Crystal Structures of Two Engineered Thiol Trypsins†

Mary E. McGrath,‡ Marjorie E. Wilke,‡ Jeffrey N. Higaki,§ Charles S. Craik,‡,§ and Robert J. Fletterick*,‡

Department of Biochemistry and Biophysics and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0448

Received May 15, 1989; Revised Manuscript Received July 27, 1989

ABSTRACT: We have determined the three-dimensional structures of engineered rat trypsins which mimic the active sites of two classes of cysteine proteases. The catalytic serine was replaced with cysteine (S195C) to test the ability of sulfur to function as a nucleophile in a serine protease environment. This variant mimics the cysteine trypsin class of thiol proteases. An additional mutation of the active site aspartate to an asparagine (D102N) created the catalytic triad of the papain-type cysteine proteases. Rat trypsins S195C and D102N,S195C were solved to 2.5 and 2.0 Å, respectively. The refined structures were analyzed to determine the structural basis for the 10^6-fold loss of activity of trypsin S195C and the 10^8-fold loss of activity of trypsin D102N,S195C, relative to rat trypsin. The active site thiols were found in a reduced state in contrast to the oxidized thiols found in previous thiol protease structures. These are the first reported structures of serine proteases with the catalytic centers of sulfhydryl proteases. Structure analysis revealed only subtle global changes in enzyme conformation. The substrate binding pocket is unaltered, and active site amino acid 102 forms hydrogen bonds to H57 and S214 as well as to the backbone amides of A56 and H57. In trypsin S195C, D102 is a hydrogen-bond acceptor for H57 which allows the other imidazole nitrogen to function as a base during catalysis. In trypsin D102N,S195C, the asparagine at position 102 is a hydrogen-bond donor to H57 which places a proton on the imidazole nitrogen proximal to the nucleophile. This tautomer of H57 is unable to act as a base in catalysis. The 100-fold diminished activity of trypsin D102N,S195C compared to trypsin S195C is most likely due to stabilization of the incorrect tautomer of H57, which is unable to form the critical thiolate–imidazolium ion pair with C195-Sγ, and to altered electrostatics at the catalytic site due to removal of the negative charge from residue 102. H57 has shifted by 0.2 Å but remains within hydrogen-bonding distance of D/N102 (2.7 Å) and S195 (3.5 Å). The sulfur nucleophile of C195 is larger than the oxygen it replaces but is not sterically hindered. In both structures, C195-Sγ may also be able to form hydrogen bonds with D193-N (3.7 Å), which is part of the oxyanion hole. The sulfur atom points away from the binding pocket compared to the oxygen atom in trypsin, and the strand of peptide chain which includes residue 193 is closer to the side chain of residue 195 in rat trypsin than it is in cow trypsin. This causes occlusion of the oxyanion hole which could account for the inefficacy of these enzymes. Other possible contributing factors to the reduced activity of the two enzymes are discussed.

Proteases mediate a vast array of biological functions including blood clotting, digestion, and fertilization. Serine and cysteine proteases are two important classes of these catalysts and are eponymously named for the amino acid which serves as the active site nucleophile. The two protease types are markedly different in amino acid sequence and tertiary structure but share a common arrangement of active site nucleophile and histidine, with aspartate in serine proteases and asparagine in cysteine proteases (Husain & Lowe, 1968). Recently, a class of viral cysteine proteases has been found to be homologous to trypsin-type serine proteases by sequence analysis (Bazan & Fletterick, 1988). The viral enzymes have a cysteine nucleophile and active site histidine and aspartate (Dougherty et al., 1989) and presumably feature the tertiary fold of the trypsin-type serine proteases (Bazan & Fletterick, 1988). Three-dimensional structures of these Cys trypsins are not yet available.

The more familiar subclass of cysteine proteases is represented by papain, which possesses a completely different tertiary fold from that of the serine proteases. However, structure comparisons have demonstrated that the side chains of the three active site residues of chymotrypsin (serine type) and papain (cysteine type) are virtually superimposable (0.8-Å rms deviation) (Garavito et al., 1977; Drenth et al., 1976). Even before this similarity in tertiary structure was known, the serine oxygen of subtilisin was converted chemically to a sulfur to mimic the thiol proteases. It was predicted that the chemical similarity between sulfur and oxygen would allow sulfur to function in the context of a serine protease active site. Sulfur is generally more nucleophilic than oxygen, and the lower pK_a of a sulfhydryl relative to that of a hydroxyl should ensure a higher concentration of active nucleophile in solution at physiological pH.

Although chymotrypsin was originally chosen to test the role of the serine nucleophile, by converting serine to alanine (Weiner et al., 1966), subtilisin became the serine protease of choice to determine if sulfur could replace oxygen as the nucleophile. This is because subtilisin has no disulfides which could be altered by chemical modification of oxygen to sulfur. Thiol subtilisin was shown to be inactive on esters and on amide substrates of subtilisin and showed low activity toward the activated ester substrate nitrophenyl acetate (Neet & Koshland, 1966; Polgar & Bender, 1966). Since formation of the covalent enzyme–substrate intermediate (acyl enzyme) was observed only with activated substrates, acylation was most likely the rate-limiting step (Polgar & Bender, 1967). Dea-

† This work was supported by NIH Grant DK 39304 to R.J.F. and NSF Grant GM 8608086 to C.S.C. J.N.H. was a recipient of NIH Postdoctoral Fellowship (National Research Service Award) GM 11598.

* Address correspondence to this author.

‡ Department of Biochemistry and Biophysics.

§ Department of Pharmaceutical Chemistry.

