Structural Origins of Substrate Discrimination in Trypsin and Chymotrypsin†

John J. Perona,‡,§ Lizbeth Hedstrom,|| William J. Rutter,⊥,# and Robert J. Fletterick*,⊥

Department of Pharmaceutical Chemistry, Department of Biochemistry and Biophysics, and Hormone Research Institute, University of California, San Francisco, California 94143, and Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02254

Received June 27, 1994; Revised Manuscript Received October 26, 1994®

ABSTRACT: Converting the specificity of trypsin to that of chymotrypsin has been shown to require the exchange of amino acids in multiple portions of the protein, including two surface loops which do not directly contact the substrate. Crystallographic analysis of two mutant trypsins possessing chymotrypsin-like specificity now reveals that these distal surface loops alter function by directly determining the structure of the primary binding site. Efficient acylation of cognate substrates correlates with a distinct backbone conformation of the conserved Gly216 residue. This amino acid is located on the surface of the specificity pocket and forms two main-chain hydrogen bonds with a nonspecific portion of substrate. By contrast, the improvement in substrate binding affinity effect by the substitution of the distal Tyr172 residue with Trp derives from structural rearrangements at the extreme base of the pocket. Together, the kinetic and crystallographic data strongly suggest that both Asp189 and Gly216 must be considered as primary determinants of substrate specificity in trypsin.

Specificity in enzyme catalysis arises as a consequence of optimal transition-state interactions made by cognate, but not noncognate, substrates. These electrostatic and apolar contacts dictate enzyme−substrate complementarity, a necessity for surmounting the activation energy barrier between the ground and transition states (Jencks, 1969). Since substrate binding sites are preformed and relatively rigid, the free energy of substrate binding can be converted to catalysis without a large entropic penalty. Rate acceleration must therefore also depend upon the ability of distal portions of the structure to stabilize the binding and catalytic residues. In this context an important question arises: does the global protein architecture play a strictly passive role by providing an inert framework for the substrate binding surface, or alternatively, does it also function in a more active manner? Such active function might include determining the detailed conformation of those amino acids which form the contact surface for the substrate. Alternatively, distal portions of the protein may determine the degree to which the substrate binding site is inherently deformable. The operation of either or both mechanisms would provide a means by which the substrate specificity profile could be understood to arise as a more distributed property of the enzyme fold.

Trypsin and chymotrypsin catalyze peptide bond cleavage by identical mechanisms (Polgar, 1989) and possess very similar tertiary structures consisting of two juxtaposed six-stranded β-barrel domains (Ruhlmann et al., 1973; Krieger et al., 1974; Matthews et al., 1967; Steitz et al., 1969). Trypsin efficiently hydrolyzes peptidyl amide substrates possessing Arg and Lys at the P1 position with a selectivity, as assessed by relative k_{cat}/K_m values, of nearly 10^6-fold relative to analogous Phe-containing substrates (Hedstrom et al., 1992; Table 1; see Figure 1, legend, for the subsite nomenclature). Conversely, chymotrypsin favors peptide substrates possessing Trp, Tyr, and Phe side chains at position P1, with an overall specificity relative to P1-Lys substrates to 10^4-fold (Hedstrom et al., 1992). Since the structures of the S1 sites of the two enzymes are very similar (Figure 1), the divergent specificities appeared to be a simple property of the local electrostatic environments (Stroud, 1974). However, replacement of the primary binding determinant Asp189 of trypsin with the analogous Ser189 of chymotrypsin fails to convert specificity and results instead in a poor nonspecific protease (Graf et al., 1988). Conversion of trypsin to the chymotrypsin-like protease Tr⟹Ch-[S1+L1+L2] requires substitution of four residues in the S1 site together with the exchange of two adjacent surface loops (loops 1 and 2; Figures 1 and 2) which do not directly contact the substrate (Hedstrom et al., 1992). Consequently, it appears that substrate specificity in trypsin and chymotrypsin provides an example in which a crucial role is played by distal portions of the protein structure.

Tr⟹Ch[S1+L1+L2] hydrolyzes substrates possessing large hydrophobic amino acids at position P1 with a similar acylation rate (k_2) relative to chymotrypsin (Table 1; Hedstrom et al., 1992) but is highly impaired in substrate binding affinity (K_s). Additional exchange of a Tyr172 to Trp in a third distal segment of the structure creates the new hybrid enzyme Tr⟹Ch[S1+L1+L2+Y172W] (Hedstrom et al., 1994a; Figure 2), which is further improved in catalytic efficiency toward chymotrypsin-like substrates by 20−50-fold. Significantly, the specificity profile of this enzyme toward substrates containing P1-Trp, Tyr, Phe, and Leu more closely mimics that of chymotrypsin (Hedstrom et al., 1994a). Extensive characterization of these and other related mutant

† Supported by the Lucille P. Markey charitable trust to Brandeis University and by NIH Grants DK21344 (to W.J.R.), DK39304 (to R.J.F.), and GM13818-03 (to J.J.P.).

‡ Department of Pharmaceutical Chemistry, University of California.

§ Present address: Department of Chemistry and Interdepartmental Program in Biochemistry and Molecular Biology, University of California at Santa Barbara, Santa Barbara, CA 93106.

|| Graduate Department of Biochemistry, Brandeis University.

⊥ Department of Biochemistry and Biophysics, University of California.

Hormone Research Institute, University of California.

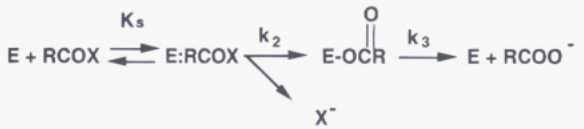
® Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-1489$09.00/0

Table 1: Kinetic Parameters of Hybrid Trypsins[a]

	K_s (M)	k_2 (s^{-1})	k_3 (s^{-1})
trypsin			
Ac-Phe-NH_2	1×10^{-3}	2.2×10^{-4}	30
suc-AAPF-*p*NA	>0.25	>0.2	36
Z-GPR-AMC	8.6×10^{-5}	319	101.5
D189S			
Ac-Phe-NH_2	0.17	0.018	39
suc-AAPF-*p*NA	0.015	0.29	33
Z-GPR-AMC	$>10^4$	>1.1	126.2
Tr⟹Ch[S1+L1+L2]			
Ac-Phe-NH_2	0.065	0.018	33
suc-AAPF-*p*NA	0.011	20	37
Tr⟹Ch[S1+L1+L2+Y172W]			
Ac-Phe-NH_2	>0.1	>0.08	38
suc-AAPF-*p*NA	5.0×10^{-4}	41	63
chymotrypsin			
Ac-Phe-NH_2	0.023	0.43	60
suc-AAPF-*p*NA	1.5×10^{-3}	850	52

