Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates

Evette S. Radisky*, Justin M. Lee†, Chia-Jung Karen Lu‡, and Daniel E. Koshland, Jr.§

Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Contributed by Daniel E. Koshland, Jr., March 15, 2006

Atomic resolution structures of trypsin acyl-enzymes and a tetrahedral intermediate analog, along with previously solved structures representing the Michaelis complex, are used to reconstruct events in the catalytic cycle of this classic serine protease. Structural comparisons provide insight into active site adjustments involved in catalysis. Subtle motions of the catalytic serine and histidine residues coordinated with translation of the substrate reaction center are seen to favor the forward progress of the acylation reaction. The structures also clarify the attack trajectory of the hydrolytic water in the deacylation reaction.

acyl-enzyme | enzyme mechanism | steady state | reaction trajectory

Serine proteases catalyze peptide bond hydrolysis in two sequential steps. In the first (acylation) reaction, the nucleophilic serine attacks the substrate scissile bond, forming first a tetrahedral intermediate and then a covalent acyl-enzyme with release of the C-terminal fragment. In the second (deacylation) reaction, a water molecule attacks the acyl-enzyme, leading to a second tetrahedral intermediate followed by release of the N-terminal fragment. This mechanism has long served as one of the classic paradigms illustrating the catalytic power of enzymes (1), yet crucial details remain controversial.

It is generally accepted that a histidine residue acts as a general base in accepting a proton to activate serine as a nucleophile, and subsequently acts as a general acid, donating the proton to the nitrogen of the peptide leaving group (1). This same histidine is also presumed to deprotonate the hydrolytic water. What is not clear is why, if the histidine is ideally situated to deprotonate serine, does the first tetrahedral intermediate not collapse back to the reactant complex with return of the proton to serine? Does the active site undergo spatial reorganization as catalysis proceeds to favor the forward progress of the reaction (2, 3)? A reaction-driven "His flip" mechanism has been proposed to address this question (4) but has met with resistance (5–8). Another longstanding question concerns the positioning and activation of the hydrolytic water molecule and the trajectory of attack in the deacylation reaction (9–13).

These questions can now be approached through the study of protein structures of intermediates along the catalytic pathway. Structures of the enzyme/substrate Michaelis complex, analogs of the tetrahedral intermediates, and acyl-enzymes can yield great insight into the atomic motions and shifting geometric relationships along the reaction coordinate. Here, we report (*i*) high resolution structures of trypsin acylated by two good peptide substrates and (*ii*) higher resolution structures of two previously solved trypsin complexes: the stable acyl-enzyme formed with the poor substrate *p*-nitroguanidinobenzoate (14) and the covalent complex formed with the inhibitor leupeptin, which mimics a tetrahedral intermediate (15). Comparisons provide insight into active site adjustments involved in catalysis and clarify the activation and attack trajectory of the hydrolytic water.

Results

Experimental Strategy: Trapping Reactive Acyl-Enzymes. It should be possible to observe acyl-enzymes formed in reaction with good substrates, under conditions where turnover is rapid, by choosing an enzyme for which the deacylation step is rate-limiting when substrate is saturating. Under these pseudosteady-state conditions, although enzyme is constantly turning over, the majority of the enzyme will be present in the acylated state because this species precedes the bottleneck (16).

Bovine trypsin was selected as the enzyme of study because the low density crystal form has particularly large solvent channels open to the active site (17, 18). Additionally, although smaller amide substrates react with rate-limiting acylation, longer oligopeptide substrates that form more specific contacts in the active site react with rate-limiting deacylation (5, 19). Through empirical experimentation, we identified two specific oligopeptide substrates, succinyl-Ala-Ala-Pro-Arg-*p*-nitroaniline (suc-AAPR-pNA) and succinyl-Ala-Ala-Pro-Lys-*p*-nitroaniline (suc-AAPK-pNA), along with pseudosteady-state conditions for observation of each acyl-enzyme in high occupancy by rapidly perfusing crystals with substrate followed by flash freezing.

The oligopeptidyl acyl-enzymes formed within the crystals were catalytically competent and underwent rapid deacylation. This competency was confirmed for the suc-AAPR-pNA substrate by soaking a crystal under identical conditions to those used for trapping the acyl-enzyme, then briefly soaking the crystal in the absence of substrate before freezing the crystal and obtaining diffraction data. The resultant structure showed no electron density for substrate or product in the active site, consistent with rapid deacylation and diffusion of product from the crystal.

Molecular Snapshots: Structures Used in Reconstruction of the Reaction Pathway. The trapping strategy described in *Experimental Strategy: Trapping Reactive Acyl-Enzymes* resulted in structures for two acyl-enzymes: suc-AAPR-trypsin [Protein Data Bank (PDB) ID code 2AGE] and suc-AAPK-trypsin (PDB ID code 2AGG). The model for suc-AAPR-trypsin was refined at 1.15 Å resolution, and displayed an acyl-enzyme occupancy of virtually 100% (Fig. 1*A*). The model for suc-AAPK-trypsin was refined at 1.28 Å resolution and showed acyl-enzyme occupancy of 70%, with the residual 30% of trypsin present in the form of free enzyme (Fig. 1*B*). Statistics for data collection and refinement of both structures are found in Table 2, which is published as supporting information on the PNAS web site.