0006-2960/89/0428-9264$01.50/0

cylation of the enzyme was reduced by only 100-fold. Because the chemical differences between sulfur and oxygen seemed too small to account for this inactivity, it was thought that a conformational change at the active site was responsible, perhaps resulting from the larger atomic radius of sulfur and the altered bond angle between the sulfur atom and neighboring active site residues (Polgar & Bender, 1967). However, an array of experiments, including optical rotatory dispersion and binding assays, failed to reveal any structural anomalies (Neet et al., 1968). The integrity of the active site was further probed with highly specific inhibitors and showed that inhibition rates were indistinguishable between subtilisin and thiol subtilisin (Tsai & Bender, 1979). These experiments indicated an intact catalytic apparatus but could not directly assay active site conformation. Crystallographic studies on thiol subtilisin at medium resolution (Alden et al., 1970) revealed no noticeable conformational changes, but this result may have been complicated by oxidation of the new thiol group. In fact, all structure determinations of cysteine proteases, to date, including high-resolution studies of papain (Kamphius et al., 1985) and actinidin (Baker, 1981), show oxidation of the catalytic thiol. This obscures interpretation of active site geometry since one must determine where the atoms would be if no extraneous oxygens were present and hinders comparisons between the thiol and serine proteases since the latter cannot be similarly oxidized.

Trypsins S195C and D102N,S195C are catalytically active [see accompanying paper (Higaki et al., 1989)] and not oxidized. Our work is the first which specifically and directly addresses the structural integrity of an active site cysteine in a serine protease. The single variant mimics the Cys trypsins in topology and active site, while the double variant was designed to test the efficacy of a papain-type active site in trypsin.

Experimental Procedures

Rat trypsins S195C and D102N,S195C were constructed as described previously (Higaki et al., 1987) and expressed in *Escherichia coli* (Evnin & Craik, 1988; Vasquez et al., 1989; Higaki et al., 1989). Both variants were isolated (Higaki et al., 1989) and were further purified by anion exchange HPLC subsequent to affinity chromatography. Approximately 2 mg of protein was then exchanged into 50 mM Tris, pH 8.0, 10 mM $CaCl_2$, and 10 mg/mL benzamidine and concentrated to 30 mg/mL (McGrath et al., 1989). Hanging-drop vapor diffusion (Hampel et al., 1968; McPherson, 1982) against 22% $MgSO_4$ produced diffraction-quality cubic crystals (space group $I23$, a = 124.4 Å) in approximately 1 week. Crystals used in data collection were 0.2–0.6 mm on an edge. Further details of structure solution and refinement pertain to trypsin D102N,S195C specifically. The trypsin S195C data were treated similarly. Diffraction data were recorded to 2.0-Å resolution with monochromatic copper Kα radiation on a Xentronics area detector (Table I). A total of 720 frames of data were collected, with an ω scan angle of 0.25° per frame. The scan time was 200 s per frame. The XENGEN software package, from Nicolet Instrument Co., was used for data reduction.

Structure Solution and Refinement. Molecular replacement provided phase information for structure determination. It was facilitated by use of a model from the trypsin D102N structure (Sprang et al., 1987). The cubic Laue group caused ambiguity in assigning a, b, and c axes, but once the proper hand was found, the data set scaled to the model at 27%. Since the orientation of active site residues was of interest, N102, H57, and C195 side chains (chymotrypsin numbering system; Hartley, 1964), as well as the solvent, benzamidine, and calcium molecules, were removed from the model. The resulting F_c values were refined (Hendrickson & Konnert, 1980) against the 12.0–2.3-Å data, for four cycles before an $F_o - F_c$ difference electron density map was calculated to locate the omitted residues. The missing residues were fitted to the density (Figure 1) with FRODO (Jones, 1981), and the new model, containing the correctly positioned catalytic triad and the benzamidine, calcium, and 90 solvent molecules, was subjected to 36 cycles of refinement with 12.0–2.2-Å data. At this point the 2.2–2.0-Å data were included. Adjustments to main-chain and side-chain atoms were made by FRODO with a $2F_o - F_c$ map on the Evans and Sutherland PS330. More solvent molecules were added, and further manual adjustments of coordinates were made during the course of an additional 24 cycles of refinement, after which an $F_o - F_c$ difference map was essentially featureless.

Data from a single crystal of trypsin D102N,S195C provided 5.4-fold redundancy of observations, and 95.8% of reflections from infinity to 2.0 Å were collected. Of these, 94.4%

Table I: Crystal and Diffraction Data for Trypsins S195C and D102N,S195C

	D102N,S195C	S195C[a]
Crystal Data		
single crystal (mm)	0.6 × 0.6 × 0.2	0.2 × 0.2 × 0.1
space group	$I23$	$I23$
cell dimensions (Å)	a = 124.4	a = 124.4
molecules per asym unit	1	1
Diffraction Data		
data measurement	Xentronics	
resolution (Å)	2.0	2.5
total observations	115538	87945
observations > 2σ	109062	74998
unique reflections	21181	14691
R_{sym}[b]	0.12	0.15
Refinement Results		
refinement (cycles)	50	16
R_{cryst}[c]	0.20	0.16
resolution (Å)	8.0–2.0	8.0–2.5
rms difference		
bond (Å)	0.015	0.018
angle (deg)	2.6	3.4
no. of atoms	1792	1813
no. of solvent molecules	117	138

[a] Data were collected and merged from three crystals. [b] $R_{sym} = \{hi-[I(h)] - I_i(h)\}/hiI_i(h)$, where $[I(h)]$ is the mean of all measurements of reflections h and $I_i(h)$ is the ith measurement of reflection h. [c] $R = \sum|F_o - F_c|/\sum F_o$.

