Studies on the Binding of Pepstatin and Its Derivatives to *Rhizopus* Pepsin by Quantum Mechanics, Molecular Mechanics, and Free Energy Perturbation Methods

B. G. Rao† and U. C. Singh*

Contribution from the Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California 92037. *Received January 14, 1991*

Abstract: The ab initio quantum mechanics, molecular mechanics, and free energy perturbation methods have been applied to study the energetics of the active site of *Rhizopus* pepsin and its interactions with several inhibitors derived from pepstatin. The studies on the Asp diad in the active site of the enzyme show that the energetics of the diad are very sensitive to small changes in the relative orientations of the diad and hence the energetic equivalence of the two charge states of the diad (arising due to the protonation of either of the two aspartates) can be easily attained by small changes in the atomic positions of the diad. Further, the studies point out that the proton possibly shuttles between the two inner oxygens of the diad. The barrier for the proton shuttle could be as low as 1.0 kcal/mol when the inner oxygen distance is around 2.5 Å and it increases with the increase in this distance. Although the present studies show that the configurations of the Asp diad distorted from planarity are lower in energy than the coplanar configuration found in the crystal structure, the latter configuration is found to be crucial for optimal inhibitor binding. This is also borne out in the calculated binding free energy differences between pepstatin and its derivatives. The calculated values obtained with the lower energy configuration of the Asp diad were found to be lower than those obtained with the Asp diad configuration found in the crystal structure, and the latter values were closer to the experimental results. For the mutation of the central statine residue of pepstatin to dehydroxystatine, the calculated free energy difference of 5.17 kcal/mol is in good agreement with the experimental value. This shows that the contribution of about 5 kcal/mol to binding from the hydroxyl group of the central statine residue is mainly due to the strong interaction of this group with the negatively charged Asp diad. The results of the other mutations on pepstatin are also in support of this view.

Introduction

The aspartic proteinases are an important class of proteolytic enzymes.[1,2] Several of these enzymes are therapeutic targets for drug design. In particular, human renin and HIV-1 proteinase are targeted for developing drugs against hypertension and AIDS, respectively. The cellular proteinases, *Rhizopus* pepsin and human renin, are monomers of 325 and 336 amino acids, respectively. The virus encoded HIV-1 proteinase is a dimer with each monomer having 99 amino acid residues. An impressive amount of structural and kinetic data on these enzymes is available in the literature. Structures of several of these enzymes and their inhibitor complexes have been determined at high resolution.[3] A large number of inhibitors have been designed which bind to these enzymes at very low concentrations, and their mode of inhibition has been studied.[4] In spite of the wealth of information on these enzymes, the following important aspects of the enzymes' action and inhibition remain ambiguous: (1) The details of the mechanism followed by the enzyme-catalyzed reaction is not clear, though a general acid–general base mechanism is presently favored.[5] (2) The nature of the acidity of the active site Asp diad is not completely understood.[6] (3) The nature of the enzyme inhibition by tight binding inhibitors like pepstatin is not clearly established.[7] The present study is concerned with the last two aspects, the details of which are discussed in the following paragraphs.

The aspartic proteinases have two catalytically important aspartic acid residues (Asp-35 and Asp-218 in *Rhizopus* pepsin) at the center of the active site cleft. The pH activity profiles of pepsin-catalyzed hydrolysis suggests pK_a values of 1.2 and 4.7 for the two Asp residues. The low pK_a is usually attributed[8] to Asp-35, and the high pK_a is assigned to Asp-218 based on the reactions of these groups with DAN (diazo-DL-norleucine methyl ester) and EPNP (1,2-epoxy-3-(*p*-nitrophenoxy)propane). The structural studies[3] show that the center of the enzyme's active site is fairly rigid, and the two carboxylate groups of the Asp diad at this center are held in almost a plane due to intricate hydrogen bonding from the neighboring residues. In *Rhizopus* pepsin,[9] the two carboxyl oxygens of Asp-35 make hydrogen bonds with N of Gly-37 and OG of Ser-38. Similarly, the two carboxyl oxygens of Asp-218 make hydrogen bonds with N of Gly-220 and OG1 of Thr-221. The hydrogen-bonding interactions between the two Asp residues in the diad and the neighboring residues are shown in Figure 1. A tightly bound water molecule found close to the Asp diad makes hydrogen bonds with all four carboxyl oxygens of the two Asp residues. Further, two additional water molecules are found to be making hydrogen bonds with the two outer oxygens of the Asp diad, respectively. Similar arrangements around the two Asp residues are also observed in the crystal structures of penicillinopepsin,[10] endothiapepsin,[11] and human renin.[12] In the structures of the two pepsins, the molecules found close to the active site Asp diad was thought to be either an NH_4^+ or an H_3O^+ ion. It is known from the crystal structures that the environments around the two catalytically important Asp residues are very much similar. Even though the lower pK_a is assigned to Asp-35 based on chemical studies, the structural data was initially interpreted[10] to suggest that the negative charge is associated with Asp-218. Such an assignment has no firm basis, and the preference of a lower pK_a for one Asp against the other cannot be inferred from any strong geometric reasons. In fact, Pearl and Blundell[11] pointed out that the environments of the two Asp residues in endothiapepsin are related by a local 2-fold axis. They suggest that the hydrogen occupies positions closer to the two Asp residues for equal times like in a dicarboxylic acid, and the abnormal pK_a values of the Asp diad are due to its behavior similar to a dicarboxylic acid.

† Present address: Vertex Pharmaceuticals Inc. Cambridge, MA 02139.

(1) Fruton, J. S. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1976**, *44*, 1.
(2) *Acid Proteases, Structure, Function and Biology*; J. Tang, Ed.; Plenum: New York, 1977.
(3) For a review, see: Davies, D. R. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 189.
(4) See: Rich, D. H. In *Proteinase Inhibitors*: Barrett, Salversen, Eds.; Elsevier Science Publishers BV: 1986; Chapter 5, p 179.
(5) (a) Hofmann, T.; Fink, A. L. *Biochemistry* **1984**, *23*, 5247. (b) Somayaji, V.; Keillor, J.; Brown, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 265.
(6) Goldblum, A. *Biochemistry* **1988**, *27*, 163, 157, 450.
(7) Rich, D. H. *J. Med. Chem.* **1985**, *28*, 263.
(8) Kiston, T. M.; Knowels, J. R. *FEBS Lett.* **1971**, *16*, 337.
(9) Suguna, K.; Bott, R. R.; Padlan, E. A.; Subramanian, E.; Sheriff, S.; Cohen, G. H.; Davies, D. R. *J. Mol. Biol.* **1987**, *196*, 877.
(10) James, M. N. G.; Sielecki, A. R. *Biochemistry* **1985**, *24*, 3701.
(11) Pearl, L. H.; Blundell, T. *FEBS Lett.* **1984**, *174*, 96.
(12) Sielecki, A. R.; Hayakawa, K.; Fujinaga, M.; Murphy, M. E. P.; Fraser, M.; Muir, A. K.; Carilli, C. T.; Lewicki, J. A.; Baxter, J. D.; James, M. N. G. *Science* **1989**, *243*, 1346.

0002-7863/91/1513-6735$02.50/0

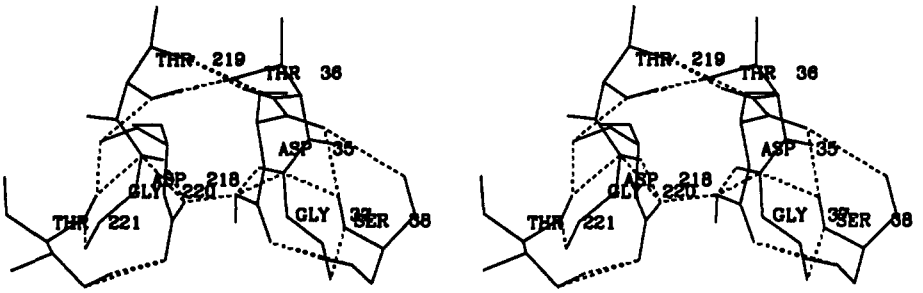


Figure 1. Structure of the active site Asp diad and the neighboring residues in the crystal structure of *Rhizopus* pepsin. The hydrogen bonds are shown in dotted lines.

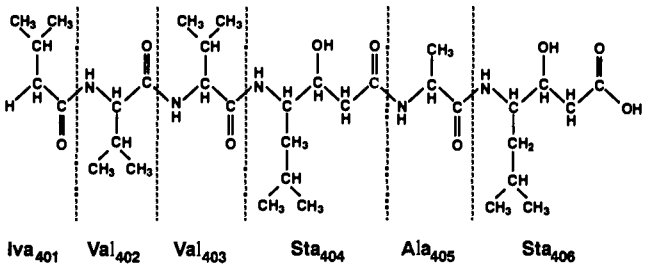


Figure 2. Structure of pepstatin. The residues of pepstatin are numbered from 401 to 406 in the figures and the text.

Experimental determination of actual location of hydrogen atoms, which is possible by neutron diffraction, is not yet reported for any of the aspartic proteinases. Theoretical approaches may be used to get important insight into the energetics of the active site, in addition to locating the energetically favorable hydrogen atom positions.

Pepstatin (Iva-Val-Val-Sta-Ala-Sta) containing a novel amino acid, statine (4(*S*)-amino-3(*S*)-hydroxy-6-methylheptanoic acid), is a tight-binding inhibitor of aspartic proteinases.[13] The structure of pepstatin is shown in Figure 2. It binds to the active site of procine pepsin with an unusually small dissociation constant of 45.7 pM.[14] Since pepstatin is a tight binding inhibitor of most aspartic proteinases and the 3(*S*)-hydroxyl group on the central statine is structurally related to a hydroxyl group in the tetrahedral intermediate formed during amide hydrolysis, it was considered a transition-state analogue inhibitor.[15] It has been established from kinetic studies that the hydroxyl group of the central statine residue is important for the tight binding of pepstatin.[16] The derivatives of pepstatin, containing a modified central statine residue lacking the hydroxyl group in *S* configuration, are weaker inhibitors of these enzymes. In fact, deletion of this hydroxyl group on the central statine residue leads to about a 4000-fold decrease in its binding to pepsin, and when the configuration of the hydroxyl group is changed from *S* to *R*, the resulting inhibitor is about 1000-fold weaker than pepstatin.[16,17] Two different roles had been attributed to the hydroxyl group of the central statine residue based on the kinetic and structural studies. The structural data on the complexes of aspartic proteinases with pepstatin[18] and other inhibitors[19,20] containing central statine residue show that the hydroxyl oxygen of the central statine is placed in the middle of the carboxylates of the two active site Asp residues, and the hydroxyl group makes multiple hydrogen bonds to the carboxylate oxygens. Therefore, it is possible that the binding of pepstatin is aided by these multiple hydrogen bond interactions in addition to its role as a transition-state analogue. The derivatives of pepstatin, lacking the hydroxyl group in *S* configuration on central statine are weaker inhibitors due to the absence of the hydrogen bond interactions with the active site Asp residues. However, it is not clear if the multiple hydrogen bond interactions alone are responsible for the difference in the binding free energy of about 5 kcal/mol between pepstatin and dehydroxypepstatin. An alternative role is also suggested for the hydroxyl group of the central statine residue by the experimental data. Since the position of the hydroxyl oxygen of the central statine in pepstatin coincides with that of the oxygen of the tightly bound water molecule in the uncomplexed enzyme, the tighter binding of pepstatin may result from the positive entropy change (of about 10–16 eu or 3–5 kcal/mol) associated with displacement of the active site bound water molecule into the bulk solvent, as suggested by Rich.[7] It has been his contention that pepstatin is a bisubstrate inhibitor, in addition to being a transition-state analogue inhibitor. Though both the roles suggested for the hydroxyl group of statine appear very reasonable, it is not clear which one of them has major contribution to the tight binding of pepstatin. In spite of the impressive amount of structural information and kinetic data, this question is not resolved satisfactorily.

A complete understanding of the nature of the active site Asp diad and inhibitor binding is very important for designing new inhibitors and elucidating the mechanism of enzyme action. In several strong inhibitors of human renin and HIV-1 proteinase, which are based on the substrates of these enzymes and contain hydroxyethylene isostere of the scissile peptide bond, the 4-(*S*)-hydroxyl group of the isostere interacting with the active site Asp diad is found to be crucial for tight binding.[21-25] It has also been shown in a recent study[26] that a single hydroxyl group of an inhibitor of adenosine deaminase contributes up to 9 kcal/mol to free energy of binding, though the nature of the interaction of the hydroxyl group with the active site of the enzyme is not yet known.

A theoretical study[6] on the acidity of the active site in aspartic proteinases has been reported recently with semiempirical cal-

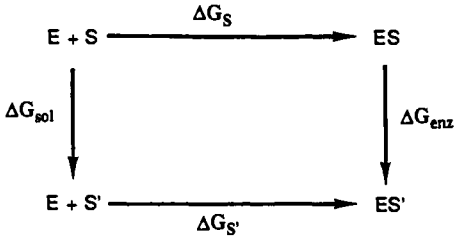
(13) Umezawa, H.; Aoyagi, T.; Moroshima, H.; Matzusaku, M.; Hamada, H.; Takeuchi, T. *J. Antibiot.* **1970**, *23*, 259.
(14) (a) Kunimoto, S. et al. *J. Antibiot.* **1974**, *27*, 413. (b) Workman, R. J.; Burkitt, D. W. *Arch. Biochem. Biophys.* **1979**, *194*, 157.
(15) Merciniszyn, J.; Hartsuck, J. A.; Tang, J. *J. Biol. Chem.* **1976**, *251*, 7088.
(16) Rich, D. H.; Sun, E.; Singh, J. *Biochem. Biophys. Res. Commun.* **1977**, *74*, 762.
(17) Rich, D. H.; Sun, E.; Ulm, E. *J. Med. Chem.* **1980**, *23*, 77.
(18) Bott, R.; Subramanian, E.; Davies, D. R. *Biochemistry* **1982**, *21*, 6956.
(19) James, M. N. G.; Sielecki, A.; Salituro, F.; Rich, D. H.; Hofmann, T. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6137.
(20) Cooper, J. B.; Foundling, S. I.; Blundell, T. L.; Boger, J.; Jupp, R. A.; Kay, J. *Biochemistry* **1989**, *28*, 8596.
(21) Dreyer, G. B.; Metcalf, B. W.; Tomaszek, T. A., Jr.; Carr, T. J.; Chandler, A. C. III; Hyland, L.; Fakhoury, S. A.; Magaard, V. W.; Moore, M. L.; Strickler, J. E.; Debouck, C.; Meek, T. D. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9752.
(22) Tomasselli, A. G.; Olson, M. K.; Hui, J. O.; Staples, D. J.; Swayer, T. K.; Heinrikson, R. L.; Tomich, C. C. *Biochemistry* **1990**, *29*, 264.
(23) Roberts, N. A.; Martin, J. A.; Kinchington, D.; Broadhurst, A. V.; Craig, J. C.; Duncan, I. B.; Galpin, S. A.; Handa, B. K.; Kay, J.; Krohn, A.; Lambert, R. W.; Merrett, J. H.; Mills, J. S.; Parkes, K. E. B.; Redshaw, S.; Ritchie, A. J.; Taylor, D. L.; Thomas, G. I.; Machin, P. J. *Science* **1990**, *248*, 358.
(24) Rich, D. H.; Green, J.; Toth, M. V.; Marshall, G. R.; Kent, S. B. H. *J. Med. Chem.* **1990**, *33*, 1288.
(25) Erickson, J.; Neidhart, D. J.; VanDrie, J.; Kempf, D. J.; Wang, X. C.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, N.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, M. *Science* **1990**, *249*, 527.
(26) Kati, W. M.; Wolfenden, R. *Science* **1989**, *243*, 1591.

culations on a model of the active site. In the present study, we use a combination of theoretical methods including quantum mechanics, molecular mechanics and free energy perturbation. First, we employed the ab initio quantum mechanics method to obtain the energetics of an acetic acid–acetate complex, the simplest model of the Asp diad. In a larger model, a water molecule or an NH_4^+ ion was included. Second, we determined the energetics of the catalytically important Asp residues inside the active site of the enzyme by treating the side chain atoms of the two Asp residues by the ab initio method and the rest of the enzyme by molecular mechanics method. Third, free energy perturbation calculations were employed to obtain the free energy of transfer of proton from oxygen atom of one Asp to that of the second Asp in the active site.