BIOCHEMISTRY

Conflict of interest statement: No conflicts declared.

Abbreviations: PDB, Protein Data Bank; *p*-NPGB, *p*-nitrophenyl guanidinobenzoate substrate; suc-AAPK-pNA, succinyl-Ala-Ala-Pro-Lys-*p*-nitroaniline; suc-AAPR-pRN, succinyl-Ala-Ala-Pro-Arg-*p*-nitroaniline.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2AGE, 2AGG, 2AH4, and 2AGI).

*Present address: Mayo Clinic Cancer Center, Jacksonville, FL 32224.

†Present address: Genentech, South San Francisco, CA 94080.

‡Present address: Georgetown University School of Medicine, Washington, DC 20057.

§To whom correspondence should be addressed. E-mail: dek@berkeley.edu.

© 2006 by The National Academy of Sciences of the USA

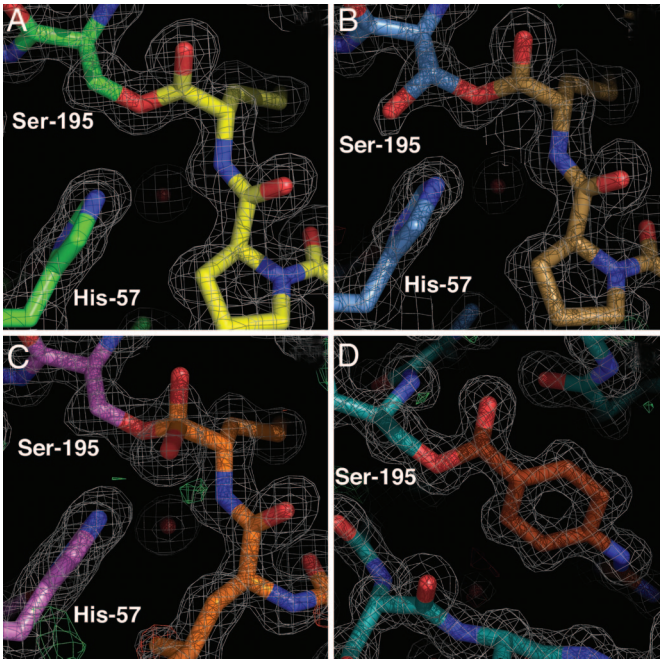

Fig. 1. Stick diagrams overlaid with $2F_o–F_c$ density maps contoured at 1 σ (gray mesh), showing covalent attachment of substrate ligands to trypsin. Oxygens are colored red and nitrogens are colored blue for all structures; carbons are color-coded differently for the enzyme and ligand residues of each structure, as detailed below. (*A*) For AAPR-trypsin, enzyme carbons are green, and substrate carbons are yellow. (*B*) For AAPK-trypsin, enzyme carbons are light blue, and substrate carbons are tan. Both conformations of Ser-195 are shown. (*C*) For the leupeptin-trypsin hemiacetal, enzyme carbons are purple, and substrate carbons are orange. Both hemiacetal conformations are shown. (*D*) For guanidinobenzoyl-trypsin, enzyme carbons are teal, and substrate carbons are brown. The guanidinobenzoyl moiety is substantially rotated in the active site compared with the other substrate ligands. Consequently, positioning of the figure to clearly display substrate density removes the His-57 side chain from the viewing slab.

To represent the noncovalent Michaelis complex, we used complexes of trypsin with protein protease inhibitors. Such complexes feature a productive, near-ideal geometry (20) and can even undergo acylation on the millisecond time scale (21). Only the existence of an even more rapid religation reaction results in a crystallographically observable prereaction complex (20). Six structures of bovine trypsin in complex with protein protease inhibitors have been deposited in the PDB, ranging in resolution from 1.5 to 2 Å (PDB ID codes 1TAW, 1PPE, 1EJM, 2BTC, 3BTK, and 1SFI) (18, 22–26). Active site geometries for all six structures superposed very closely with respect to all functional groups; their consensus represents a snapshot of the trypsin Michaelis complex.

The complex between trypsin and leupeptin, an oligopeptidyl aldehyde inhibitor, serves as a reasonable model of a tetrahedral intermediate. Structures for the leupeptin-trypsin complex have been reported previously at resolutions of 1.7–1.8 Å (15). For the purposes of the present analysis, we collected data and refined the structure to 1.14 Å resolution, using the same protocols used for the acyl-enzyme structures (Table 2; PDB ID code 2AGI). The hemiacetal has two possible orientations, one in which the oxygen atom faces the oxyanion hole, mimicking the carbonyl-derived oxygen of the tetrahedral intermediate, and a second in which the oxygen faces the active site His-57. In this second orientation, the hemiacetal oxygen can be inferred to mimic the position of an amine leaving group in the acylation reaction, or the attacking water-derived oxygen of the tetrahedral intermediate in the deacylation reaction. We modeled both orientations as alternative conformations and found ≈15% occupancy in the oxyanion hole and ≈85% occupancy of the position adjacent to His-57 (Fig. 1*C*).