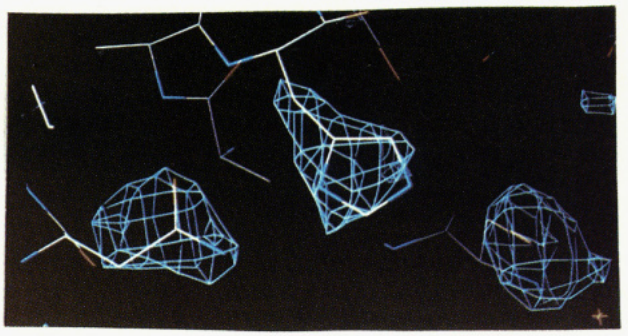

FIGURE 1: Difference Fourier map ($F_o - F_c$) at the catalytic site of trypsin D102N,S195C using 12.0–2.3-Å data. Side chains of amino acid residues H57, D102, and S195 had been removed from the model structure. The new model was refined against the observed data before a difference electron density map was calculated. The density for the three active site residues is therefore unbiased by the model at these positions. Electron density for N102, H57, and C195 was used to build in coordinates for the three residues, and these are also shown.

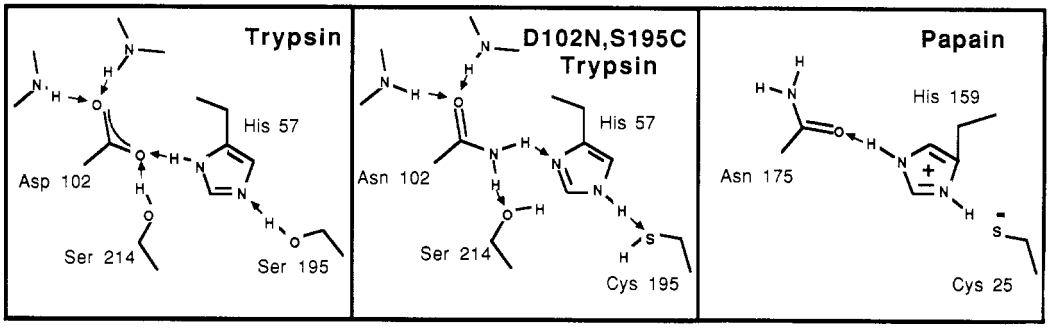

FIGURE 2: (Left) Hydrogen-bonding network at the active site of cow trypsin causes the H57 imidazole proton to be proximal to D102. This allows H57 to act as a base and abstract a proton from S195-Oγ during catalysis. (Center) In D102N,S195C trypsin, hydrogen bonds with the backbone amides of A56 and H57 force a reversal of the hydrogen-bonding network and stabilize a tautomer of H57 which carries a proton on Nϵ2 and cannot function as a base. (Right) In papain N175 at the active site forms only one hydrogen bond. Thus, it can act as a hydrogen-bond acceptor to H159 and draws the imidazole proton away from the nucleophile C25-Sγ enabling H159 to assist in catalysis. H159 interacts with N175 via H159-Nδ1, unlike the serine proteases which have the imidazole flipped over and thus use H57-Nϵ2. The tethering of the imidazole ring of H159 at Cβ and Nϵ2 in papain allows it to swivel in response to different oxidation states of C25-Sγ and different steps in the catalytic mechanism (Baker & Drenth, 1987).

of the reflections were greater in magnitude than twice the standard deviation (Table I).

Results and Discussion

Overall Structure and Substrate Binding Pocket. The 2.0- and 2.5-Å structures of S195C,D102N and S195C rat trypsins in space group *I*23 exhibit only minor deviations from the trypsin D102N structure in space group $P2_12_12_1$ (rms deviation = 0.3 Å), which has been shown to be almost identical with bovine trypsin (Sprang et al., 1987).

Folding and removal of the signal peptide have occurred correctly. Introduction of the cysteine results in 13 sulfhydryls which must be arranged to form six disulfides and one unpaired sulfhydryl. The mutation does not appear to interfere with folding of the enzyme. An *E. coli* signal peptidase has removed the signal sequence correctly since electron density for the N-terminal residue clearly shows it to be the expected isoleucine (Higaki et al., 1989). In trypsin, the nitrogen of this Ile forms an essential salt bridge with D194 which stabilizes the substrate binding pocket (Fehlhammer et al., 1977). Trypsinogen lacks this salt bridge since D194 forms a hydrogen bond with H57. This blocks the active site and inhibits substrate binding. Both variants exhibit the active trypsin conformation in which I16-N forms a 2.5-Å salt bridge with D194–Oδ1.

The integrity of the specificity pocket is evident in its ability to bind tightly to affinity resins (immobilized pancreatic trypsin inhibitor and benzamidine); in its K_m for Z-Gly-Pro-Arg–AMC, which is essentially the same as trypsin (Higaki et al., 1989); and in its three-dimensional structure. In trypsins S195C and D102N,S195C, the pocket is occupied by the inhibitor benzamidine. D189, which is located at the bottom of the substrate binding pocket and is a primary determinant of substrate specificity, forms a pair of hydrogen bonds with the amidinium group (N1 and N2) of benzamidine (N1–Oδ1 = 3.2 Å; N2–Oδ2 = 3.0 Å). In both structures, benzamidine also interacts with the carbonyl oxygens of binding pocket residues S190 and S214, Oγ of S214, and a water molecule, via N1 and N2. The contacts between benzamidine and residues lining the pocket are as previously described for bovine trypsin (Krieger et al., 1974; Bode & Schwager, 1975).