To understand the importance of the interactions between the hydroxyl group of pepstatin and the enzyme active site, we have mutated the central statine residue of pepstatin into several of its derivatives by the free energy perturbation method. The most important mutation simulated in this study is that of the central statine to dehydroxystatine, which determines the difference in the free energies of binding between the two inhibitors. It is assumed in these calculations that pepstatin analogue, containing a dehydroxystatine in the place of central statine, binds in the active site of the enzyme in the same way as pepstatin does. These calculations, therefore, do not include the entropic contribution to binding arising due to the displacement of the active site bound water molecule. If the later contribution is significant as postulated by Rich,[7] the calculated free energy difference should be smaller than the experimental value of about 5 kcal/mol. If the calculated value is closer to the experimental value, then we may infer that the major contribution to the free energy of binding is enthalpic and the entropic contribution caused by the displacement of water is not significant. The free energy perturbation method, employed in the present study, has been tested in a variety of applications involving calculation of free energy differences.[27] We have used the free energy perturbation method extensively for various studies, including hydrophobic hydration, nonaqueous solvation, and enzyme inhibition and mechanism.[28] Several other groups have also employed this method successfully to determine difference in the free energy of binding between two closely related inhibitors of an enzyme.[29] Therefore, we feel that a detailed study on pepstatin binding to aspartic proteinase will be helpful in resolving some of the questions discussed earlier and also provide new insights.

Methods

The present studies employed several computational methods including (1) ab initio quantum mechanics, (2) molecular mechanics, (3) molecular dynamics, and (4) free energy perturbation methods in addition to graphics modeling. Graphics modeling was carried out on Evans and Sutherland PS 390 system with MOGLI software.[30] For ab initio quantum mechanical calculations, the QUEST program[31] was employed. Molecular mechanics, molecular dynamics, and free energy perturbation calculations were carried out with BORN, NEWTON, and GIBBS modules of the AMBER program.[32] The details of these methods and their implementations into AMBER programs have been described in earlier publications.[33] Additionally, the SIVA module was used for the optimization calculations inside the enzyme active site. With the SIVA module, which is a combination of the ab initio program, QUEST, and the molecular mechanics program, BORN, it is possible to treat certain parts of the system by quantum mechanics and the other parts by molecular mechanics. Since the main thrust of the present calculations is the determination of free energy differences, a brief description of the free energy perturbation method is given below.



$$\Delta\Delta G_{bind} = \Delta G_{S'} - \Delta G_S = \Delta G_{enz} - \Delta G_{sol}$$

Figure 3. The thermodynamic cycles used in the computation of $\Delta\Delta G_{bind}$.

Free Energy Perturbation Method. In the free energy perturbation approach, the free energy difference between two states of a system is computed by transforming one state into the other by changing a single coupling parameter in several steps. For instance, the two states, A and B represented by Hamiltonians, $\mathcal{H}_A$ and $\mathcal{H}_B$ are coupled by a dimensionless parameter, λ, as

$$\mathcal{H}_\lambda = \lambda\mathcal{H}_A + (1 - \lambda)\mathcal{H}_B \qquad 0 \le \lambda \le 1 \tag{1}$$

When $\lambda = 1$, $\mathcal{H}_\lambda = \mathcal{H}_A$, and when $\lambda = 0$, $\mathcal{H}_\lambda = \mathcal{H}_B$. Therefore, the state A can be smoothly transformed to state B by changing the value of λ in small increments, $\Delta\lambda$, such that the system is in equilibrium at all values of λ. At intermediate values of λ, the state is a hypothetical mixture of A and B.

The Gibbs free energy change at constant pressure due to the perturbation of the Hamiltonian from $\mathcal{H}_\lambda$ to $\mathcal{H}_{\lambda'}$ is given by

$$\Delta G_\lambda = -\frac{1}{\beta} \ln \langle \exp(-\beta\Delta\mathcal{H}_\lambda)\rangle_0 \tag{2}$$

where $\beta = \frac{1}{RT}$. The average of exp $(-\beta\mathcal{H}_\lambda)$ is computed over the unperturbed ensemble of the system. If the perturbation from $\lambda = 1$ to $\lambda = 0$ is carried over N intervals, then the total free energy change equals the sum over all these intervals

$$\Delta G = \sum_{i=1}^{N} \Delta G_{\lambda_i} \tag{3}$$

It is a standard practice to make use of the thermodynamic cycle[34] described in Figure 3 in order to determine the difference in the free energy of binding, $\Delta\Delta G_{bind}$ between an inhibitor, S, and another inhibitor, S′, with an enzyme, E. In this figure, ΔG_S and $\Delta G_{S'}$ represent the free energies of binding of inhibitors, S and S′, respectively. Further, ΔG_{sol} represents the difference between the solvation free energies of the two inhibitors, and the difference in the free energies of binding of the two inhibitors in the enzyme active site are represented by ΔG_{enz}. Since free energy is a state function, it is possible to write the following equation from the consideration of the scheme given in Figure 3. It is straightforward and much easier to determine ΔG_{enz} and ΔG_{sol} terms compared to the ΔG_S and $\Delta G_{S'}$ terms. Therefore, we determined the ΔG_{enz} and ΔG_{sol} terms by mutating one inhibitor into another in the active site of the enzyme and in water, respectively, in two separate simulations to compute $\Delta\Delta G_{bind}$.

$$\Delta\Delta G_{bind} = \Delta G_{S'} - \Delta G_S = \Delta G_{enz} - \Delta G_{sol} \tag{4}$$

(27) For a review of recent applications of free energy perturbation methods, see: (a) Beveridge, D. L.; DiCupa, F. M. *Ann. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431. (b) van Gunsteren, W. F. *Protein Eng.* **1988**, *2*, 5. (c) Richards, W. G.; King, P. M.; Reynolds, C. A. *Protein Eng.* **1989**, *2*, 319. (d) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184. (e) Kollman, P. A.; Merz, K. M., Jr. *Acc. Chem. Res.* **1990**, *23*, 246. (f) Karplus, M.; Petsko, G. A. *Nature* **1990**, *347*, 631.

(28) (a) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1989**, *111*, 3125. (b) Rao, B. G.; Singh, U. C. *J. Am. Chem. SOc.* **1990**, *112*, 3811. (c) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1991**, *113*, 4381. (d) Ramnarayan, K.; Rao, B. G.; Singh, U. C. *J. Chem. Phys.* **1990**, *92*, 7057. (e) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574. (f) Rao, S.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551. (g) Singh, U. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4280. (h) Singh, U. C.; Benkovic, S. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9519.

(29) (a) Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830. (b) Hwang, J. K.; Warshel, A. *Biochemistry* **1987**, *26*, 2669. (c) Fleischman, S. H.; Brooks III, C. L. *Proteins* **1990**, *7*, 52.

(30) MOGLI Molecular Graphics, Evans & Sutherland Computer Corporation, Salt Lake City, UT, 1987.

(31) QUEST (Version 1.0); Singh, U. C.; Kollman, P. A. University of California: San Francisco, 1986.

(32) AMBER (Version 3.3) is a fully vectorized version of AMBER (Version 3.0); Singh, U. C.; Weiner, P. K.; Caldwell, J. W.; Kollman, P. A. University of California: San Francisco, 1986. AMBER (Version 3.3) also includes coordinate coupling and intra/interdecomposition.

(33) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607; also see references cited therein.

(34) Tembe, B.; McCammon, J. A. *Comput. Chem.* **1982**, *8*, 281.

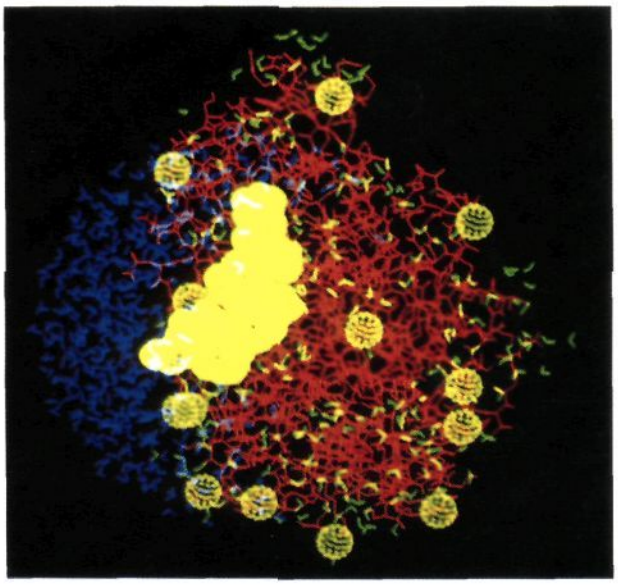


Figure 4. The structure of the model of the enzyme, rhizopus pepsin, obtained from the crystal structure and modified for the calculations as described in the text. The middle yellow region is the active site, the small yellow spheres are the Cl^- counter ions, the green molecules are the crystallographic water molecules, and the blue ones are the water molecules added to partially solvate the active site region.

Computational Details

Models of the Enzyme and Enzyme–Inhibitor Complexes. The studies on *Rhizopus* pepsin and its inhibitor binding were carried out with the structures of this enzyme and its inhibitor complex described in the 2apr[9] and the 3apr[35] files in the Brookhaven protein data bank (pdb), respectively. These two structures are very high-resolution structures of the enzyme and enzyme–inhibitor complexes and have several hundred water molecules in and around the enzyme. However, some simple modifications of these structures were needed for appropriate construction of the model systems for simulations. The 2apr structure has been modified in two different ways, and these new models are termed as model 1 and model 2. The modified version of the 3apr structure is termed as model 3. The details of each of these models are described below.

Model 1 is constructed from the 2apr pdb file. Since this enzyme is active at acid pH, the Asp and Glu residues were made neutral except for a few Asp residues which possibly form salt bridges with the positively charged Lys or Arg residues. In the active site, one Asp residue was made neutral and the other left negatively charged. All the Lys and Arg residues were positively charged. This enzyme does not have any histidine residue. To compensate for the net positive charge on the enzyme molecule, negatively charged counter ions were placed near the positively charged residues. Cl^- ions were used as negatively charged counter ions in this study. In total 13 such ions were required. It has been the practice in our laboratory[36] to balance the charges on the surface of the protein by adding counter ions since such a treatment of the system mimics the actual protein environment better. The importance of including counter ions and solvent molecules for computer simulations of enzyme systems has been emphasized in a recent study by Hagler and co-workers[37] also. Since one of the active site Asp residues is charged in the active pH range, the net charge on the protein is −1. The amino end and carboxy end of the enzyme were appended with an acetyl group and an amide group, respectively. Some of the water molecules found in the crystal structure are found to be far away from the enzyme surface. Water molecules farther than 5 Å from the surface of the enzyme were discarded, and the rest (about 230) of these water molecules were retained for our calculations. Further, the active site region of the enzyme was solvated by a cap of approximately 250 water molecules with a radius of 22 Å. This model is shown in Figure 4. This model has been used primarily for studying the nature of the active site Asp diad. The same number of the counter ions and the water molecules were used in the two models of enzyme–inhibitor complexes described below.

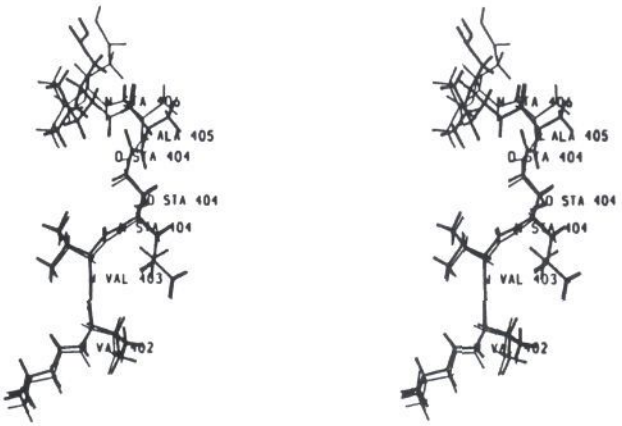


Figure 5. The superimposition of the two models of pepstatin. The molecule shown in thin lines is from the crystal structure, and the molecule shown in thick lines is from model 2.

Model 2 is the structure of the enzyme complexed with pepstatin. This model was obtained from visual docking of pepstatin into the active site of the enzyme (the 2apr structure). When this project was started in October 1987 in our laboratory, the 3apr pdb file was not available to us. We attempted to model pepstatin in an extended conformation into the enzyme active site by following the description given in the paper by Bott et al.[18] We found it difficult to follow the description given in the paper, because the primary structure used for the enzyme in the paper was not the same as that in the 2apr file. For instance, it was stated in this paper that the Ala-405 carbonyl oxygen of the inhibitor hydrogen bonds to the Tyr-196 OH. But no Tyr residue is found at position 196 in the 2apr file. Some of the descriptions given in the paper were found to be incorrect due to the lack of knowledge on the correct primary structure of the enzyme at the time of the first refinement of the structure. This was corrected later in the revised 2apr pdb file. We were not aware of this situation when this modeling was carried out. Because of the ambiguity in the structural description, the orientations of the side chains of the last two residues of pepstatin could not be modeled with confidence. We placed the terminal Sta-406 side chain approximately into the S3′ binding pocket of the enzyme. Later, when this model of pepstatin was compared with the model obtained in the crystal structure, the two structures were found very similar except for the difference in the orientation of the peptide bond between the middle Sta-404 and the following Ala-405 residues. The comparison of the two structures is shown in Figure 5. It may be noted from the figure that the conformations of the two models of pepstatin are very much similar, and the match between the first four residues of the models is also good. Though the side chains of the last two residues of the two models occupy almost the same regions, the match between them is not that good. This is mainly due to difference in the orientation of the peptide bond between Sta-404 and Ala-405. The incorrect orientation of the peptide bond in the model also led to the omission of hydrogen bond interactions of this peptide with Gly-37 and Gly-78 residues of the enzyme. This is depicted in the structure of the enzyme–pepstatin complex shown in Figure 6. As the 3apr coordinates were made available in the protein data bank, we made use of the new information to make the following new model with the correct conformation of pepstatin.

