Catalytic Role of Proton Transfers in the Formation of a Tetrahedral Adduct in a Serine Carboxyl Peptidase†

Haobo Guo,‡ Alexander Wlodawer,§ Toru Nakayama,|| Qin Xu,‡ and Hong Guo*,‡

Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, Maryland 21702, and Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 6-6-11 Aoba-yama, Sendai 980-8579, Japan

Received March 8, 2006; Revised Manuscript Received May 8, 2006

ABSTRACT: Quantum mechanical/molecular mechanical molecular dynamics and 2D free energy simulations are performed to study the formation of a tetrahedral adduct by an inhibitor *N*-acetyl-isoleucyl-prolyl-phenylalaninal (AcIPF) in a serine-carboxyl peptidase (kumamolisin-As) and elucidate the role of proton transfers during the nucleophilic attack by the Ser278 catalytic residue. It is shown that although the serine-carboxyl peptidases have a fold resembling that of subtilisin, the proton transfer processes during the nucleophilic attack by the Ser residue are likely to be more complex for these enzymes compared to the case in classical serine proteases. The computer simulations demonstrate that both general base and acid catalysts are required for the formation and stabilization of the tetrahedral adduct. The 2D free energy maps further demonstrate that the proton transfer from Ser278 to Glu78 (the general base catalyst) is synchronous with the nucleophilic attack, whereas the proton transfer from Asp164 (the general acid catalyst) to the inhibitor is not. The dynamics of the protons at the active site in different stages of the nucleophilic attack as well as the motions of the corresponding functional groups are also studied. It is found that the side chain of Glu78 is generally rather flexible, consistent with its possible multifunctional role during catalysis. The effects of proton shuffling from Asp82 to Glu78 and from Glu32 to Asp82 are examined, and the results indicate that such proton shuffling may not play an important role in the stabilization of the tetrahedral intermediate analogue.

General acid−base transition-state stabilization (*1*−*5*) is believed to be one of the most important strategies that enzymes use to catalyze chemical reactions. Although the existence of general acid−base catalysis has been known for a long time, a detailed understanding of the proton transfers associated with catalysis and the motions of the corresponding functional groups is still lacking (*4*). In many of the cases, the activation barrier for a primary reaction of bond breaking and making between nonhydrogen atoms is found to be lowered through one or more secondary proton-transfer processes, leading to rate acceleration (*1*−*5*). This strategy seems to be used by a variety of enzymes including sedolisins (serine-carboxyl peptidases), a recently characterized family of proteolytic enzymes.

Sedolisins have a fold resembling that of subtilisin and a maximal activity at low pH (*6*−*16*). The defining features of this family are a unique catalytic triad, Ser-Glu-Asp (Ser278-Glu78-Asp82 for kumamolisin-As; see Figure 1), as well as the presence of an aspartic acid residue (Asp164 for kumamolisin-As) that is an equivalent of Asn155 of subtilisin, a residue that creates the oxyanion hole in the classical serine protease. Previous experimental and theoretical studies (*17*−*23*) showed that the serine residue is the catalytic nucleophile, whereas the nearby Glu in the catalytic triad (Glu78 for kumamolisin-As) is likely to act as the general base to accept the proton from Ser and assist in the nucleophilic attack. One prediction from computer simulations (*23*) that is of considerable interest is that sedolisins may use the Asp residue (i.e., the residue that replaces Asn155 of subtilisin) as a general acid catalyst to stabilize the tetrahedral intermediate, leading to the rate acceleration for the enzyme-catalyzed reaction. This mechanism is in contrast to the case of serine proteases for which the oxyanion-hole electrostatic interaction is used for stabilizing the tetrahedral intermediate during catalysis. Thus, the first step of the sedolisin-catalyzed reactions, the nucleophilic attack by the Ser residue on the substrates, seems to involve both the general base and acid catalysts. Nevertheless, the exact role of these two proton transfers during the formation of the tetrahedral intermediate and adduct is still not clear.

Some structural differences at the active sites for different members of the sedolisin family have been reported. For kumamolisin and kumamolisin-As, the catalytic triad is linked to another Glu residue (Glu32), which in turn forms a hydrogen bond to Trp129 (Figure 1A) (*20*−*23*). In contrast,

† This work was supported in part by the Center of Excellence for Structural Biology and UT-ORNL Science Alliance, the University of Tennessee, the Petroleum Research Fund (to H.G.), and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research (to A.W.).

* To whom correspondence should be addressed. Phone: (865) 974-3610. Fax: (865) 974-6306. E-mail: hguo1@utk.edu.

‡ University of Tennessee.

§ National Cancer Institute at Frederick.

|| Tohoku University.