To the Michaelis complexes, leupeptin complex, and reactive acyl-enzymes, which traced out the catalytic pathway for good substrates, we added the structure of a stable acyl-enzyme formed with a poor substrate, enabling analysis of structural features that help to determine deacylation rates. The structure of guanidinobenzoyl-trypsin has been reported previously at 1.8–2.0 Å resolution (14); for the purposes of the present analysis, we collected data and refined the structure to 1.13 Å resolution (Table 2; PDB ID code 2AH4). This adduct could not be formed by soaking *p*-nitrophenyl guanidinobenzoate substrate (*p*-NPGB) into the low density form trypsin crystals, possibly because of the subtle active site distortion required for binding and acylation of this nonpeptidyl substrate. Instead, crystallization of GB-trypsin was accomplished in the presence of *p*-NPGB (Fig. 1*D*).

The Acylation Reaction: Coordinated Motion of Serine, Histidine, and the Substrate Reaction Center. To discover whether active site adjustments occur as catalysis proceeds, we superposed structures for a typical trypsin/inhibitor complex (the trypsin/APPI complex), the leupeptin-trypsin complex, and the suc-AAPR-trypsin acyl-enzyme. These three structures, representing a progression from Michaelis complex to tetrahedral intermediate to acyl-enzyme, allow reconstruction of a probable reaction coordinate for the acylation reaction (Fig. 2*A*). The largest structural adjustments center on the reactive carbonyl carbon, which moves 1.3 Å downward and to the left (as viewed in the figure) to form the covalent tetrahedral intermediate, and then 0.5 Å back and upward to form the acyl-enzyme. Throughout this translation, the motion of the carbonyl oxygen is minimized through interactions with the oxyanion hole, whereas the Arg side chain remains nearly static. Movements of the enzyme catalytic residues are subtle but significant: the Ser-195 side chain rotates 13° about the Cα–Cβ bond and the His-57 side chain swivels slightly in the plane of the ring upon formation of the covalent bond. The result is to shift the His-57 $N_{\varepsilon 2}$, from its initial favorable distance for activation of the serine, 0.6 Å nearer to the amine leaving group and attacking water for subsequent reaction steps. In the leupeptin-trypsin structure, the distance between His-57 $N_{\varepsilon 2}$ and the major conformational position of the hemiacetal oxygen is 2.74 Å. Because this hemiacetal oxygen is analogous to the leaving group nitrogen in the acylation reaction, His-57 is well placed to donate a proton to the leaving group. These very high resolution structures clearly demonstrate that for trypsin, the 180° rotation of His-57 about the Cβ-Cγ bond proposed by Ash *et al.* (4) is not necessary to position His-57 for productive breakdown of the tetrahedral intermediate.

Another type of histidine movement is portrayed in our electron density maps. AAPR-trypsin shows trace evidence of a second conformation for the His-57 side chain (Fig. 2*D*); however, this second conformation accounted for <10% occupancy, and modeling it did not improve refinement statistics, so it was omitted from the model. In AAPK-trypsin, His-57 displays two conformations with roughly equal occupancy (Fig. 2*E*). The second conformation, in which His-57 is rotated around Cα–Cβ out of H-bonding distance to water S-25 and away from Ser-195, appears to represent a nonproductive fluctuation in acyl-enzyme structure, although an alternative possibility is that this conformation corresponds to the 30% proportion of free enzyme present in the crystal. The conformational flexibility of His-57 apparent in our maps may be involved in other steps of the reaction pathway, such as initial substrate binding or product release.

The Deacylation Reaction: Trajectory of the Hydrolytic Water. In the AAPR-trypsin and AAPK-trypsin acyl-enzymes, we observed a well defined water molecule, S-25 (Fig. 2 *D* and *E*), positioned