[a] Mechanistic kinetic parameters for trypsin, chymotrypsin, D189S, and the hybrid trypsins Tr⟹Ch[S1+L1+L2] and Tr⟹Ch[S1+L1+L2+Y172W]. Data are taken from Hedstrom et al. (1994a) and Perona et al. (1994). K_s, k_2, and k_3 represent measures of substrate binding affinity, acylation rate, and deacylation rate, respectively. They are derived from steady-state kinetic constants obtained on equivalent pairs of amide and ester substrates, *suc*-AAPF-*p*NA and the ester substrate *suc*-AAPF-SBzl, Ac-Phe-NH_2 and the ester Ac-F-*p*NP (Hedstrom et al., 1992, 1994a), and Z-GPR-AMC and the ester Z-GPR-SBzl (Perona et al., 1994), according to the standard serine protease mechanism (Zerner et al., 1964):



For reactions in which acylation is the rate-determining step, $K_m = K_s$ and $k_{cat} = k_2$. If deacylation is rate-determining, then $K_m = K_s[k_{cat,ester}/(k_2 + k_{cat,ester})]$ and $k_{cat} = k_2 k_{cat,ester}/(k_2 + k_{cat,ester})$.

trypsins has also shown that the key kinetic determinant of substrate specificity is the rate of the acylation reaction (Hedstrom et al., 1994b). Ground-state binding plays only a minor role in discriminating among different P1-substrate side chains.

Understanding the molecular mechanisms by which the three distal elements influence substrate specificity in trypsin will provide insight into the larger question of the role of the protein scaffold in enzyme catalysis. Consequently, we determined the crystal structures of Tr⟹Ch[S1+L1+L2] and Tr⟹Ch[S1+L1+L2+Y172W] complexed with the inhibitor *suc*-AAPF-CMK[1] at 2.0 and 2.1 Å resolution, respectively. From the data it appears that the structural basis of discrimination in the acylation step is the ability of loop 2 to uniquely specify the conformation of the conserved Gly216, which forms two main-chain hydrogen bonds with the P3 residue of the substrate. By contrast, the additional substitution Y172W propagates structural changes into the base of the S1 binding pocket, resulting in improved binding affinity toward peptidyl chymotryptic-like substrates. The data describe a direct correlation between structural and kinetic parameters for amide substrate hydrolysis: improvements in either of the mechanistic kinetic parameters k_2 or K_s occur as a consequence of conformational changes in residues which directly contact substrate.

[1] Abbreviations: *suc*-AAPF-CMK, succinyl-Ala-Ala-Pro-Phe chloromethyl ketone; SBTI, soybean trypsin inhibitor; BPTI, bovine pancreatic trypsin inhibitor; *suc*-AAPF-SBzl, succinyl-Ala-Ala-Pro-Phe thiobenzyl ester; *suc*-AAPF-*p*NA, succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide; MES, 2-(*N*-morpholino)ethanesulfonic acid; Hepes, *N*-(2-hydryxoyethyl)piperazine-*N*′-2-ethanesulfonic acid; *suc*-AAPF-AMC, succinyl-Ala-Ala-Pro-Phe-aminomethylcoumarin; Ac-Phe-NH_2, *N*-acetylphenylalaninamide; Ac-Phe-*p*NP, *N*-acetylphenylalanine *p*-nitrophenyl ester; APPI, serine protease inhibitor domain of the β-amyloid precursor protein; Z-GPR-AMC, *N*-carbobenzoxy-L-Gly-Pro-Arg-aminomethylcoumarin; Z-GPR-SBzl, *N*-carbobenzoxy-Gly-Pro-Arg thiobenzyl ester.

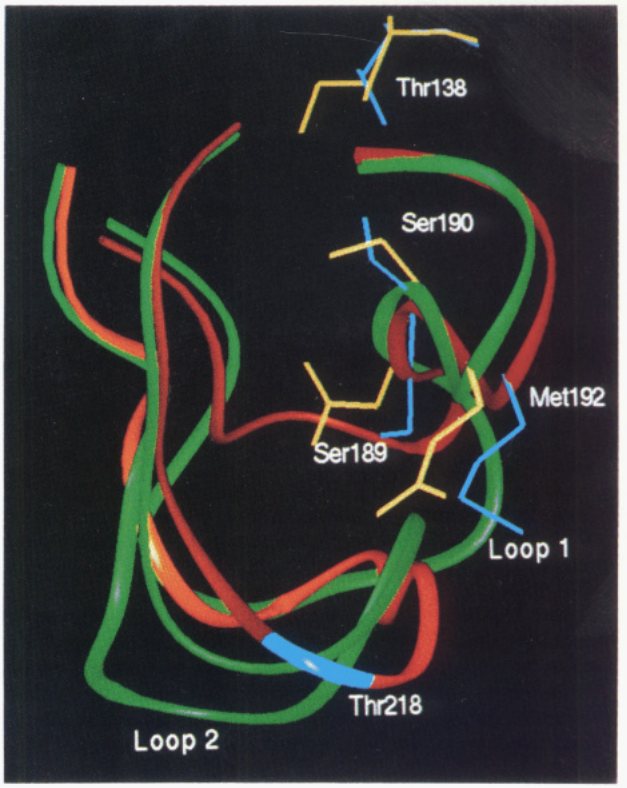


FIGURE 1: Superposition of the primary substrate binding pockets of trypsin (green) and chymotrypsin (red). The view is directly into the pocket as presented to substrates; residue 189 is at the base of the site. Labels refer to side chains of chymotrypsin. Amino acids in the S1 site which differ in identity or conformation between the two enzymes are shown in yellow (trypsin) and blue (chymotrypsin). Thr218 represents an insertion in the squence of chymotrypsin relative to trypsin. The walls of the S1 site which interact directly with substrate (top) are very similar, but the structures of loops 1 and 2 (bottom), which do not contact substrate, are divergent. Nomenclature for the substrate amino acid residues is P*n*, ... P2, P1, P1′, P2′, ..., P*n*′, where P1−P1′ denotes the hydrolyzed bond. S*n*, ..., S2, S1, S1′, S2′, ..., S*n*′ denote the corresponding enzyme binding sites (Schechter & Berger, 1968).

MATERIALS AND METHODS

Enzymes and Inhibitors. Rat anionic trypsinogen variants Tr⟹Ch[S1+L1+L2] and Tr⟹Ch[S1+L1+L2+Y172W] were produced in a *Saccharomyces cerevisiae* expression system by fusing the trypsinogen coding sequence to the α-factor leader peptide, as described (Phillips et al., 1990; Hedstrom et al., 1992, 1994a). The zymogen forms were purified by chromatography on Toyopearl 650M (Supelco) prior to activation by enterokinase treatment. The activated enzymes were further purified by affinity chromatography on SBTI Sepharose (Sigma Chemical Co.) (Hedstrom et al., 1994a). *Suc*-AAPF-CMK was purchased from Enzyme Systems Products (Livermore, CA).