Model 3 uses 3apr structure of the enzyme–inhibitor complex. The inhibitor described in this pdb file is a reduced peptide inhibitor. Pepstatin is modeled by superimposing it on the reduced peptide inhibitor in the enzyme–inhibitor complex. The conformation of the newly modeled pepstatin and its interactions with the enzyme active site are very much similar to those in the crystal structure of *Rhizopus* pepsin–pepstatin complex.[38] This model of the enzyme–inhibitor complex, highlighting the inhibitor conformation and the binding pockets of the inhibitor side chains, is shown in Figure 7. A stereodiagram of the pepstatin model in the active site of the enzyme is shown in Figure 8. It may be seen in this model that the carbonyl oxygen of Sta-404 and the amide nitrogen of Ala-405 make hydrogen bonds to the amide nitrogen of Gly-78 and carbonyl oxygen of Gly-37, respectively. These two hydrogen bonds were absent in the earlier model due to incorrect orientation of the peptide bond between Sta-404 and Ala-405. As in models 1 and 2, this model has about 250 water molecules from the crystal structure and a cap of about 250 water molecules centered around the oxygen atom of

(35) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7009.
(36) See refs 28e and 28f.
(37) Avbelj, F.; Moult, J.; Kitson, D. H.; James, M. N. G.; Hagler, A. T. *Biochemistry* **1990**, *29*, 8658.

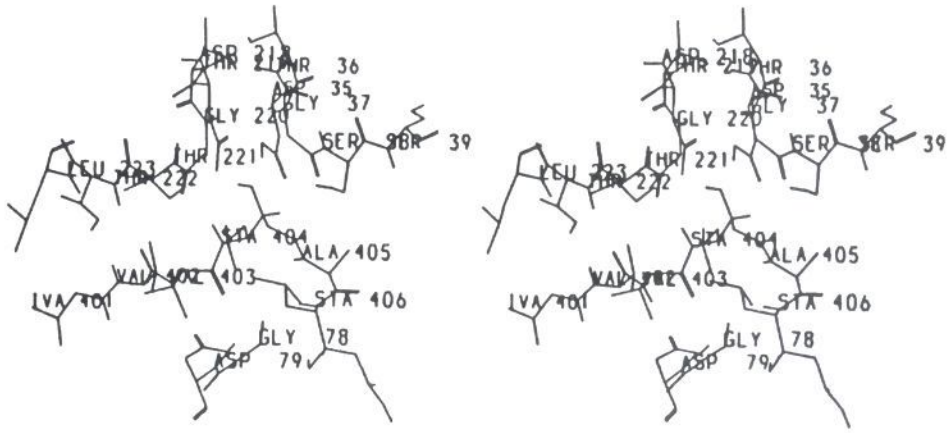


Figure 6. Structure of the enzyme–pepstatin complex around pepstatin in model 2.

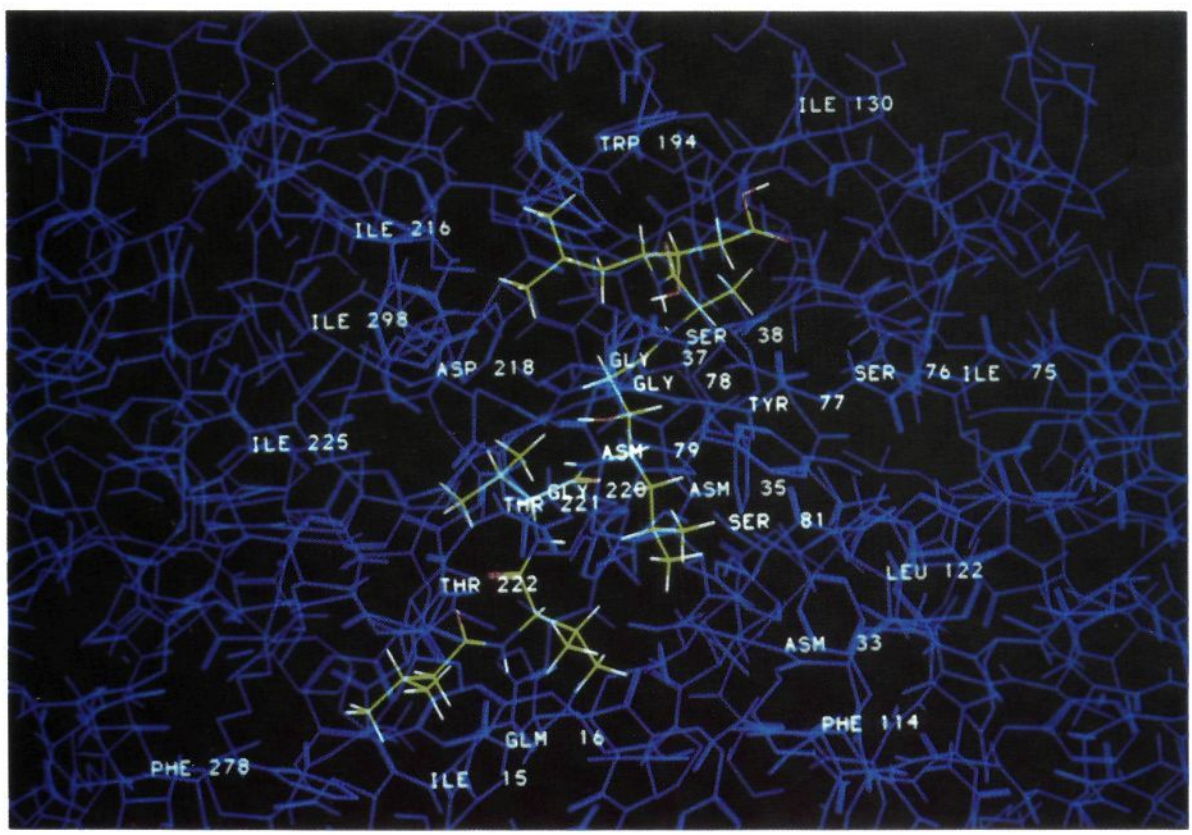


Figure 7. The structure of pepstatin in the active site of the enzyme in model 3. The pepstatin molecule is color coded, while the enzyme is shown in blue. The enzyme residues forming the binding pockets for the side chains of the inhibitor residues are labeled in white.

the central statine residue of the inhibitor. Charged residues were treated as in the earlier models.

For molecular mechanics, molecular dynamics, and free energy perturbation calculations on these models, standard AMBER force field parameters[39] were used. A united atom model representation was used for the residues of the enzyme. For the inhibitor residues, all-atom model was used. In most calculations, all-atom model was used for the two active site Asp residues. The partial charges for the novel amino acids were determined from electrostatic potential[40] generated by the ab initio method with 6-31G* basis set. TIP3P parameters[41] were used for all water molecules.

(38) Suguna, K.; Padlan, E. A.; Bott, R.; Boger, J.; Davies, D. R., results to be published.
(39) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *196*, 765.
(40) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.

Gas-Phase Calculations. Geometries of the model complexes were optimized by single-point calculations with the 6-31G basis set. The model of acetic acid–acetate complex is shown schematically in Figure 9. Four different sets of calculations were done on the model systems. In the first set, a perfectly planar configuration was used as the starting geometry for ab initio calculations. In the second set, the geometry of the side chains of the active site Asp diad in the *Rhizopus* pepsin crystal structure was used as the model. In the third set, a water molecule was included in the model complex. In the fourth set, an NH_4^+ ion was included in place of the water molecule. In each of these calculations, the HOD position was moved from OD2 to OD3 in several intervals, and the geometries were optimized at each position of HOD.

Optimization of the Enzyme Active Site. For geometry optimization of the two states of the active site of the enzyme, the models 1 and 3 structures were minimized with BORN in two stages. First, the positions of the counter ions, the inhibitor and the solvent molecules in the enzyme system, were minimized for 2000 cycles with conjugate gradient method. Second, the whole system was minimized with conjugate gradient method for another 2000 cycles. A cut-off distance of 12 Å was used for non-

Table I. Ab Initio Charges[a] for the Side-Chain Atoms of the Asp Diad at the Starting and Intermediate Geometries

atom	starting	intermediate	atom	starting	intermediate
CB	−0.529 51	−0.531 02	CB	−0.424 02	−0.431 09
HB1	0.169 31	0.161 48	HB1	0.092 76	0.100 09
HB2	0.169 31	0.161 48	HB2	0.092 76	0.100 09
CG	0.940 24	0.952 66	CG	0.960 41	0.965 81
OD1	−0.766 38	−0.806 69	OD3	−0.800 39	−0.785 79
OD2	−0.690 59	−0.709 91	OD4	−0.870 20	−0.872 97
HOD	0.539 60	0.579 16			

[a] All values are in eu.

bonded interactions. The side-chain atoms of the two active site Asp residues were treated by quantum mechanics, and the rest of the system was treated by molecular mechanics. The interaction between the quantum mechanical and the molecular mechanical parts of the system was treated by molecular mechanics. In these calculations, the geometries of the active site were optimized for only two proton positions on each of the aspartates. The geometries were not optimized for any intermediate hydrogen positions. In the first model (model 1), an NH_4^+ ion was included in the active site, and, in the second model (model 3), pepstatin was included in the active site. During the initial minimization with BORN, the distance between the two inner oxygens of the Asp diad got shortened from 2.97 to 2.57 Å. Therefore, in a different set of calculations with model 3, the initial minimization was done on the system excluding the Asp-35 and Asp-218 residues.

Mutation of the Proton Position in the Active Site. For calculating the free energy difference between the two states of the enzyme with protonation of Asp-35 or Asp-218, the hydrogen position from one Asp to the other Asp is perturbed through an intermediate state by the free energy perturbation method. The intermediate state was obtained from the gas-phase calculations on acetic acid–acetate complex described earlier. The partial charges for all atom models of the two Asp residues at different proton positions were obtained from ab initio calculations on the optimized structures of acetic acid–acetate complex. The electrostatic potential method implemented in the QUEST program was used for calculation of charges at the level of 6-31G* basis set. These charges for the side chain atoms of the two Asp residues at the initial state and an intermediate state are given in Table I. The charges for the diad in the final state are the same as that of the initial state since the Asp diad in the two states is almost equivalent.

For free energy perturbation calculations, the system was minimized additionally for 100 cycles with SHAKE[42] after the two stages of minimization described earlier. Then the system was equilibrated for 10.0 ps at constant temperature (300 K) and pressure (1 atm) with a time step of 0.002 ps. During equilibration and subsequent perturbation runs, the residues within 22 Å from the center of the active site (taken to be the position of the hydroxyl oxygen of the central statine residue of the inhibitor) were treated as the active part of the system, and the rest was frozen. The active part of the system interacts with the frozen part in the normal manner. A constant dielectric of 1 was used for all simulations. A cut-off distance of 12 Å was used for solute–solvent and solvent–solvent nonbonded interactions. All solute–solute nonbonded interactions were included. The water molecules in the cap were restrained by a force constant of 0.7 kcal/Å. The distance between the two inner oxygens of the Asp diad was restrained to its value in the minimized structure by a force constant of 200 kcal/Å to facilitate the proton transfer. The free energy perturbation calculations were carried out with the window method. For mutation of proton position from one Asp to the intermediate position between the two Asp residues, 11 or 25 windows were used with 1.0 ps of equilibration followed by 1.0 ps of data collection at each window. In the next run, the proton was transferred from the intermediate position to the inner oxygen of the second Asp residue in 11 or 25 windows with 1.0 ps of equilibration and 1.0 ps of data collection. This process was repeated for the reverse direction as well. In a second set of simulations, an ammonium ion was placed near the Asp diad in the active site, and the proton position was mutated in forward and backward directions as described above. The same process was repeated for a third set of simulations, wherein a pepstatin molecule was included in the active site instead of an ammonium ion.

Mutations on Pepstatin Inside the Enzyme. The inhibitor, water, and the counter ion positions were initially minimized for 4000 cycles, and then the enzyme was included for minimization for another 4000 cycles, followed by 100 cycles of minimization with SHAKE. During equilibration and subsequent perturbation runs, the residues within 22 Å from the center of the active site were treated as the active part of the system, and the rest were frozen. The active part of the system interacts with the frozen part in the normal manner. A 12 Å nonbonded cut-off was used except for the inhibitor, for which an infinite nonbonded cut-off was used. A constant dielectric of 1 was used for all simulations. The system was initially equilibrated for 10.0 ps at constant temperature (300 K) with a time step of 0.001 ot 0.002 ps. The free energy perturbation calculations were carried out with the window method. In most simulations, the mutations were achieved either in 101 or 51 windows with 0.4 ps of equilibration and 0.4 ps of data collection at each window.

Mutations on Pepstatin in Water. The calculation for each mutation was repeated in water to determine the difference in the free energies of solvation and to complete the thermodynamic cycle described in Figure 3. For mutations in water, the inhibitor was placed at the center of a rectangular water box. Each system was minimized in four stages. First, the solvent around the solute was minimized for 500 cycles with the steepest descent method. Second, it was minimized for the next 2000 cycles with the conjugate gradient method. Third, the whole system was minimized with the conjugate gradient method for another 1000 cycles followed by a minimization of 100 cycles with SHAKE. Then the system was initially equilibrated for 10 ps at constant temperature (300 K) and pressure (1 atm) with a time step of 0.002 or 0.001 ps. During minimization, equilibration, and the subsequent perturbation runs, periodic boundary conditions were applied. A constant dielectric of 1 was used for all simulations. For nonbonded interactions, a cut-off distance of 8 or 12 Å was employed. For each mutation, 51 or 101 windows were employed with 0.4 ps equilibration followed by 0.4 ps of data collection at each window. The calculations repeated with a 12-Å nonbonded cut-off resulted in the values which differed by less than 1% of the earlier results. Some mutations were simulated by keeping the inhibitor at the center of a large "water drop" instead of a rectangular water box. Several other input control parameters of the GIBBS program were varied to find out their effect on the calculated results. The details of these simulations are given in the Results section.

Table II. Ab Initio Distances and Energies for the Planar Configuration of Acetic Acid–Acetate Complex[a]

OD2–OD3	OD2–HOD	OD3–HOD	E (au)	ΔE (kcal/mol)
2.45	1.03	1.43	−454.864 80	0.000
2.39	1.07	1.33	−454.864 38	0.261
2.35	1.14	1.23	−454.863 65	0.721
2.35	1.19	1.18	−454.863 54	0.789
2.38	1.28	1.12	−454.863 95	0.532
2.45	1.43	1.03	−454.864 80	0.000

[a] The values in the first three columns are the interatomic distances in Å. The ΔE values in the last column are calculated by taking the first ab initio energy as the reference energy.