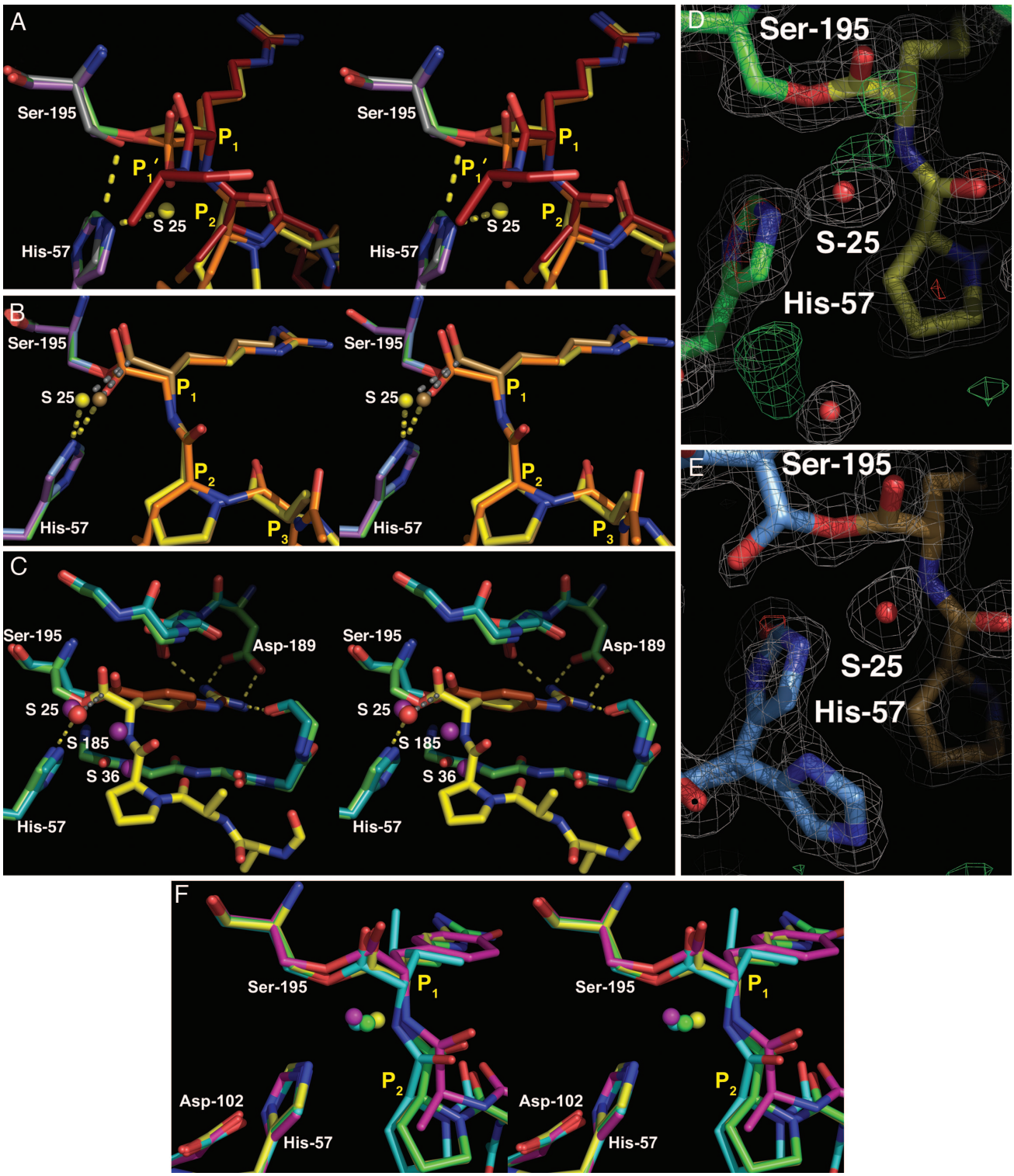

Fig. 2. Structural views and superpositions focusing on enzyme and substrate residues and water molecules involved in reaction. (*A*) Stereoviews comparing AAPR-trypsin (enzyme carbons, green; substrate carbons and water molecule, yellow) and leupeptin-trypsin (enzyme carbons, purple; substrate carbons, orange) with a representative trypsin/inhibitor Michaelis complex (enzyme carbons, gray; inhibitor carbons, brick red) allow reconstruction of a probable reaction coordinate for the acylation reaction. (*B*) Stereoviews comparing AAPR-trypsin, AAPK-trypsin (enzyme carbons, light blue; substrate carbons and water molecule, tan), and leupeptin-trypsin structures allow reconstruction of a probable reaction coordinate for water attack in the deacylation reaction. (*C*) Stereoviews comparing AAPR-trypsin and guanidinobenzoyl-trypsin (enzyme carbons, teal; inhibitor carbons, brown) reveal differences in attack trajectory possibly responsible for the great differences in reactivity between these substrates. (*D* and *E*) AAPR-trypsin (*D*) and AAPK-trypsin (*E*), stick diagrams overlaid with $2F_o$–F_c density maps contoured at 1 σ (gray mesh) and F_o–F_c maps scaled at 3 σ (green mesh) reveal dual conformations of His-57, as well as well defined density for water S-25. (*F*) Stereoviews of superimposed acyl-enzyme structures of different serine proteases, demonstrating a conserved position for the proposed hydrolytic water. Carbons and the attacking water molecule are colored differently for each structure: suc-AAPR-trypsin (green), suc-AAPK-trypsin (yellow), Ac-NPI-elastase [aqua; PDB ID code 1GVK (28)], and γ-chymotrypsin [magenta; PDB ID code 2GCT (29)].

Table 1. Geometric parameters for attack of potential hydrolytic water molecules

Parameter	Water—$C_{carbonyl}$, Å	Water—His-57 $N_{\varepsilon 2}$, Å	Ser-195 Oγ—His-57 $N_{\varepsilon 2}$, Å	His-57 $N_{\delta 1}$—Asp-102 $O_{\delta 2}$, Å	θ_x,* °	θ_y,* °	Δ,† Å
AAPR-trypsin							
Attack by S25	3.03	2.92	3.04	2.75	102.1	83.8	0.08
AAPK-trypsin							
Attack by S25	2.44	2.77	2.93	2.77	94.7	97.1	0.20
Leupeptin-trypsin‡	1.40	2.74	3.04	2.75	89.3	100.1	0.40
GB-trypsin							
Attack by S25	3.90	2.78	3.82	2.78	129.3	53.9	0.04
Attack by S36	3.99	3.64			82.6	113.3	
Attack by S185	3.47	3.37			87.6	68.1	

*See ¶ for θ_x and θ_y definitions.