Catalytic Residues. The position of N102 in trypsins S195C and D102N,S195C is the same as found previously for this residue (Sprang et al., 1987). The N102 side chain forms hydrogen bonds of 2.8 and 2.9 Å with the backbone amides of A56 and H57, respectively (Table II and Figure 2). Since N102-Nδ2 and the amides are hydrogen-bond donors and cannot come within 3.4 Å of each other, N102-Oδ1 must point toward the amides. A consequence of this orientation is that N102-Nδ2 forms a hydrogen bond with H57. This shifts the imidazole proton from the nitrogen near N102 to the nitrogen near C195, forming a tautomer of histidine which is unable to act as a base in catalysis. In trypsin and S195C trypsin, one D102-Oδ forms hydrogen bonds with the backbone amides while the other Oδ accepts a hydrogen bond from H57. The resultant hydrogen-bonding network promotes the catalytically favored histidine tautomer with the lone pair of electrons near the nucleophile of residue 195. The inability of the histidine to function as a base has been postulated to account for the 10^4-fold loss of activity of the D102N variant (Sprang et al., 1987; Craik et al., 1987). Furthermore, the loss of the negative charge of D102 changes the electrostatic potential at the active site which probably hinders catalysis (Warshel et al., 1989; Soman et al., 1989). As in trypsin and trypsin S195C, residue 102 forms a hydrogen bond with S214-Oγ. In trypsin D102N,S195C the OH of D102 has been replaced with Nδ1, and the direction of the hydrogen bond is reversed. The two hydrogen bonds formed by N102-Nδ2 restrict its mobility and probably account for the low temperature factor of this atom (B = 5.0 Å^2).

In the papain-type sulfhydryl proteases, the only hydrogen bond formed by the asparagine is to the catalytic histidine (2.8 Å). This bond length is invariant in trypsin, trypsin S195C, trypsin D102N, and trypsin D102N,S195C, although its direction is reversed in the latter two. In trypsin, the D102–H57 bond maintains H57 in the catalytically correct conformer and tautomer (Sprang et al., 1987; Craik et al., 1987). Placement

Table II: Interatomic Distances at the Active Sites of Trypsin, Actinidin, and Trypsins S195C and D102N,S195C

amino acid residue	atoms	distance (Å)		
		D102N,-S195C[a]	trypsin	actinidin
N/D102	Oδ1–A56-N	2.8	2.9	
N/D102	Oδ1–H57-N	2.9	2.8	
N/D102	Nδ2–H57-Nδ2	2.7	2.7	2.8
N/D102	Nδ2–S214-Oγ	2.9	2.8	
H57	Nϵ2–C/S195-Sγ/Oγ	3.5	3.0	3.3
H57	Nϵ2–C/S195-Cβ	3.4	3.2	3.8
C/S195	benzamidine-C4	3.9	3.7	5.4
C/S195	APPA-C9[b]	0.9	1.7	
C/S195	S/Oγ–C/S195-N	3.1	3.1	
C/S195	S/Oγ–G193-N	3.7	4.7	

[a] The distances in S195C are all within 0.2 Å of those found for D102N, S195C. [b] APPA = (*p*-amidinophenyl)pyruvate.

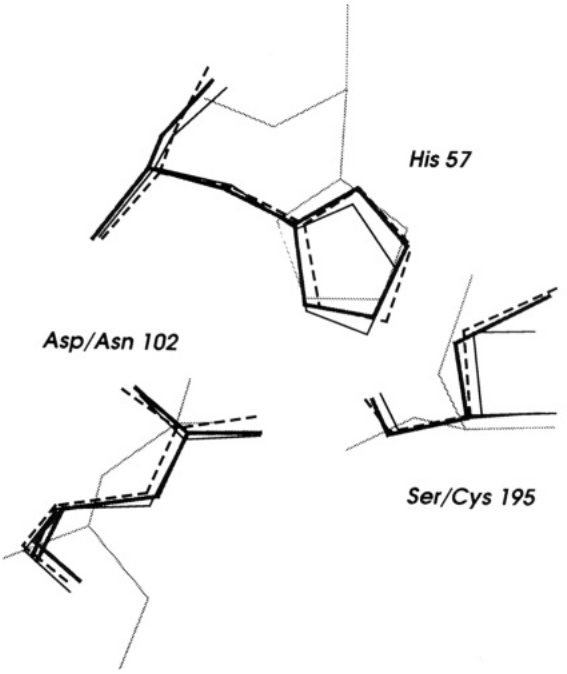

FIGURE 3: Superposition of the three active site residues of cow trypsin, actinidin, and trypsins S195C and D102N,S195C, using the program INSIGHT (Biosym, Inc.), shows that the active sites form very similar constellations. The nucleophilic C195-Sγ in both variants occupies an intermediate position between that found for C25-Sγ in actinidin and S195-Oγ in cow trypsin. The shaded, thin line represents actinidin; the solid black line is cow trypsin; the dashed line is trypsin D102N,S195C; and the heavy, shaded line is trypsin S195C.

of the larger sulfur atom next to the histidine in these variants has caused the plane of the ring to move away from the nucleophile by 0.2 Å (Figure 3). This shift does not affect the H57–N102 hydrogen-bond length and results in a H57–C195 distance of 3.5 Å. This is in good agreement with a study of 273 sulfur–nitrogen hydrogen bonds in various proteins which revealed that 25% of them were between 3.45 and 3.55 Å (L. Gregoret and S. Rader, unpublished results). Since the orientation of N102 forces H57-Nϵ2 to be a proton donor in the double variant, and the C195 sulfur is probably protonated, no hydrogen bond is possible between these two atoms. However, a hydrogen bond may form in trypsin S195C, since H57-Nϵ2 is a proton acceptor and the interatomic distance (3.5 Å) is the same as the sum of the van der Waals radii for sulfur and nitrogen. The distance is also reasonable for a hydrogen bond when compared to those in papain-type proteases (3.3 Å in actinidin; 3.6 Å in papain) (Kamphuis et al., 1985). The temperature factor (B) for H57 is greater than that of other histidines in the double variant. For atoms Nϵ2 and Nδ1, B values are 17 and 18 Å^2 for the H57, while the mean value for the other three trypsin histidines is 14 Å^2. This may reflect slightly greater mobility for H57, perhaps as it attempts to accommodate C195. However, it is clear that the catalytic histidine is ordered and positioned for catalysis. Though the chemistry of N102 prevents formation of the thiolate–imidazolium ion pair in the double variant, geometrical constraints do not. Inability to form this ion pair is a major reason for the low activity of trypsin D102N,S195C.

