Structural Analyses on Intermediates in Serine Protease Catalysis*

Received for publication, January 17, 2006, and in revised form, April 11, 2006 Published, JBC Papers in Press, June 5, 2006, DOI 10.1074/jbc.M600495200

Bin Liu‡, Christopher J. Schofield‡, and Rupert C. Wilmouth‡

From the ‡School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore and the †Department of Chemistry, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, United Kingdom

Although the subject of many studies, detailed structural information on aspects of the catalytic cycle of serine proteases is lacking. Crystallographic analyses were performed in which an acyl-enzyme complex, formed from elastase and a peptide, was reacted with a series of nucleophilic dipeptides. Multiple analyses led to electron density maps consistent with the formation of a tetrahedral species. In certain cases, apparent peptide bond formation at the active site was observed, and the electron density maps suggested production of a cis-amide rather than a trans-amide. Evidence for a cis-amide configuration was also observed in the noncovalent complex between elastase and an α1-antitrypsin-derived tetrapeptide. Although there are caveats on the relevance of the crystallographic data to solution catalysis, the results enable detailed proposals for the pathway of the acylation step to be made. At least in some cases, it is proposed that the alcohol of Ser-195 may preferentially attack the carbonyl of the cis-amide form of the substrate, in a stereoelectronically favored manner, to give a tetrahedral oxyanion intermediate, which undergoes N-inversion and/or C=N bond rotation to enable protonation of the leaving group nitrogen. The mechanistic proposals may have consequences for protease inhibition, in particular for the design of high energy intermediate analogues.

Following from their role in the development of mechanistic and structural enzymology, serine proteases, and more generally enzymes catalyzing hydrolysis or related reactions proceeding via an acyl-enzyme complex, have been shown to be of importance in human diseases. Serine proteases, including thrombin, factor X, and elastase, are targets for medicinal chemistry (1, 2). One important class of inhibitors is designed to mimic a tetrahedral intermediate, but until recently there has been little three-dimensional structural information on the early intermediates in catalysis.

In outline, the accepted mechanism for serine protease involves nucleophilic attack by the alcohol of Ser-195 onto the amide carbonyl of the peptide substrate to form a first tetrahedral intermediate. This intermediate collapses to form an ester, or acyl-enzyme complex, and the COOH-terminal product fragment then leaves the active site. Hydrolysis of the Ser-195-linked ester bond then occurs via a second tetrahedral intermediate, to form the NH2-terminal product fragment and return the active site to its resting state. Much evidence has accumulated in support of this general mechanism, including the role of His-57 in enabling acid/base catalysis, and it has been extended to other types of hydrolysis reaction, including esterases. Although there has been extensive structural work on covalently bound inhibitors and some on substrate analogues (3), there is relatively little direct crystallographic evidence for the key intermediates, and questions remain as to the precise stereoelectronic pathway of the reaction, including the timing and nature of the requisite proton transfers (4).

Following from studies on nonpeptide substrates (3, 5, 6), which were inconclusive with respect to the positioning of peptide substrates at the active site, crystallographic analyses were reported on the acyl-enzyme complex formed between porcine pancreatic elastase (PPE) and human β-caseinomorphin-7 (YPFVEPI, BCM7) (7). They revealed the heptapeptide bound to the active site in an apparently productive manner and with the isoleucine covalently linked to Ser-195. A well defined water molecule, hydrogen-bonded to His-57, appeared well positioned for nucleophilic attack onto the ester bond, but reaction was not observed presumably due to protonation of His-57. Separate experiments supported productive binding by demonstrating that BCM7 with two additional alanine residues (YPFVEPIAA) was cleaved by PPE between the isoleucine and alanine. Cryocystallographic analyses of pH-jumped PPE:BCM7 co-crystals were used to study intermediates in the deacylation pathway (8). Although probably not reflecting a fully homogeneous population, the observed changes in the electron density maps implied formation of a tetrahedral intermediate during hydrolysis of the acyl-envelope complex.

Although there are various structures of protein or peptide inhibitor complexes that span the peptide binding site, there is no reported structural information on the first tetrahedral intermediate or the enzyme-substrate complex prior to its formation. Such information would be useful in the design of inhibitors that mimic these intermediates. To investigate the
acylation mechanism, we employed a strategy in which nucleophilic dipeptides were soaked into crystals of the PPE-BCM7 acyl-enzyme complex with the aim of forming early intermediates in reverse to the normal catalytic direction. Here we report cryocrystallographic analyses on complexes produced by this methodology; the results suggest new mechanistic ideas that may be useful in inhibitor design.

MATERIALS AND METHODS

Peptide Synthesis—Ala-Ala, Ala-Gly, Arg-Phe, and Asp-Phe were purchased from Sigma. All other peptides were synthesized by solid phase synthesis, applying the Fmoc (N-(9-fluorenylmethoxycarbonyl) protection strategy. Peptides were purified by reverse-phase chromatography with a Luna 5 μm C18 column (Phenomenex). They were eluted with an acetonitrile gradient and identified by electrospray ionization mass spectrometry. The peptide-containing fractions were lyophilized and stored at −80 °C.

Crystallization—The PPE-BCM7 complex was crystallized as described (7). The PPE-NH2-Asn-Pro-Ile-COOH (NPI) and PPE-MFLE crystals were obtained using the same methodology as employed for PPE-N-acetyl-Asn-Pro-Ile-COOH (AcNPI) (9) (MFLE was at a saturating concentration). For the acylation solution for 30–60 min; either solid dipeptide (−2 mg) was added directly to 0.25 mM sodium acetate (pH 5.0) (5 μl) to form a saturated solution, or the dipeptide concentration was increased stepwise to reach a final concentration of −20–50 mg ml−1 in 0.25 mM sodium acetate (pH 5.0) (5 μl). The crystals were then briefly soaked in cryoprotectant solution (25% (v/v) glycerol, 0.25 mM sodium acetate (pH 5.0)) and rapidly frozen in liquid nitrogen. The crystals were then mounted on the x-ray generator and held at −140 °C using a dry nitrogen gas stream (X-Stream; Rigaku). After collecting a full data set, the crystal was removed from the nitrogen gas stream and immediately placed in the pH jump solution (25% (v/v) glycerol, 0.25 M CHES (pH 9.0)) for 30–60 s. The high pH value was used to provide fast triggering of the reaction. PPE has maximal activity at pH 8.5. If the crystal survived then it was rapidly transferred back to the nitrogen gas stream and refrozen at the same temperature (−140 °C). After another set of data were collected, this pH jump procedure was attempted a second time. If successful, a further data set was collected. Error-scaled distance difference matrices (10) were used to analyze the changes in the electron density maps shown in the figures are representative examples. In total, over 60 x-ray diffraction data sets were collected and refined.