10.1021/bi060461i CCC: $33.50
Published on Web 07/07/2006

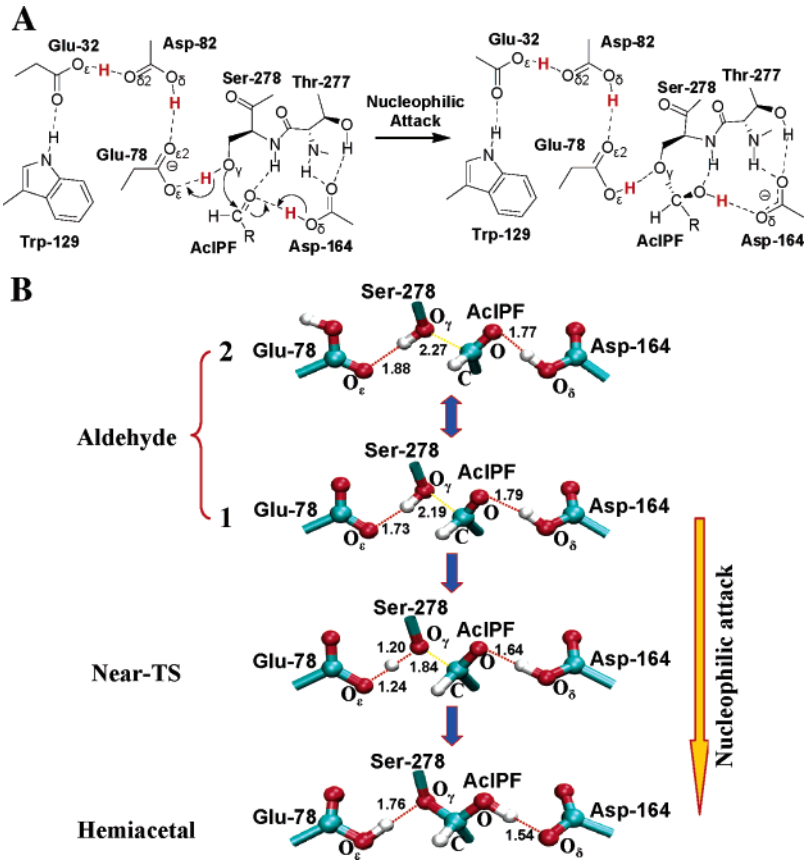

FIGURE 1: (A) Schematic diagram of the process of nucleophilic attack by Ser278 during the formation of the tetrahedral adduct (hemiacetal) complex from the aldehyde complex in the active site of kumamolisin-As. The four protons that are the subject of the present study are colored in red. For the aldehyde complex, the proton between Glu78 and Asp82 was found to be either located on Asp82 or Glu78 during the QM/MM MD simulations (see Figure 4A and B), whereas the proton between Glu32 and Asp82 was either located on Glu32 or involved in a low-barrier hydrogen bond (Figure 4B). Only the configuration of the aldehyde complex from which the nucleophilic attack can start is given here. For the other structures, see Figures 1B and 4B. (B) Average active-site structures for the atoms involved in nucleophilic attack in the aldehyde and hemiacetal complexes, respectively, obtained from the QM/MM simulations. The average structure for the system near the TS from the rare event simulations (Figure 5B) is also given. The key proton-transfer processes between Ser278 and Glu78 and between Asp164 and the ligand during the nucleophilic attack are shown. There are two possible protonation states for Glu78 for the aldehyde complex, and the average structures for both protonation states are given (designated as 1 and 2; also see the caption of Figure 1A). The structure for each of the protonation states was found to be well populated during the MD simulations with a free energy difference of only 0.5 kcal/mol (Figure 4B). Therefore, the change from 2 to 1 is not considered a separate step of the enzyme-catalyzed reaction. The nucleophilic attack by Ser278 starts from 1, for which Glu78 is unprotonated and able to accept the proton from Ser278 (*23*).

these additional interactions are absent in sedolisin (PSCP), another member of the sedolisin family for which the X-ray structures have also been solved (*17*−*19*). The exact role of these two additional residues in the catalytic mechanism of kumamolisin and kumamolisin-As is still not clear, although it was suggested that they might facilitate proton delocalization during the nucleophilic attack by the serine residue (*21*).

In this article, we examine the proton transfers and the motions of the functional groups in kumamolisin-As during the formation of the tetrahedral adduct (hemiacetal) by an inhibitor *N*-acetyl-isoleucyl-prolyl-phenylalaninal (AcIPF[1]) (Figure 1). It is expected that the proton transfers and interactions that stabilize the tetrahedral adduct may also be used by the enzyme to stabilize the tetrahedral intermediate during the enzyme-catalyzed reaction. We demonstrate that the proton transfer processes in this recently characterized family of enzymes are likely to be more complex compared to the case of serine proteases during the nucleophilic attack by the Ser residue (Ser278). Indeed, both general base and acid catalysts are required for the stabilization of the tetrahedral adduct and, presumably, the tetrahedral intermediate as well during the enzyme-catalyzed reaction. The 2D quantum mechanical/molecular mechanical (QM/MM) free energy simulations further show that the proton transfer from Ser278 to Glu78 is synchronous with the nucleophilic attack in the sense that the degree of completion of the proton transfer from Ser278 to Glu78 is similar to that of the formation of the hemiacetal complex. Therefore, for the structures near the transition state of the nucleophilic attack, the average position of the proton between Ser278 and Glu78 is close to the middle of these two residues. In contrast, the proton transfer from Asp164 to the ligand has not really started as the system approaches the transition state; it occurs mainly at a later stage of the nucleophilic attack. The side chain of Glu78 is found to be rather flexible at different stages of the nucleophilic attack (e.g., in the aldehyde and

[1] Abbreviations: QM/MM, quantum mechanical/molecular mechanical; MD, molecular dynamics; AcIPF, *N*-acetyl-isoleucyl-prolyl-phenylalaninal; PMF, potential of mean force; SCC-DFTB, self-consistent-charge density-functional tight-binding; TS, transition state; WHAM, weighted histogram analysis method.

tetrahedral adduct complexes). This flexibility seems to be consistent with a possible multifunctional role of Glu78, for example, it may act as a general base to accept a proton from Ser278 and as a general acid to protonate the tetrahedral intermediate during catalysis in a way that is similar to the behavior of the corresponding residue His57 in the serine proteases. The role of proton shuffling from Asp82 to Glu78 and from Glu32 to Asp82 is also examined, and it is shown that such proton shuffling may not have a significant energetic effect on the first step of the reaction.

METHODS

A fast semiempirical density-functional method (self-consistent-charge density-functional tight-binding or simply SCC-DFTB) (*24*) implemented in the CHARMM32b program (*25*) was used for the QM/MM molecular dynamics (MD) and free energy (potential of mean force or PMF) simulations. The initial coordinates for the simulations were obtained from the crystal structure (pdb ID: 1SIO, resolution 1.8 Å) of kumamolisin-As complexed with the inhibitor AcIPF (*N*-acetyl-isoleucyl-prolyl-phenylalaninal) that is covalently bound to Ser278, the catalytic nucleophile (*20*). The crystals for the sedolisin−tetrahedral adduct (hemiacetal) complexes were normally grown at acidic conditions (e.g., pH ~3−4) (*18*−*22*). The stochastic boundary MD method (*26*) was used for the QM(SCC-DFTB)/MM simulations. The system was separated into a reaction zone and a reservoir region; the reaction zone was further divided into the reaction region and the buffer region. The reaction region was a sphere with radius R of 16 Å, and the buffer region had R equal to 16 Å $\leq R \leq$ 18 Å. The reference center for partitioning the system was chosen to be the O_γ atom of Ser278 (Figure 1). The inhibitor AcIPF and the side chains of Glu32, Asp82, Glu78, Ser278, and Asp164 were treated by QM (a total of 92 QM atoms) and the rest of the system by MM. The all-hydrogen potential function (PARAM22) (*27*) was used for MM atoms. The link-atom approach (*28*) available in the CHARMM program (*25*) was used to separate the QM and MM regions. A modified TIP3P water model (*29*, *30*) was employed for the solvent, and explicit water spheres with a radius of 18 Å centered on the O_γ atom of Ser278 were used to solvate the system using the standard procedure in the CHARMM program. The resulting system contains 3453 atoms (2700 enzyme atoms and 251 water molecules, including 98 crystal water molecules).