Table III. Ab Initio Distances and Energies for the Acetic Acid–Acetate Complex in the Crystal Structure Geometry[a]

OD2–OD3	OD2–HOD	OD3–HOD	E (au)	ΔE (kcal/mol)
2.95	0.99	1.98	−454.854 75	0.00
2.88	1.08	1.81	−454.851 99	1.73
2.86	1.26	1.61	−454.840 63	8.86
2.89	1.45	1.45	−454.831 04	14.87
2.92	1.63	1.29	−454.835 31	12.20
2.95	1.96	0.99	−454.859 25	−2.82

[a] The values in the first three columns are the interatomic distances in Å. The ΔE values in the last column are calculated by taking the first ab initio energy as the reference energy.

Results

Energetics of the Active Site. In the first set of calculations, a planar configuration of the acetic acid–acetate complex with interoxygen distances closer to the distances between the carboxylate oxygens of Asp-35 and Asp-218 residues of *Rhizopus* pepsin was considered. The distance between the two inner oxygen atoms, OD2 and OD3, which share the proton, HOD between them, is shortened from 2.9 to 2.47 Å after the initial optimization. The energies of the complex for different proton positions are given in Table II. The energies of the initial and the final states are equal (−454.864 80 au), and the transition state (at −454.863 54 au) is about 0.8 kcal/mol higher than the initial state. The geometries of the complex in the initial and final states were found

(42) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

Table IV. Ab Initio Distances and Energies for the Acetic Acid–Acetate–Water Complex[a]

OD2–OD3	OD2–HOD	OD3–HOD	E (au)	ΔE (kcal/mol)
2.97	0.98	2.57	−530.864 11	0.00
2.89	1.06	1.83	−530.868 34	2.65
2.88	1.43	1.45	−530.838 87	15.84
2.95	1.96	0.98	−530.863 43	0.43

[a] The values in the first three columns are the interatomic distances in Å. The ΔE values in the last column are calculated by taking the first ab initio energy as the reference energy.

Table V. Ab Initio Distances and Energies for the Acetic Acid–Acetate–Ammonium Complex[a]

OD3–HOD	E (au)	ΔE (kcal/mol)
1.50	−511.606 07	0.0
1.41	−511.606 37	−0.18
1.32	−511.604 33	1.09
1.23	−511.601 16	3.08
1.01	−511.597 29	5.51

[a] The values in the first column are the interatomic distances in Å. The ΔE values in the last column are calculated by taking the first ab initio energy as the H reference energy.

to be equivalent, and the two carboxylates were found to be perfectly coplanar in all optimized geometries with different proton positions.

In the second set of calculations, we used the starting geometry from the crystal structure and constrained the distance between the two inner oxygens at 2.97 Å (closer to the crystallographic distance) during all optimizations. The results of these optimizations at different proton positions are given in Table III. The energy increases from −454.854 75 to −454.831 04 au and then decreases to −454.859 25 au. In the crystal structure, the two carboxylates are not perfectly coplanar, and hence their interaction with each other is not symmetric. When the geometry of the complex was optimized for different proton positions by moving the proton from OD2 to OD3 in small steps, the resulting geometries at each successive step were found to have increasing distortion from planarity. The two carboxylates in the transition state and the final state were almost perpendicular to each other, whereas they were almost coplanar in the initial state. The energy of the transition state is about 15 kcal/mol higher than the initial state. This barrier is substantially higher than the barrier found in the earlier case when the distance between the two oxygens is much smaller (2.45 Å).

In the crystal structures of aspartic proteinases, a tightly bound water molecule is found near the Asp diad, which could also be either a hydronium ion or an ammonium ion. The hydrogen bonding between this molecule and the carboxylate oxygens may possibly affect the relative energies of the complex at different proton positions. To examine this possibility, we included a water molecule near the acetic acid–acetate complex in the third set of calculations. The results of different optimizations are summarized in Table IV. The energy of the acetic acid–acetate–water complex increases from −530.864 11 to −530.838 87 au and then decreases to −530.863 43 au. The transition state is again about 15 kcal/mol higher than the initial state. Though the interaction of the complex with the water molecule decreases the energies of the complex at different proton positions, the relative energies of the different configurations of the complex are not affected to any significant extent. Our attempts to optimize the acetic acid–acetate complex with an ammonium ions in the place of the water molecule resulted invariably in transfer of a proton form the ammonium ion to acetate ion when the OD2–OD3 distance was constrained at 2.97 Å. This problem did not arise when the OD2–OD3 distance was not constrained, but the OD2–OD3 distance then got shortened from 2.97 to 2.45 Å. The results of these optimizations are given in Table V. In this configuration, the energy of the system increases almost linearly as the position of the proton is moved from the initial state (−511.606 07 au) to the final state (−511.597 29 au).

The results of optimization inside the active site by treating the side-chain atoms of the two Asp residues by quantum mechanics and the rest of the enzyme and the surrounding water molecules (including the substrate) by molecular mechanics are summarized in Table VI. In the first set of calculations, an ammonium ion was included in the active site, and, in the next two sets of calculations, pepstatin was included in the active site. For the first two sets of calculations, one with an ammonium ion and another with pepstatin in the active site, the system was minimized initially by conjugate gradient method with BORN module, and then the SIVA module was used for the combined quantum mechanics and molecular mechanics optimization. The initial minimization caused the two inner oxygens of the Asp diad to approach to less than 2.5 Å, whereas the crystallographic distance is 2.97 Å. To avoid this, the two Asp residues were excluded during the initial minimization for the third set of calculations on the enzyme–pepstatin complex. For the first set of calculations, with ammonium ion at the active site, there is a preference for the charged Asp-35 over the charged Asp-218 by 3.42 kcal/mol. The major contribution to the difference in the energies of the two configurations comes from the surroundings. In fact, the ab initio energies show that the charged Asp-218 configuration is preferred over the charged Asp-35 by about 1.9 kcal/mol. The second set of calculations, which includes pepstatin in the active site, shows that the charged Asp-218 is preferred over the charged Asp-35 configuration by about 5.8 kcal/mol. Again the major contribution to the difference in energies of the two configurations comes from the surroundings. In the third set of calculations, in which the distance between the two inner oxygens of the active site Asp carboxylates was closer to the crystallographic distance, the preference for the charged Asp-218 configuration was observed. However, the difference in energy between the two configurations is only 1.3 kcal/mol, which is much smaller than the value obtained in the second calculation. Moreover, the ab initio energies for the two configurations are higher than those found in the second set of calculations. Higher ab initio energies are expected since the distance between the two oxygens is larger in this case. However, the energies due to the surrounding are also higher when the distance between the two oxygens is larger.

Table VI. Ab Initio Distances and Energies from SIVA Optimization[a]

model	O(D35)–O(D218)	O(D218)–HOD	O(D35)–HOD	E_{qm}	E_{mm}	E_{total}	ΔE (kcal/mol)
NH_4^+							
D35	2.35	1.30	1.05	−454.858 23	−0.367 19	−455.225 42	
D218	2.39	1.00	1.38	−454.855 15	−0.375 72	−455.230 87	−3.42
pepstatin							
D35	2.41	1.41	1.03	−454.860 04	−0.277 46	−455.137 51	
D218	2.47	1.00	1.52	−454.857 78	−0.270 51	−455.128 29	5.79
pepstatin							
D35	2.85	2.41	0.99	−454.836 30	−0.253 23	−455.089 53	
D218	2.96	0.98	2.43	−454.836 98	−0.250 48	−455.087 46	1.30

[a] D35 or D218 under model in the first column represents the neutral residue (Asp-35 or Asp-218) in the model. The active site ligand (NH_4^+ or pepstatin) used in the model is also included. The interatomic distances are in Å, and the E values are in atomic units. E_{qm} is the ab initio energies of the side chains of the Asp diad. E_{mm} is the molecular mechanics energies from the enzyme surroundings. ΔE values are calculated by taking the first ab initio energy as the reference energy.

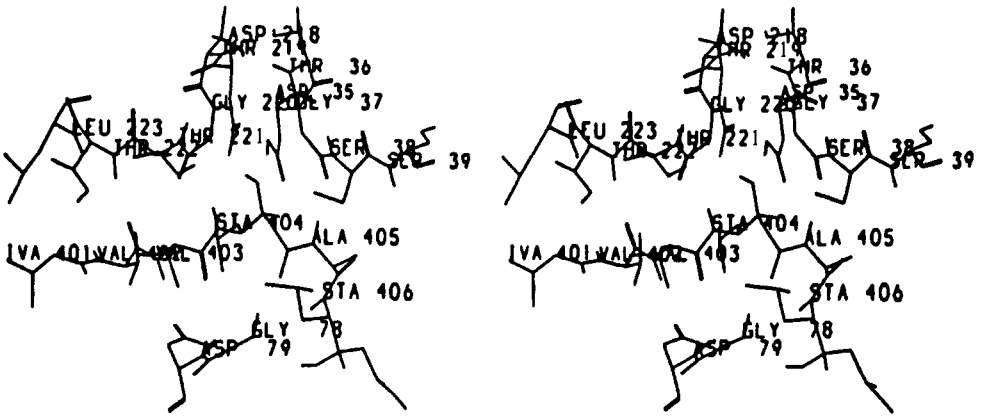


Figure 8. Structure of the enzyme–pepstatin complex around pepstatin in model 3.

Table VII. Free Energy Changes for the Transfer of the Proton[a]

simulation	ligand	ΔG_1	ΔG_2	ΔG_{tot}
D218 → D35	NH_4^+	1.16	4.63	5.79
D218 → D35	NH_4^+	0.51	7.06	7.57
D218 → D35	NH_4^+	1.21	6.68	7.89
D218 → D35	pepstatin	1.26	−0.44	0.82
D35 → D218	pepstatin	0.82	−0.69	0.13

[a] All ΔG values are in kcal/mol and are the averages of the forward and backward values with a maximum hysteresis of 0.2 kcal/mol.

Free Energy Perturbation on Proton Position. The results of free energy perturbation calculations for the transfer of proton from one inner oxygen to the other oxygen of the Asp diad in the active site are given in Table VII. The calculations with an ammonium ion in the active site were repeated several times with different time steps of mutation. The first set of calculations was done for 22 ps for the two steps of the transfer of the proton: (1) from the initial state to the intermediate state and (2) from the intermediate state to the final state. The second set was repeated with different equilibrated configuration, and the third set was repeated with 50 ps for each of the two steps. The transfer of the proton from Asp-218 to the intermediate state gives rise to a change in free energy (ΔG_1) of 1.16, 0.51, and 1.21 kcal/mol, respectively, for the three runs, and the free energy change (ΔG_2) for the transfer of the proton from the intermediate state to Asp-35 is 4.63, 7.06, and 6.68 kcal/mol. The corresponding ΔG_{tot} values are 5.79, 7.57, and 7.89 kcal/mol. When pepstatin is present in the active site instead of an ammonium ion, the mutation of proton position from Asp-218 to the intermediate state gives rise to a free energy change (ΔG_1) of 1.26 kcal/mol, and in the second step of mutation from the intermediate state to Asp-35 the free energy change (ΔG_2) is −0.44 kcal/mol. These calculations were repeated for the proton transfer in the reverse direction from Asp-35 to Asp-218, and the ΔG_1 and ΔG_2 values are 0.82 and −0.69 kcal/mol, respectively. The ΔG_{tot} values are 0.82 and 0.13 kcal/mol, respectively, for the forward and reverse transfers of the proton.

Enzyme–Pepstatin Complex on Minimization and Equilibration. The structure of the active site of the enzyme after minimization is shown in Figure 10. The minimized structure is very much similar to the starting structure shown in Figure 1 except for small changes in the atomic positions. The most significant change is the shortening of the distance between the two inner oxygens of the Asp diad from 2.97 to about 2.6 Å. It may also be noticed that the coplanarity of the two carboxylates is distributed to a small extent, and the amide hydrogen of Gly-220, which is involved in hydrogen-bonding interaction with Asp-35 in the starting structure, now points toward the inner oxygen of the charged Asp-218. The interaction of the inhibitor with the enzyme active site is, however, strengthened in the minimized structure. This may be seen from Figure 11, showing the minimized structure of the enzyme–pepstatin complex around pepstatin and the crucial contact distances between pepstatin and the enzyme active site listed in Table VIII.

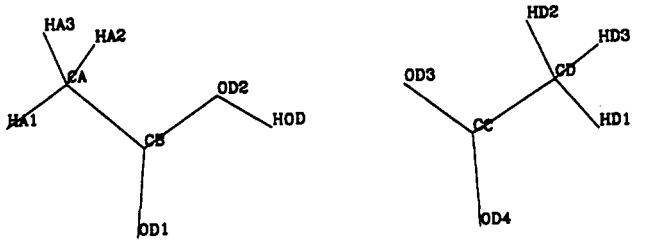


Figure 9. Schematic of the acetic acid–acetate ion complex used for the ab initio optimization studies. The geometry shown in this figure is that of the side chains of the active site Asp diad of *Rhizopus* pepsin found in the crystal structure.

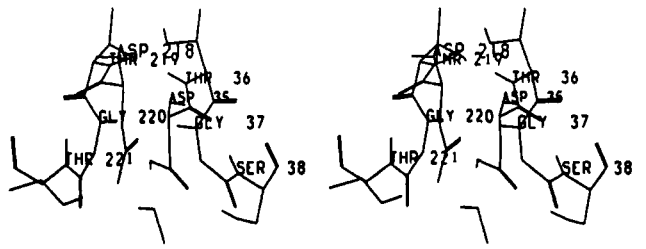


Figure 10. Structure of the active site Asp diad and the neighboring residues obtained after minimization.

On equilibrating the minimized structure by molecular dynamics simulation, the planarity of the two carboxylates of the active site Asp diad is further disturbed. The structure of the Asp diad and the neighboring residues found in the enzyme–pepstatin complex at the end of 10 ps of equilibration is shown in Figure 12. The two carboxylates are almost perpendicular to each other in this structure. The protonated carboxyl oxygen of the neutral Asp-35 residue points into the middle of the carboxylate group of Asp-218. As a result of the change in the orientation of the two Asp residues, there are significant changes in the hydrogen-bonding interactions between the two carboxylates of the Asp diad and the neighboring residues. In particular, the amide nitrogen of Gly-37, which hydrogen bonds to the inner oxygen of Asp-35 in the crystal structure, is closer to the outer oxygen in the equilibrated structure. The outer oxygen of Asp-35 gets hydrogen bonded to amide nitrogen of Ser-38 in addition to its side-chain hydroxyl group. The changes around the Asp diad affect the interactions of the inhibitor with the enzyme also. The equilibrated structure of the enzyme–pepstatin complex around pepstatin is shown in Figure 13. The contact distances between the atoms of the inhibitor and the enzyme active site making crucial interactions are given in Table VIII. It may be noted that the central hydroxyl group of pepstatin forms only a single hydrogen bond with the outer oxygen of Asp-218. In the crystal structure, it appears to make hydrogen bond interactions with all four oxygens of the diad. The weak hydrogen bond between N of Sta-404 (of the inhibitor) and O of Gly-220 is weakened on minimization and equilibration. On the other hand, N of Sta-404

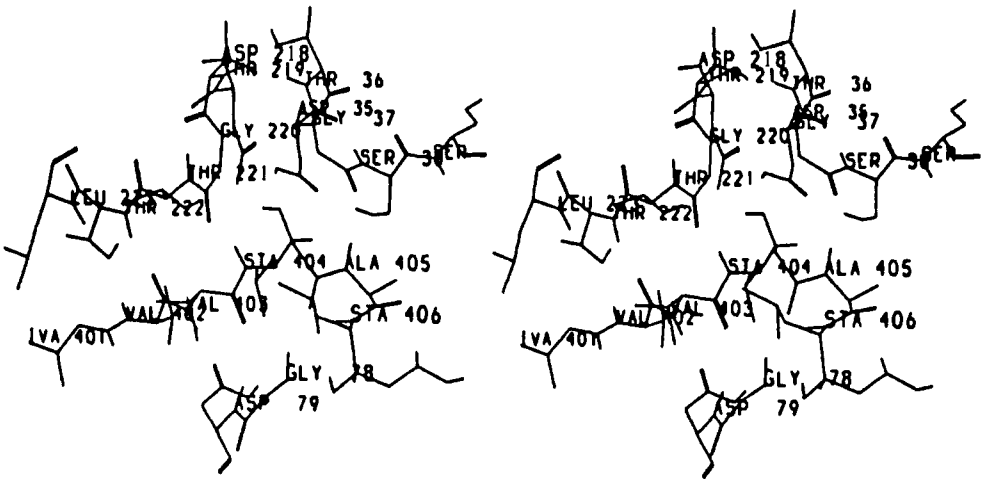


Figure 11. Structure of the enzyme–pepstatin complex around pepstatin after minimization.