Data Collection—All data (except that for the AcNPI-AA structure) were collected at either −140 °C (for the acylation studies) or −160 °C (for the PPE-NP and PPE-MFLE structures) using a Rigaku MicroMax-007 x-ray generator operating at 40 kV and 20 mA and a Rigaku R-Axis IV++ image plate detector. Data for the AcNPI-AA structure were collected at −173 °C at beamline X06SA at the Swiss Light Source, Villingen-PSI, Switzerland at a wavelength of 0.98 Å using a MarResearch CCD detector. Data were processed with the programs MOSFLM and SCALA (CCP4 suite (11)). A structure of native PPE (originally phased using the 1.65 Å structure of Meyer et al. (12)) was used as the starting model in all refinements. For all data sets, a total of 5% of the reflections were randomly selected to provide a test set for the \( R_{free} \) calculations (13). The structures were refined using REFMAC 5.0 and the CCP4 suite version 5. For the tetrahedral and \( E+S \) complexes, additional topology parameters were introduced via a LINK entry, but the geometric restraints were kept deliberately weak. The side chains of the peptides were modeled using fragments of electron density and the positions of the side chains in the hybrid squash inhibitor structure (14) as a guide. These atoms were not refined against crystallographic data (an occupancy of 0.01 was used to correct the geometry) and were intended only as a guide to the possible location of the side chains. Models and electron density maps were displayed with the program O (15). All figures were created using PyMOL (16).

Kinetics—Kinetic assays and numerical analyses were performed as described previously (17). A Varian Cary 300 UV-visible spectrometer was used for recording absorbance measurements.

RESULTS

Work with different types of substrate and inhibitor on the specificity of the PPE S’ subsites (which bind residues on the COOH-terminal side of the cleavage site) implies that the specificity in these sites may change depending upon the context (11–13). We thus chose a range of II-dipeptides to investigate as nucleophiles for reaction with the PPE-BCM7 complex: NH2-Ala-Ala-COOH, NH2-Ala-Gly-COOH, NH2-Lys-Ser-COOH, NH2-Lys-Ala-CONH2, NH2-Arg-Phe-COOH, and NH2-Asp-Phe-COOH. The dipeptides were soaked into PPE-BCM7 co-crystals at pH 5 for 30–60 min (Table 1). The crystals were then frozen (−140 °C), and x-ray diffraction data were collected. All co-crystals were then subjected to an attempted pH jump via exposure to pH 9 buffer. The crystals were warmed to room temperature, their pH was raised by buffer exchange (pH 5 to pH 9), and then they were refrozen. This technique enabled an additional x-ray diffraction data set to be collected on the same crystal that was used for the initial data collection at pH 5. It was found that a slightly higher temperature than usually employed during data collection (−140 °C instead of −160 °C) aided this procedure, perhaps by making the freeze-thaw process less stressful to the crystals. Nonetheless, the pH-jumped crystals frequently cracked or dissolved during the pH jump, giving a relatively low success rate. Occasionally, a second immersion in pH 9 buffer and subsequent refreezing was also possible; this allowed the collection of a third data set on the same crystal. These experiments were repeated several times for each dipeptide, and the electron density maps shown in the figures are representative examples. In total, over 60 x-ray diffraction data sets were collected and refined.
### Table 1

Data collection and refinement statistics

Values in parentheses are for the highest resolution shell.