It has been experimentally observed that an increase in pH would cause the breaking of the covalent bond between Ser278 and the inhibitor in the sedolisin−hemiacetal complexes, leading to the formation of the original enzyme−aldehyde complex. This observation is consistent with the fact that sedolisins have a maximal activity at low pH (*6*−*16*) because the loss of the ability to stabilize the tetrahedral intermediate may reduce activity. It also indicates that the tetrahedral adduct is only stable in acidic conditions. To generate the models for the active form of the enzyme (i.e., at low pH), the models were manually modified by adding protons on certain Asp and Glu residues on the basis of the information of the local environments of these residues. The initial structure for the entire stochastic boundary system was optimized by an adopted basis Newton−Rhaphson (ABNR) method. The system was gradually heated from 50 to 310 K in 30 ps and equilibrated at 310 K for 40 ps. A 1-fs time step was used for the integration of the equations of motion, and the coordinates were saved every 10 fs for analyses. The LD frictional constants were 250 ps^{-1} for the protein atoms and 62 ps^{-1} for the water molecules. The bonds involving hydrogen atoms in the MM region were fixed by the SHAKE algorithm (*31*). Then, 1.2 ns QM/MM MD simulations were performed for the aldehyde and tetrahedral adduct complexes. The average structure from the simulations for the tetrahedral adduct complex was then used to test whether the experimental structure could be well reflected from the computer model and simulations. It was found that the average structure is rather close to the experimental structure obtained at low pH (*23*), suggesting that the model used in this study should be reasonable. For instance, the distances between the nonhydrogen atoms for the key hydrogen bonds in the active site are very close to those in the experimental structure with an average deviation of only about 0.1 Å (see Figure 1 of ref 23 (*23*)). Certain parts of the average active-site structures for the enzyme complexes obtained from the MD and free energy simulations (see below) are given in Figure 1B.

The umbrella sampling method (*32*) implemented in the CHARMM program along with the weighted histogram analysis method (WHAM) (*33*, *34*) was applied to determine the changes of the free energy (PMF) for the nucleophilic attack and proton transfers. The 1D free energy curves for the wild-type enzyme with the proton fixed on Asp82 or Glu32 using the SHAKE algorithm (*31*) were obtained for the process of the interconversion between the tetrahedral adduct and the aldehyde complexes; the free energy curve for the Asp164Ala mutant was determined as well without the use of the SHAKE algorithm. The distance $\xi = r(C-O_\gamma)$ was used as the reaction coordinate for the 1D free energy simulations. Ten windows along the reaction coordinate were used between the aldehyde ($\xi \approx 2.2$ Å) and the hemiacetal ($\xi \approx 1.5$ Å) complexes. The 1D free energy curves obtained with the SHAKE algorithm for the wild-type enzyme were compared with the earlier results (*23*) without the use of the SHAKE algorithm, and the effects of the proton shuffling from Asp82 to Glu78 and from Glu32 to Asp82 on the free energy profile are examined. The 2D free energy maps were also determined by umbrella sampling (*32*). In addition to $\xi = r(C-O_\gamma)$, the second reaction coordinate in the 2D free energy simulations is the proton-transfer coordinate between Asp164 and the ligand or between Ser278 and Glu78. The distance differences $\xi_1 = r[H-O(AcIPF)] - r[H-O_\delta(D164)]$ and $\xi_2 = r[H-O_\epsilon(E78)] - r[H-O_\gamma(S278)]$ were used as the reaction coordinates for these two proton-transfer processes, respectively. Here, $r[H-O_\epsilon(E78)]$ and $r[H-O_\gamma(S278)]$ are the distances of the proton to the oxygen atoms of Glu78 (O_ϵ) and Ser278 (O_γ), respectively. A similar definition was formed for $r[H-O(AcIPF)]$ and $r[H-O_\delta(D164)]$. The 2D free energy maps were obtained by the WHAM algorithm (*34*), with a bin size of 0.05 × 0.05 $Å^2$. Fifty windows were used in the 2D free energy calculations with force constants in the range of 100−800 kcal·mol^{-1}·$Å^{-2}$ for ξ and 50−300 kcal·mol^{-1}·$Å^{-2}$ for ξ_1 and ξ_2, respectively. For the 1D free energy simulations, the force constants are in the range of 100−800 kcal·mol^{-1}·$Å^{-2}$.

The fluctuations of some important distances related to the proton motions for the aldehyde and hemiacetal com-

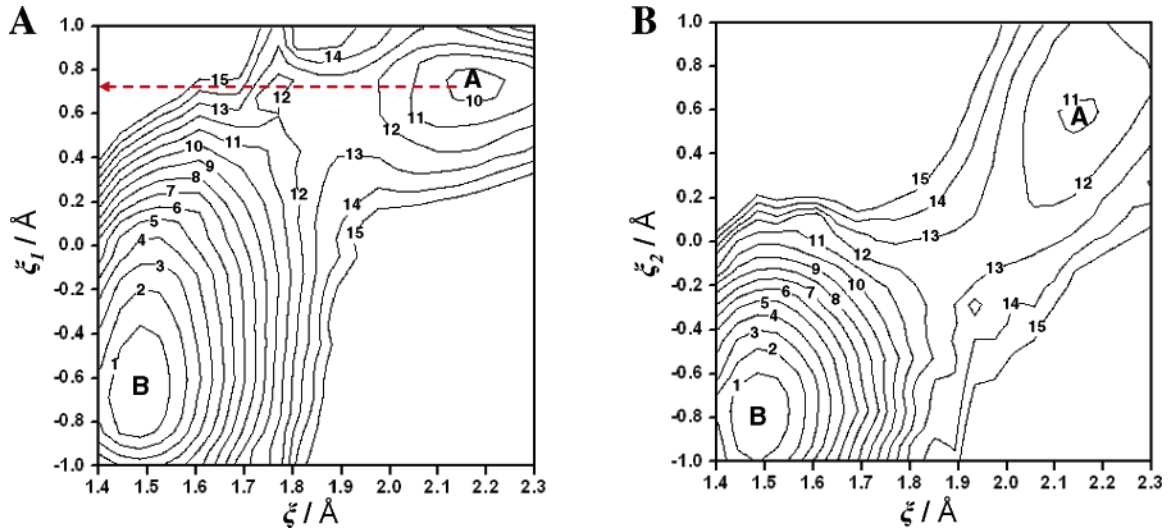

FIGURE 2: Two-dimensional free energy contours for the nucleophilic attack and the proton transfer between AcIPF and Asp164 or between Glu78 and Ser278 (in kcal/mol). Point A designates the aldehyde complex and point B the hemiacetal complex. The hemiacetal complex is ~10 kcal/mol more stable than the aldehyde complex from both free energy simulations. (A) Free energy contour for the nucleophilic attack (reaction coordinate: $\xi = r(C-O_\gamma)$) and the proton transfer between O_δ of Asp164 and O of AcIPF (reaction coordinate: $\xi_1 = r[H-O(AcIPF)] - r[H-O_\delta(D164)]$). The dashed arrow indicates a hypothetical reaction pathway in which the general acid catalysis by Asp164 is not involved during the nucleophilic attack, even though Asp164 could still provide electrostatic stabilization of the tetrahedral adduct. (B) Free energy contour for the nucleophilic attack and the proton transfer between O_γ of Ser278 and O_ϵ of Glu78 (reaction coordinate: $\xi_2 = r[H-O_\epsilon(E78)] - r[H-O_\gamma(S278)]$).