The sulfur of C195 is not crowded at the active site, and both variants show the same conformer for this residue. Although sulfur has an atomic radius 0.4 Å larger than oxygen, here it is not involved in any close contacts. C195-Sγ most closely approaches H57-Nϵ2 (3.5 Å). Superposition of the cow trypsin (Bode & Schwager, 1975), actinidin (Baker & Dodson, 1980), trypsin S195C, and trypsin D102N,S195C active sites indicates a position for the nucleophile in the latter two which is intermediate between that found in serine and sulfhydryl proteases (Figure 3). In trypsin, the hydroxyl points out from Cβ toward H57 and the substrate binding pocket, while in actinidin the sulfur appears to point away from the histidine. Both variants show a rotation of 30° about C195 Cα–Cβ such that Sγ points slightly away from the histidine and lies closer to the oxyanion hole. The oxyanion hole is a structural element of the serine protease active site which uses the backbone amide groups of G193 and S195 to stabilize the negatively charged oxygen in the tetrahedral intermediate during catalysis (Henderson et al., 1971). Thus, in our two structures C195-Sγ has moved to a medial position between H57-Nϵ2 and G193-N (Table II).

It is possible that C195-Sγ forms hydrogen bonds with both H57-Nϵ2 and G193-N, thereby affecting catalysis by interfering with the necessary bonding interaction with H57-Nϵ2 and by sterically occluding the oxyanion hole. Only minimal mobility (0.2 Å) of the sulfur would be required to form the hydrogen bond to G193-N. This is in agreement with our structure which shows a well-ordered C195-Sγ with a temperature factor of 15 Å^2. Superposition of rat and cow trypsins reveals that the oxyanion hole is closer to the active site nucleophile in the rat enzyme. Specifically, the region of the polypeptide chain from residue 192 to residue 194 appears to have shifted by at least 0.5 Å. This may be a real difference between trypsins from the two species, since all four rat variants for which we have refined structures show this phenomenon (one binding pocket and three active site variants, unpublished results). We cannot rule out the possibility that the differences we observe between cow and rat trypsins are structural consequences of amino acid substitution, since we do not have a wild-type rat trypsin structure for comparison. A consequence of the position of the oxyanion hole, and of the rotation of the nucleophile in both variants, is that while the distance from S195-Oγ to G193-N in benzamidine-inhibited cow trypsin (Bode & Schwager, 1975) is 4.7 Å, it is 3.5 Å in S195C trypsin and 3.7 Å in D102N,S195C trypsin. Thus, the oxyanion hole is blocked by C195-Sγ in both structures and probably cannot easily accommodate the oxyanion which is present transiently during catalysis (Figure 4). This architectural problem could be the major reason for the diminished activity of both variants.

Orientation of the nucleophile with respect to the substrate binding pocket was determined by measuring the distance from benzamidine C4, at the top of the substrate binding pocket, to C195-Sγ, in both variants. The interatomic distance increased by 0.2 Å as a result of this mutation. The orientation of the inhibitor (amidinophenyl)pyruvate (APPA) was previously determined in trypsin (Walter & Bode, 1983) and was modeled into the active sites of the variants. C195-Sγ would closely approach the carbonyl carbon (1.0 Å). Therefore, crowding of the substrate at the active site may contribute to low substrate turnover in these variants, although it is not clear whether C195-Sγ would adjust to accommodate a substrate. Structure determination of a trypsin S195C–APPA complex is in progress.

Trypsins S195C and D102N,S195C provide the first view of catalytically active sulfhydryl protease active sites. In all previous sulfhydryl protease structures, interpretation of interatomic distances and conformations has been hampered by modification of the nucleophilic sulfur. It is oxidized to sulfinic (SO_2) and sulfonic (SO_3) acids in structures of actinidin and papain, respectively (Baker & Drenth, 1987). Activatable papain, in which the thiol forms a disulfide with a free cysteine,