Crystallization. Enzymes were incubated with a 10−50-fold molar excess of *suc*-AAPF-CMK in a solution containing 50 mM Hepes (pH 8.0), 10 mM $CaCl_2$, and 100 mM NaCl at 25 °C. Aliquots were removed at several time

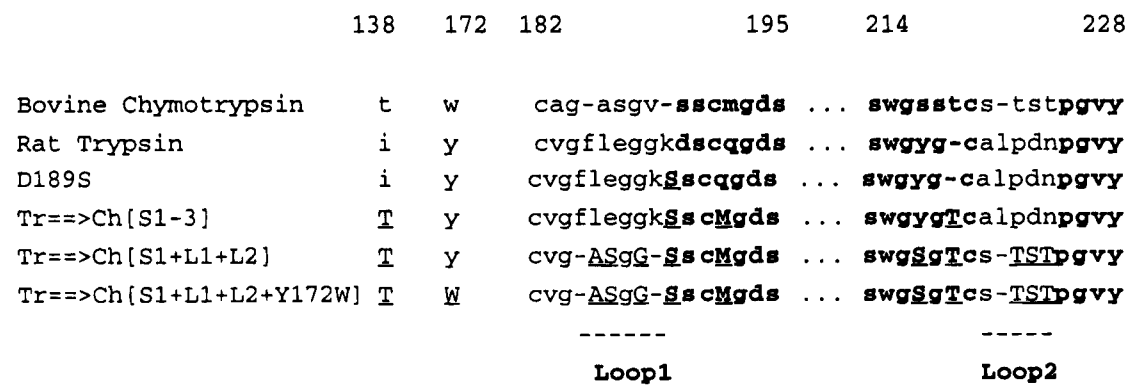


FIGURE 2: Alignment of trypsin mutants showing the amino acid substitutions present in the various hybrid enzymes. Capital letters indicate mutations. Amino acids located in the S1 site are in boldface. Dashed lines indicate the amino acids within each of the surface loops 1 and 2.

Table 2: Summary of Crystallographic Data Collection and Refinement

enzyme	Tr⇒Ch[S1+L1+L2]	Tr⇒Ch[S1+L1+L2+Y172W]
space group	$P3_221$	$P3_221$
cell dimensions (Å)	$a = 92.3, c = 62.0$	$a = 92.1, c = 61.8$
V_m (Å^3/Da)	3.2	3.2
resolution (Å)	2.0	2.1
total obs	70260	55138
unique obs	30768	28085
% complete	89.1	82.8
R_{merge}[a] (%)	9.7	8.2
R_{cryst}[b] (%)	19.3 (6.0−2.0 Å)	18.5 (6.0−2.1 Å)
rms bonds (Å)	0.011	0.012
rms angles (deg)	2.8	2.8

[a] $R_{merge} = (\sum_h \sum_i |\langle F_h \rangle - F_{hi}|)/(\sum_h F_h)$, where $\langle F_h \rangle$ is the mean structure factor magnitude of i observations of symmetry-related reflections with Bragg index h. [b] $R_{cryst} = (\sum_h \sum_i ||F_o| - |F_c||)/(\sum |F_o|)$, where F_o and F_c are the observed and calculated structure factor magnitudes.

intervals and the enzymes diluted into an identical solution which also contained excess amounts of the substrates *suc*-AAPF-SBzl or *suc*-AAPF-*p*NA (Hedstrom et al., 1994b). Incubation was continued until no residual activity was observed. The inactivated enzymes were dialyzed into solutions containing 1 mM HCl and 10 mM $CaCl_2$ and concentrated by Centricon centrifugation (Amicon) to 10−15 mg/mL. Crystals were obtained by hanging-drop vapor diffusion from solutions containing 8−12 mg/mL enzyme, 9−15% PEG 8000, 10 mM $CaCl_2$, and 50 mM MES (pH 5.8−6.4) (final concentrations) at 4 °C and were stabilized prior to data collection in mother liquor containing 20% PEG 8000. Precession photography was used to identify the space group and unit cell dimensions of the crystals (Table 2). Remarkably, these crystals are isomorphous with crystals of rat trypsin variants complexed with the 58 amino acid protein inhibitor BPTI, despite the very large difference in the size of the inhibitors as well as alteration of the growth conditions (Perona et al., 1993a).

Structure Determinations. X-ray diffraction amplitudes were collected on an *R*-axis imaging plate with a Rigaku rotating anode source; data were reduced using the accompanying software packages. Rigid-body, positional, and *B*-factor refinement was accomplished via the program X-PLOR (Brunger et al., 1987). For the determination of the structure of Tr⇒Ch[S1+L1+L2], the initial model used to provide phase information was the structure of rat trypsin D189S complexed with BPTI (Perona et al., 1994) with the deletion of loops 1 and 2 (Figure 1), all inhibitor atoms, and all enzyme side chains/solvent molecules present in the S1 binding pocket. The crystal structure of Tr⇒Ch[S1+L1+L2+Y172W] was determined using the refined structure of Tr⇒Ch[S1+L1+L2] as the initial model, from which the *suc*-AAPF-CMK inhibitor, all side chains and solvent molecules in the S1 site, and the side chain of Tyr172 were removed. Parameters for refinement of the covalent chloromethyl ketone inhibitor interactions were obtained from the crystal structure of human neutrophil elastase complexed with the analogous inhibitor *suc*-Ala-Ala-Pro-Ala chloromethyl ketone (Navia et al., 1989). Model-building and structural analysis was done using the INSIGHT (Dayringer et al., 1986) and FRODO (Jones, 1985) molecular graphics programs implemented on Evans & Sutherland and Silicon Graphics workstations. Structure superpositions were also carried out with the program OVRLAP (Rossman & Argos, 1976). Final refinement statistics are reported in Table 2.

Computation of Simulated Annealing Omit Maps. To assess the accuracy of the structural models in the region in and adjacent to the primary binding pocket, we computed simulated annealing omit maps for each structure using the program X-PLOR (Brunger et al., 1990). For both Tr⇒Ch[S1+L1+L2] and Tr⇒Ch[S1+L1+L2+Y172W] a 10.0 Å sphere centered at Gly188 was defined as the omitted region, and atoms in a 3 Å shell about this sphere were harmonically restrained to prevent artificial movements. Very tight harmonic restraints were applied to water molecules. Simulated annealing was carried out beginning at an initial temperature of 1000 K with cooling in increments of 25° to a final temperature of 300 K. One hundred cycles of postional refinement were performed following the annealing procedures. Maps were calcualted using calculated amplitudes and phases derived from these coordinates, with all atoms inside the defined 10.0 Å sphere still omitted.