†Δ represents the displacement of the carbonyl carbon out of the plane defined by Ser-195 Oγ, the carbonyl oxygen, and Cα.

‡For leupeptin-trypsin, the major conformation of the hemiacetal oxygen represents the attacking nucleophile for purposes of calculation.

within 3 Å of both the activating His-57 $N_{\varepsilon 2}$ and the acyl carbonyl carbon, with a geometry suggesting that this water is poised for hydrolytic attack (Table 1). Water S-25 is absent in the leupeptin-trypsin structure, displaced by the major conformation of the hemiacetal oxygen, lending further support to the idea that S-25 is the hydrolytic water.

Comparison of the AAPR-trypsin, AAPK-trypsin, and leupeptin-trypsin structures allows reconstruction of a probable reaction coordinate for water attack in the deacylation reaction (Fig. 2*B*). Only subtle adjustments are needed to form the tetrahedral intermediate, with the attacking water and the carbonyl carbon moving 1.6 Å nearer to each other to form the covalent bond. Several geometric parameters suggest that AAPK-trypsin is further advanced along the reaction coordinate than AAPR-trypsin, providing a midway snapshot along the trajectory of attack (Table 1). In AAPR-trypsin, water S-25 is 3 Å from the carbonyl carbon and approaches from the side nearer Ser-195 Oγ, rather than the substrate Arg Cα (θ_x = 102°, where 90° would represent a symmetrical approach).¶ In the AAPK-trypsin structure, water S-25 has moved closer to the acyl carbonyl by 0.6 Å and nearer to the midline bisecting the acyl bond plane (θ_x = 95°). In addition, the carbonyl carbon has become more pyramidal, as evidenced by a 0.2 Å out-of-plane displacement of the carbonyl carbon and a 13° increase in θ_y. From the AAPK-trypsin structure to the tetrahedral intermediate (as represented by leupeptin-trypsin), the attacking hydroxyl moves an additional 1 Å nearer to the carbonyl carbon, forming the covalent bond and reaching the midline bisecting the former acyl bond plane (θ_x = 89°); the carbon center is now fully tetrahedral.

"Good" Vs. "Bad" Acyl-Enzymes: Structural Determinants of Reactivity. Whereas the turnover of trypsin with suc-AAPR-pNA or suc-AAPK-pNA in solution is rapid [k_{cat} = 91 s^{-1} and 41 s^{-1}, respectively (E.S.R., J.M.L., D.E.K., unpublished work)], the deacylation step in reaction with *p*-NPGB proceeds more than six orders of magnitude more slowly [$k_{cat} = 3.40 \times 10^{-5}$ s^{-1}; (27)]. Comparison of AAPR-trypsin and GB-trypsin structures reveals differences in active site orientation potentially responsible for the great differences in deacylation rate (Fig. 2*C*). As discussed in *The Deacylation Reaction: Trajectory of the Hydrolytic Water* and shown in Table 1, water S-25 is positioned favorably for hydrolyzing AAPR-trypsin. In GB-trypsin, S-25 is located in a similar position, but because the shorter guanidinobenzoyl chain is pulled more deeply into the specificity pocket to interact with Asp-189, the acyl bond is tilted and translated such that S-25 is not optimally aligned to attack the carbonyl. Two alternative waters, S-185 or S-36, might be candidates for the hydrolytic water with this substrate, but neither is optimally positioned for attack, and neither is near enough to His-57 for facile deprotonation (Table 1). The lack of a water appropriately positioned both for activation and for an optimal attack trajectory is likely responsible for the very slow hydrolysis of GB-trypsin.

Time-resolved Laue structures of GB-trypsin led Singer *et al.* (10) to propose a particular hydrolytic water molecule for this acyl-enzyme. Their proposed Wat-1082 corresponds, within 0.3 Å, to our S-36. Given that Singer *et al.* observed Wat-1082 only upon increasing the pH within a crystal to trigger hydrolysis, we were surprised to see the strong water peak at this position in our structure. Of the three potentially hydrolytic water molecules in our GB-trypsin structure, S-36 is positioned at the furthest distance from both the carbonyl carbon and His-57 $N_{\varepsilon 2}$, but maintains a more favorable trajectory of attack, being the only water for which θ_y approaches or exceeds 90°. If S-36 is indeed the hydrolytic water, it would suggest that the geometry of attack is more critical than proximity to an activating base for promoting the reactivity of this nucleophile.