|                | PPE-BCM7-RF | PPE-BCM7-KS | PPE-BCM7-DF | PPE-BCM7-KANH2_1 | PPE-BCM7-RF (2) first pH jump | PPE-BCM7-RF (2) second pH jump | PPE-AcNP1-AA | PPE-ENP1-AA | PPE-MFLE | PPE-NPI |
|----------------|-------------|-------------|-------------|------------------|-----------------------------|-----------------------------|---------------|-------------|-----------|---------|
| **Experimental condition** | pH 5 (45 min) | pH 5 (60 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) | pH 5 (30 min) |
| **Resolution (Å)** | 1.70 (1.79-1.70) | 1.70 (1.79-1.70) | 1.60 (1.69-1.60) | 1.60 (1.69-1.60) | 1.60 (1.69-1.60) | 1.70 (1.79-1.70) | 1.90 (2.00-1.90) | 1.80 (1.90-1.80) | 1.70 (1.79-1.70) | 1.60 (1.69-1.60) | 1.80 (1.90-1.80) |
| **P2_12_2_1 cell dimensions (Å)** | | | | | | | | | | | |
| a | 50.19 | 50.23 | 49.88 | 50.41 | 50.04 | 50.65 | 50.54 | 51.38 | 50.51 | 50.46 | 50.08 | 50.15 |
| b | 57.88 | 57.94 | 57.58 | 57.91 | 57.70 | 57.59 | 57.78 | 57.59 | 57.67 | 57.66 | 57.35 | 57.87 |
| c | 74.60 | 74.77 | 74.14 | 74.73 | 74.36 | 74.52 | 74.76 | 74.76 | 73.87 | 74.23 | 74.15 | 74.53 |
| **Total no. of observations** | 82,936 | 86,085 | 93,777 | 125,012 | 99,200 | 70,197 | 83,824 | 60,398 | 66,057 | 87,649 | 100,669 | 88,860 |
| **Unique reflection set** | 24,541 | 24,646 | 28,611 | 29,238 | 29,030 | 20,837 | 24,742 | 17,944 | 20,470 | 24,840 | 27,749 | 19,783 |
| **Completeness (%)** | 99.9 (100) | 99.9 (100) | 99.0 (93.9) | 99.0 (97.7) | 99.8 (99.5) | 99.9 (100) | 99.9 (100) | 99.4 (99.0) | 99.5 (100) | 99.9 (99.9) | 96.1 (91.9) | 95.8 (93.1) |
| **Rmerge (%)** | 3.9 (13.7) | 4.3 (13.0) | 29.6 (91.3) | 6.2 (20.4) | 4.6 (18.1) | 4.6 (16.9) | 6.3 (32.8) | 8.6 (46.0) | 3.4 (14.4) | 4.2 (20.8) | 3.0 (6.1) |
| **Average B-factors** | 20.60 | 19.36 | 25.07 | 22.89 | 25.43 | 30.06 | 27.84 | 33.38 | 17.45 | 16.46 | 21.40 | 12.91 |
| **Root mean square deviations** | 0.017 | 0.019 | 0.014 | 0.015 | 0.016 | 0.019 | 0.019 | 0.012 | 0.019 | 0.013 | 0.012 | 0.012 |
| **Bond lengths (Å)** | 1.593 | 1.717 | 1.491 | 1.508 | 1.575 | 1.746 | 1.746 | 1.274 | 1.662 | 1.371 | 1.413 | 1.286 |
| **Bond angles (degrees)** | 16.8/20.5 | 18.6/22.0 | 17.3/19.7 | 17.0/20.1 | 17.7/21.1 | 17.4/21.2 | 17.3/22.0 | 17.2/22.6 | 17.1/21.2 | 16.3/19.4 | 18.2/20.5 | 17.7/22.2 |
| **Protein Data Bank code** | 2BD2 | 2BD4 | 2B4 | 2BD3 | 2BD7 | 2BD5 | 2BD8 | 2BD9 | 2BD4 | 2BD5 | 2BD8 | 2BD9 |
| **Figures** | 1a and 2a | 1b and 6a | 2b | 2c and 6b | 2d | 3a | 3b | 3c | 4a | 4b | 4c | 6e |
is observed to bind close to the S1’ subsite in native elastase (12) but not in the PPE-BCM7 acyl-enzyme structure. A disadvantage of the modified technique was that only short soaking times were possible before the crystals dissolved. In the case of the Ala-Ala and Ala-Gly dipeptides, only weak density was observed in the S1’ subsites despite high concentrations of peptide being used in the soaking solution (70–180 ms). For all of the other chosen dipeptides (KS, KANH2, RF, and DF), reasonably good occupancy was observed for the main chain atoms in the S’ subsites (the average B factors were 36–40 Å²) (Table 1), but the side chains were almost invariably disordered, consistent with the selectivity studies (14, 18, 19) that imply the absence of defined side-chain binding sites in the S’ subsites. The occupancy of BCM7 in the S subsites was concomitantly observed to decrease on soaking.

For most dipeptides (see below for exceptions), apparently continuous electron density was observed that linked the amino terminus of the dipeptide, the side-chain oxygen of Ser-195, and the COOH-terminal carbon of BCM7 (Figs. 1a and 2, a–d). In each case, the unrestrained electron density surrounding the carbon derived from the carbonyl group of the isoleucine of BCM7 was apparently tetrahedral (with R stereochemistry), implying sp³ hybridization. Given the reproducibility of the observation of this species, BCM7 and the dipeptide main chain atoms (but not the side chains) were refined into the structures. The observation of electron density that could be refined for a tetrahedral intermediate both with different crystals and different dipeptides, indicates that the observations of an apparent tetrahedral intermediate are not artifactual. For example, two separate crystals both soaked with Arg-Phe, one for 45 min and the other for 50 min (Figs. 1a and 2, a and d, respectively) demonstrated similar very strong electron density in the region of Ser-195.

In all cases, the assigned oxyanion was located in the oxyanion hole within hydrogen bonding distance of the amide nitrogen of Gly-193 and Ser-195 (3.2 and 2.4 Å, respectively). Significantly, an N–H–N hydrogen bond between the nitrogen atoms (but not the side chains) were refined into the structures. The observation density that could be refined for a tetrahedral intermediate both with different crystals and different dipeptides, indicates that the observations of an apparent tetrahedral intermediate are not artifactual. For example, two separate crystals both soaked with Arg-Phe, one for 45 min and the other for 50 min (Figs. 1a and 2, a and d, respectively) demonstrated similar very strong electron density in the region of Ser-195.

In all cases, the assigned oxyanion was located in the oxyanion hole within hydrogen bonding distance of the amide nitrogen of Gly-193 and Ser-195 (3.2 and 2.4 Å, respectively). Significantly, an N–H–N hydrogen bond between the α-amino nitrogen atom of the P1’ residue and the amide nitrogen of Gly-193 (2.7 Å) was apparent. This implies that the amide nitrogen of Gly-193 in these structures has a bifurcated hydrogen bond and that the lone pair of the nitrogen atom of the P1’ residue (which would consequently have S-stereochemistry) is involved in this hydrogen bond, whereas the hydrogen atom on the nitrogen projects toward Ser-195. The atoms surrounding the tetrahedral carbon atom (Oy of Ser-195, Co of BCM7 Ile, and the oxyanion) and the nitrogen atom of the P1’ residue (Co of the P1’ residue, a hydrogen atom, and a lone pair) lie in a fully staggered conformation.