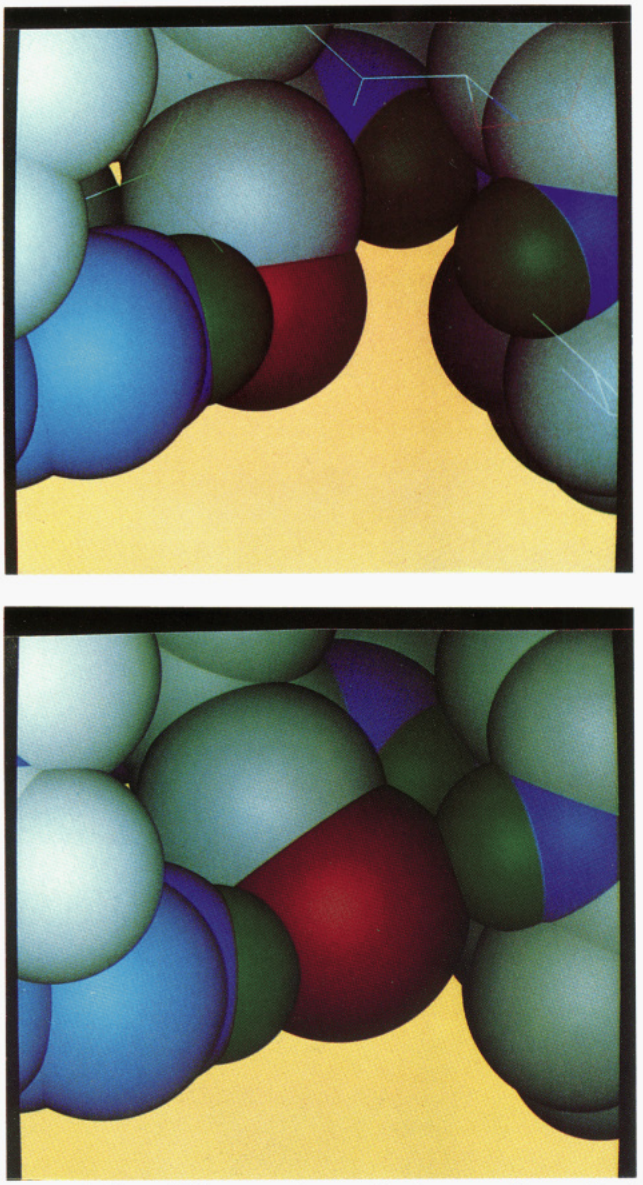

FIGURE 4: Space-filling computer graphics (INSIGHT, Biosym, Inc.) representations of catalytic amino acid residues 57 and 195 and the backbone amides of residues 193 and 195. The latter two form most of the oxyanion hole in trypsin. (Top) In cow trypsin, S195-Oγ (red) forms a hydrogen bond with H57 (blue, at left) via Nϵ2. The sulfur is 4.7 Å from G193-N (blue, at right; green is the amide hydrogen), and the oxyanion which forms during catalysis has easy access to the two backbone amides of residues 193 and 195. (Bottom) In trypsins S195C and D102N,S195C, C195-Sγ has rotated to a medial position between H57-Nϵ2 and G193-N. The C195-Sγ distance to G193-N is only 3.7 Å, which occludes the oxyanion hole. The distance from the nucleophilic atom of residue 195 to the 195 backbone amide (blue, at top; green is the amide proton) is unchanged in cow trypsin and the two variants (3.1 Å).

is also evident in crystal structures (Lowe, 1975). One ramification of these modifications is rotation of the histidine imidazole either above or below the sulfur, when they are, ostensibly, coplanar in the catalytically active thiol–imidazolium ion pair. Oxidation also produces small changes in the orientation of the sulfur (Drenth, 1976). A medium-resolution structure of thiol subtilisin may also have been oxidized (Alden et al., 1970). Although the structural perturbations caused by oxidation are minimal, the positional differences between the nucleophiles in trypsin S195C or D102N,S195C and papain and between those in trypsin and papain are also small, and may be an artifact of the inactive papain. Sulfhydryl proteases are thought to be readily oxidized because of the proximity of the oxyanion hole, which can accommodate one oxygen. The prealigned hydrogen-bonding donors of the oxyanion hole are positioned to "solvate" the oxyanion, and so the reorganization energy normally required for introducing an oxygen into this environment is greatly reduced (Warshel, 1978). Trypsin, likewise, has an oxyanion hole, but it is situated differently in the active site and may not provide such ready stabilization. Also, blockage of the oxyanion hole in trypsins S195C and D102N,S195C may protect C195-Sγ from oxidation during enzyme purification and X-ray data collection. Oxidation of C195-Sγ has not occurred to a detectable extent during purification of the two variants, and $F_o - F_c$ maps show no oxygens or extra density in the region of C195.

Although trypsin and papain at first appear quite similar with respect to their active sites and catalytic mechanisms, closer inspection reveals important differences in mechanism which stem from disparities in oxygen and sulfur chemistry and which are exploited and compensated for by electrostatics at the respective active sites (Polgar & Asboth, 1986). In papain the catalytic cysteine resides at the N-terminus of an α helix. The pK_a of the nucleophile is affected by the positive charge at that end of the helix dipole. Hol et al. (1978) suggests that this environment favors donation of the nucleophile proton to histidine. Other local groups, possibly including D158 (Angelides & Fink, 1979) and solvent (Rullmann et al., 1989), contribute to stabilization of the thiolate–imidazolium ion pair and are also determinants of the unusual pK_as reported for histidine (8.5) and cysteine (4.0) at this site (Lavery et al., 1983; Brocklehurst et al., 1983). A striking demonstration of the importance of the charge effect is that although thiol proteases actinidin and papain have very different isoelectric points, the electrostatic potentials at their active sites work similarly to stabilize the thiolate–imidazolium ion pair (Pickersgill et al., 1988). A similar arrangement of residues which would promote ionization of the sulfur is not present in trypsin. Therefore, although C195-Sγ and H57-Nϵ2 are within hydrogen-bonding distance of each other in S195C trypsin, their chemical environment is not conducive to forming the ion pair necessary for catalysis. A recent comparison of electrostatics at the active sites of cationic and anionic trypsins showed that charge is conserved in this region (Soman et al., 1989) and thus may be designed to assist catalysis only when oxygen is the nucleophile.

Another reason for the inefficacy of the variants may be a difference in the catalytic mechanisms of trypsin and papain. In the former, protonation of the substrate leaving group follows nucleophilic attack, while in the latter protonation occurs first (Howard & Kollman, 1988). This sequence of events is important for a cysteine protease because thiolate is a better leaving group than alkoxide and forms a weaker bond with carbon (C–S = 272 kJ/mol, C–O = 360 kJ/mol) and will therefore tend to dissociate first from an acyl enzyme intermediate. Thus, prior protonation of the leaving group ensures its rapid dissociation from the intermediate. It may occur because the asparagine–histidine hydrogen bond in papain involves imidazole nitrogen ϵ2 instead of δ1 (Figure 2), and so the histidine Cβ–Cγ bond is collinear with the asparagine Nδ2–histidine Nϵ2 bond. This allows the imidazole ring to swivel about these two attachment points. In catalysis, the histidine could help deprotonate the sulfhydryl and then rotate 30° around Cβ–Cγ to donate the proton to the substrate leaving group (Drenth et al., 1976). In the serine proteases, the imidazole of H57 is tethered such that it cannot rotate in the same manner, and such movement has not been implicated

in the catalytic mechanism. Trypsins D102N,S195C and S195C feature the sulfur but do not have the machinery for preprotonation and thus cannot mitigate return of the tetrahedral intermediate to the Michaelis complex.