plexes as well as in the structures near the transition state (TS) for the nucleophilic attack (i.e., $r(C-O_\gamma)$ ~1.84 Å) were monitored. For the distance fluctuations in the aldehyde and hemiacetal complexes, the trajectories from the 1.2 ns simulations were used; the initial model for the aldehyde complex was produced from the PMF simulations (see above). For the fluctuations in the structures near the TS, the distances were obtained from a trajectory for which a harmonic constraint with a force constant of 1200 kcal·mol^{-1}·Å^{-2} was used on $r(C-O_\gamma)$ at 1.84 Å to force the system to sample the structures in the conformational space near the TS. The average structure from the simulations is also given in Figure 1B (Near-TS). In addition to the protons locations between Asp164 and the ligand and between Ser278 and Glu78, the positions of the protons between Asp82 and Glu78 and between Glu32 and Asp82 were also monitored. The 2D free energy contour for the proton transfers between Asp82 and Glu78 and between Glu32 and Asp82 in the aldehyde complex was generated from the probability density $\rho(\xi_3,\xi_4)$ based on the 1.2 ns MD simulations and the following formula (*41*−*42*):

$$W(\xi_3,\xi_4) = -k_B T \ln \rho(\xi_3,\xi_4).$$

Here k_B is the Boltzmann's constant, and T is the temperature (310 K). The coordinates used to monitor the proton positions are $\xi_3 = r[H-O_\delta(D82)] - r[H-O_\epsilon 2(E78)]$, and $\xi_4 = r[H-O_\epsilon(E32)] - r[H-O_\delta 2(D82)]$.

RESULTS AND DISCUSSION

Nucleophilic Attack by Ser278 and General Acid−*Base Catalysis.* The 2D free-energy contour maps for the nucleophilic attack and one of the key proton-transfer processes involved in the general acid−base catalysis are given in Figure 2. Figure 2A shows the free energy change along $\xi = r(C-O_\gamma)$ (the reaction coordinate for the nucleophilic attack) and $\xi_1 = r[H-O(AcIPF)] - r[H-O_\delta(D164)]$,

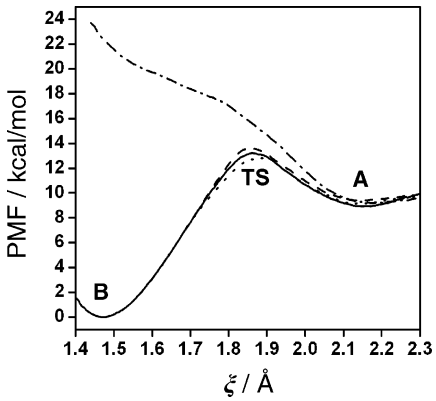

FIGURE 3: One-dimensional free energy curves as functions of $\xi = r(C-O_\gamma)$. Solid line, the curve for the wild-type enzyme obtained from ref 23; dot−dashed line, the Asp164Ala mutant; dashed line, the wild-type enzyme with the proton fixed on O_ϵ of Glu32 by using the SHAKE algorithm (Figure 1A); dotted line, the wild-type enzyme with the proton fixed on O_δ of Asp82 by using the SHAKE algorithm.

whereas Figure 2B gives the free energy map for the system changing along $\xi = r(C-O_\gamma)$ and $\xi_2 = r[H-O_\epsilon(E78)] - r[H-O_\gamma(S278)]$. ξ_1 and ξ_2 are the reaction coordinates for the two proton-transfer processes between Asp164 and AcIPF (involved in the general acid catalysis (*23*)) and between Ser278 and Glu78 (involved in the general base catalysis (*16*−*23*)), respectively. Therefore, these free energy maps allow us to determine the relationship between the energetics of the nucleophilic attack and each of the proton-transfer processes. Figure 2 shows that on the basis of 2D free energy simulations the hemiacetal complex (Point B in Figure 2A and B) is ~10 kcal/mol more stable than the aldehyde complex (Point A), consistent with the earlier results from the 1D free energy calculations (*23*) (also see Figure 3). The relative stabilities for the aldehyde (A) and hemiacetal (B) complexes were found to be similar from two separate 2D free energy simulations (Figure 2A and B). This is to be

expected because the two free energy simulations examine the same process of the nucleophilic attack in going from A to B along with the corresponding proton transfers. The only difference is that Figure 2A monitors the free energy change along ξ_1 in addition to that along ξ, whereas Figure 2B monitors the free energy change along ξ_2 and ξ.

Figure 2A demonstrates that the proton transfer from Asp164 to AcIPF is not synchronous with the nucleophilic attack process in the sense that the degree of completion of the proton transfer from Asp164 to AcIPF is different from that of the formation of the tetrahedral adduct complex. Indeed, the proton transfer significantly lags behind the nucleophilic attack, and there is little progress of the proton transfer before the substrate complex reaches $\xi \sim 1.84$ Å (i.e., the area near the TS) along the free energy valley of Figure 2A. The proton transfer mainly occurs at a later stage of the nucleophilic attack. This proton transfer from Asp164 to AcIPF is expected to play a very important role in stabilizing the tetrahedral adduct and intermediate. For instance, it has been demonstrated from 1D QM/MM free energy simulations in an earlier study (*23*) that the hemiacetal complex became considerably less stable than the aldehyde complex when the proton transfer from Asp164 to AcIPF was prevented by the MM treatment of Asp164. This may also be seen by using the 2D free energy map of Figure 2A. If we follow the hypothetical reaction pathway indicated by the dashed arrow (in which proton transfer does not occur and Asp164 does not act as a general acid), the hemiacetal complex becomes as much as 5 kcal/mol less stable than the aldehyde complex, consistent with the earlier results obtained with the MM treatment of Asp164 (*23*).

In contrast to the effect of the proton transfer between Asp164 to AcIPF, which is not synchronized with the primary reaction, the proton transfer from Ser278 to Glu78 was found to be synchronous with the nucleophilic attack, as demonstrated in Figure 2B. Indeed, the free energy valley in Figure 2B is essentially along the diagonal of the map, and in the area near the TS, the average position of the proton between Ser278 and Glu78 is close to the middle of the two residues (i.e., ξ_2 is close to the zero), leading to the formation of a low-barrier hydrogen bond between them. As can be seen from Figure 2B, this synchronized reaction pathway is likely to make the nucleophilic attack more efficient than the nonsynchronized processes. The energetic effect of the formation of the low-barrier hydrogen bond in the lowering of the activation barrier (*37*−*39*) has to be investigated further in a separate study because it is necessary to separate this effect from the other effects contained in the free energy simulations.