The residues from BCM7 were located in almost identical position to that observed in the reported PPE-BCM7 acyl-enzyme structure obtained at pH 5, with good occupancy being observed for the P1 residue and poorer occupancy for the P2, P3, and P4 residues. A similar observation on the relative occupancy of the P1 residue compared with the P2–P4 residues was made in the previous crystallographic analyses on the PPE-BCM7 acyl-enzyme complex (8). The residues derived from the nucleophilic dipeptide were observed to be in a similar location as the P’ residues observed in the crystal structure of a complex between PPE and a 28-residue noncovalently bound and nonhydrolyzed peptide inhibitor that spanned the PPE active site (14). The location of the main chain of the dipeptides was clearly identifiable through the position of the carbonyl group of the peptide bond linking the two residues of the dipeptide. There was an apparent hydrogen bond between the amide nitrogen of the P2’ residue and the main chain carbonyl of Thr-41 but insufficient electron density for the side chains of the dipeptide to be refined. The effectiveness of the soaking to remove sulfate from the crystals was demonstrated by examining the average B factor of the conserved (among all PPE crystal structures) sulfate ion remote from the active site. In the PPE-BCM7 acyl-enzyme structure, the B factor was 17 Å². This increased to an average of 62 Å² for the four structures shown in Fig. 2.

Formation of an Apparent Enzyme-Substrate Complex—For one of the dipeptides (Lys-Ser), significantly different electron density from that described above was observed at a 60-min soaking time (at pH 5). This density did not apparently correspond to a tetrahedral species; nor was there continuous electron density extending from the side chain oxygen of Ser-195 (Fig. 1b). However, there was an apparent link between the carbonyl derived from the isoleucine of BCM7 and the NH₂-terminal α-amino group derived from the dipeptide lysine. The geometry was as expected for an enzyme-substrate intermediate except that the scissile peptide bond appeared to be in the cis configuration. The carbonyl oxygen of the putative new amide bond linking the BCM7 isoleucine and dipeptide lysine was located within the oxyanion hole (2.9 and 2.5 Å from the amide nitrogens of Gly-193 and Ser-195, respectively). The alcohol oxygen atom of the Ser-195 side chain was apparently well positioned for attack onto the nascent amide with an angle of ~100° onto the plane of the peptide bond, in accordance with the optimal geometry proposed by Bürgi et al. (20) for nucleophilic addition to a carbonyl group. When the geometry of the scissile bond was forcibly refined with the trans conformation, the atoms of new amide bond did not fit well within the “omit” map (Fig. 1c), and the difference density maps (Fig. 1d) showed negative density in the region of the trans peptide bond after refinement. When the apparent cis-ES complex was refined into data from the Arg-Phe (45-min) tetrahedral intermediate (Oy of Ser-195 was removed to make the map clearer), positive density could be observed in the region of the “tetrahedral” carbon and negative density in the region of the cis-amide bond (Fig. 1e).

The same PPE-BCM7 KS-derived crystal was then removed from the stream of cryocooled nitrogen gas, immersed in pH 9 buffer for 30 s, and refrozen, and an additional x-ray diffraction data set was collected. In this case, the electron density revealed a linkage between Ser-195 and the P1 isoleucine and P1’ lysine residues and closely resembled the tetrahedral species observed for the other dipeptide substrates (Fig. 3a).

pH Jump Experiments on the Assigned Tetrahedral Structures—A pH jump experiment was also carried out on the PPE-BCM7-RF crystal soaked for 50 min. After the first pH jump (pH 9 for 30 s), the electron density (Fig. 3b) was almost identical to that of the pH 5 structure (Fig. 2d), except for a slight reduction in occupancy for the dipeptide (Arg-Phe). After the second exposure to pH 9 buffer (28 s), the electron density

Intermediates in Serine Protease Catalysis
Intermediates in Serine Protease Catalysis
density at the active site had changed, and there was no longer continuous electron density for Ser-195 (Fig. 3c). In this case, the density appeared to resemble the putative enzyme-substrate complex observed in the PPE/BCM7-KS structure (Fig. 1b). From the $mF_o-DF_c$ "omit" map, the occupancy of the dipeptide appeared to have increased relative to the first pH jump, but the quality of the electron density for the peptides was noticeably worse in the $2mF_o-DF_c$ map, and hence they were not refined into the structure.

**BCM7 Analogues**—In order to test whether the observations were limited to BCM7 complexes, two analogues, AcNPI and NPI, were then studied. AcNPI is a competitive inhibitor of PPE ($K_i = 0.8$ mM), and an atomic resolution (0.95 Å) structure of AcNPI complexed with PPE confirmed that this peptide formed an almost identical acyl-enzyme intermediate as BCM7, including the presence of an apparently productively positioned water molecule, and demonstrated that the ester bond was not distorted from planarity (9). A PPE:AcNPI complex crystal was soaked with a saturated solution of Ala-Ala for 30 min using the same methodology as for the PPE-BCM7 crystals, and an x-ray diffraction data set was collected. The electron density (Fig. 4a) clearly showed a tetrahedral conformation around Ser-195 analogous to that observed in the PPE-BCM7-dipeptide crystals (Figs. 1a and 2, a–d). This result demonstrates that the formation of the tetrahedral intermediate is not dependent on the presence of a specific sequence for the peptide bound in the S subsites.

**NPI** was a very weak inhibitor of PPE, supporting the proposal that a positive charge at the P4 position reduces the binding affinity to the active site; the drop in the IC$_{50}$ value from 25 mM to 7.5 mM after a 5-min preincubation of the enzyme was suggestive of a slow binding process. PPE was co-crystallized with NPI, and an x-ray diffraction data set was collected. The electron density implied that there was no