RESULTS

Structure of Tr⇒Ch[S1+L1+L2] Complexed with Suc-AAPF-CMK. Peptidyl chloromethyl ketones such as *suc*-AAPF-CMK are covalent irreversible inactivators of serine proteases (Powers & Harper, 1986; Hedstrom et al., 1994b). In the complex formed with Tr⇒Ch[S1+L1+L2] the inhibitor occupies the S1−S4 sites on the surface of the enzyme and forms covalent bonds with both the catalytic Ser195 and His57 residues (Figure 3). The interactions of *suc*-AAPF-CMK with Ser195 and His57 of Tr⇒Ch[S1+L1+L2] are identical to those made by the analogous *suc*-AAPA-CMK inhibitor complexed to human neutrophil elastase (Navia et al., 1989). Similar observations were made in an early study

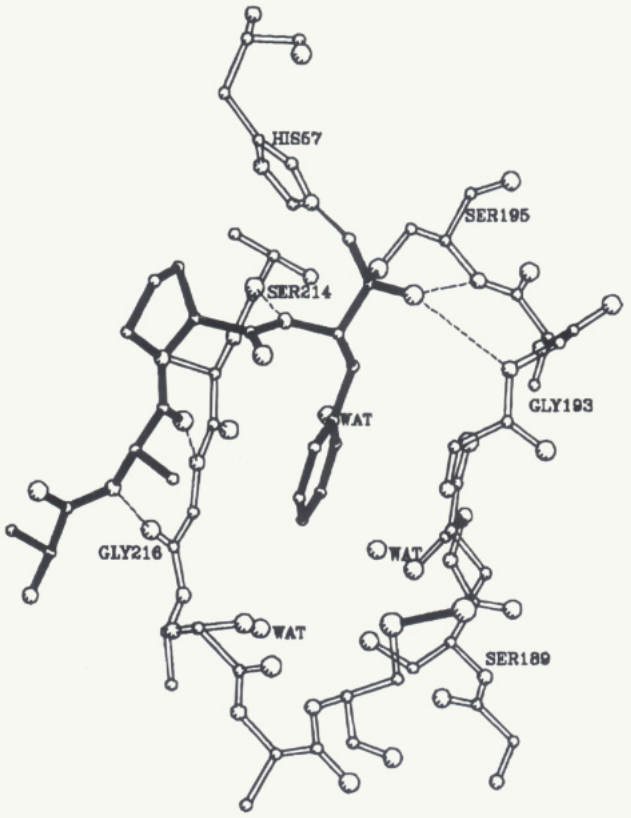


FIGURE 3: Interactions of *suc*-AAPF-CMK with Tr⟹Ch[S1+L1+L2]. Polypeptide backbone atoms of the S1 site are shown in open lines, and the inhibitor is drawn in filled lines. Dotted lines indicate the five hydrogen bonds formed in the complex. The filled line near the lower right indicates the disulfide bond formed between Cys191 and Cys220. Unconnected open circles represent water molecules (WAT). Covalent bonds formed by the side chains of Ser195 and His57 with the reactive portion of the inhibitor are shown in thin filled lines.

of inhibitor binding to subtilisin BPN′ (Poulos et al., 1976). At positions P2−P4 the binding of *suc*-AAPF-CMK and *suc*-AAPA-CMK to Tr⟹Ch[S1+L1+L2] and human neutrophil elastase diverges somewhat to accommodate differences in backbone structures of the enzymes at positions 214−216. However, in both complexes one hydrogen bond is formed between the P1 amide nitrogen and the carbonyl oxygen of the conserved Ser214, and two are formed in an antiparallel β-sheet fashion between the P3 residue of the inhibitor and residue Gly216 (Figure 3; Navia et al., 1989). The keto oxygen atom forms two fairly long (3.2 Å) hydrogen bonds with the oxyanion hole amide nitrogen groups at Gly193 and Ser195 in both enzyme complexes. The succinyl moiety which extends beyond the P4 residue is disordered in the structure described here.

The Phe ring at position P1 of *suc*-AAPF-CMK is sandwiched in the S1 site of Tr⟹Ch[S1+L1+L2] between backbone atoms at positions 215−216 and 191−192. The Phe side chain binds in an orientation similar to that observed in the crystal structure of chymotrypsin complexed with the inhibitor Leu-Phe trifluoroketone (Brady et al., 1990; Figure 4). In the structure of Tr⟹Ch[S1+L1+L2] the aromatic ring occupies a slightly shallower position in the pocket, most probably owing to the more constrained interactions of the chloromethyl ketone moiety with Ser195 and His57. Three ordered water molecules are found to bind within the S1 site as well (Figures 3 and 4). K_{inact}/K_i values for *suc*-AAPF-CMK inhibition of a series of trypsin mutants with partial

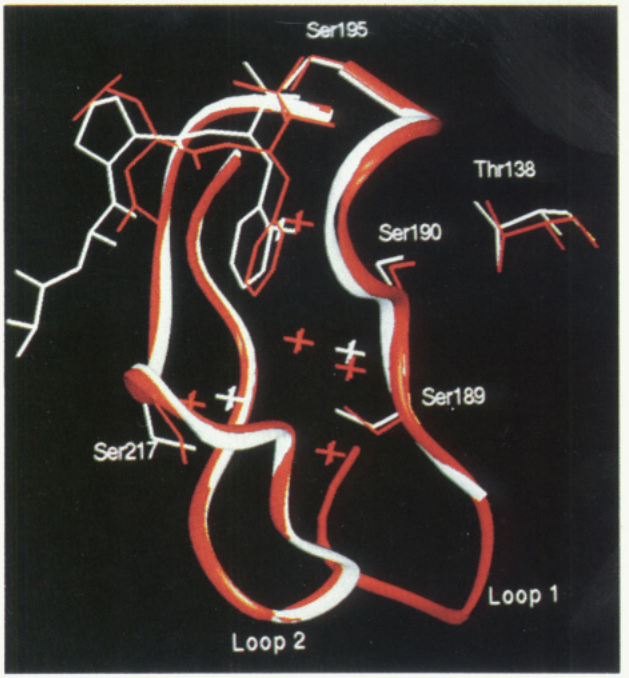


FIGURE 4: Superposition of the S1 sites of chymotrypsin complexed with Leu-Phe trifluoroketone (red) and of Tr⟹Ch[S1+L1+L2] complexed with *suc*-AAPF-CMK (white). The absence of the chain in loop 1 of the hybrid enzyme indicates disorder. Red and white crosses indicate ordered water molecules present in the chymotrypsin and Tr⟹Ch[S1+L1+L2] S1 sites, respectively.

chymotrypsin-like activities correlate very well with k_{cat}/K_m values for hydrolysis of the analogous *suc*-AAPF-AMC substrates (Hedstrom et al., 1994b). This shows that the inhibitor functions as a transition-state analog and suggests that the interactions observed in the crystal structure reflect true substrate interactions in the transition-state complex.