Figure 3 displays some 1D free energy curves as functions of $\xi = r(C-O_\gamma)$. One of the important differences at the active sites for different members of the sedolisin family is that for kumamolisin and kumamolisin-As the catalytic triad is linked to another Glu residue (Glu32) that in turn forms a hydrogen bond to Trp129 (Figure 1A) (*20*−*23*). These additional interactions are absent in some other members of the sedolisin family, including sedolisin (*17*−*19*). The exact role of these two additional residues in the catalytic mechanism of kumamolisin and kumamolisin-As is still not clear. One possible role for these two additional residues is to facilitate proton delocalization during the nucleophilic attack by the serine residue. Therefore, it would be of interest to examine whether preventing proton shuffling from Asp82 to Glu78 or from Glu32 to Asp82 would lead to a different free energy profile for nucleophilic attack. Figure 3 shows that the free energy profiles for the cases in which the proton was fixed on O_ϵ of Glu32 or on O_δ of Asp82 using the SHAKE algorithm are essentially the same as the profile without the use of the SHAKE algorithm (*23*). Thus, the results suggest that proton shuffling from Asp82 to Glu78 or from Glu32 to Asp82 may not have a significant effect on the nucleophilic attack process. Figure 3 also shows that the mutation of Asp164 to Ala makes the tetrahedral adduct considerably less stable, suggesting that the tetrahedral adduct may not be able to form in the absence of the carboxyl group from Asp164. This theoretical prediction may explain why our attempts to grow the crystals and determine the experimental structures of the tetrahedral adduct complexes for D164N and D164A were not successful. The mutation of D164 is also expected to significantly reduce activity. For instance, it has been observed in an experimental study that the Asp164Ala mutant of kumamolisin did not have any measurable proteolytic activity (*21*).

Motions of Protons and Functional Groups. The positions of the four protons (colored in red in Figure 1A) in the aldehyde complex during the QM/MM MD simulations were monitored in Figure 4A (two protons between Glu78 and Ser278 and between AcIPF and Asp164, respectively) and B (two protons between Glu32 and Asp82 and between Asp82 and Glu78, respectively), whereas the positions of the protons in the tetrahedral adduct (hemiacetal) complex were monitored in Figure 4C and D. Every point in each figure represents one structure of the complex that was saved every 10 fs during the 1.2 ns MD simulations (i.e., a total of 120 000 points in each figure representing 120 000 structures obtained from the simulations). The presence of broad distributions for the protons in these figures is mainly due to the fact that the hydrogen bond distances undergo fluctuations during the QM/MM MD simulations. The existence of some relatively large deviations along $\xi_2 = r[H-O_\epsilon(E78)] - r[H-O_\gamma(S278)]$ in Figure 4A (to relatively large values in the positive direction) is due to the flexibility of the Glu78 side chain, which underwent some rotations during the MD simulations and, therefore, increased the distance differences (Figure 5). A similar argument can be made for the deviations along ξ_2 in Figure 4C (where the deviations are in the negative direction because the proton is on Glu78).

Figure 4B shows that the aldehyde complex can adopt structures with two of the protons (i.e., those between Asp82 and Glu78 and between Glu32 and Asp82) located at different positions. The structures around the two areas (1 and 2) are all well populated, suggesting that they should have similar stabilities; the structure with a low-barrier hydrogen bond (LBHB, the area 2 in Figure 4B) is somehow more populated and more stable. Indeed, the free energy calculations show that the free energy difference for the two different configurations is only 0.5 kcal/mol. Therefore, the transition between 1 and 2 is not considered as a separate step of the enzyme-catalyzed reaction in this study. The existence of a LBHB in the substrate analogue complex observed here is in contrast to the case of serine proteases, where LBHBs were found in transition state analogue complexes (*37*−*39*). It has been a subject of debate as to whether the LBHBs in serine proteases might play an