![FIGURE 1. Stereoviews of the active site of PPE complexed with BCM7 and Arg-Phe (soaked for 45 min) (a) and Lys-Ser (soaked for 60 min) (b) at pH 5, showing the refined "tetrahedral structure" and "cis-amide enzyme-substrate complex," respectively. The atoms of the enzyme are shown in green, and the atoms of the peptides are shown in gold. All atoms shown in gray were modeled but not refined against crystallographic data (see "Materials and Methods"). The $2mF_o-DF_c$ electron density maps for the enzyme are shown in beige and contoured at $1.5\sigma$. The $mF_o-DF_c$ "omit" electron density maps for the peptides are shown in blue and contoured at $2\sigma$. c, close-up stereoview showing the same "omit" map as displayed for b and overlaid with the Lys-Ser E+S complex forcibly refined in the trans conformation (shown in purple). This trans-amide conformation is similar to that found in the hybrid squash inhibitor-PPE structure (14). d, Close-up stereoview showing the $mF_o-DF_c$ difference density maps (contoured at $\pm 2.8\sigma$) after refinement of the trans-Lys-Ser E+S complex (shown in purple) into the data from the Lys-Ser structure (b). e, close-up stereoview showing the difference density maps (contoured at $\pm 3\sigma$) after refinement of the cis-E+S complex (shown in purple) from the Lys-Ser structure (b) into the data from the Arg-Phe structure (a). Positive density is shown in blue, and negative density is shown in red. The resolution of all maps is 1.7 Å.](#)

**FIGURE 3.** Views of the active site of PPE complexed with BCM7 and different dipeptides at pH 5 showing the refined tetrahedral structures for Arg-Phe (soaked for 45 min) (a), Asp-Phe (soaked for 30 min) (b), Lys-Ala-NH$_2$ (soaked for 30 min) (c), and a second Arg-Phe crystal (soaked for 50 min) (d). The color scheme and contouring levels for the atoms and maps are as in Fig. 1a. The resolution of the structures is 1.6 Å (except for a, which is 1.7 Å).
covalent linkage between Ser-195 and the COOH-terminal carboxylate carbon of the isoleucine of NPI, and the conformation of the Ser-195 side chain was identical to that in native PPE (see Fig. 6e). As with the pH jump studies, electron density for the isoleucine of NPI was clearly visible, but the P2 proline and P3 asparagine residues were in lower occupancy. Although this compound only binds poorly to PPE, this data set may reflect a position adopted by a COOH-terminal carboxylate product in the PPE active site; such an intermediate was not observed in the previous work on the deacylation of the acyl-enzyme. This difference reveals that there are subtle differences between peptides in their interactions with the PPE active site and that carboxylate product release may in fact be more efficient with (some) longer peptides, perhaps reflecting the fact that many or most of the PPE substrates in vivo are likely to be proteins.

We then soaked the PPE-NPI crystals with Ala-Ala at pH 5 and collected an x-ray diffraction data set. As before, the COOH terminus of NPI was not covalently linked to Ser-195, and the dipeptide was clearly visible in the S1’–S2’ subsites in reasonably good occupancy. However, it was bound in the reverse direction (C to N rather than N to C) to the other dipeptides examined (Fig. 4b). The clear location of the carbonyl group of the peptide bond between the two alanines allowed assignment of the dipeptide as binding (at least predominantly) in the reverse direction. “Reverse binding” of peptides in the PPE active site has been observed previously (21), and therefore it seems that peptides containing amino acids with small side chains are flexible with regard to their binding mode in the S′′ subsites. The fact that no tetrahedral intermediate was observed implies the necessity for formation of a productive acyl-enzyme intermediate before nucleophilic attack by the dipeptide can occur.

**α₁-Antitrypsin-derived Peptide**—

α₁-Antitrypsin, a serpin, is a natural inhibitor of elastase (2). The reactive center loop in α₁-antitrypsin is cleaved by elastase, leading to a large conformational change and the formation of a stable complex that inactivates the protease. Mass spectrometric and kinetic analyses have shown that the reactive center loop sequence MFLEAIPM, corresponding to residues P8–P1, forms a stable complex with PPE and gives rise to apparent uncompetitive inhibition (22). Systematic truncation of this peptide demonstrated that a shorter tetrapeptide MFLE (residues P8–P5 of the reactive center loop) retained uncompetitive inhibitory activity ($K_i = 0.37 \text{ mM}$). The MFLE tetrapeptide was then co-crystallized with PPE. The electron density in the active site (Fig. 4c) was also consistent with the formation of an enzyme-substrate complex with the methionine being located in the S1 subsite and FLE located in the S1’–S3’ subsites. The P1 methionine residue of MFLE was located further out of the S1 subsite than the corresponding Ile residue in the other structures, possibly due to the longer length of Met compared with Ile; density corresponding to the methylene carbon atoms of the methionine side chain was visible. Significantly, the conformation of the amino acids of the inhibitor was similar to that obtained for the BCM7-KS E:S complex, including the apparent cis configuration of the scissile P1–P1’ amide bond, indicating that such a configuration can be accessed by binding of an intact
peptide as well as by reaction of an acyl-enzyme intermediate with a dipeptide (e.g. KS).

**DISCUSSION**

**Tetrahedral Intermediates**—Spectroscopic evidence for oxyanionic tetrahedral intermediates in the serine protease mechanism has been reported (for a review, see Ref. 23). Several crystallographic studies have examined the conformation of various tetrahedral adducts formed by the reaction of unnatural substrate analogues and inhibitors. A crystal structure of γ-chymotrypsin showed that a peptide (an autolysis product, assigned as Pro-Gly-Ala-Tyr) had bound to the S subsites and formed an acyl-enzyme intermediate (24). When this experiment was repeated in hexane, rather than aqueous solution, the density around Ser-195 appeared tetrahedral and was modeled as such (25). However, the identity of the likely peptide(s) in the S’ subsites was unknown; based on analysis of the electron density, an aspartyl residue was modeled into the S1’ subsite. It was suggested that multiple peptides may have existed within the crystallization buffer, several of which then reacted with the acyl-enzyme intermediate. The conformation of the assigned chymotrypsin tetrahedral carbon is similar to that in the structures reported here, but in the chymotrypsin study, the refined angles around the tetrahedral carbon are distorted, with bond angles of 87 and 121°. The aspartic acid residue that was built into the S1’ subsite of γ-chymotrypsin also appears not optimally positioned for productive binding with the side chain lying in the location of the main chain of the P’ peptide. In contrast, it is likely that the reproducibly obtained structures presented here, with defined peptides in both the S and S’ subsites, more closely resemble the structure of the first tetrahedral intermediate.