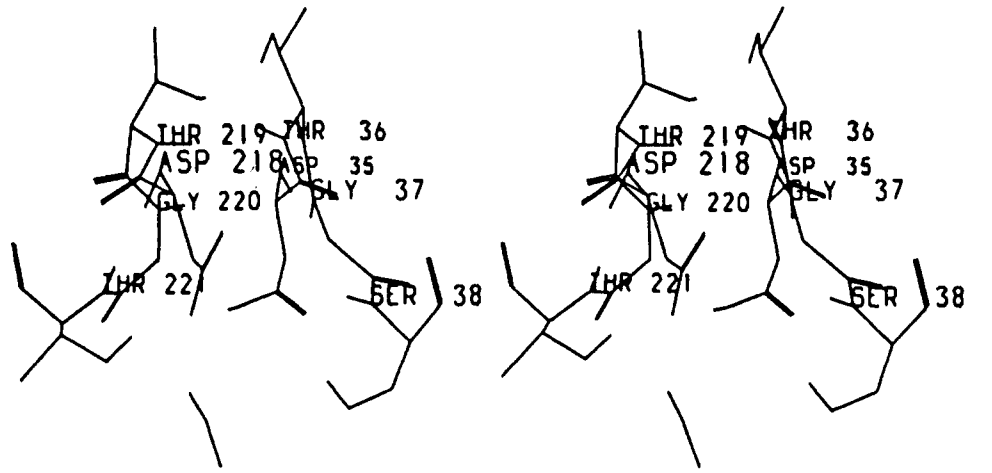


Figure 12. Structure of the active site Asp diad and the neighboring residues obtained after equilibration.

interacts weakly with the hydroxyl of Thr-221 and the outer oxygen of Asp-218. Similarly, the interactions of O of Sta-404 and N of Ala-405 of the inhibitor with N of Gly-78 and O of Gly-37 of the enzyme, respectively, were also weakened in the minimized and the equilibrated structures.

The crystallographic studies point out that the configuration of the active site diad, which is almost planar, is maintained in all structures, and it is not disturbed even on inhibitor binding. The carboxylates of these two Asp residues appear to be as rigid as the backbone atoms in the protein chain. However, the results obtained in the present studies with ab initio optimizations, molecular mechanics minimizations, and molecular dynamics show that the planar configuration is not the most stable configuration from the considerations of the energetics. This was also the apparent reason for not optimizing the crystallographic positions of the active site by Goldblum[6] in his study on the acidity of the active site. It must be emphasized here that one of the Asp residues is negatively charged and the other is neutral. Such a situation breaks down the 2-fold C_2 symmetry of the active site described in the crystal structures. However, the planar configuration of the two carboxylates of the Asp diad appears to be crucial for the maintenance of the integrity of the active site and its interaction with the inhibitor. For the symmetry of the active site to be maintained even when the energetics are taken into consideration, the proton, HOD, should be shared equally between the two Asp residues. This may be realized if the proton shuttles between the two inner oxygens of the diad with high frequency. The results described in an earlier section suggested that the proton may shuttle between the two inner oxygens with a small barrier of about 1 kcal/mol. Such a situation, however, cannot be treated in the molecular mechanics and molecular dynamics calculations with the existing force fields.

Table VIII. Hydrogen Bond Contacts between Pepstatin and the Enzyme[a]

pepstatin	enzyme	model 3	min	equi	min_{wc}	$equi_{wc}$
O(Iva-401)	OG1(Thr-222)		4.66	4.90	4.43	
N(Val-402)	OG1(Thr-222)	2.90	2.88		3.17	4.10
O(Val-402)	N(Thr-222)	2.96	2.88	3.71	3.02	3.67
O(Val-402)	OG1(Thr-222)		3.37	4.82	3.28	4.07
N(Val-403)	OD1(Asp-79)	2.91	2.99		3.73	
N(Val-403)	OG1(Thr-221)		4.60	4.51	4.41	4.43
O(Val-403)	N(Gly-78)	3.17	3.23	3.73	3.88	2.99
O(Val-403)	N(Asp-79)	3.25	3.49	3.25	3.46	
O(Val-403)	OD2(Asp-79)		3.44	3.82	3.23	4.31
N(Sta-404)	OD2(Asp-218)		4.82	4.34	3.71	4.14
N(Sta-404)	O(Gly-220)	2.90	3.54		3.74	3.46
N(Sta-404)	OG1(Thr-221)	3.42	4.05	3.05	3.73	3.46
OH(Sta-404)	OD1(Asp-35)	3.13	2.87	4.28	3.18	3.27
	OD2(Asp-35)	2.68	2.66	3.29	2.71	2.87
	OD1(Asp-218)	2.67	2.79	2.82	2.66	2.81
	OD2(Asp-218)	2.94	2.61	4.32	2.87	2.93
O(Sta-404)	N(Gly-78)	2.68	2.67	2.89	2.81	2.83
N(Ala-405)	O(Gly-37)	2.76	2.90		2.85	3.87
O(Ala-405)	NE1(Trp-194)	2.80			2.94	4.68
N(Sta-406)	O(Ser-76)	4.46	4.06	3.57	4.65	4.09
OA(Sta-406)	N(Gly-78)		3.13	3.40		3.56

[a] The distances greater than 5 Å are not shown. The data given in the last two columns represented by min_{wc} and $equi_{wc}$ are obtained with constraints on the active site Asp diad.

Two approaches have been adopted in the present study to keep the configuration of the active site Asp diad closer to that found in the crystal structure. In the first approach, the two Asp residues were included in the list of the rigid part of the system so that the crystallographically determined configuration is maintained

throughout the dyanmics run. In the second approach, four interoxygen distances of the Asp diad are constrained to their crystallographic values by treating these distances as virtual bonds.[43] These four distances are (1) OD1 of Asp-35 with OD1 of Asp-218, (2) OD1 of Asp-35 with OD2 of Asp-218, (3) OD2 of Asp-35 with OD1 of Asp-218, and (4) OD2 of Asp-35 with OD2 of Asp-218. The freezing of the two active site Asp residues in the first approach might result in the arrest of the movement of the backbone atoms of the residues on either side of the frozen residues, which may be considered too strong a constraint during molecular dynamics simulations. On the other hand, only the movement of the two side chains of the Asp diad is partially arrested in the second approach, and the rest of the system is unaffected. Therefore, the second approach appears to be better than the first one. However, the two approaches give almost identical results in the sense that the configurations of the two Asp residues and the neighboring residues are preserved as found in the crystal structure during the equilibration and the perturbation runs. In the equilibrated structure with constraints, the hydrogen bonding between the N atom of Gly-37 and Gly-220 and the two inner oxygens of the Asp diad and between the Ser-38 and Thr-221 hydroxyl groups and the two outer oxygens of the diad is maintained as in the starting structure shown in Figure 1. The strong interaction of the central hydroxyl of pepstatin with the active site aspartates is also maintained as in model 3 structure shown in Figure 8. The contact distances between the inhibitor backbone atoms and the enzyme active site are given in Table VIII. These distances are shorter in most cases than those obtained from minimization and equilibration done without the active site constraints. It should be of interest to note here that N of Sta-404 and N of Ala-405 make only weak interactions with the enzyme active site even in the structures obtained with the constraints. Moreover, the N of Sta-404 is involved in multiple interactions with the enzyme instead of making a single strong interaction as suggested in the crystal structure.

Mutations on Pepstatin in Water. The results of different mutations on pepstatin in water are given in Table IX. In all the simulations except for the last simulation, 19, the mutations on pepstatin were done on the groups at the C(3) atom of the central statine residue. The first eight simulations determined the free energy change for the mutation of statine, Sta, to dehydroxystatine, dSta. In this mutation, the *S*-hydroxyl group of Sta was mutated to a hydrogen. This mutation was carried out with different simulation conditions to check the dependence of the calculated free energy changes on these conditions. The eight simulations were started from three different configurations of the pepstatin solutions. Simulations 1 and 2 used pepstatin in a rectangular box of 759 water molecules. Simulations 3–7 used pepstatin in a larger box of 1086 water molecules. The last simulation used a sphere of 1012 water molecules centered around pepstatin molecule. The simulations were repeated with different simulation conditions to see their effect on the calculated results. The simulation conditions were varied by (1) including the electrostatic decoupling, (2) changing the number of perturbation windows (the $\Delta\lambda$ value), (3) removing the center of motion of the solute during dynamics, and (4) including all the solute–solute and solute–solvent nonbonded interactions without any cut-off distance. The results seem to be affected only marginally by the variation of these simulation conditions. The average ΔG_{sol} value obtained from all eight simulations for this mutation is 3.01 kcal/mol, and the standard deviation is 0.29 kcal/mol. The maximum error in these calculated values is, therefore, about 10%. In simulation 1, the free energy run was decoupled into electrostatic and van der Waals contributions, which are 2.20 and 0.69 kcal/mol, respectively. The ΔG_{sol} value (2.89 kcal/mol) for this simulation is very close to the average value (3.01 kcal/mol). Simulation 2, with a smaller simulation time of 30.6 ps, gave 3.09 kcal/mol, which is very close to the average value. Comparison of simulations 3 and 4, which used 51 and 101 windows, respectively, shows that the free energy values for this mutation are not affected by the number of windows. However, the free energy values of these two simulations are lower than the average value. It must be noted that no nonbonded cutoff was employed for solute–solute interactions in these two simulations, whereas such a cut-off was employed for simulations 5–7. The application of periodic boundary conditions to solute–solvent interactions in simulation 5 reduces the value of simulation 6 from 3.42 to 3.07 kcal/mol. The latter value is close to the average value. In all the above simulations, the center of mass of the solute was reset to the initial position every two steps during the dynamics run. In simulation 7, the solute center of mass was not disturbed by any resetting. The structure at the end of the 40-ps perturbation run showed that the solute was still at the center of the box. The free energy value from this simulation is very close to that from the earlier simulation. It appears that the resetting of the solute's center of mass does not affect the results to any significant extent. Resetting of the solute center of mass is recommended when the nonbonded cut-off distance is greater than the distance between the solute surface and the side of the box. In such a situation, the periodic boundary conditions for the solute–solvent simulations should also not be enforced. The last simulation used a sphere of water shell around the solute, and no periodic boundary conditions were used. This condition is equivalent to that used for the simulations of the enzyme–inhibitor complex. Therefore, the value for ΔG_{sol} from this simulation should be more appropriate for calculating the $\Delta\Delta G$ value. The value obtained in this simulation is very close to the average value and the values from the previous two simulations. Since, the simulation with a periodic box gives a ΔG_{sol} value at constant temperature and constant pressure, it can be compared directly with the experimental value for the free energy of solvation. Moreover, if the simulations with water shell are not properly set up, the solute sometimes escapes partially from the water shell. Therefore, the simulations for all the other mutations were done with periodic boxes.

The next two simulations, 9 and 10, determine the differences in the free energies of solvation between two analogues of pepstatin; in the first analogue, rSta, the central statine has the hydroxyl group in the *R* configuration at the C(3) atom, and the second analogue is dSta, which lacks the hydroxyl group. In simulation 9, a hydroxyl group was grown on dSta in the *R* configuration, while the hydroxyl group of rSta was mutated to a hydrogen atom in simulation 10. The average of the ΔG_{sol} values from these two simulations is 3.58 kcal/mol, which is close to the value obtained for the mutation of the 3(*S*)-OH group of Sta in the first eight simulations.

The simulations 11 and 12 give the solvation free energy difference between pepstatin and its analogue in which the central Sta is substituted by a dihydroxy analogue, hSta. As in the earlier case, in simulation 11, a hydroxyl group was grown on Sta in pepstatin in the *R* configuration, while the *pro-R* hydroxyl group of hSta was mutated to a hydrogen atom in simulation 12. The average of the ΔG_{sol} values from these two simulations is 4.27 kcal/mol. In the next two simulations 13 and 14, the solvation free energy difference between the rSta and hSta substituted analogues of pepstatin is determined. In simulation 13, a hydroxyl group was grown on the *pro-R* side at the C(3) atom of central statine, while the *pro-S* hydroxyl group of dihydroxy analogue of pepstatin was mutated to a hydrogen atom in simulation 14. The average of the ΔG_{sol} values from these two simulations is 4.77 kcal/mol. The two sets of simulations determine the cost of removing a hydroxyl group either from the *pro-R* or *pro-S* position of the dihydroxy analogue of pepstatin, and the two values of ΔG_{sol}

(43) The virtual bond is set up between two nonbonded atoms to constrain the distance between the two atoms to an initial distance. This is achieved by substituting the nonbonded interactions between the two atoms by a bond-stretching term with a force constant of 200 kcal/Å. The interaction potential between the two atoms and the rest of the system is not affected.

(44) Miller, M.; Schneider, J.; Sathyanarayana, B. K.; Toth, M. V.; Marshall, G. R.; Clawson, L.; Selk, L.; Kent, S. B. H.; Wlodawer, A. *Science* **1989**, *246*, 1149.

(45) DeVault, D. *Quantum Mechanical Tunneling in Biological Systems*, 2nd ed.; Cambridge University Press: 1984.

(46) Meot-Ner, M.; Sieck, L. W. *J. Am. Chem. Soc.* **1986**, *108*, 7525.

(47) Rich, D. H.; Bernatowicz, M. S.; Agarwal, N. S.; Kawai, M.; Salituro, F. G. *Biochemistry* **1985**, *24*, 3165.

are, therefore, close to each other.