Discussion

Authenticity of Acyl-Enzyme Structures. A number of structures of acyl-enzymes with peptide substrates have been reported; however, the acyl-enzyme has generally been trapped by the use of a nonphysiologic low pH, by the use of a poor substrate with an exceptionally long-lived acyl-enzyme, or both (12, 28, 29). Such structures are necessarily imperfect approximations of the true intermediate in the cleavage of a favorable substrate. Acyl-enzymes of moderately reactive small ester substrates have been trapped by using cryo-crystallographic methods (30, 31), but these small substrates lack the full range of enzyme-substrate interactions, and resulting precise orientation, that result in rapid catalysis with a specific peptide substrate. Several structures have revealed autoproteolytic mixtures in the active site attributed to acyl-enzymes in partial occupancy (29, 32–34), but such interpretation is not straightforward because the density maps represent mixed species. Here, we have succeeded in trapping authentic acyl-enzymes under favorable reaction conditions through careful choice of enzyme, substrates, and soaking conditions to favor pseudosteady-state conditions under which the acyl-enzyme accumulates, and we have

¶θ_x and θ_y, diagrammatically defined in Fig. 3 *Lower*, which is published as supporting information on the PNAS web site, together describe the attack trajectory of a hydrolytic water in the deacylation reaction. θ_x is the angle between (*i*) the plane containing the attacking water and the acyl-enzyme carbonyl carbon and oxygen and (*ii*) the plane containing the enzyme Ser-195 γ-oxygen, the acyl-enzyme carbonyl oxygen, and the acyl-enzyme α-carbon. θ_y is the angle defined by the attacking water, the acyl-enzyme carbonyl carbon, and the acyl-enzyme carbonyl oxygen. Similar angles were previously delineated for the acylation attack trajectory (20).

obtained unambiguous density maps at very high resolution. Pseudosteady-state techniques for observing reaction intermediates crystallographically, more typically involving the use of flow cells, have been used successfully for other enzymes including xylose isomerase (35) and tryptophan synthase (36). To our knowledge, the present paper is the first application of the steady-state trapping of Petsko and Ringe to the structural study of the serine hydrolase mechanism. The approach holds great appeal, because the information obtained bears directly on the productive active site geometries and attack trajectories of trypsin and other serine proteases.

Conservation of the Hydrolytic Water. It must be acknowledged that no static x-ray structure can provide absolute proof that a given crystallographically observed water is "the" water. We observe a chemically plausible hydrolytic water, S-25, in two acyl-enzymes formed with good substrates under conditions conducive to reaction. Viewing these two structures as snapshots along a reaction coordinate suggests an equally plausible trajectory of attack for this potential hydrolytic water molecule. Further support for S-25 as the hydrolytic water is lent by structural comparisons with other acyl-enzymes formed with other serine proteases, trapped via different strategies. Comparison with structures of (*i*) elastase complexed with a peptide substrate selected for its particularly long-lived acyl-enzyme (28) and (*ii*) γ-chymotrypsin at pH 2, found to be a stable acyl-enzyme formed with autoproteolysis products (29), reveals that these structures also contain active site water molecules at positions very similar to S-25 (Fig. 2*F*). The conserved position of this water molecule in the active sites of these enzymes suggests a conserved trajectory of attack by the hydrolytic water in the chymotrypsin family of serine proteases.

Conservation of Nucleophilic Attack Trajectory. Geometric parameters can be compared between the acylation reaction (Table 3, which is published as supporting information on the PNAS web site) and the deacylation reaction (Table 1). The acylation and deacylation reaction trajectories share a very similar geometry, including approach of the nucleophile from an angle close to 90° with respect to both θ_x and θ_y. For both reactions, His-57 $N_{\varepsilon 2}$ is positioned optimally to deprotonate the oxygen nucleophile, whether from the Ser-195 Oγ or from water, and through both reactions, His-57 $N_{\delta 1}$ remains in close H-bonding distance to Asp-102.

Insights into Catalysis. The present reconstruction of atomic motions involved in the acylation reaction suggests a reaction coordinate dominated by translation of the carbonyl carbon between leaving group and attacking nucleophile, accompanied by small but significant concerted adjustments in the catalytic Ser and His residues. Our findings join a growing cohort of crystallographic analyses of other enzymes catalyzing nucleophilic displacements and substitutions, including ribosyltransferases (37, 38), glycosidases (39), and protein farnesyltransferase (40). A common thread appears to be an economy of atomic motion, with the recurring theme of migration of the electrophilic carbon center between a fixed leaving group and nucleophile (37). For the trypsin deacylation reaction, again we observe minimal atomic motion, but in this case the movement of the attacking nucleophilic water molecule is featured more prominently. The productive arrangement of catalytic residues in the serine protease active site is optimized not only to position and activate both Ser-195 Oγ and the hydrolytic water at appropriate times in the catalytic cycle but also to subtly shift the position of His-57 as the cycle progresses, to favor catalysis.