The structure of Tr⟹Ch[S1+L1+L2] shows that the backbone of the exchanged loop 2 (residues 217−224) adopts a conformation nearly identical to that of chymotrypsin (Figure 4). Loop 2 appears to be less well ordered than in either trypsin or chymotrypsin, as suggested by high thermal factors of 35−45 Å^2. However, the electron density at a contour level of 1.0σ in maps computed from amplitudes ($2F_o - F_c$) is continuous and permits complete modeling of the backbone and unambiguous specification of the locations of the peptide group oxygen atoms. The increased disorder in this loop relative to chymotrypsin is also manifested in the absence of interpretable electron density beyond C$^\beta$ for residues Ser221, Thr222, and Ser223 within loop 2; each of these amino acids was therefore modeled as alanine. However, inside the S1 pocket, the side chains of Ser189, Ser190, and Thr138 adopt chymotrypsin-like rotamers, with Ser190 rotated out of the binding pocket (Figure 4). Three water molecules are bound in the pocket; two of these adopt positions nearly identical to those found in the structure of chymotrypsin complexed with Leu-Phe trifluoroketone. While two conformations were found for the side chain of Met192 at the lip of the pocket, one of these is identical to that observed in chymotrypsin; refinement of the two conformers gave relative occupancies of 60−40%. Also similarly to chymotrypsin, residues 147−149 which are adjacent to loop 1 aer disordered in Tr⟹Ch[S1+L1+L2].

In contrast to these similarities the structure of Tr⟹Ch-[S1+L1+L2] also reveals the presence of a severe lesion

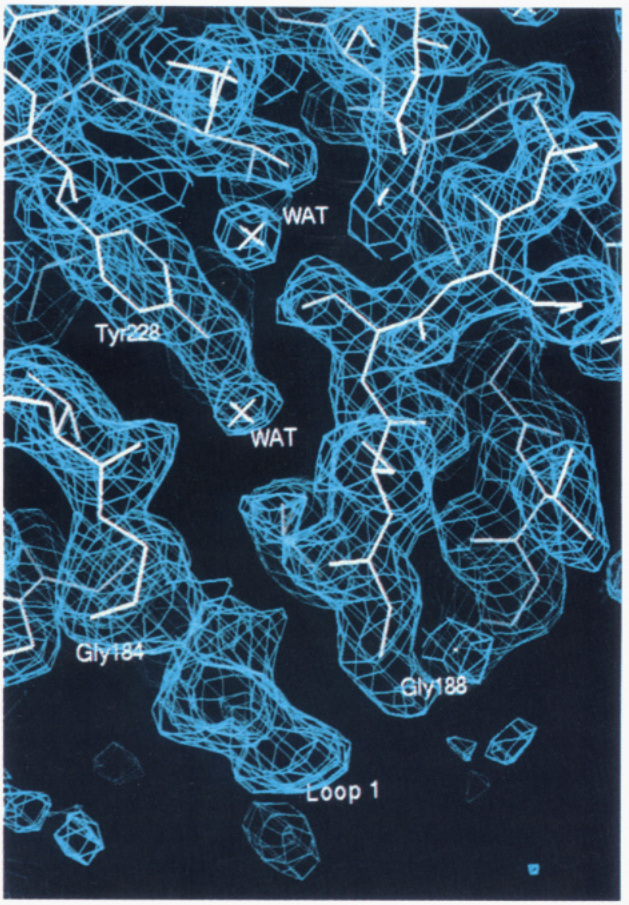


FIGURE 5: Simulated annealing OMIT map of Tr⟹Ch[S1+L1+L2] compared with coefficients $2F_o - F_c$, using the program X-PLOR (Brunger et al., 1987). The map is contoured at 1.0σ and is calculated over the resolution range 2.1−20 Å (see Materials and Methods). SDS gel electrophoresis of carefully washed and dissolved crystals of both Tr⟹Ch[S1+L1+L2] and Tr⟹Ch-[S1+L1+L2+Y172W] shows that the polypeptide chains are intact (data not shown); therefore, the interruptions in continuity of the electron density within loop 1 (bottom) do not arise from autoproteolysis (see also Figure 7).

adjacent to the S1 site, which is not found in chymotrypsin. Residues Ala185, Ser186, and Gly187 of loop 1 are disordered and not visible in electron density maps in the hybrid enzyme (Figures 4 and 5). Gly184 and Gly188 at the two boundaries of the disordered segment adopt conformations similar to chymotrypsin. Loop 1 interacts with loop 2 as well as with solvent molecules within the S1 site in both trypsin and chymotrypsin. Not surprisingly, therefore, the positions of ordered water molecules also do not correspond well with those observed in the crystal structure of chymotrypsin complexed with Leu-Phe trifluoroketone (Figure 4). The detailed side-chain conformers of the loop 2 amino acids Ser217 and Pro225 differ as well; each of these amino acids is conserved in bovine trypsin and chymotrypsin. Additionally, the side chain of the inserted Thr219 in Tr⟹Ch[S1+L1+L2] is rotated approximately 180° in $\chi 1$ relative to chymotrypsin (Figure 4).

It is clear from comparison of the crystal structures of trypsin, chymotrypsin, and Tr⟹Ch[S1+L1+L2] that, while loop 1 does not interact directly with the substrate, it does provide important interactions which stabilize the structure forming the base of the S1 site. Therefore, it appears likely that the disorder of this loop is at least in part responsible for the weak binding affinity exhibited by Tr⟹Ch[S1+L1+L2] toward chymotrypsin-like substrates (Table 1). Further, the disorder of loop 1 also strongly suggests that the high acylation rate of this hybrid enzyme toward chymotrypsin-specific substrates must arise solely from the exchange of loop 2. The requirement for the additional exchange of loop 1 (Hedstrom et al., 1992) to improve the catalytic efficiency of Tr⟹Ch[S1+L1+L2] is likely to be removal of hindrance to the adoption of a chymotrypsin-like conformation by loop 2. Modeling of a trypsin-like loop 1 into Tr⟹Ch-[S1+L1+L2] results in severe steric clashes with the chymotrypsin-like loop 2.