Conclusions

Trypsins S195C and D102N,S195C show no global structural perturbations, and the side chains of the three active site residues superimpose well with both trypsin and papain. However, two features of the active sites of the variants may be responsible for much of the catalytic inefficacy. The replacement of Asp with Asn in trypsin D102N,S195C forces a tautomer of the catalytic histidine which is unable to function as a base in catalysis. Also, the substitution of a polar, uncharged residue for a negatively charged one at the active site disturbs the local electrostatic potential. The second structural problem is the position of C195-Sγ. It appears to obstruct the oxyanion hole and may form a hydrogen bond with G193-N. This arrangement could lower the stabilization of the oxyanion during catalysis, thus adversely affecting substrate turnover. Other differences between serine and cysteine proteases, such as active site electrostatics and subtle differences in catalytic mechanisms, undoubtedly contribute to the low activity of the two variants.

These results present a structural basis for the reduced activity of the engineered thiol trypsins. Although the conserved active site geometry of the serine and cysteine proteases is required, it is not sufficient for highly efficient catalytic hydrolysis of peptide bonds. Future directions must focus on the subtle yet significant differences between the convergent active sites.

Acknowledgments

We thank Stuart Tugendreich for technical assistance with the figures and John Vasquez for critical reading of the manuscript.

References

Alden, R. A., Wright, C. S., Westall, F. C., & Kraut, J. (1970) *Structure Function Relationships of Proteolytic Enzymes* (Desnuelle, P., Neurath, H., & Ottesen, M., Eds.) p 173, Munksgaard, Copenhagen.

Angelides, K. J., & Fink, A. L. (1978) *Biochemistry 17*, 2659–2668.

Baker, E. N. (1981) in *Structural Studies on Molecules of Biological Interest* (Dodson, G. G., Glusker, J. P., & Sayre, D., Eds.) pp 339–349, Clarendon Press, Oxford.

Baker, E. N., & Dodson, E. J. (1980) *Acta Crystallogr., Sect. A 36*, 559.

Baker, E. N., & Drenth, J. (1987) In *Biological Macromolecules and Assemblies* (Jurnak, F. A., & McPherson, A., Eds.) Vol. 3, pp 312–369, Wiley, New York.

Bazan, J. F., & Fletterick, R. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7872–7876.

Bode, W., & Schwager, P. (1975) *J. Mol. Biol. 98*, 693–717.

Brocklehurst, K. (1983) *Biochem. J. 213*, 559–560.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., & Rutter, W. J. (1985) *Science 228*, 291–297.

Craik, C. S., Roczniak, S., Largman, C., & Rutter, W. J. (1987) *Science 237*, 909–913.

Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F., & Fletterick, R. J. (1989) *Virology 172*, 302–310.

Drenth, J. (1976) in *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. 3, Proteins (Fasman, G. D., Ed.) pp 356–357, CRC Press, Cleveland, OH.

Drenth, J., Enzing, C. M., Kalk, K. H., & Vassies, J. C. A. (1976) *Nature 264*, 373–377.

Drenth, J., Kalk, K. H., & Swen, H. M. (1976) *Biochemistry 15*, 3731–3738.

Evnin, L. B., & Craik, C. S. (1988) *Enzyme Eng. 542*, 61–74.

Fehlhammer, H., Bode, W., & Huber, R. (1977) *J. Mol. Biol. 111*, 415–438.

Garavito, R. M., Rossman, M. G., Argos, P., & Eventoff, W. (1977) *Biochemistry 16*, 5065–5071.

Hampel, A., Labanauskas, M., Conners, P. G., Kirkegard, L., RajBhandary, U. L., Sigler, P. B., & Bock, R. M. (1968) *Science 162*, 1384–1387.

Hartley, B. S. (1964) *Nature 201*, 1284–1291.

Henderson, R., Wright, C., Hess, G., & Blow, D. (1971) *Cold Spring Harbor Symp. 36*, 63–70.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Rameseshan, S., & Venkatesan, K., Eds.) p 13.01, Indian Academy of Sciences, Bangalore.

Higaki, J. N., Gibson, B. W., & Craik, C. S. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 615–621.

Higaki, J. N., Evnin, L. B., & Craik, C. S. (1989) *Biochemistry* (preceding paper in this issue).

Hol, W. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443–446.

Howard, A. E., & Kollman, P. A. (1988) *J. Am. Chem. Soc. 110*, 7195–7200.

Husain, S. S., & Lowe, G. (1968) *Biochem. J. 108*, 855–859.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A., & Drenth, J. (1984) *J. Biol. Chem. 179*, 233–256.

Krieger, M., Kay, L. M., & Stroud, R. M. (1974) *J. Mol. Biol. 83*, 209–230.

Lavery, R., Pullman, A., & Wen, Y. K. (1983) *Int. J. Quantum Chem. 24*, 353–371.

Lowe, G. (1976) *Tetrahedron Rep. 32*, 291–302.

McGrath, M. E., Fletterick, R. J., & Craik, C. S. (1989) *Biotechniques 7*, 246–247.

McPherson, A. (1982) *Preparation and Analysis of Protein Crystals*, p 96, Wiley, New York.

Neet, K. E., & Koshland, D. E. (1966) *Proc. Natl. Acad. Sci. U.S.A. 56*, 1606–1611.

Neet, K. E., Nanci, A., & Koshland, D. E. (1969) *J. Biol. Chem. 243*, 6392–6401.

Pickersgill, R. W., Goodenough, P. W., Sumner, I. G., & Collins, M. E. (1988) *Biochem. J. 254*, 235–238.

Polgar, L., & Bender, M. L. (1966) *J. Am. Chem. Soc. 88*, 153–154.

Polgar, L., & Bender, M. L. (1967) *Biochemistry 6*, 610–620.

Polgar, L., & Asboth, B. (1986) *J. Theor. Biol. 121*, 323–326.

Rullman, J. A. C., Bellido, M. N., & van Duijnen, P. T. (1989) *J. Mol. Biol. 206*, 101–108.