The next four simulations, 15, 16, 17, and 18, involve the methylated analogues of pepstatin and determine the contribution of a methyl group on either the *pro-S* or *pro-R* side of the C(3) atom of the central statine in pepstatin to the free energy of solvation. When the methyl group is in the *pro-R* position, the hydroxyl group is in the *pro-S* position and the analogue is termed as mSta. The other stereomer, which is termed as mrSta, has the methyl group in the *pro-S* position and the hydroxyl group in the *pro-R* position. In simulations 15 and 16, a methyl group of the mSta analogue was mutated to a hydrogen atom. The average of the ΔG_{sol} values from these two simulations is −0.21 kcal/mol. In the simulations 17 and 18, the methyl group of mrSta was mutated into a hydrogen atom. The average of the ΔG_{sol} values from these two simulations is −0.11 kcal/mol. It appears that the solvation free energy of pepstatin is almost unaffected by the addition of a methyl group in either configuration at the C(3) atom of the central statine residue.

The last simulation 19, gives the change in the free energy of solvation when Val-403 of pepstatin is mutated to Ala-403. The ΔG_{sol} for this mutation is −0.39 kcal/mol which suggests that pepstatin is slightly less soluble than pepstatin analogue having a shorter hydrophobic side chain. The experimental values for the free energies of solvation of pepstatin and its analogues used in these calculations are not available in the literature for comparison with the calculated values. The calculated ΔG_{sol} values from all the simulations, however, appear to be reasonable, and these values are used for the calculation of $\Delta\Delta G_{bind}$ values for different inhibitors.

Mutations on Pepstatin Inside the Enzyme. The free energy differences obtained for different mutations on pepstatin complexed with the enzyme are given in Table X. Seven different mutations were carried out, and each of these were done with different starting structures and simulation conditions. In particular 10 different simulations were carried out for the Sta → dSta mutation since this mutation is central to the question being addressed in this study.

Mutation of 3(*S*)-OH Group. This mutation was carried out with models 2 and 3. Simulations 6–10 with model 3 were carried out with constraints on the active site diad.

Simulations 1 and 2. These two simulations used model 2 as the starting structure. Some of the difficulties, we experienced in preparing in this model were described earlier. In the first model structure with the 2apr coordinates, pepstatin itself was modeled incorrectly. The Ala-405 residue was a D-amino acid instead of an L-amino acid. The simulation 1 used the wrong structure of pepstatin, and the simulation 2 used the correct structure.

The structure around the active site of the enzyme obtained after equilibration of the structure in model 2 was found to be quite similar to the structure described in Figure 12. In this structure, the two carboxylates are distorted from the planarity. The inner oxygen of Asp-35 makes strong hydrogen bonds with the two oxygens of Asp-218. The outer oxygen of Asp-35 makes hydrogen bonds with both N and OG of Ser-38. Similarly, both N and OG1 of Thr-221 are hydrogen bonded to an outer oxygen of Asp-218. As a result, the interactions of the carbonyl oxygens of Gly-37 and Gly-220 with the inhibitor amide nitrogens at Sta-404 and Ala-405 are drastically weakened. The interactions between the backbone atoms of the first three residues of pepstatin and the enzyme are present as described in the crystal structure. The orientations of the last two residues (Ala-405 and Sta-406) and their interaction with the enzyme active site are found to be different. For instance, the side chain of the Sta-406 is oriented toward the solvent instead of the interior of the active site. The hydrogen-bonding interactions of the Sta-404 carbonyl oxygen and Ala-405 amide nitrogen with the active site are absent in the structure of this simulation. The central Sta hydroxyl group makes only a single strong hydrogen bond with the Asp diad. The hydroxyl oxygen is at a distance of 2.91 Å from the outer oxygen of Asp-218 and is at a distance greater than 3.2 Å from the other three oxygens. The Sta → dSta mutation in this configuration results in a free energy change of 1.29 kcal/mol. The interaction energy (IE) between the inhibitor and the rest of the system increases from −132.4 to −127.2 kcal/mol.

In simulation 2, the configuration at the CA atom of Ala-405 of pepstatin was changed from *R* to *S* so that pepstatin contains L-Ala-405 instead of D-Ala-405 residue. Pepstatin with L-Ala-405 was remodeled into the active site (in the structure obtained after equilibration in simulation 1) by superimposing it on the old inhibitor. The new pepstatin structure in the complex was minimized by holding the enzyme rigid and allowing the inhibitor and the cap water molecules to explore possible minimum energy conformations. The resulting structure was equilibrated for 10 ps. The equilibrated structure around the active site of the enzyme is very close to the structure given in Figure 12 and also closer to the structure obtained in the previous simulation. But the interaction of the last two residues of pepstatin are different from the structure of the last simulation but similar to that described in Figure 13. The important interactions of this structure are discussed below. The two carboxyl groups of the Asp diad remain almost perpendicular to each other, and the distance between the two inner oxygens is around 2.60 Å. The hydrogen bonds between the inner oxygen of Asp-35 and N of Gly-37 and between the outer oxygen of Asp-35 and OG of Ser-38 are absent. Instead N of Ser-38 forms a weak hydrogen bond with the outer oxygen of Asp-35 (3.70 Å). The N of Gly-220 and the inner oxygen of Asp-218 remain well-defined, but the outer oxygen of Asp-218 makes hydrogen bonds with both N and OG1 of Thr-221 (2.81 and 2.74 Å) in addition to its interaction with the hydroxyl oxygen of central Sta. Though the orientations of the last two residues (Ala-405 and Sta-406) inside the enzyme active site are correct in this structure, their interaction with the active site is not significantly improved from the previous simulation. The hydrogen-bonding interactions of the Sta-404 carbonyl oxygen and Ala-405 amide nitrogen with the active site are still absent in this structure. However, O of Sta-404 interacts strongly with OG of Ser-38 (2.90 Å), and as a result the interaction of the later with the OD1 of Asp-35 is weaked (4.41 Å). The interactions of the first three residues and the enzyme active site remain well-defined as in earlier cases. The hydroxyl oxygen of Sta-404 makes a single hydrogen bond with the outer oxygen of Asp-218 (2.84 Å). The free energy change for Sta → dSta mutation in this configuration is 2.88 kcal/mol, and the IE decreases from −106.7 to −125.0 kcal/mol.

Simulations 3–5. These simulations used model 3 for the mutation of Sta → dSta. The structures of the enzyme–inhibitor complex obtained after minimization and equilibration have been described earlier. The two carboxylates of the active site Asp diad are almost perpendicular to each other in the equilibrated structure (Figure 12). The proton is positioned almost perpendicular to the plane of the carboxylate group of Asp-35. The distance between the two inner oxygens of the diad is 2.60 Å. For simulation 3, the free energy change for the mutation of the hydroxyl group of Sta in this configuration is 3.88 kcal/mol, and the IE decreases from −126.5 to −130.6 kcal/mol.

The next two simulations used the enzyme–pepstatin complex structure obtained from the optimization of model 3 by a combined method of quantum mechanics and molecular mechanics with the SIVA module of the AMBER program. The details of the optimization were described in an earlier section on the acidity of the active site of the enzyme. In this optimization, the side-chain atoms of the two Asp residues in the active site were treated by quantum mechanics, while the rest of the system was treated by molecular mechanics. This optimized structure was equilibrated for 20 and 30 ps to generate two input structures for the two perturbation simulations. The two structures look very similar to the structure described in Figure 13 except for some minor differences. The two carboxylates are oriented perpendicular to each other in these structures also. The hydroxyl group of the central statine residue makes only a single strong interaction with the Asp diad. The distance between the hydroxyl oxygen and the closest oxygen of the diad (OD2 of Asp-218) is about 2.90 Å. The free energy changes in the two simulations are almost equal. In the first simulation, the ΔG_{enz} is 2.05 kcal/mol and IE increases

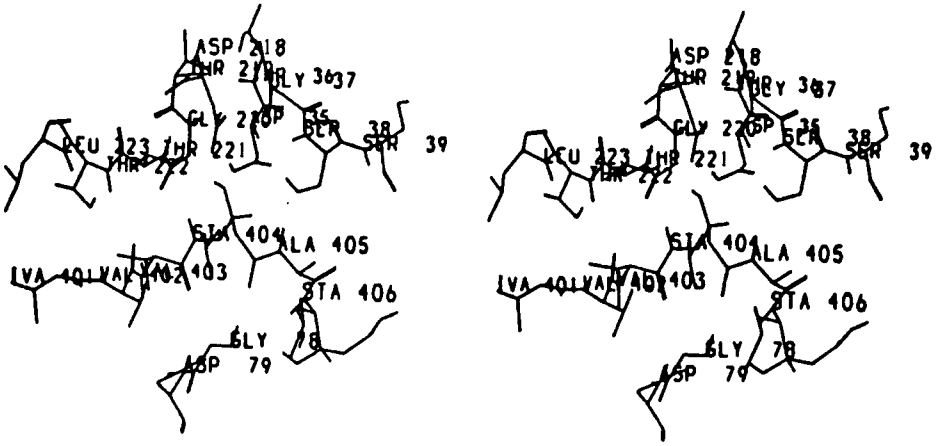


Figure 13. Structure of the enzyme–pepstatin complex around pepstatin after equilibration.

from −141.5 to −131.7 kcal/mol, and in the second simulation, the ΔG_{enz} is 2.22 kcal/mol and IE increases from −142.4 to −128.0 kcal/mol.

Simulations 6–10. Since the configuration of the active site Asp diad is considerably distorted from the configuration observed in the crystal structure and thereby weakening its interaction with the inhibitor, the following five simulations were carried out with model 3 with constraints to keep the Asp diad conformation closer to planarity as found in the crystal structure. In simulation 6, the structure was first minimized without any constraints, which resulted in the shortening of the two inner oxygens from 3.0 to 2.6 Å. However, the two carboxylates remained coplanar. This configuration of the two Asp residues was frozen during the equilibration and perturbation runs. The free energy change for the mutation, ΔG_{enz}, in this configuration is 6.63 kcal/mol. The IE increases from −134.4 to −130.4 kcal/mol during the mutation.

To see the effect of larger inner oxygen distances found in the crystal structure on the free energy change for this mutation, the geometry of the active site Asp diad was frozen during the minimization for simulation 7. The rest of the enzyme, inhibitor, and the surrounding counter ions and water molecules were minimized. The configuration of the two Asp residues was kept frozen during the equilibration and the perturbation runs also. The free energy change, ΔG_{enz}, for the mutation in this configuration is 9.02 kcal/mol. The IE increases from −144.4 to −127.8 kcal/mol during the mutation.

In the next simulation (no. 8), the coplanarity of the two carboxylates of the Asp diad was maintained by constraining the four interoxygen distances to their crystallographic values. The free energy change for the mutation, ΔG_{enz}, in this configuration is 9.23 kcal/mol. The IE increases from −145.0 to −123.8 kcal/mol during the mutation. This simulation was repeated with the same equilibrated structure twice in simulations 9 and 10. In simulation 9, 101 windows were used for the perturbation run instead of 51 windows. This doubled the simulation time from 40.8 to 80.8 ps. The ΔG_{enz} value for this simulation is 7.85 kcal/mol, and the IE increases from −145.0 to −126.1 kcal/mol. In simulation 10, the perturbation run was carried out without the active site constraints. This simulation used 51 windows. The ΔG_{enz} value is 2.43 kcal/mol, and the IE increases from −142.8 kcal/mol to −116.1 kcal/mol. This ΔG_{enz} value obtained in this simulation is much smaller than the values obtained from the simulation with active site constraints (simulations 11–14) and is closer to the values obtained from the simulations 1–10, which did not use constraints during equilibration or perturbation runs.

Mutation of 3(*R*)-OH Group. Simulations 11 and 12. The two simulations determine the contribution of the hydroxyl group of the central Sta residue of pepstatin in the *R* configuration. This was achieved in two different ways by mutating dSta → rSta and rSta → dSta. In both the simulations, no constraints on the active site Asp diad were used. In simulation 11, a hydroxyl group was grown in *R*-configuration at the C(3) atom of dehydroxystatine. The starting structure used for this simulation is that obtained at the end of the perturbation run in simulation 1. Since this simulation used model 2 for the enzyme–inhibitor complex without any active site constraints, the Asp diad was distorted in the same manner as described in simulation 1. The ΔG_{enz} value for mutation of dSta → rSta in this configuration is −3.17 kcal/mol, and the IE increases from −145.8 to −134.2 kcal/mol during the mutation. In simulation 12, the hydroxyl group in the *R*-configuration was mutated into a hydrogen. This simulation used model 3 in which the positions of the hydroxyl group and hydrogen atom attached to the C(3) atom of the central statine residue were swapped. In this structure, the *R*-OH group of the central statine residue makes only single hydrogen-bonding interaction with one of the carboxyl oxygens of Asp-35. The ΔG_{enz} value for mutation of rSta → dSta in this configuration is 2.97 kcal/mol, and the IE increases from −131.1 to −123.4 kcal/mol during the mutation.

Mutations on Dihydroxy Analogue. Simulations 13–16. These simulations determine the contribution of a second hydroxyl group at the C(3) atom of central statine to the free energy of binding of the inhibitor in the active site of the enzyme. The simulation 13 mutated the hydrogen at the C(3) atom into a hydroxyl group with model 2 as the starting structure. The ΔG_{enz} value for this mutation is −4.64 kcal/mol, and the IE decreases from −128.4 to −142.3 kcal/mol. The next simulation 14 used the coordinates obtained at the end of the perturbation run in the previous simulation (no. 13), and the hSta was mutated back to Sta in 61.2 ps. The coordinates were minimized and equilibrated for 10 ps before starting the perturbation run. The ΔG_{enz} value for this mutation is 5.64 kcal/mol, and the IE remains almost constant at −147.0 kcal/mol. The next simulation, 15, used model 3 as the starting structure, which has been modified such that the C(3) atom of central statine residue has the second hydroxyl group in the *pro-R* configuration as well. This structure obtained after equilibration looks very similar to the structure described in Figure 13. Additionally, the *R*-OH group is in hydrogen bond contact with the outer oxygen of Asp-35. The ΔG_{enz} value for the mutation of hSta → Sta in this configuration is 5.22 kcal/mol, and the IE decreases from −119.8 to −127.6 kcal/mol.

The ΔG_{enz} values obtained from the three different simulations for the mutation of hSta → Sta are close to each other. Therefore, for hSta → rSta mutation, only one simulation (no. 16) was carried out. This simulation used the coordinates obtained at the end of the perturbation run in simulation 13, and these coordinates are the same as those used in simulation 14. The ΔG_{enz} for mutation of hSta → rSta in this structure is 6.52 kcal/mol, and the IE increases from −147.0 to −139.4 kcal/mol.

Mutations on the C(3)-Methylated Analogues. Simulations 17–19. These simulations determine the effect of substitution of a methyl group in the place of the hydrogen attached to the C(3) of Sta or rSta of the inhibitor. The new inhibitors were modeled into model 3 structure by simple substitution, and the resulting structures were minimized and equilibrated before the mutation runs. No constraints on the active Asp diad were imposed during these simulations. The presence of a *pro-R* methyl group at the

C(3) of central statine did not affect the interaction of the *S*-OH group and other interactions between the inhibitor and the enzyme active site. The simulation 17 determines the contribution of a methyl group at the C(3) of the central statine residue in *pro-R* configuration to the free energy of binding. The ΔG_{enz} value for this mutation, (mSta → Sta), is 1.25 kcal/mol, and the IE increases from −127.4 to −122.3 kcal/mol during the mutation.