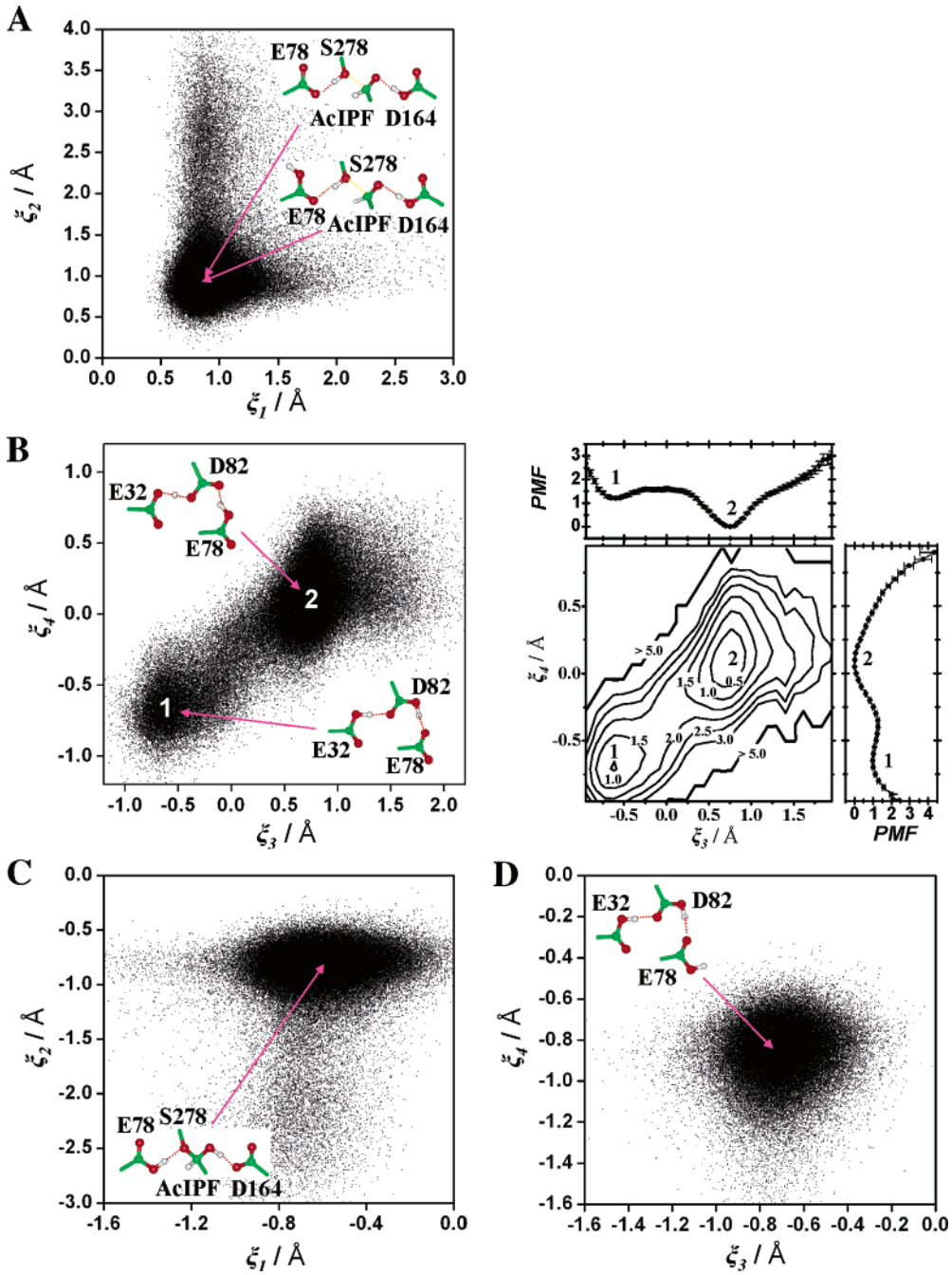

FIGURE 4: Proton positions in the active site during 1.2 ns QM/MM MD simulations for the aldehyde and hemiacetal complexes. The distance difference (in Å) from the proton to the two oxygen atoms involved in the corresponding interactions is used to monitor the proton position between each pair of the residues and ligand. For instance, $\xi_1 = r[H-O(AcIPF)] - r[H-O_\delta(D164)]$ is the difference for the distances of the proton to the O of AcIPF and the O_δ of Asp164, respectively (Figure 1B). (A) Proton positions between AcIPF (O) and Asp164 (O_δ) and between Glu78 (O_ϵ) and Ser278 (O_γ) in the aldehyde complex. The coordinates used to monitor the proton positions during the simulations are $\xi_1 = r[H-O(AcIPF)] - r[H-O_\delta(D164)]$ and $\xi_2 = r[H-O_\epsilon(E78)] - r[H-O_\gamma(S278)]$. (B) Left panel: proton positions between Asp82 (O_δ) and Glu78 ($O_\epsilon 2$) and between Glu32 (O_ϵ) and Asp82 ($O_\delta 2$) in the aldehyde complex. The coordinates used to monitor the proton positions are $\xi_3 = r[H-O_\delta(D82)] - r[H-O_\epsilon 2(E78)]$ and $\xi_4 = r[H-O_\epsilon(E32)] - r[H-O_\delta 2(D82)]$. Two different distributions were observed (shown as areas 1 and 2 in the Figure). One of them contains a low-barrier hydrogen bond between Glu32 and Asp82 with the distributions centered at $\xi_4 \sim 0$ Å (i.e., area 2 in the plot), suggesting that the proton spent about equal time on each of the two oxygen atoms. Right: free energy contour generated from the structural distributions from the MD simulations (i.e., the distributions shown on the left). The projections of the free energy along ξ_3 and ξ_4 are also given with the error bars estimated based on ref 43 (*43*). (C) The proton positions between AcIPF and Asp164 and between Glu78 and Ser278 in the hemiacetal complex during the simulations. The negative ξ_1 and ξ_2 values for the points suggest that the two protons are located on AcIPF and Glu78, respectively, after nucleophilic attack (also see Figure 1B). (D) Protons positions between Asp82 and Glu78 and between Glu32 and Asp82 in the hemiacetal complex. The negative ξ_3 and ξ_4 values for the points suggest that the two protons are located on Asp82 and Glu32, respectively, consistent with the schematic Figure of the hemiacetal complex plotted in Figure 1A (right).

important role in transition state (TS) stabilization (*37*−*39*). Although a large stabilizing effect for the aldehyde complex was not observed as a result of the LBHB formation in kumamolisin-As, care must be exercised in the interpretation of the results because the formation of the LBHB is not an isolated event, and there are other structural changes in the

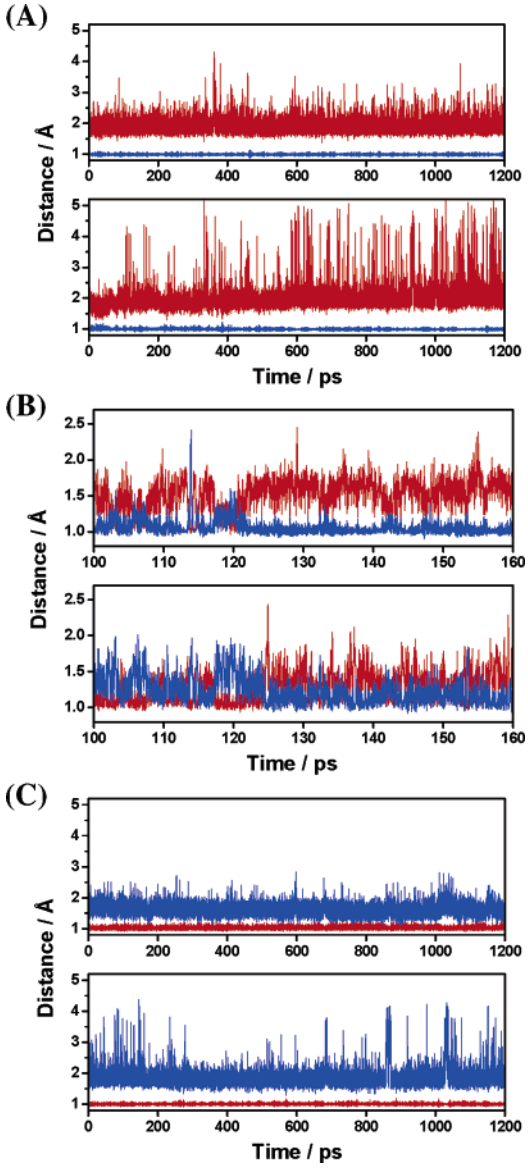

FIGURE 5: Fluctuations of some important distances related to proton motions in the aldehyde, near-TS, and hemiacetal complexes obtained from the trajectories of the QM/MM MD simulations. See Figure 1A for atom designations. The distances are in Å. (A) Aldehyde complex from the 1.2 ns MD simulations. Top panel: the fluctuations of the H−O_δ(D164) and H−O(AcIPF) distances as functions of time. Red, *r*[H−O(AcIPF)]; blue, *r*[H−O_δ(D164)]. Bottom panel: the fluctuations of the H−O_γ(S278) and H−O_ϵ-(E78) distances as functions of time. Red, *r*[H−O_ϵ(E78)]; blue, *r*[H−O_γ(S278)]. (B) Complex obtained from a rare event MD simulation in the area near the transition state (near-TS). A harmonic constraint with a force constant of 1200 kcal/mol/Å^2 was added at *r*(C−O_γ) ∼1.84 Å to force the system to sample the structures in the conformational space near the TS of the nucleophilic attack. These distances were obtained from the trajectory of a 60-ps productive run after a 100-ps equilibration. Top panel: the fluctuations of the H−O_δ(D164) and H−O(AcIPF) distances as functions of time. Red, *r*[H−O(AcIPF)]; blue, *r*[H−O_δ(D164)]. Bottom panel: the fluctuations of the H−O_ϵ(E78) and H−O_γ(S278) distances as functions of time. Red, *r*[H−O_ϵ(E78)]; blue, *r*[H−O_γ-(S278)]. The average distances of the proton to Glu78 and Ser278 are similar, indicating that a low-barrier hydrogen bond may be formed at the transition state. (C) Hemiacetal complex from the 1.2 ns MD simulations. Top panel: the fluctuations of the H−O_δ-(D164) and H−O(AcIPF) distances as functions of time. Red, *r*[H−O(AcIPF)]; blue, *r*[H−O_δ(D164)]. Bottom panel: the fluctuations of the H−O_γ(S278) and H−O_ϵ(E78) distances as functions of time. Red, *r*[H−O_ϵ(E78)]; blue, *r*[H−O_γ(S278)].