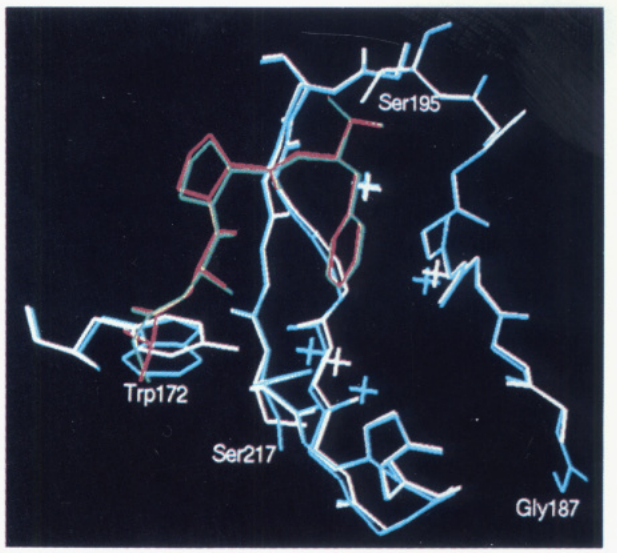


FIGURE 6: Superposition of the S1 sites of Tr⟹Ch[S1+L1+L2] (white) and Tr⟹Ch[S1+L1+L2+Y172W] (light blue). Crosses represent water molecules in the respective pockets. Binding of *suc*-AAPF-CMK to Tr⟹Ch[S1+L1+L2] and to Tr⟹Ch[S1+L1+L2+Y172W] is shown in red and green, respectively.

Structure of Tr⟹Ch[S1+L1+L2+Y172W]. The observation that Tr⟹Ch[S1+L1+L2] efficiently catalyzes the chemical step of acylation, while being unable to bind substrate well in the ground state, was explained in terms of the following model: the enzyme exists as two conformations in equilibrium, the predominant one of which is inactive and possesses a deformed S1 pocket. A minor species exists which possesses an active chymotrypsin-like S1-site structure that can be trapped by high concentrations of substrate (Hedstrom et al., 1994a). This model suggests that the additional substitution Y172W improves the catalytic efficiency of Tr⟹Ch[S1+L1+L2] by stabilizing the S1 binding site. Increased stability is judged by an increased affinity of proflavin binding as well as improved catalytic efficiencies toward several substrates (Table 1; Hedstrom et al., 1994a,b). Tr⟹Ch[S1+L1+L2+Y172W] retains the high acylation rate toward peptidyl chymotryptic substrates manifested by chymotrypsin and Tr⟹Ch[S1+L1-L2]. In trypsin, the side chain of residue Tyr172 points toward the S1 site and makes hydrogen bonds with main-chain atoms of Tyr217 and Pro225 at the two ends of loop 2 (Figures 2 and 6). In chymotrypsin the larger Trp side chain at this position donates a hydrogen bond to Pro225 but cannot function as a hydrogen bond acceptor.

The crystal structure of Tr⟹Ch[S1+L1+L2+Y172W] shows that *suc*-AAPF-CMK binds in an orientation nearly identical to that observed in the structure of Tr⟹Ch-[S1+L1+L2] (Figure 6). The only difference is in the

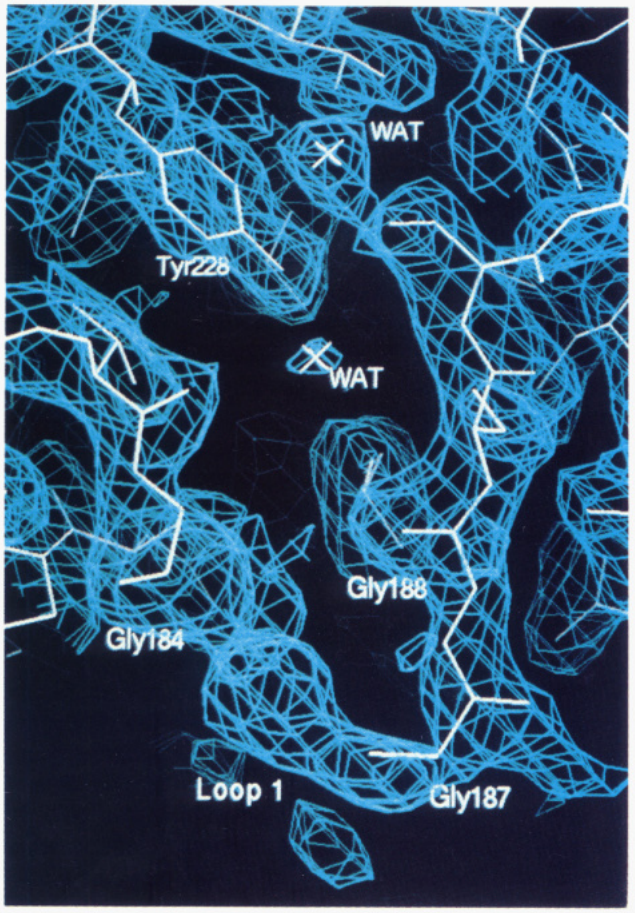


FIGURE 7: Simulated annealing OMIT map of Tr⟹Ch[S1+L1+L2+Y172W] computed with coefficients $2F_o - F_c$, using the program X-PLOR (Brunger et al., 1987). The map is contoured at 1.0σ and is calculated over the resolution range 2.1−20 Å (see Materials and Methods).

conformation of the alanine residue at position P4: in Tr⟹Ch[S1+L1+L2+Y172W] rotation of a main-chain dihedral angle of the inhibitor occurs to exchange the positions of the main-chain peptide nitrogen atom and side-chain methyl group (Figure 6). This movement permits a van der Waals contact between the nitrogen atom and the newly introduced Trp172 side chain. The other interactions of the side chain of Trp172 are identical to those observed in chymotrypsin (Brady et al., 1990). Three structural changes accompany the exchange of Tyr172 for Trp (Figure 6). First, the side chain of Ser217 rotates by 100° about the $\chi 1$ rotamer angle, adopting an orientation identical to that seen in chymotrypsin. Second, a small but significant reconfiguration of water molecules occurs to more closely mimic the solvent structure at the base of the S1 site in chymotrypsin. Simple steric considerations show that the presence of the larger Trp172 indole group causes a displacement of adjacent waters, which then also requires a reorientation of Ser217. Thermal factors in loop 2 are not significantly lowered relative to Tr⟹Ch[S1+L1+L2], although the side-chain hydroxyl group of Ser223 is now visualized in electron density maps.

A third consequence of the substitution Y172W is an improvement in the degree of order in loop 1. This effect is seen clearly in a comparison of simulated annealing omit maps calculated over this region for each of the two hybrid enzymes (Figures 5 and 7). Residue Gly187 in loop 1 is well ordered in the new enzyme, and the improved stabilization of the loop is also evident in considerably better density for residues Ala185 and Ser186. While the map indicates the approximate location of these two residues, the density does not reveal the positions of either main-chain carbonyl groups or side-chain atoms (Figure 7). Refinement of the structure following inclusion of these two residues in the model did not significantly improve the quality of the electron density; therefore, neither amino acid was included in the final model. The improved order observed crystallographically in identical ligand binding states confirms the hypothesis made on the basis of improved proflavin binding. Tr⟹Ch[S1+L1+L2+Y172W] is an improved chymotrypsin-like protease because the substitution Y172W increases the stability of the S1 binding pocket. It is important to note that both stability and specificity are improved: the increased order of the pocket is manifested in rearrangements which create an improved resemblance to chymotrypsin. Notably, the mutation Y172W improves the stability of residues at a distance of some 10 Å, across the full breadth of the S1 pocket. However, the mechanism by which this propagation occurs is not evident from a comparison of the crystal structures.