The simulations 18 and 19 determine the contribution of the methyl group at the C(3) atom of the central Sta residue in *pro-S* position when the hydroxyl group is in the *pro-R* position. In this structure of the enzyme–inhibitor complex, the *pro-S* methyl group of the inhibitor is positioned at the center of the two carboxylates of the Asp diad, and the *pro-R* hydroxyl group is in hydrogen bond contact with the outer oxygen of Asp-35. The ΔG_{enz} value for this mutation (mrSta → rSta) is −0.10 kcal/mol. The IE increases from −131.1 to −133.5 kcal/mol during the mutation. In simulation 27 also, the model 3 coordinates were used, except that the Asp-35 was charged and the Asp-218 was neutral. In all earlier simulations, Asp-218 was the charged residue. This change was made because it appeared from the equilibrated structure in the earlier simulation that the *pro-R* hydroxyl of the inhibitor will have stronger interaction with the Asp diad if Asp-35 is charged. The ΔG_{enz} value for the mrSta → rSta mutation in this structure is −1.38 kcal/mol. The IE remains almost the same (about −133 kcal/mol) at the beginning and the end of the mutation.

Simulation 20. The last simulation determines the contribution of the hydrophobic side chain of the Val-3 residue of pepstatin in the P_2 subsite. The starting coordinates used for this simulation were taken from simulation 4, which used the model 2 coordinates. The mutation of Val-405 → Ala at this site gives rise to a ΔG_{enz} value of −0.88 kcal/mol. The IE decreases from −128.6 to −145.4 kcal/mol during the mutation.

Discussion

Acidity of the Active Site. The results of the geometry optimization studies on the model systems in the gas phase show that the geometry observed for the carboxylates of the two active site Asp residues is not the most stable configuration. The two inner oxygens prefer to be at a distance of about 2.5–2.6 Å in the absence of the enzyme environment rather than 2.9 Å observed in the crystal structure. Moreover, the planarity of the two carboxylates is distorted in the optimized structures. The energy of the acetic acid–acetate complex is very sensitive to the change in the distance between the inner oxygens and the change in the relative orientations of the two carboxylates. It is also found that the water molecule near the complex, which makes a strong hydrogen bond with the acetate ion, does not affect the relative energies of the complex with different proton positions. The calculated ab initio energies for the Asp diad inside the enzyme are higher than that in the gas phase. However, the contribution of the surroundings lowers the energy substantially in each case. Nevertheless, it is interesting to note that the total energies of the diad inside the enzyme complexed with pepstatin are lower when the distance between the two inner oxygens are smaller. The present results, therefore, suggest that the larger distance between the two inner oxygens observed in the crystal structure is not necessarily the lowest energy configuration of the enzyme active site. It may be of interest here to note that this distance is 2.6 Å in HIV-1 proteinase.[44] The barrier for proton transfer between the two Asp residues of the diad is about 15 kcal/mol, when the distance between the two inner oxygens of the Asp diad is about 2.9 Å. This barrier is smaller (0.8 kcal/mol), when the distance is about 2.45 Å.

The results of optimization of the Asp residues inside the enzyme complexed with an ammonium ion or pepstatin show that the charged Asp-35 is preferred in the presence of ammonium ion, and charged Asp-218 is preferred in the presence of pepstatin. These results are in agreement with the earlier conjectures[10] made from the examination of the structure of penicillinopepsin. The difference in energies of the two states is about 3.5 kcal/mol in the presence of ammonium ion and about 5.0 kcal/mol in the presence of pepstatin, when the distance between the two inner oxygens of the Asp diad is between 2.4 and 2.5 Å. When the distance between the two inner oxygens is around 2.90 (closer to the crystallographic distance), the energy difference is lower (1.3 kcal/mol), though the energies of the two states of the diad are higher in this case than in the earlier cases with smaller interoxygen distances. It may be noted here that the barrier to proton transfer is higher when the interoxygen distance is larger even though the energy difference between the two states is smaller. This has been brought out by the results from the gas-phase study as well as the free energy perturbation study. These results show that the energies of the system are very sensitive to the relative positioning of the two aspartates. Even small changes in the relative orientation of the two carboxylates affect these energies. Though the active site Asp residues are found to be rigid inside the enzyme, small changes in the relative positions or orientations of the two aspartates are expected to occur due to the dynamic nature of the enzyme. Therefore, the energetics of the system, which are dynamic in nature, get affected to different extents by the dynamics of the enzyme. Different substrates also seem to affect the energetics to different extents. This suggests that both the charge states may reach energetically equivalent configurations due to the dynamics of the enzyme.

In the presence of an ammonium ion, the change in the free energy with the mutation of the proton position is about 7.0 kcal/mol, suggesting that the initial state is stable compared to the final state. It may be noted that both the ΔG_1 and ΔG_2 values are positive, and the latter is quite large. However, the ΔG_2 values are expected to be similar to the ΔG_1 values in magnitude and opposite in sign if the intermediate state were a transition state. We found that when the proton is transferred from Asp-218 to Asp-35 in the presence of an ammonium ion, the proton is positioned in almost a perpendicular orientation with respect to the carboxylate plane (Figure 12), whereas the proton in the initial structure is positioned in a cis configuration (Figure 1). The same trend was noticed in the gas-phase optimization study on the acetic acid–acetate–ammonium ion complex, in which case also the energy increased by about 5 kcal/mol from the initial state to the final state. The reason for this monotonous increase is obviously the inherent difference in the energy of a carboxyl group with proton in different conformations. The cis conformation is the lowest in energy and the high-energy conformation found in the final state is due to the distortion of the coplanar conformation of the two carboxylates. Such a distortion of the two carboxylates is caused both in the gas phase and in the enzyme active site. It is also possible that the equilibrium configuration of the final state is not attained during the perturbation due to inadequate sampling. This problem is probably accentuated by the presence of ammonium ion. For the calculations in the presence of pepstatin, such a problem does not manifest, and the ΔG_1 and ΔG_2 values are almost equal and carry opposite sign. This suggests that a transition state of about 1.0 kcal/mol exists for the transfer of proton from one site to the other site. The ΔG_{tot} values for both the forward and reverse transfers of the proton are very small compared to the ΔG_{tot} values obtained in the presence of an ammonium ion suggesting that the two charge states are almost equivalent in the presence of pepstatin. These results support the view that the proton is shared equally by the two Asp residues in the active site, and the proton can shuttle between the two sites because of the low-energy barrier of about 1.0 kcal/mol. The same barrier of about 1.0 kcal/mol is observed for the gas-phase optimization calculations of the acetic acid–acetate complex in planar configuration.

The present results are in agreement with those of an earlier study[6] on a model of the active site of an aspartic proteinase in the gas phase on the following three points: (1) There is little preference for one Asp to be ionized over the other inside the active site complexed with a neutral ligand. (2) A water molecule in the active site does not alter the relative energies of the two states of the diad in the gas phase. (3) The neighboring residues affect the energies of the two states of the diad inside the enzyme. (4) The energies of the two states are very sensitive to changes in their interoxygen distances. In addition to these important conclusions,

Table IX. Free Energy Changes for the Mutations in Water[a]

no.	mutation	Δt (ps)	tot t (ps)	ΔG_{sol}	IE_{start}	IE_{end}
1	Sta → dSta	0.002	82.0	2.89	−123.2	−110.8
2	Sta → dSta	0.002	40.8	3.09	−110.3	−118.1
3	Sta → dSta	0.001	40.8	2.57	−124.8	−130.5
4	Sta → dSta	0.001	50.5	2.67	−123.3	−123.6
5	Sta → dSta	0.001	40.8	3.42	−125.3	−112.9
6	Sta → dSta	0.001	40.8	3.07	−125.3	−122.6
7	Sta → dSta	0.001	40.8	3.27	−125.3	−121.5
8	Sta → dSta	0.001	40.8	3.08	−113.4	−128.0
9	dSta → rSta	0.002	40.8	−3.77	−115.8	−118.6
10	rSta → dSta	0.002	40.8	3.40	−119.8	−113.4
11	Sta → hSta	0.002	40.8	−4.54	−122.9	−134.4
12	hSta → Sta	0.001	40.8	4.00	−149.1	−120.1
13	hSta → rSta	0.002	60.6	4.83	−105.6	−132.0
41	hSta → rSta	0.001	40.8	4.71	−149.8	−122.2
15	mSta → Sta	0.002	40.8	−0.05	−104.1	−118.1
16	mSta → Sta	0.001	40.8	−0.37	−134.5	−127.0
17	mrSta → rSta	0.002	40.8	−0.11	−112.7	−125.0
18	mrSta → rSta	0.001	40.8	−0.12	−115.0	−143.9
19	Val-330 → Ala	0.002	40.8	−0.39	−111.5	−121.9

[a] The values in the last three columns are in kcal/mol. ΔG_{sol} values are the averages of the forward and backward values with a maximum hysteresis of 0.1 kcal/mol or less. IE values have a root mean square deviation of about 5 kcal/mol.

Table X. Free Energy Changes for the Mutations Inside the Enzyme[a]

no.	mutation	model	Δt (ps)	tot t (ps)	ΔG_{enz}	IE_{start}	IE_{end}
1	Sta → dSta	2	0.002	65.5	1.29	−132.4	−127.2
2	Sta → dSta	2	0.001	50.4	2.88	−106.7	−125.0
3	Sta → dSta	3	0.002	40.8	3.88	−126.5	−130.6
4	Sta → dSta	3	0.002	40.8	2.05	−142.5	−131.7
5	Sta → dSta	3	0.002	40.8	2.22	−141.4	−128.0
6	Sta → dSta	3	0.001	40.8	6.63	−134.4	−130.4
7	Sta → dSta	3	0.001	40.8	9.02	−144.4	−127.8
8	Sta → dSta	3	0.001	40.8	9.23	−145.0	−123.8
9	Sta → dSta	3	0.001	80.8	7.85	−145.0	−126.1
10	Sta → dSta	3	0.001	40.8	2.43	−142.8	−116.1
11	dSta → rSta	2	0.002	40.8	−3.17	−145.8	−134.2
12	rSta → dSta	3	0.002	40.8	2.97	−131.1	−123.4
13	Sta → hSta	2	0.002	40.8	−4.64	−128.4	−142.3
14	hSta → Sta	2	0.002	61.2	5.64	−147.0	−147.3
15	hSta → Sta	3	0.002	40.8	5.22	−119.8	−127.6
16	hSta → rSta	2	0.002	60.6	6.52	−147.0	−139.4
17	mSta → Sta	3	0.001	40.8	1.25	−127.4	−122.3
18	mrSta → rSta	3	0.001	40.8	−0.10	−131.1	−133.5
19	mrSta → rSta	3	0.001	40.8	−1.38	−133.4	−133.0
20	Val-330 → Ala	2	0.002	40.8	−0.88	−128.6	−145.4

[a] The values in the last three columns are in kcal/mol. ΔG_{enz} values are the averages of the forward and backward values with a maximum hysteresis of 0.1 kcal/mol or less. IE values have a root mean square deviation of about 5 kcal/mol.

our study shows that the barrier to proton transfer is higher if the distance between the two oxygens is close to 2.9 Å and is much lower (about 1.0 kcal/mol) if the distance is close to 2.5 Å. We suspect that the carboxylates of the two Asp residues in the active site assume the configuration found in the gas phase to facilitate proton transfer due to the dynamic nature of the protein. It is also possible that the proton shuttles between the two oxygens by quantum mechanical tunneling,[45] which the present study does not address.

Inhibitor Binding. The results of minimization and equilibration on the enzyme–inhibitor complexes showed that the interaction between the enzyme and the inhibitor were drastically affected because of the changes in the configuration of the active site Asp diad. It has also been shown that the coplanar configuration of the Asp diad carboxylates is crucial for the optimal interactions between the enzyme and the inhibitor. Therefore, the results of the mutations of the central statine residue of pepstatin inside the enzyme active site obtained without constraints on the Asp diad were found to be different from those obtained with the constraints. The $\Delta\Delta G_{bind}$ values for different mutations were computed from the ΔG_{enz} and ΔG_{sol} values obtained in different simulations. These values are summarized in Table XI. For each mutation, the ΔG_{sol} values obtained in several mutations are averaged, and the average values are presented in the table. The ΔG_{enz} values obtained with or without constraints and those obtained with different starting structures are listed as different entries in this table.

The free energy change for the mutation of the central statine to dehydroxystatine in water, ΔG_{sol}, is equal to 3.01 kcal/mol. For the same mutation inside the active site of the enzyme, two different values for free energy change, ΔG_{enz}, were obtained for the two different sets of calculations done with and without constraints on the active Asp residues. The ΔG_{enz} values are 2.46 and 8.18 kcal/mol, respectively, for the mutations done with and without constraints. The standard deviations calculated for the two sets of ΔG_{enz} values are 0.87 and 1.20 kcal/mol, respectively. These two different values lead to $\Delta\Delta G_{bind}$ of −0.55 and 5.17 kcal/mol, respectively. The latter value is closer to the experimental value of about 5.0 kcal/mol.