Methods

Bovine trypsin (product no. T-1426; Sigma) was dissolved at 15–35 mg/ml in a buffer containing 100 mM Bis Tris (pH 6.0), 0.3 M ammonium sulfate, 6 mM calcium chloride, and 60 mM benzamidine. Crystals were grown at room temperature in 10-μl hanging drops over a reservoir of 100 mM Bis Tris (pH 6.0) with 1.6–2.4 M ammonium sulfate. Under these conditions, crystals grew within 1 week and were found to be exclusively of a low-density orthorhombic form reported in ref. 17. To remove benzamidine, crystals were soaked three times in fresh solution containing 100 mM Bis Tris (pH 6.0), 2.5 M ammonium sulfate, and 6 mM calcium chloride. Subsequently, crystals were transferred into fresh soaking solution containing 1 mM leupeptin (Sigma), 10 mM suc-AAPR-pNA (Bachem), or 10 mM suc-AAPK-pNA (Bachem). The leupeptin soak was at room temperature for 3 hours. Soaking conditions with substrates were optimized to produce the highest possible acyl-enzyme occupancy with the least occupancy of contaminating species, evaluated from preliminary density maps from unrefined x-ray data solutions. For suc-AAPR-pNA, the substrate soak was carried out for 60 min at 4°C. For suc-AAPK-pNA, deacylation was not sufficiently rate-limiting at pH 6, resulting in a mixture of species; the acyl-enzyme was instead trapped by performing the substrate soak in 100 mM sodium citrate (pH 4.0), 2.5 M ammonium sulfate, and 6 mM calcium chloride for 60 min at 4°C. All crystals described above were very briefly dipped in a cryoprotectant composed of the soaking solution with the addition of 30% dextrose, and then flash frozen in liquid N_2.

For *p*-NPGB, no acyl-enzyme density was observed after substrate soaks of preformed crystals. Instead, crystallization conditions were as described above but with 5 mM *p*-NPGB present in each crystallization drop. In addition, better acyl-enzyme occupancy was obtained at pH 5. Crystals formed in the presence of *p*-NPGB were all of the higher density orthorhombic form observed previously for GB-trypsin (14). These crystals were briefly dipped in a cryoprotectant composed of 100 mM sodium citrate (pH 4.6), 2.5 M ammonium sulfate, 20 mM calcium chloride, and 30% dextrose, and then flash frozen in liquid N_2.

Synchrotron x-ray data were collected at 100 K by using an Area Detector Systems Corporation (ADSC) Quantum 210 charge-coupled device detector at Advanced Light Source beam line 8.3.1 (Lawrence Berkeley National Laboratory, Berkeley, CA). The automation package ELVES (41) was used to direct the programs MOSFLM (42), for indexing and integration, and SCALA (43), for scaling and merging the reflections. The suc-AAPR-trypsin structure was solved by molecular replacement using EPMR (44), using as the search model the structure of bovine β-trypsin in the low density crystal form [PDB ID code 1TLD (17)], with solvent molecules omitted. A test set composed of 5% of the total reflections, assigned at random, was excluded from refinement to allow calculation of the free R factor. The model was improved through alternating cycles of manual rebuilding (including addition of the acylating substrate) by using the interactive graphics program O (45), automated refinement using REFMAC5 (46), and automated solvent addition using ARP/WARP (47). The covalent bond between trypsin Ser-195 and the acylating substrate was defined with an ideal distance of 1.45 Å, but restraints on this distance were very loose throughout refinement, and no restraints were placed on bond angles surrounding the ester linkage, in order that this feature of the model should be refined with maximal weight on the x-ray data. Initial refinement used isotropic B factors. After convergence of the initial refinement stage and visual inspection of all waters, anisotropic displacement parameters were introduced, hydrogen atoms were added in the riding positions, and refinement of the model was continued to convergence using REFMAC5. Refine-

ment of the suc-AAPK-trypsin and leupeptin trypsin structures used as a starting model the suc-AAPR-trypsin model after the initial stage of refinement, with solvent and acylating substrate omitted. The free R test sets for these structures were inherited from suc-AAPR-trypsin, and the same refinement protocol was followed. The GB-trypsin structure was solved by molecular replacement using EPMR, by using as the search model the previously solved GB-trypsin structure [PDB ID code 1GBT, (14)] with solvent and the acylating guanidinobenzoyl moiety omitted. The same refinement protocol described above was followed. Model validation was carried out using PROCHECK (43) and WHATCHECK (48). Superpositions were performed by using the program GEM (49) for comparisons within the bovine trypsin structures or PYMOL (50) for comparisons with other serine proteases. PYMOL was used to generate the structure figures.

This work was supported in part by National Institutes of Health Grant DK09765 (to D.E.K.). The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory (Berkeley, CA).