A comparison of the structures proposed as closely reflecting the first tetrahedral intermediate (TI) described here with the putative structure of the second TI that was previously described (8) (Fig. 6d), indicate an almost identical conformation around the tetrahedral carbon, with the oxyanion occupying a similar position in the oxyanion hole. In superimposed structures, the two oxygen atoms bound to the central carbon in the tetrahedral species reported here (“first TI”), and the oxygen atom and nitrogen atom in the previously reported tetrahedral species (“second TI”) are both approximately equidistant from the carbonyl oxygen of the acyl-enzyme intermediate. The evidence for the tetrahedral species presented here is more extensive than that obtained for the TI in the hydrolysis step, where it is likely that more than one species was present, but the close resemblance between the two structures supports the previous assignment.

The distance of the assigned oxyanion to the amide nitrogen of Gly-193 is longer (3.1–3.2 Å in the first TI and 2.9 Å in the second TI) than that to the amide nitrogen of Ser-195 (2.4–2.5 Å in the first TI and 2.6 Å in the second TI). The distance to the amide nitrogen of Ser-195 is longer in the acyl-enzyme complex (2.8 Å) relative to the TIs probably in part because the angle Oγ-C=O (sp² hybridization) is larger than Oγ-C-O’ (sp³ hybridization). In each case, the relatively short distance between the oxyanions in the tetrahedral intermediates and the amide nitrogen of Ser-
Intermediates in Serine Protease Catalysis

**FIGURE 5.** Two proposed forms of the first tetrahedral intermediate (A and B) interchangeable via N-inversion and/or rotation about the C–N bond. In A, the P1' nitrogen lone pair projects away from His-57, and in B it projects toward it. In intermediate B, R and R' refer to either hydrogen or the P1' residue, depending on whether N-inversion has occurred. The amide nitrogen of Ser-195, which forms part of the oxyanion hole, is not shown for clarity; from the viewpoint shown, it lies behind the plane of the picture. Note that steric constraints mean that the conformation in which the nitrogen lone pair is exactly coplanar with the C–O(Ser-195) bond is unlikely.

195 compared with that of Gly-193 is notable (26). Since the N–H–N hydrogen bond between Gly-193 and the P1' nitrogen atom has better geometry than the N–H–O hydrogen bond between Gly-193 and the tetrahedral oxyanion, it is possible that the latter has relatively less significance in stabilizing the oxyanion. The fact that the oxygen atom of the carbonyl group of the ES complex (like the acyl-enzyme intermediate) can also form good hydrogen bonds to these two amide nitrogens supports the suggestion that polarization of the amide or ester carbonyl and consequent electrophiilic catalysis is likely to be an important driving force for the mechanism. This is in agreement with the view of Warshel and Russell (27) that electrostatic interactions are of prime importance in the serine protease mechanism.

Despite the collection of many data sets, there was no clear evidence for a productively bound PPE-BCM7 acyl-enzyme intermediate with unreacted dipeptide present in the S' subsites. This suggests that under these soaking conditions in the presence of excess dipeptide, the equilibrium of the addition reaction in the PPE crystals lies toward the TI rather than the unreacted dipeptide. The acyl-enzyme intermediates formed from BCM7 or AcNPI appear stable at pH 5 in the crystal except when a peptide substrate is present. In the latter case, the first TI appears to be the predominant structure, suggesting that nucleophilic addition by the NH₂-terminal peptide amine occurs more readily than that by the hydrolytic water molecule. Although the majority of the structures for the assigned tetrahedral intermediate were obtained at pH 5, the pH-jumped structures of PPE-BCM7-KS and PPE-BCM7-RF (Fig. 3, a and b) suggest that a similar conformation may occur at higher pH.

**Conformational Changes during Catalysis**—It has been suggested that during the course of serine protease catalysis in solution there may be significant movements in some regions of the active site, particularly the side chains of His-57 and Ser-195 (28), and recent high resolution work on a protease-inhibitor complex suggested slight movements in the position of His-57 between the acyl-enzyme and ES complexes (29). Previous crystallographic work on the deacylation step implied only a small movement for the side chain of His-57 during deacylation (8); movement of His-57 was coupled via a hydrogen bond network (His-57–Asp-102–Ser-214) to a small movement in the Ser-214–Val-216 β-strand away from the peptide substrate, thus potentially weakening the hydrogen bonding interactions that bind the substrate. This is a possible mechanism for synchronizing hydrolysis of the acyl-enzyme intermediate with product release, and support for this proposal has come from modeling studies (30). The refined pH jump technique in the current work enabled use of the same crystal for data collection before and after the change in pH, allowing possible conformational changes in the enzyme to be observed more clearly.

Although at the resolution of the reported structures, small differences may not be apparent, the results indicate that the overall conformational changes, particularly in the region of the active site, between the acyl-enzyme intermediate, first TI, and ES complex are small. Unlike in the pH jump experiments with acylating lactam inhibitors (31), no significant movement of His-57 was observed. Further, the hydrogen bonding patterns argue against the proposed ring flip of the imidazole ring of His-57 (28). Overall, the observations support the concept of a relatively rigid active site architecture into which a more flexible substrate (at least in the region of the scissile bond) can bind; catalysis is thus apparently mediated by relatively small structural changes.

