromi, I.; Tronchet, J. M. J. *J. Phys. Chem.* **1997**, *101*, 3554.

(24) (a) Fraenkel, G.; Niemann, C. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 688. (b) Gillespie, R. J.; Birchall, T. *Can. J. Chem.* **1963**, *41*, 149.

(25) Gustchina, A.; Weber, I. T. *FEBS Lett.* **1990**, *269*, 269.

atoms with just slight changes in the interatomic distances. Similarly, the two O···H−O hydrogen bonds in **CGTS** and **SGTS2** are already established in **CGR** and **SGINT**, respectively, in a premature form (see Figures 3 and 4). On the basis of these observations together with the relatively small variation of entropy along the catalytic pathways, we argue that Apases optimize their catalytic ability by prepositioning the hydrogen-bonding groups in the enzyme−substrate complex or in an intermediate appearing in the catalytic mechanism and thus minimizing the entropic penalty for forming hydrogen bonds in the subsequent transition state. A similar mechanistic feature has been proposed also for the catalytic action of the serine proteases.[26]

Finally, we observe two different elements of the catalytic actions working in the mechanisms under consideration. The formation of multiple hydrogen bonds is likely to be a key feature in the GA/GB mechanism. As can be seen in Figure 3, four hydrogen bonds between the two Asp residues and the substrate including a catalytic water molecule are involved in **CGR** and **CGTS**, with those in the latter stronger than those in the former. Similarly, three hydrogen bonds are established in all stationary-state structures of the stepwise GA/GB mechanism. Two of them in **SGR** are relatively weak with associated O···H distances of 1.938 and 2.301 Å, while those formed in the two transition structures **SGTS1** and **SGTS2** and the tetrahedral intermediate **SGINT** have short O···H distances, all within 1.82 Å (see Figure 4). Judging from these geometrical features, we expect that, during the catalytic action of the APases through the GA/GB route, active site groups would be aligned in such a way that stabilizes specifically the transition states and the intermediate, rather than the reactant. This causes the lowering of the activation barrier and thus leads to the rate enhancement. The use of this kind of multiple interaction, such as hydrogen bonds, instead of relying on a single strong interaction, has been recognized often in various enzymatic reactions,[27] and the idea is being utilized in the design of enzyme inhibitors, synthetic receptors, and catalysts.[28]

The importance of stabilization of transition states by hydrogen bonds in the GA/GB mechanism may be more easily appreciated when the free energy profile is compared to that of an uncatalyzed reaction, namely, the stepwise reaction of formamide with a water molecule which lacks such hydrogen bond stabilizations. For this reaction the first and second free energies of activation were predicted to be 36.0 and 33.3 kcal/mol, respectively, at the MP2/6-31G**//RHF/3-21G** level of theory.[29] Therefore, in the uncatalyzed reaction both steps may affect the overall hydrolysis rate. This implies that Apases should be able to catalyze both steps: the first and second activation barriers of the stepwise GA/GB mechanism are lowered by 15.8 and 13.5 kcal/mol, respectively, relative to those of the uncatalyzed reaction.

On the other hand, the most significant factor in the nucleophilic mechanism is the initial protonation of the N_s atom, which induces the pyramidalization at N_s. The positive charge developed at C_s in the transition state **CNTS**, caused by the N_s-protonation and the rupture of the C_s-N_s bond, is then stabilized by a nucleophilic attack of the carboxylate group of an Asp residue. Thus, in contrast to the GA/GB mechanism, Apases are likely to catalyze the amide hydrolysis by direct participation in the reaction as a reactant in the nucleophilic mechanism.

Conclusion

We have investigated the mechanistic features and energetics of Apase-catalyzed peptide hydrolysis based on quantum chemical calculations on a model reaction system consisting of a formamide and acetate/acetic acid pair. A key result of the present study is that either the nucleophilic or GA/GB mechanism may be operative in an actual catalytic reaction, depending on the structural specificity of the peptide and the active site of the enzyme. For the model reaction system under consideration, the nucleophilic mechanism is energetically more favored if one considers only the potential energy profile along the IRC. However, when the entropic contribution due to nuclear motions transverse to the IRC (including the zero-point vibrations) is counted together, the GA/GB mechanism is predicted to be the preferred one, although the difference in the free energy of activation is just 1.9 kcal/mol. This shows that inclusion of the entropic factor is essential in the theoretical study of Apase catalysis.

In contrast to the previous theoretical studies which did not include the effects of electron correlation,[8,9] concerted and stepwise pathways are predicted to be almost equally favored in the GA/GB mechanism. On the other hand, the present study shows that a concerted reaction pathway would be preferred in the nucleophilic mechanism. For all reaction pathways that have been considered, the protonation of the leaving amine group is found to be essential for surmounting the activation barrier leading to peptidic C−N bond cleavage, and it is the primary step in the concerted nucleophilic mechanism. The hydrogen-bonding groups appearing in the transition state structures are pre-aligned in their respective incipient minimum energy structure, thereby minimizing the entropic penalty for the establishment of hydrogen bonds in the transition states. In the GA/GB mechanism, the formation of multiple hydrogen bonds seems to be a key element of the catalytic action whereas in the nucleophilic mechanism Apases are likely to catalyze the amide hydrolysis by a direct participation in the reaction as a reactant. Further study is needed to understand the effects of active-site environments and solvents on the energetics and geometrical features of the catalytic pathways.

Acknowledgment. This work was supported by a grant from the Center for Molecular Catalysis of the Korea Science and Engineering Foundation.

JA992849P

(26) (a) Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219. (b) Henderson, R. *J. Mol. Biol.* **1970**, *54*, 341. (c) Ruhlmann, A.; Kukla, D.; Schwager, P.; Bartels, K.; Huber, R. *J. Mol. Biol.* **1973**, *77*, 417. (d) Blow, D. M.; Janin, J.; Sweet, R. M. *Nature* **1974**, *249*, 54. (e) Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.*. **1981**, *78*, 4046.

(27) (a) Fierke, C. A.; Jencks, W. P. *J. Biol. Chem.* **1986**, *261*, 7603. (b) First, E. A.; Fersht, A. R. *Biochemistry* **1995**, *34*, 5030 and references therein.

(28) (a) Kato, Y.; Conn, M. M.; Rebek, J., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1208. (b) Jeong, K. S.; Rebek, J., Jr. *J. Am. Chem. Soc.* **1988**, *110*, 3327. (c) Huang, C.-Y.; Cabell, L. A.; Anslyn, E. V. *J. Am. Chem. Soc.* **1994**, *116*, 2778. (d) Zimmerman, S. C.; VanZyl, C. M.; Hamilton, G. S. *J. Am. Chem. Soc.* **1989**, *111*, 1373. (e) Perreault, D. M.; Chen, X.; Anslyn, E. V. *Tetrahedron* **1995**, *51*, 353.

(29) Oie, T.; Loew, G. H.; Burt, S. K.; Binkley, J. S.; MacElroy, R. D. *J. Am. Chem. Soc.* **1982**, *104*, 6169.