system. Figure 4A shows that both ξ_1 and ξ_2 values for the protons are positive, suggesting that the corresponding protons are located on Asp164 and Ser278, respectively (i.e., *r*[H−O_δ(D164)] and *r*[H−O_γ(S278)] are smaller than *r*[H−O(AcIPF)] and *r*[H−O_ϵ(E78)], respectively). This interpretation is consistent with the average structure shown in Figure 1B before the nucleophilic attack. Figure 4C and D show that the values of ξ_1, ξ_2, ξ_3 and ξ_4 for the protons are all negative in the hemiacetal complex, indicating that Glu32 (O_ϵ), Asp82 (O_δ), Glu78 (O_ϵ), and the ligand are protonated, respectively, after the nucleophilic attack.

The fluctuations of the distances related to the proton motions between D164 and AcIPF in the aldehyde, near-TS, and hemiacetal complexes are plotted as functions of time in the top panels of Figure 5A, B, and C, respectively. These Figures demonstrate how the H−O_δ(D164) and H−O(AcIPF) distances (the blue and red lines, respectively) change as the system changes from the aldehyde (Figure 5A) to the hemiacetal complex (Figure 5C). The top panel of Figure 5B shows that as the system proceeds to the area near the TS from the aldehyde complex, the proton remains essentially on Asp164 (blue line), consistent with the 2D free energy map in Figure 2A. However, there are transient formations of the low-barrier hydrogen bond between Asp164 and the ligand (e.g., at ∼120 ps) or even the H−O(AcIPF) covalent bond. The fluctuation of the H−O_δ-(D164) distance here is larger than that in the aldehyde complex (blue line in the top panel of Figure 5A), indicating that the covalent bond has been weakened as the system approaches the TS; as expected, the average distance for the corresponding hydrogen bond (red line) is also shorter than that in the aldehyde complex. In the hemiacetal complex (top panel of Figure 5C), the proton has transferred from Asp164 to the ligand. It is of interest to note that the fluctuation of the H−O(AcIPF) distance seems to be larger than that of other O−H covalent bonds in the aldehyde and hemiacetal complexes, suggesting that this covalent bond is probably weaker. Consistent with this suggestion, the hydrogen bond between the ligand and Asp164 seems to be rather strong in the hemiacetal complex.

The bottom panels of Figure 5A, B, and C show the fluctuations of the distances related to the proton motions between S278 and E78 in the aldehyde, near-TS, and hemiacetal complexes, respectively. It is interesting to see that the side chain of Glu78 seems to be rather flexible in both aldehyde and hemiacetal complexes, as indicated by the large fluctuations of the corresponding distances (the red line in the bottom panel of Figure 5A and the blue line in the bottom panel of Figure 5C). A flexibility for the equivalent residue (His57) in serine proteases has also been suggested (*40*). Because Glu78 may play a similar multi-functional role as His57 does in serine proteases (e.g., acting as a general base to accept a proton from Ser278 and a general acid to protonate the tetrahedral intermediate in producing the leaving group and acyl-enzyme), the flexibility of its side chain might be of importance. The bottom panel of Figure 5B shows that the distances of the proton to Ser278 and Glu78 are similar and overlap with each other, leading to the formation of a low-barrier hydrogen bond between these two residues near the transition state.

CONCLUSIONS

It has been shown from the QM/MM MD and free energy simulations that the proton-transfer processes during the nucleophilic attack by the Ser residue (Ser278) are likely to be more complex for this recently characterized family of enzymes compared to the case of classical serine proteases. It has been demonstrated that both general base and acid catalysts are required for the formation and stabilization of the tetrahedral adduct. Although this study examined the general acid−base mechanism in the formation of the tetrahedral adduct, the same mechanism may also be used by the enzyme to stabilize the tetrahedral intermediate during the enzyme-catalyzed reaction. The computer simulations further showed that the proton transfer from Ser278 to Glu78 (the general base catalyst) is synchronous with nucleophilic attack, whereas the proton transfer from Asp164 (the general acid catalyst) to the inhibitor is not. The dynamics of the protons at different stages of nucleophilic attack in the active site and the motions of the related functional groups were also studied. It was found that the side chain of Glu78 is generally rather flexible, supporting its multifunctional role during catalysis as proposed for His57 in serine proteases. The effects of proton shuffling from Asp82 to Glu78 and from Glu32 to Asp82 are examined, and the results indicate that such proton shuffling may not play an important role in the stabilization of the tetrahedral intermediate analogue.

ACKNOWLEDGMENT

We thank Professor Martin Karplus for the gift of the CHARMM program.