1. Voet, D. & Voet, J. G. (2004) *Biochemistry* (Wiley, New York), 3rd Ed.
2. Bizzozero, S. A. & Dutler, H. (1981) *Bioorg. Chem.* **10,** 46–62.
3. Bachovchin, W. W. (1986) *Biochemistry* **25,** 7751–7759.
4. Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G. & Bachovchin, W. W. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 10371–10376.
5. Hedstrom, L. (2002) *Chem. Rev.* (*Washington, D.C.*) **102,** 4501–4524.
6. Topf, M. & Richards, W. G. (2004) *J. Am. Chem. Soc.* **126,** 14631–14641.
7. Topf, M., Varnai, P., Schofield, C. J. & Richards, W. G. (2002) *Proteins* **47,** 357–369.
8. Ishida, T. & Kato, S. (2003) *J. Am. Chem. Soc.* **125,** 12035–12048.
9. Henderson, R. (1970) *J. Mol. Biol.* **54,** 341–354.
10. Singer, P. T., Smalas, A., Carty, R. P., Mangel, W. F. & Sweet, R. M. (1993) *Science* **259,** 669–673.
11. Perona, J. J., Craik, C. S. & Fletterick, R. J. (1993) *Science* **261,** 620–622.
12. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J. & Schofield, C. J. (1997) *Nat. Struct. Biol* **4,** 456–462.
13. Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J. & Hajdu, J. (2001) *Nat. Struct. Biol.* **8,** 689–694.
14. Mangel, W. F., Singer, P. T., Cyr, D. M., Umland, T. C., Toledo, D. L., Stroud, R. M., Pflugrath, J. W. & Sweet, R. M. (1990) *Biochemistry* **29,** 8351–8357.
15. Kurinov, I. V. & Harrison, R. W. (1996) *Protein Sci.* **5,** 752–758.
16. Petsko, G. A. & Ringe, D. (2000) *Curr. Opin. Chem. Biol.* **4,** 89–94.
17. Bartunik, H. D., Summers, L. J. & Bartsch, H. H. (1989) *J. Mol. Biol.* **210,** 813–828.
18. Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R. & Brady, R. L. (1999) *J. Mol. Biol.* **290,** 525–533.
19. Christensen, U. & Ipsen, H. H. (1979) *Biochim. Biophys. Acta* **569,** 177–183.
20. Radisky, E. S. & Koshland, D. E., Jr. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 10316–10321.
21. Radisky, E. S., Kwan, G., Karen Lu, C. J. & Koshland, D. E., Jr. (2004) *Biochemistry* **43,** 13648–13656.
22. Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A. & Kossiakoff, A. A. (1997) *Protein Sci.* **6,** 1806–1824.
23. Bode, W., Greyling, H. J., Huber, R., Otlewski, J. & Wilusz, T. (1989) *FEBS Lett.* **242,** 285–292.
24. Grzesiak, A., Helland, R., Smalas, A. O., Krowarsch, D., Dadlez, M. & Otlewski, J. (2000) *J. Mol. Biol.* **301,** 205–217.
25. Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P. & Smalas, A. O. (1999) *Acta Crystallogr. D* **55 (Pt. 1),** 139–148.
26. Helland, R., Otlewski, J., Sundheim, O., Dadlez, M. & Smalas, A. O. (1999) *J. Mol. Biol.* **287,** 923–942.
27. Chase, T., Jr., & Shaw, E. (1969) *Biochemistry* **8,** 2212–2224.
28. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R. & Schofield, C. J. (2002) *J. Biol. Chem.* **277,** 21962–21970.
29. Dixon, M. M., Brennan, R. G. & Matthews, B. W. (1991) *Int. J. Biol. Macromol.* **13,** 89–96.
30. Alber, T., Petsko, G. A. & Tsernoglou, D. (1976) *Nature* **263,** 297–300.
31. Ding, X., Rasmussen, B. F., Petsko, G. A. & Ringe, D. (1994) *Biochemistry* **33,** 9285–9293.
32. Dixon, M. M. & Matthews, B. W. (1989) *Biochemistry* **28,** 7033–7038.
33. Blanchard, H. & James, M. N. (1994) *J. Mol. Biol.* **241,** 574–587.
34. Schmidt, A., Jelsch, C., Ostergaard, P., Rypniewski, W. & Lamzin, V. S. (2003) *J. Biol. Chem.* **278,** 43357–43362.
35. Farber, G. K., Glasfeld, A., Tiraby, G., Ringe, D. & Petsko, G. A. (1989) *Biochemistry* **28,** 7289–7297.
36. Schneider, T. R., Gerhardt, E., Lee, M., Liang, P. H., Anderson, K. S. & Schlichting, I. (1998) *Biochemistry* **37,** 5394–5406.
37. Schramm, V. L. & Shi, W. (2001) *Curr. Opin. Struct. Biol.* **11,** 657–665.
38. Fedorov, A., Shi, W., Kicska, G., Fedorov, E., Tyler, P. C., Furneaux, R. H., Hanson, J. C., Gainsford, G. J., Larese, J. Z., Schramm, V. L. & Almo, S. C. (2001) *Biochemistry* **40,** 853–860.
39. Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. (2001) *Nature* **412,** 835–838.
40. Long, S. B., Casey, P. J. & Beese, L. S. (2002) *Nature* **419,** 645–650.
41. Holton, J. & Alber, T. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 1537–1542.
42. Leslie, A. G. W. (1992) *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography* (Daresbury Laboratories, Warrington, U.K.).
43. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr. D. Biol. Crystallogr.* **50,** 760–763.
44. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999) *Acta Crystallogr. D* **55 (Pt. 2),** 484–491.
45. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47 (Pt. 2),** 110–119.
46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D.* **53,** 240–255.
47. Perrakis, A., Morris, R. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463.
48. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. (1996) *Nature* **381,** 272 (lett.).
49. Fauman, E. B., Rutenber, E. E., Maley, G. F., Maley, F. & Stroud, R. M. (1994) *Biochemistry* **33,** 1502–1511.
50. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).