Analysis of the kinetic parameters for amide substrate hydrolysis by Tr⟹Ch[S1+L1+L2+Y172W] shows that the improved catalytic efficiency may be realized in either of the mechanistic parameters K_s or k_2 (Table 1; Hedstrom et al., 1994a). For the substrate *suc*-AAPF-*p*NA, the improvement is manifested in a 20-fold improvement in the binding affinity, K_s. Therefore, in this case the structure at the base of the specificity pocket in chymotrypsin determines binding affinity. The improved catalytic efficiency toward Ac-Phe-NH_2 is manifested as an improvement in k_2 rather than in a lowered K_s. An improved k_2 arises because the increasing enzyme−substrate complementarity required to surmount the energy barrier between the ground and transition states is favored by increased order of the S1 site. Since the smaller substrate Ac-Phe-NH_2 has few contacts with the enzyme outside of the S1 site, it requires this pocket to be more rigid so that a high proportion of the intrinsic binding energy potentially available for translation into a high catalytic rate is in fact realized (Hedstrom et al., 1994a).

DISCUSSION

Structural Determinants of Substrate Specificity. High acylation rates k_2 in serine protease catalysis require two essential factors: (i) intact catalytic machinery and (ii) accurate positioning of the substrate scissile bond relative to the oxyanion hole and Ser195/His57 couple. The crystal structures of trypsin, chymotrypsin, trypsin D189S (Perona et al., 1994), Tr⟹Ch[S1+L1+L2], and Tr⟹Ch[S1+L1+L2+Y172W] show that differences in the detailed conformation of the S1 substrate binding pocket do not affect the structure of the catalytic triad or of the S1′ enzyme site which binds the leaving group of the substrate. This indicates that differences in acylation rates among these enzymes are very likely to reflect their respective abilities to accurately position the substrate scissile bond, allowing investigation of the role of distal portions of the enzyme structures in this function. Further, the correlation between catalytic acitivity and inhibition by boronate or chloromethyl ketone inhibitors (Hedstrom et al., 1994b) indicates that the transition states must be very similar in each of the five enzymes and that their catalytic mechanisms are identical.

A further crucial feature of these enzymes is that substrate occupancy of the S1 binding site alone confers only modest specificity: chymotrypsin hydrolyzes Ac-F-NH_2 only 100-fold more efficiently than does trypsin (Hedstrom et al., 1994a). The ability of chymotrypsin to hydrolyze *suc*-AAPF-AMC and *suc*-AAPF-*p*NA nearly 10^6-fold more rapidly than trypsin is thus a consequence of binding the P2−P4 positions of the substrate in such a way that the scissile bond at P1 is accurately positioned adjacent to the Ser195/His57 catalytic couple. Chymotrypsin, Tr⟹Ch-[S1+L1+L2], and Tr⟹Ch[S1+L1+L2+Y172W] each are capable of translating the S2−S4 site binding energy of peptidyl P1-Phe substrates into very high acylation rates, while trypsin and D189S are compromised by 10^2−10^4-fold. Since exchange of the S1 site of chymotrypsin into trypsin fails to confer high k_2 values toward peptidyl P1-Phe substrates (Hedstrom et al., 1992), and since loop 1 is disordered in Tr⟹Ch[S1+L1+L2], acceleration of acylation by binding of the P2−P4 residues must occur via loop 2. Therefore, loop 2 of chymotrypsin confers upon trypsin the ability to properly orient the scissile bond of chymotrypsin-like substrates with respect to the catalytic machinery.

We carried out a comparative analysis of the structures of trypsin, chymotrypsin, D189S, Tr⟹Ch[S1+L1+L2], and Tr⟹Ch[S1+L1+L2+Y172W] in an attempt to identify conserved structural differences which might control substrate bond positioning. The main-chain conformation of the conserved Gly216 residue differs in trypsin and chymotrypsin; further, it adopts a chymotrypsin-like structure in both hybrid enzymes (Figure 8). D189S, which cannot efficiently acylate *suc*-AAPF-pNA (Table 1; Hedstrom et al., 1992), possesses a trypsin-like conformation at Gly216 (Figure 8). In all the enzymes, two antiparallel β-strand hydrogen bonds are formed by Gly216 with the P3 residue of the inhibitor. Comparison of these five structures therefore suggests that these hydrogen bonds are used to promote accurate scissile bond positioning in a discriminatory way in trypsin and chymotrypsin. If the backbone conformation of Gly216 is trypsin-like, then the hydrogen-bonding geometry of the P3−S3 interaction results in misorientation of the scissile bond for P1-Phe-containing substrates. Therefore, even though the P3 hydrogen bonds may well form with favorable distance and linearity criteria in these noncognate complexes, the free energy made available is not translated into a high value of k_2. Gly216 can function as a specificity determinant despite being conserved, because the differing structures of loop 2 in trypsin and chymotrypsin maintain it in distinct conformations.

In addition to these distal Gly216 interactions it is clear that in trypsin the S1 pocket also plays an important role. Trypsin variants D189S and D189G/G226D, in which the negatively charged carboxylate at the base of the pocket is deleted or relocated, are adversely affected in both K_s and k_2 toward P1-Arg and P1-Lys substrates (Table 1; Perona et al., 1994). Asp189 thus plays a pivotal role in both the binding and catalytic steps of trypsin hydrolysis and provides assistance in positioning the substrate scissile bond. Since D189S also does not possess high k_2 values toward peptidyl P1-Phe-containing substrates, it seems likely that a trypsin-like backbone conformation at Gly216 is not in itself sufficient for accurate scissile bond positioning, regardless of the identity of the P1 substrate residue. We expect that the inability of D189S to acylate *suc*-AAPF-AMC at high levels is not due to the lack of a chymotrypsin-like structure at the base of the pocket. Chymotrypsin appears not to require precise interactions at this position to achieve high levels of k_2, since a wild-type rate is attained by Tr⟹Ch-[S1+L1+L2] despite its partly disordered S1 site (Table 1; Hedstrom et al., 1992).