**Comparison with Transition State Analogue Inhibitors**—Comparison of the conformation of the assigned tetrahedral species described in this paper with the structures of serine proteases with various proposed intermediate analogues reveals differences. The structure of PPE complexed with a peptidyl boronic acid (32) showed that its peptide residues were located in the S subsites in an almost identical location to the residues of BCM7 in the PPE-BCM7 structures. One of the oxygen atoms of the boronic acid is located in the oxyanion hole, as observed for the PPE-BCM7 acyl-enzyme structure. The other oxygen atom of the boronic acid is much closer to the side chain of His-57 than is the P1'-derived nitrogen in the TI structures reported here. The crystal structure of a complex between chymotrypsin and a peptidyl trifluoromethyl ketone inhibitor (33, 34) also revealed a tetrahedral structure adjacent to Ser-195. The hemiketal oxygen atoms in this structure are very close to the location of the boronic acid oxygen atoms in the PPE complex. A similar tetrahedral conformation was observed in the complexes between chymostatin and *Streptomyces griseus* protease A (35), a peptidyl boronic acid inhibitor and chymotrypsin (36), and leupeptin and trypsin (37). The fact that in the TI structures described here the oxygen atoms (or oxygen and nitrogen atoms) are equidistant between the carbonyl oxygen in the acyl-enzyme intermediate gives a new view of the conformation of the tetrahedral intermediate.

**Geometry of the Scissile Amide Bond**—The possibility of a cis-peptide conformation being preferred (at least in some cases) for the scissile amide in the enzyme-substrate complex
FIGURE 6. Left, stereoviews of the active site of PPE, showing the view from crystal structures that potentially reflect the intermediates. a, the apparent enzyme-substrate complex, PPE-BCM7-Lys-Ser at pH 5, the same as in Fig. 1b. b, the apparent tetrahedral structure from PPE-BCM7-Lys-Ala-NH$_2$ at pH 5, the same as in Fig. 2c. c, the acyl-enzyme complex (8), PPE-BCM7 at pH 5 (Protein Data Bank code 1HAX). d, the apparent tetrahedral structure from previous work (8), PPE-BCM7 jumped to pH 9 for 1 min (Protein Data Bank code 1HAZ). e, PPE-NPI at pH 5 as a possible analogue of an enzyme-product complex. The color scheme and contouring levels for the atoms and maps are as in Fig. 1a, except that the contour level of the $mF_o$-$DF$ "omit" electron density maps for c and d was 2.5 $\sigma$. Right, serine protease mechanism showing the hydrogen bonds (dashed lines) to the amide nitrogens of Gly-193 and Ser-195 forming the oxyanion hole and between $\text{Ne}_2$ of His-57 and Ser-195 and the hydrolytic water molecule.
Intermediates in Serine Protease Catalysis

has not been widely discussed (aside from the possibility of a strained \( \omega \) dihedral in a D-Ala-D-Ala peptidase (38)). It is likely that the binding energy of a peptide or polypeptide substrate to the active site would be sufficient to enable a trans/cis isomerization (\(~10 \text{ kJ mol}^{-1}\) for Xaa-non-Pro bonds (39)). It also seems probable that an ES complex in which the scissile bond is cis would be closer in structure to the conformation of the first tetrahedral species observed in this work than the corresponding trans configuration (Fig. 1c). Previous structures of peptide inhibitors, which form stable noncovalent complexes with serine proteases, have an intact peptide bond in a trans conformation spanning the S and S’ subsites of the active site (14), including the high resolution structure of a trypsin inhibitor complex (29). It is possible that these peptides are inhibitors in part because the conformation of their scissile amides is bound in a nonproductive trans form. It is notable that \( \beta \)-lactams, which contain a conformationally restrained amide bond, are inhibitors of a variety of nucleophilic serine enzymes, including the D-Ala-D-Ala transpeptidases of bacterial cell wall biosynthesis and serine proteases. Also of note is a crystal structure of the Gyra intein (not complexed with a protease), which revealed a cis conformation at its scissile peptide bond (40).

Presently it is unclear precisely why different peptides may bind as trans or cis amides at the position of the scissile bond. The available evidence suggests that residues other than those at the P1 and P1’ positions are involved in determining this variable. The Bowman-Birk inhibitors have a relatively rigid reactive site loop that adopts a fixed “canonical” conformation with a trans peptide bond at the scissile bond position. They also contain a highly conserved cis-proline at the P3’ position that is necessary for biological activity (41). This rigidity may prevent trans/cis isomerization of the scissile peptide bond and hence allow these proteins to function as inhibitors.

It should be noted that there is no evidence presented here to imply that the cis configuration of the scissile amide is necessarily the only productive route; it has been argued that catalysis can proceed via more than one stereoelectronically viable arrangement (42). However, assuming that cis-trans interconversion does not occur during the experiments, the pH jump experiments do imply that the observed conformation for the enzyme-amide complex is productive, at least in the crystalline state used here. This preference for a cis amide may also occur in the case of ester substrates where the cis/trans isomerization has a lower energy barrier (the thermodynamically favored (Z)-conformation for esters is analogous to a trans-amide bond).

A comparison of the structure of the putative enzyme-substrate complex and acyl-enzyme intermediate shows a significant difference in the location of the amide and ester carbonyl groups, respectively. However, the apparent angle of nucleophilic attack suggested by the structures is almost the same (close to the Bürgi-Dunitz angle of 100°), implying that there is some flexibility in achieving this stereoelectronically preferred arrangement in the active site, perhaps reflecting the fact that in vivo elastases probably are required to hydrolyze a range of substrates.

Stereo electronic Implications—Assuming the correct assignment of the hydrogen bond donor-acceptors between the P1’ nitrogen atom of the first TI and the amide nitrogen of Gly-193, the lone pair of the P1’ nitrogen faces away from His-57. This arrangement is predicted by Deslongchamps’ stereoelectronic theory (43); following nucleophilic attack of Ser-195, the nitrogen lone pair will be antiperiplanar to the C–O bond between Oy of Ser-195 and the tetrahedral carbon (Fig. 5). In this conformation, it is probably not possible for the leaving group P1’ nitrogen to accept a proton directly from Ne2 of His-57, which in the accepted mechanism has been protonated upon formation of the first TI. One possibility is that a solvent molecule donates a proton to the nitrogen atom (4). However, it seems more likely that if this conformation is on the catalytic path, then the bonds about the nitrogen rearrange (by inversion and/or rotation) such that the nitrogen lone pair is directed toward Ne2 of His-57 as shown in Fig. 5. As well as promoting C–N bond cleavage by protonation, this arrangement would relatively disfavor C–O cleavage compared with C–N cleavage, since the nitrogen lone pair is not antiperiplanar with respect to the C–O(Ser-195) bond. It should be noted that even after rearrangement, the distance between the P1’ nitrogen and Ne2 of His-57 may be too large for hydrogen bond formation; in this case, indirect proton transfer, possibly via Oy of Ser-195, may be required.