Soman, K., Yang, A.-S., Honig, B., & Fletterick, R. J. (1989) *Biochemistry* (in press).

Sprang, S., Standing, T., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W. J., & Craik, C. S. (1987) *Science 237*, 905–909.

Tsai, I.-H., & Bender, M. L. (1979) *Biochemistry 18*, 3764–3768.

Vasquez, J. R., Evnin, L. B., Higaki, J. N., & Craik, C. S. (1989) *J. Cell. Biochem.* (in press).

Walter, J., & Bode, W. (1983) *Hoppe-Seyler's Z. Physiol. Chem. 364*, 949–959.
Warshel, A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5250–5254.
Warshel, A., Naray-Szabo, G., Sussman, F., & Hwang, J.-K. (1989) *Biochemistry 28*, 3629–3637.
Weiner, H., White, W. N., Hoare, D. G., & Koshland, D. E., Jr. (1966) *J. Am. Chem. Soc. 88*, 3851.

Pituitary Multicatalytic Proteinase Complex. Specificity of Components and Aspects of Proteolytic Activity†

Marian Orlowski* and Charlene Michaud
Department of Pharmacology, Mount Sinai School of Medicine of the City University of New York, New York, New York 10029

Received April 24, 1989; Revised Manuscript Received June 29, 1989

ABSTRACT: The 700-kDa multicatalytic proteinase complex from bovine pituitaries separates in polyacrylamide gel electrophoresis under dissociating and reducing conditions into 11 components with molecular masses ranging from 21 to 32 kDa. No higher molecular mass components were detected. A rabbit polyclonal antibody raised against the complex recognizes five immunoreactive components. As reported previously, the complex exhibits three distinct proteolytic activities designated as chymotrypsin-like, trypsin-like, and peptidylglutamyl-peptide hydrolyzing activities. All three activities are rather rapidly inactivated by 3,4-dichloroisocoumarin, a general serine protease inhibitor, however, the pseudo-first-order rate constants of inactivation of the three components differ within a wide range, with the chymotrypsin-like activity being most sensitive to inhibition. The peptidylglutamyl-peptide hydrolyzing activity is greatly activated by low concentrations of sodium dodecyl sulfate and fatty acids and seems to constitute the main component responsible for degradation of protein substrates. In addition to cleaving bonds on the carboxyl side of glutamyl residues, this activity also cleaves, albeit at a slower rate, bonds on the carboxyl side of hydrophobic residues; however, the secondary specificity of this component is clearly different from the chymotrypsin-like activity. Heparin selectively activates the chymotrypsin-like activity. The complex cleaves rapidly both native and dephosphorylated β-casein in a reaction greatly accelerated by low concentrations of sodium dodecyl sulfate. The nature of proteolytic products, and also the rate of formation of acid-soluble, ninhydrin-reactive products, is different for the phosphorylated and dephosphorylated form of β-casein, indicating that the degree of phosphorylation influences the rate and pattern of proteolysis. Lysozyme, human serum albumin, and phosphorylase *b* are also degraded but at a slower rate. Irreversible inhibition of the chymotrypsin-like activity leads to a marked activation of the trypsin-like activity, indicating interactions between components of the complex.

Previous work in this laboratory led to the identification in bovine pituitaries of a high molecular mass protein (700 kDa) which exhibited three proteolytic activities toward synthetic substrates containing hydrophobic, basic, and acidic amino acid residues in the P_1 position (Wilk et al., 1979; Wilk & Orlowski, 1980, 1983). Evidence was presented that the protein represents a multicomponent complex of nonidentical subunits and that the three activities are associated with distinct components of the complex (Orlowski & Wilk, 1981; Wilk & Orlowski, 1983). This led us to propose the name "multicatalytic proteinase complex" (MPC)[1] in recognition that the protein apparently represents a multienzyme complex, reminiscent of other multienzyme complexes known to occur in cells (Orlowski & Wilk, 1981, 1988; Wilk & Orlowski, 1983). Subsequent work in many laboratories showed that the complex is widely distributed in animal tissues, that it is apparently a constant component of all eucaryotic cells, and that it is composed of monodisperse particles with a cylindrical or disk shape, as shown by electron microscopy (Ray & Harris, 1985, 1987; Kopp et al., 1986; Dahlman et al., 1988; Rivett, 1989). Recent reports indicate that the complex is identical with the previously observed 19S ribonucleoprotein particles present in all eucaryotic cells (Shelton et al., 1970) for which the name "prosome" was proposed (Schmid et al., 1984). It is now clear that the MPC constitutes a major extralysosomal proteolytic system; however, it has not yet been firmly established whether other functions attributed to prosomes, such as repression of mRNA translation (Schmid et al., 1984), pre-tRNA 5′ processing endonuclease activity (Castaño et al., 1986), and aminoacyl transferase activity (Shelton et al., 1970), are indeed an integral part of the complex, nor is it certain whether small cytoplasmic RNA is indeed a constant part of the complex (Arrigo et al., 1988; Falkenburg et al., 1988).

† This work was supported by Grant DK 25377 from the National Institutes of Health.

* Address correspondence to this author at Box 1215, Department of Pharmacology, Mount Sinai School of Medicine, Fifth Ave. and 100th St., New York, NY 10029.

[1] Abbreviations: Cbz, benzyloxycarbonyl; Boc, *tert*-butoxycarbonyl; Bz, benzoyl; 2NA, 2-naphthylamide; DFP, diisopropyl fluorophosphate; DTT, dithiothreitol; Glt, glutaryl; MPC, multicatalytic proteinase complex; PAGE polyacrylamide gel electrophoresis; PCMB, *p*-mercuribenzoate; pNA, *p*-nitroanilide; SDS, sodium dodecyl sulfate; TCA, trichloroacetic acid.

0006-2960/89/0428-9270$01.50/0