The large difference between the ΔG_{enz} values obtained with and without the constraints on the active site Asp residues is obviously a result of the differences in the active site configurations in the vicinity of the hydroxyl group of the central statine residue of the inhibitor obtained with and without the constraints. When no constraints were applied on the active site Asp diad, drastic changes in the relative orientations of the two residues were observed. As a result, the orientations of the neighboring residues interacting with the diad were also affected, and the interactions

Table XI. Calculated and Experimental Binding Free Energy Differences[a]

no.	mutation	ΔG_{enz}	ΔG_{sol}	$\Delta\Delta G_{bind}$	expt
1	Sta → dSta	2.46	3.01	−0.55	4.9
		8.18	3.01	5.17	
2	rSta → dSta	3.07	3.58	−0.51	−0.8
3	hSta → Sta	5.16	4.27	0.89	0.3
4	hSta → rSta	6.52	4.77	1.75	4.4
5	mSta → Sta	1.25	−0.21	1.46	−5.6
6	mrSta → rSta	−0.10	−0.11	0.01	2.5
		−1.38	−0.11	−1.27	
7	Val-403 → Ala	−0.88	−0.39	−0.49	

[a] All values are in kcal/mol. The experimental values were taken from ref 47.

between the active site and the inhibitor were weakened. Therefore, the free energy change for the mutation of the hydroxyl group of pepstatin in this configuration is 2.46 kcal/mol, which is less than the free energy change (3.01 kcal/mol) obtained for the same mutation in water. These values suggest that the dehydroxypepstatin binds to *Rhizopus* pepsin tighter by about −0.55 kcal/mol than pepstatin. This is contradictory to the experimental results. A negative value for $\Delta\Delta G_{bind}$ for this mutation may be interpreted to suggest that experimental $\Delta\Delta G_{bind}$ of about 5.0 kcal/mol is mainly due to the increase in entropy due to the displacement of the tightly bound water molecule near the Asp diad into the bulk solvent. It also means that there is no major enthalpic contribution to the free energy of binding due to the interaction of the hydroxyl group with the negatively charged Asp diad. Such an interpretation is not necessarily correct in view of the weakening of the interaction of the hydroxyl group with the Asp diad due to the distortion of the active site configuration in the equilibrated structure as discussed earlier. We feel that the ΔG_{enz} value is drastically underestimated in this calculation due to the distortions in the active site. It has been estimated from different studies[39,46] that the enthalpy of association of an alcohol molecule with a formate or an acetate anion is between −17 and −20 kcal/mol. Therefore, the results of the first set of calculations may not correctly represent the free energy change for the mutation of pepstatin inside the active site. In the second set of simulations, the active site is not distorted on equilibration, and the statine hydroxyl group is in hydrogen bonding distance with all four oxygens of the Asp diad, though it is closest to the outer oxygen of Asp-218. The ΔG_{enz} value for the mutation of the hydroxyl group in this configuration is 8.18 kcal/mol, giving rise to a $\Delta\Delta G$ of 5.17 kcal/mol. This value is in good agreement with the experimental value of 5.0 kcal/mol. These results, therefore,

suggest that the difference in binding between pepstatin and dehydroxypepstatin is solely due to the interaction of the hydroxyl group with the enzyme active site. In a recent study, it has been found that a single hydroxyl group contributes up to 9 kcal/mol for binding of a transition-state analogue inhibitor of adenosine deaminase. It was suggested that a single hydrogen bond between the hydroxyl group and a charged Asp or Glu residue at the active site under favorable circumstances can enhance the binding affinity of an enzyme inhibitor by a very large factor. It appears that the favorable situation for the strong interaction of the statine hydroxyl group with the negatively charged Asp diad is provided by the enzyme structure which keeps the two carboxylates of the Asp residues in a somewhat rigid configuration. As seen earlier, no strong interaction between the hydroxyl group and the enzyme are possible if the two carboxylates are not in coplanar configuration. The second set of calculations, therefore, clearly establishes that the major contribution to the difference in binding of pepstatin and dehydroxypepstatin is due to the strong interaction between the Sta hydroxyl group and the active site Asp residues. Hence, it is not necessary to presume that the tighter binding of pepstatin over its analogue lacking the *S*-OH group is due to the displacement of the active site water molecule by pepstatin. The later explanation was advanced based on the kinetic data that the active site bound water molecule (which is displaced when pepstatin binds to the enzyme) is not displaced when dehydroxypepstatin binds to the enzyme. This is not ascertained by direct structural studies. No structural results are reported for any aspartic proteinase complexed with the inhibitors lacking the *S*-OH group. On the contrary, the structure of *Rhizopus* pepsin complexed with a reduced peptide inhibitor shows[35] that the water molecule bound between the two Asp residues is displaced by the inhibitor, though no group of the inhibitor occupies the position of the water molecule. Therefore, it is possible that the water molecule is displaced even when a dehydroxypepstatin binds to the enzyme. If dehydroxypepstatin does not displace the water molecules, then its binding to the enzyme must be even more weaker than what is suggested by the present calculations because the presence of the water molecule in the active site may cause steric hindrance to optimal binding of dehydroxy pepstatin. Though the results of the present study suggests that dehydroxypepstatin may also displace the active site water molecule, a definite proof for this view may be found only by a structural study.

The mutation of the hydroxyl group positioned in the *R* configuration at the C(3) atom of the central statine residue of the inhibitor gives rise to $\Delta\Delta G_{bind}$ of −0.51 kcal/mol, which is very close to the experimental value of −0.8 kcal/mol. It may be noted here that the starting configuration of the inhibitor in the enzyme–inhibitor complex is the same as that of pepstatin. The crystal structure of *Rhizopus* pepsin or any other aspartic proteinase complexed with the pepstatin analogue containing rSta in place of the central Sta residue is not yet available. The agreement between the experimental and calculated values suggest that the modeled configuration of the inhibitor is correct. In this configuration the hydroxyl group of the central rSta residue makes only one hydrogen bond interaction with the outer oxygen of Asp-35. This interaction is not affected to any great extent even when the constraints are not applied to the active site diad. Therefore, the results of the mutation of the hydroxyl group in the *R* configuration are almost equal with or without the constraints on the active site Asp diad. It should also be noted that no water molecule is placed near the active site when this inhibitor is bound. (We suspect that the results may not be affected to any significant extent even if the water molecule is present.) On the other hand, the good agreement obtained between the experimental and calculated results may be taken as a support for the absence of the active site water molecule when the inhibitor is bound.

The above two mutations together determine the difference in the free energies of binding between pepstatin and the analogue with the hydroxyl group in the *R* configuration. The same difference may be obtained by growing the hydrogen atom on the C(3) atom of central statine to a hydroxyl group and then mutating the hydroxyl group on the *pro-S* side to a hydrogen to get the *R* hydroxy analogue of pepstatin. The intermediate inhibitor in this scheme is a dihydroxy analogue instead of the dehydroxy analogue of pepstatin of the first two mutations. The hSta → Sta mutation involves the hydroxyl group on the *pro-R* side of the inhibitor. Since this hydroxyl group is involved affectively in one hydrogen-bonding interaction with the Asp-35 and it is not implicated in the displacement of any water molecule, the calculated $\Delta\Delta G_{bind}$ (0.89 kcal/mol) is quite close to the experimental value (0.3 kcal/mol). On the other hand, the mutation of the *pro-S* hydroxyl group of hSta in the next mutation gives $\Delta\Delta G_{bind}$ of about 1.58 kcal/mol which is lower than the experimental value of 4.4 kcal/mol. It must be noted that the calculations for mutation did not use any constraints on the Asp diad, and the underestimated value in this calculation is a result of the distortions of the active site configuration.

Rich and co-workers[47] studied the binding of methylated analogues of pepstatin to pepsin and obtained some very surprising results. They found that when the hydrogen at the C(3) atom of the central Sta is substituted by a methyl group, its inhibition goes down by a factor of about 12000. On the other hand, if the positions of the hydroxyl group and the methyl group are interchanged, the new inhibitor with the hydroxyl group in the *pro-R* position binds to pepsin almost as strongly as pepstatin. These workers attributed the strong binding affinity of the second inhibitor to the ability of the methyl group (in the *pro-S* position) to displace the active site bound water but provided no explanation for the highly reduced potency of the first inhibitor with the hydroxyl group on the *pro-S* side. In fact, this inhibitor has the ability to displace the water molecule as well as interact with the active site Asp diad, and it should, therefore, bind as strongly as pepstatin. The present simulation studies show that the inhibitor with the methyl group in the *pro-R* position binds to *Rhizopus* pepsin better by about 1.5 kcal/mol than pepstatin, which is in total disagreement with the experimental value of −5.6 kcal/mol. The experimental results point out the possibility that the *pro-R*-methyl group of the inhibitor may interfere with the optimal binding of the inhibitor due to steric reasons. Such steric interference between the inhibitor and the active site of *Rhizopus* pepsin is not observed in the model used in the present study, and, therefore, a positive ΔG_{enz} value was obtained for this mutation. Since the structures of the two enzymes in the active site region are expected to be similar, it is not clear why the *prp-R*-methylstatine-containing inhibitor is such a weaker inhibitor of pepsin. It is possible that the inhibitor binds in a conformation which is different from that modeled in the present study. When the methyl group is present in the *pro-S* configuration, it is positioned between the two carboxylates of the Asp diad in our model. A mutation of the methyl group in the hydrophilic site give rises to ΔG_{enz} of −0.10 and −1.38 kcal/mol in the two sets of the calculations, respectively. The $\Delta\Delta G_{bind}$ values for the two simulation are 0.01 and −1.27 kcal/mol, respectively. Though the calculated results appear reasonable, they are in disagreement with the experimental results. The calculated results suggest that pepstatin with central rSta is as good as or better than the methylated rSta analogue, while the experimental results suggest the opposite. As in the earlier case, the disagreement causes a suspicion that the actual binding modes of the two inhibitors are different from the binding modes modeled for this simulation. Since no structural information on the binding of these inhibitors are available, the disagreement between the experimental and the calculated results can not be explained without further studies. The present calculations, however, find it very intriguing that the methylated analogue of pepstatin with the methyl group in *pro-R* configuration is so much weaker (by about −5.6 kcal/mol) than pepstatin.

Conclusions

The configuration of the side chains of the active site Asp diad found in the crystal structures of *Rhizopus* pepsin is not of the lowest energy. The two carboxylates of the diad prefer a nonplanar configuration rather than a planar configuration. Moreover, a distance of about 2.5 Å between the two inner oxygens is preferred

instead of the distance of 2.97 Å found in the crystal structure. The present study points out that the energetics of the diad are sensitive to small changes in the relative orientation and the interatomic distances of the diad. It also points out the possibility that the proton shuttles between the two inner oxygens of the diad as the barrier for this shuttling is quite small when the inner oxygen distance is short. This may also contribute to the rigidity of the coplanar configuration of the diad. We find that the coplanar configuration of the diad is crucial for optimal binding of pepstatin. However, it has not been possible to keep the diad configuration coplanar during molecular mechanics and molecular dynamics calculations with the existing force field. We feel that the force field needs improvement for the special case of the Asp diad. One difficulty for parameterization of the interaction potential for this case is the limitation of the molecular mechanics approach to model the shuttling of the proton between the two oxygens. We are currently looking at different approaches to model this special part of the enzyme. We are also examining more closely the mobility of the Asp diad by more extensive molecular dynamics simulations of the enzymes with water or other substrates in the active site.

In the absence of improved force fields to mimic the configuration of the Asp diad, the use of constraints on the diad in our calculations gave results which are in good agreement with the experiments. Our results point out that the contribution of about 5 kcal/mol to binding from the hydroxyl group of the central statine residue is mainly due to the strong interaction of this group with the negatively charged Asp diad. It implies, therefore, that entropic contribution to the binding due to the displacement of the active site bound water molecule may not be significant. This raises the possibility that the water molecule is displaced even when pepstatin analogue lacking the hydroxyl group in *S* configuration binds to the enzyme. This may only be confirmed by further experimental studies. We feel that it is an important issue to be resolved because of its implications for designing strong inhibitors of aspartic proteinases and also for understanding their mechanism of action.

Acknowledgment. All the calculations were carried out on the Cray XMP/116-se supercomputer at the Research Institute of Scripps Clinic. We are grateful to Dr. Richard A. Lerner for providing the computational facilities and encouragement. This research is partially supported by grants from NIH (RO1-GM 39410) and Miles Inc.

Registry No. Asp, 56-84-8; pepsin, 9001-75-6; hydrogen ion, 12408-02-5; pepstatin, 26305-03-3; dehydroxypepstatin, 134486-20-7.

Ab Initio Calculations of the Relative Strengths of the π Bonds in Acetylene and Ethylene and of Their Effect on the Relative Energies of π-Bond Addition Reactions

Athanassios Nicolaides and Weston Thatcher Borden*

Contribution from the Department of Chemistry, University of Washington, Seattle, Washington 98195. *Received December 19, 1990*

Abstract: Energies at the CI-SD/6-31G** level have been calculated for acetylene, ethylene, ethane, and the vinyl and ethyl anions, radicals, and cations (classical and hydrogen-bridged). From these energies, the differences between the π-bond energies in acetylene and ethylene and between the energies of addition of H^-, $H^\bullet$, H^+, and H_2 to these π bonds have been computed. Also reported are the calculated differences between the ionization potentials and between the electron affinities of the vinyl and ethyl radicals and between the energies required for formation of vinyl and ethyl anions, radicals, and cations from, respectively, ethylene and ethane. Excellent agreement is found between the calculated energies differences and those derived from the most recent experimental data. The differences between the energies of addition reactions to the π bonds in acetylene and in ethylene are analyzed, and the effect of the nonlinear geometry of vinyl radical on the π-bond energy of acetylene is discussed.

The fact that the heat of hydrogenation of one π bond in acetylene exceeds that of the π bond in ethylene by 9 kcal/mol is commonly attributed to a weaker π bond in the former molecule than in the latter.[1,2] As discussed in the Appendix, this explanation, in terms of a weaker π bond in acetylene than in ethylene, appears to find support in comparison of the adiabatic C–C bond dissociation energies (BDEs) of acetylene, ethylene, and ethane. However, the C–C triple bond length (1.20 Å) is shorter than the C–C double bond length (1.34 Å), indicating that the p orbitals that form the π bonds overlap better in alkynes than in alkenes. The better overlap between the p–π orbitals in alkynes than in alkenes suggests[3] that each π bond in acetylene is, in fact, *stronger* than the π bond in ethylene.

As part of our explorations of the factors that affect the relative energies of π bonds[4] and the ease of addition reactions to them,[4a,c,5] we have performed ab initio calculations in order to resolve this apparent paradox. We have computed the relative energies of the π bonds in acetylene and ethylene and also the relative energies of the species formed by addition of a prototypical electrophile (H^+), radical ($H^\bullet$), and anion (H^-). The computational results reported here also provide information about the relative ease of formation of vinyl and ethyl cations, radicals, and anions from, respectively, ethylene and ethane.

Computational Methodology

Ab initio calculations of absolute bond dissociation energies (BDEs) usually require large basis sets and inclusion of high levels

(1) Streitwieser, A., Jr.; Heathcock, C. H. *Introduction to Organic Chemistry*, 3rd ed.; Macmillan: New York, 1985; pp 288–98.
(2) McMurry, J. *Organic Chemistry*, 2nd ed.; Brooks/Cole: Pacific Grove, CA, 1984; pp 229–240.
(3) Streitwieser, A., Jr. *Molecular Orbital Theory for Organic Chemists*; Wiley: New York, 1961; pp 11–16. Borden, W. T. *Modern Molecular Orbital Theory for Organic Chemists*; Prentice-Hall: Englewood Cliffs, NJ, 1975; pp 9–11. Salem, L. *Electrons in Chemical Reactions*; Wiley: New York, 1982; p 19 and refernces therein. Albright, T. A.; Burdett, J. K.; Whangbo, M. H. *Orbital Interactions in Chemistry*; Wiley: New York, 1985; p 8.

(4) (a) Sun, H.; Hrovat, D. A.; Borden, W. T. *J. Am. Chem. Soc.* **1987**, *109*, 5275. (b) Hrovat, D. A.; Sun H.; Borden, W. T. THEOCHEM. **1988**, *163*, 51. (c) Hrovat, D. A.; Borden, W. T. *J. Am. Chem. Soc.* **1988**, *110*, 4710. (d) Wang, S. Y.; Borden, W. T. *J. Am. Chem. Soc.* **1989**, *111*, 7282. (e) Coolidge, M.; Borden, W. T. *J. Am. Chem. Soc.* **1990**, *112*, 1704. (f) Hammons, J. H.; Coolidge, M.; Borden, W. T. *J. Phys. Chem.* **1990**, *94*, 5468. Hammons, J. H.; Hrovat, D. A.; Borden, W. T. *J. Phys. Org. Chem.* **1990**, *3*, 635.
(5) Getty, S. D.; Borden, W. T. *J. Am. Chem. Soc.* **1991**, *113*, 4334.

0002-7863/91/1513-6750$02.50/0