REFERENCES

1. Fersht, A. (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*, 2nd ed., W. H. Freeman and Company, New York.
2. Jencks, W. P. (1972) General acid−base catalysis of complex reactions in water, *Chem. Rev. 72*, 705−718.
3. Kirby, A. J. (1997) Efficiency of proton-transfer catalysis in models and enzymes, *Acc. Chem. Res. 30*, 290−296.
4. Schowen, K. B., Limbach, H.-H., Denisov, G. S., and Schowen, R. L. (2000) Hydrogen bonds and proton transfer in general-catalytic transition-state stabilization in enzyme catalysis, *Biochim. Biophys. Acta 1458*, 43−62.
5. Bernasconi, C. F. (1992) The principle of non-perfect synchronization, *Adv. Phys. Org. Chem. 27*, 119−238.
6. Oda, K., Sugitani, M., Fukuhara, K., and Murao, S. (1987) Purification and properties of a pepstatin-insensitive carboxyl proteinase from a gram-negative bacterium, *Biophys. Acta 923*, 463−469.
7. Oda, K., Takahashi, T., Tokuda, Y., Shibano, Y., and Takahashi, S. (1994) Cloning, nucleotide sequence, and expression of an isovaleryl pepstatin-insensitive carboxyl proteinase gene from *Pseudomonas* sp. 101, *J. Biol.* Chem. *269*, 26518−26524.
8. Oyama, H., Abe, S., Ushiyama, S., Takahashi, S., and Oda, K. (1999) Identification of catalytic residues of pepstatin-insensitive carboxyl proteinases from prokaryotes by site-directed mutagenesis, *J. Biol. Chem. 274*, 27815−27822
9. Oda, K., Ogasawara, S., Oyama, H., and Dunn, B. M. (2000) Subsite preferences of pepstatin-insensitive carboxyl proteinases from prokaryotes: kumamolysin, a thermostable pepstatin-insensitive carboxyl proteinase, *J. Biochem. (Tokyo) 128*, 499−507.
10. Tsuruoka, N., Nakayama, T., Ashida, M., Hemmi, H., Nakao, M., Minakata, H., Oyama, H., Oda, K., and Nishino, T. (2003) Collagenolytic serine-carboxyl proteinase form *Alicyclobacillus sendaiensis* strain NTAP-1: purification, characterization, gene cloning, and heterologous expression, *Appl. Environ. Microbiol. 69*, 162−169.
11. Lee, B. R., Furukawa, M., Yamashita, K., Kanasugi, Y., Kawabata, C., Hirano, K., Ando, K., and Ichishima, E. (2003) Aorsin, a novel serine proteinase with trypsin-like specificity at acidic pH, *Biochem. J. 371*, 541−548.
12. Nishii, W., Kuriyama, H., and Takahashi, K. (2003) The *Physarum polycephalum php* gene encodes a unique cold-adapted serine-carboxyl peptidase, physarolisin II, *FEBS Lett. 246*, 340−344.
13. Wujek, P., Kida, E., Walus, M., Wisniewski, K. E., and Golabek, A. A. (2004) *N*-glycosylation is crucial for folding, trafficking, and stability of human tripeptidyl-peptidase I, *J. Biol. Chem. 279*, 12827−12839.
14. Golabek, A. A., Wujek, P., Walus, M., Bieler, S., Soto, C., Wisniewski, K. E., and Kida, E. (2004) Maturation of human tripeptidyl-peptidase I *in vitro*, *J. Biol. Chem. 279*, 31058−31067.
15. Walus, M., Kida, E., Wisniewski, K. E., and Golabek, A. A. (2005) Ser475, Glu272, Asp276, Asp327, and Asp360 are involved in catalytic activity of human tripeptidyl-peptidase I, *FEBS Lett. 579*, 1383−1388.
16. Wlodawer, A., Li, M., Gustchina, A., Oyama, H., Dunn, B. M., and Oda, K. (2003) Structure and enzymatic properties of the sedolisin family of serine-carboxyl peptidases, *Acta Biochim. Pol. 50*, 81−102.
17. Wlodawer, A., Li, M., Gustchina, A., Oyama, H., Oda, K., Beyer, B. B., Clemente, J., and Dunn, B. M. (2004) Two inhibitor molecules bound in the active site of *Pseudomonas* sedolisin: a model for the bi-product complex following cleavage of a peptide substrate, *Biochem. Biophys. Res. Commun. 314*, 638−645.
18. Wlodawer, A., Li, M., Gustchina, A., Uchida, K., Oyama, H., Dunn, B. M., and Oda, K. (2001) Carboxyl proteinase from *Pseudomonas* defines a novel family of subtilisin-like enzymes, *Nat. Struct. Biol. 8*, 442−446.
19. Wlodawer, A., Li, M., Gustchina, A., Dauter, Z., Uchida, K., Oyama, H., Goldfar, N. E., Dunn, B. M., and Oda, K. (2001) Inhibitor complexes of the *Pseudomonas* serine-carboxyl proteinase, *Biochemistry 40*, 15602−15611.
20. Wlodawer, A., Li, M., Gustchina, A., Tsuruoka, N., Ashida, M., Minakata, H., Oyama, H., Oda, K., Nishino, T., and Nakayama, T. (2004) Crystallographic and biochemical investigations of kumamolisin-As, a serine-carboxyl peptidase with collagenase activity, *J. Biol. Chem. 279*, 21500−21510.
21. Commellas-Bigler, M., Fuentess-Prior, P., Maskos, K., Huber, R., Oyama, H., Uchida, K., Dunn, B. M., Oda, K., and Bode, W. (2002) The 1.4 Å crystal structure of kumamolysin: a thermostable serine-carboxyl-type proteinase, *Structure 10*, 865−876.
22. Commellas-Bigler, M., Maskos, K., Huber, R., Oyama, H., Oda, K., and Bode, W. (2004) 1.2 Å crystal structure of the serine carboxyl proteinase pro-kumamolisin: structure of an intact pro-subtilase, *Structure 12*, 1313−1323.
23. Guo, H. B., Wlodawer, A., and Guo, H. (2005) A general acid−base mechanism for the stabilization of a tetrahedral adduct in a serine-carboxyl peptidase: a computational study, *J. Am. Chem. Soc. 127*, 15662−15663.
24. Cui, Q., Elstner, M., Kaxiras, E., Frauenheim T., and Karplus, M. (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method, *J. Phys. Chem. B 105*, 569−585.
25. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
26. Brooks, C. L., III, Brunger, A., and Karplus, M. (1985) Active site dynamics in protein molecules: A stochastic boundary molecular-dynamics approach, *Biopolymers 24*, 843−865.
27. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Bunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., III, Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586−3616.
28. Field, M. J., Bash, P. A., and Karplus, M. (1990) A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations, *J. Comput. Chem. 11*, 700−733.

29. Jorgensen, W. L., Chandrasckhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
30. Neria, E., Fisher, S., and Karplus, M. (1996) Simulation of activation free energies in molecular systems, *J.* Chem. Phys. *105*, 1902−1921.
31. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes, *J. Comput. Phys. 23*, 327−341.
32. Torrie, G. M., and Valleau, J. P. (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the subcritical Lennard-Jones fluid, *Chem. Phys. Lett. 28*, 578−581.
33. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman P. A., and Rosenberg, J. M. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules: 1. The method, *J. Comput. Chem. 13*, 1011−1021.
34. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1995) Multidimensional free-energy calculations using the weighted histogram analysis method, *J. Comput. Chem. 16*, 1339−1350.
35. Hedstrom, L. (2002) Serine protease mechanism and specificity, *Chem. Rev. 102*, 4501−4523.
36. Polgár, L. (2005) The catalytic triad of serine peptidases, *Cell. Mol. Life Sci. 62*, 2161−2172.
37. Cleland, W. W., and Kreevoy, M. M. (1994) Low-barrier hydrogen-bond and enzymatic catalysis, *Science 264*, 1887−1890.
38. Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994) A low-barrier hydrogen bond in the catalytic triad of serine proteases, *Science 264*, 1927−1930.
39. Warshel, A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites, *J. Biol. Chem. 273*, 27035−27038.
40. Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G., and Bachovchin, W. W. (2000) Unusual ^{1}H NMR chemical shifts support (His) C$^{\epsilon}$1—H...O═C H-bond: Proposal for reaction-driven ring flip mechanism in serine protease catalysis, *Proc. Nat. Acad. Sci. U.S.A. 97*, 10731−10735.
41. Brooks, C. L., III, Karplus, M., and Pettitt, B. M. (1988) *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, John Wiley & Sons, New York.
42. van Gunsteren, W. F., Daura, X., and Mark, A. E. (2002) Computation of free energy, *Helv. Chim. Acta 85*, 3113−3129.
43. Kobrak, M. N. (2003) Systematic and statistical error in histogram-based free energy calculations, *J. Comput. Chem. 24*, 1437−1446.

BI060461I