Although in some nonoptimal cases, mechanistic variants may occur (e.g. the slow decylation rates observed with certain stable acyl-enzyme complexes (44, 45)), the current structures (Fig. 6) support an optimal mechanism for efficient serine protease catalysis in which every step occurs in a stereoelectronically favored manner and in which there is minimal movement of the histidine residue that enables general acid/base catalysis. These mechanistic proposals also have consequences for inhibitor design, since peptides that bind in a stable trans configuration may be regarded as high energy intermediate analogues that do not form a tetrahedral intermediate. It is also notable that the activity of many naturally occurring serine proteases is regulated by peptides/proteins without resorting to the non-peptidic functional groups (e.g. boronic acids and trifluoromethylketones) commonly used in man-made inhibitors of serine proteases; a factor in the inhibitory mechanisms of these natural compounds may be the configuration of the amides that span the active site.

REFERENCES

1. Steinmetzer, T., and Sturzebecher, J. (2004) Curr. Med. Chem. 11, 2297–2321
2. Stoller, J. K., and Aboussouan, L. S. (2005) Lancet 365, 2225–2236
3. Ding, X., Rasmussen, B. F., Petsko, G. A., and Ringe, D. (1994) Biochim. Biophys. Acta 1205, 9285–9293
4. Bizzozero, S. A., and Butler, H. (1981) Bioorg. Chem. 10, 46–62
5. Dixon, M. M., Brennan, R. G., and Matthews, B. W. (1991) Int. J. Biol. Macromol. 13, 89–96
6. Blanchard, H., and James, M. N. G. (1994) J. Mol. Biol. 241, 574–587
7. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) Nat. Struct. Biol. 4, 456–461
8. Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J., and Hajdu, J. (2001) Nat. Struct. Biol. 8, 689–694
9. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) I. Biol. Chem. 277, 21962–21970
10. Schneider, T. R. (2002) Acta Crystallogr. Sect. D 58, 195–208
Intermediate in Serine Protease Catalysis

11. Bailey, S. (1994) *Acta Crystallogr. Sect. D* **50**, 760–763
12. Meyer, E., Cole, G., Radhakrishnan, R., and Epp, O. (1988) *Acta Crystallogr. Sect. B* **44**, 247–254
13. Jones, A. T., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
14. Wright, P. A., Wilmouth, R. C., Clifton, I. J., and Schofield, C. J. (2000) *Biochem. J.* **351**, 335–340
15. Brünger, A. T. (1992) *Nature* **355**, 472–475
16. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA
17. Wright, P. A., Wilmouth, R. C., Clifton, I. J., and Schofield, C. J. (2000) *Biochem. J.* **351**, 247–254
18. Yennawar, N. H., Yennawar, H. P., and Farber, G. K. (1994) *Biochemistry* **33**, 7326–7336
19. Gerlt, J. A., Kreevoy, M. M., Cleland, W. W., and Frey, P. A. (1997) *Chem. Biol.* **4**, 259–267
20. Warshel, A., and Russell, S. (1986) *J. Am. Chem. Soc.* **108**, 6569–6579
21. 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) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10371–10376
22. Fodor, K., Harmat, V., Neutze, R., Szilágyi, L., Graf, L., and Katona, G. (2006) *Biochemistry* **45**, 2114–2121
23. Topf, M., and Richards, W. G. (2004) *J. Am. Chem. Soc.* **126**, 14631–14641
24. Wright, P. A., Wilmouth, R. C., Clifton, I. J., and Schofield, C. J. (2000) *Biochem. J.* **351**, 335–340
25. Takahashi, L. H., Radhakrishnan, R., Rosenfield, R. E., and Meyer, E. F. (1989) *Biochemistry* **28**, 7610–7617
26. Brady, K., Wei, A., Ringe, D., and Abeles, R. H. (1990) *Biochemistry* **29**, 7600–7607
27. Neidhart, D., Wei, Y., Cassidy, C., Lin, J., Cleland, W., and Frey, P. A. (2001) *Biochemistry* **40**, 2439–2447
28. Debaere, L. T. J., and Brayer, G. D. (1985) *J. Mol. Biol.* **183**, 89–103
29. Bone, R., Shenvi, A. B., Kettner, C. A., and Agard, D. A. (1987) *Biochemistry* **26**, 7609–7614
30. Kurinov, I. V., and Harrison, R. W. (1996) *Protein Sci.* **5**, 752–758
31. Kelly, J. A., Knox, J. R., Zhao, H. C., Frère, J. M., and Ghysen, J. M. (1989) *J. Mol. Biol.* **209**, 281–295
32. Stewart, D. E., Sarkar, A., and Wampler, J. E. (1990) *J. Mol. Biol.* **214**, 253–260
33. Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R., and Sacchettini, J. C. (1998) *Nat. Struct. Biol.* **5**, 31–36
34. Brauer, A. B. E., Domingo, G. J., Cooke, R. M., Matthews, S. J., and Leatherbarrow, R. J. (2002) *Biochemistry* **41**, 10608–10615
35. Topf, M., Varnai, P., Schofield, C. J., and Richards, W. G. (2002) *Proteins* **47**, 357–369
36. Deslongchamps, P. (1975) *Tetrahedron* **31**, 2463–2490
37. Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D. W., Clifton, I. J., Pritchard, G. J., Aplin, R. T., and Schofield, C. J. (1998) *Biochemistry* **37**, 17506–17513
38. Knight, W. B., Chabin, R., and Green, B. (1992) *Arch. Biochem. Biophys.* **296**, 704–708