This inability of Gly216 to confer accurate scissile bond positioning allows it to function as a negative specificity determinant, which renders trypsin better able to discriminate against noncognate substrates. Improved selectivity is acheived against substrates not containing P1-Lys or Arg because only these two amino acids can make favorable electrostatic interactions with Asp189 to enhance the intrinsically inaccurate positioning by Gly216. In trypsin, the

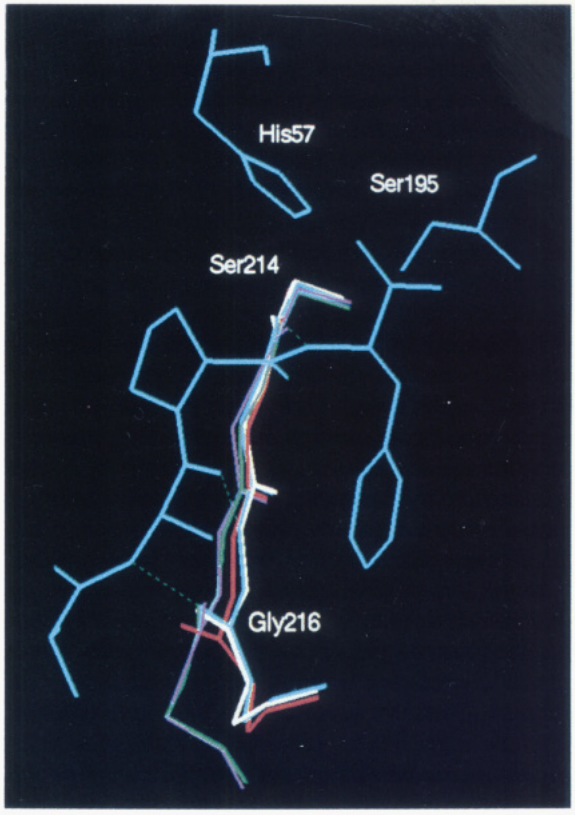


FIGURE 8: Superpositions of the structures of chymotrypsin complexed with Leu-Phe trifluoroketone (red; Brady et al., 1990), wild-type rat trypsin complexed with benzamidine (green; M. E. McGrath and R. J. Fletterick, unpublished), D189S complexed to BPTI (purple; Perona et al., 1994), Tr⟹Ch[S1+L1+L2] (white), and Tr⟹Ch[S1+L1+L2+Y172W] (light blue) along the outside rim of the S1 binding site. Binding of *suc*-AAPF-CMK (light blue) to Tr⟹Ch[S1+L1+L2] is shown with the three hydrogen bonds to Ser214 and Gly216 in dotted green lines. The distinct conformations of Tr and of D189S at Gly216, relative to the other three enzymes, are best seen in the orientations of the carbonyl oxygen atom. The hydrogen bond at this position formed by trypsin and by D189S may be weaker owing to the decreased linearity of the substrate P3 N−H- - -O216 interaction. Comparisons of the crystal structures of wild-type and mutant rat and bovine trypsins complexed to benzamidine (no interaction with Gly216), BPTI and APPI (ground-state analogs; Perona et al., 1993a), and a tripeptide boronic acid inhibitor (transition-state analog; R. M. Stroud, personal communication) show that there is no change in the conformation of Gly216. The conformation of this residue is also unchanged in all available chymotrypsin structures, both unliganded and complexed to ground-state inhibitors. These comparisons suggest that Gly216 of both trypsin and chymotrypsin does not alter conformation when bound to different inhibitors and validates the structural comparison of enzymes in differing inhibitor binding states.

Asp189 electrostatic interactions with the distal portions of the P1-Arg or Lys side chains thus function in concert with the Gly216 hydrogen bonds. The structural basis for the 10^6-fold preference of trypsin for P1-Arg/Lys relative to P1-Phe substrates is then twofold: (i) the presence of Asp189 lowers K_s by 10^2-10^3-fold for P1-Arg/Lys relative to P1-Phe substrates (Table 1; Perona et al., 1993b, 1994); (ii) the scissile bonds of P1-Phe substrates are misaligned relative to the catalytic machinery, because the Phe side chain cannot interact favorably with Asp189 to provide the required assistance to positioning by Gly216. This misalignment is reflected in acylation rates which are 10^2-10^4-fold higher to P1-Arg/Lys vs P1-Phe substrates (Table 1; Hedstrom et al., 1994a; Perona et al., 1994).

Mechanism of Conversion of Binding Energy into Catalytic Rate. The precise mechanism(s) by which binding energy is made available for acceleration of catalysis has (have) remained obscure for most enzymes. Possible mechanisms have been outlined by Jencks (1987); these include induced fit, nonproductive binding of smaller relative to larger substrates, substrate destabilization by strain or other mechanisms, and entropy reduction. In the case of trypsin and chymotrypsin, a plausible rationale for the low acylation rates exhibited by single-residue cognate substrates is that the absence of the P2−P4 distal interactions permits increased sampling of unproductive vibrational modes and bond rotations in the vicinity of the scissile bond. We suggest then that the free energy gained by formation of the P3−Gly216 hydrogen bonds with cognate substrates is used directly to reduce the internal entropy of the scissile bond. In contrast, hydrogen bonds with Gly216 formed by noncognate substrates may fix the scissile bond in a position which is either less easily attacked by Ser195-Oγ or which results in imprecise placement of the oxyanion relative to the Gly193 and Ser195 backbone amides (or both). While it is not yet possible to provide a definitive resolution of this question, we note that there is no evidence for the existence of any conformational change upon substrate binding in trypsin or chymotrypsin, such as would be expected in induced-fit or possibly in strain-based mechanisms. The use of distal substrate interactions to improve the rate of a chemical step in catalysis, rather than to improve substrate binding, has also been observed in elastase in coenzyme A transferase (Thompson & Blout, 1970, 1973; Fiercke & Jencks, 1986).

The specific increase in acylation rate for peptidyl vs single-residue cognate substrates by chymotrypsin and the two hybrid enzymes is 10^3-10^4-fold (Table 1; Hedstrom et al., 1992, 1994a), corresponding to a free energy difference of 4−6 kcal/mol. This energetic contribution arises as a consequence of new interactions at the S2−S4 sites, as well as improved interactions resulting from reduced flexibility of the P1 residue. The amount of free energy made available by the formation of one unchanged hydrogen bond has been estimated as 0.5−1.8 kcal/mol (Fersht et al., 1985) and 1.5 kcal/mol (Grobelny et al., 1989). If these estimates are accurate, then a significant proportion of the additional catalytic free energy made available by binding of the peptidyl substrates may not arise directly from the P3 hydrogen bonds but instead from improved polar and hydrophobic interactions of the P1 residue. The function of a specific Gly216 backbone conformation would then be to access this free energy by driving fixation of the scissile bond into precise alignment with catalytic groups.

Direct Correlation of Structural and Mechanistic Parameters. An important feature of the structure−function analysis is the direct correlation of structural parameters with changes in the mechanistic kinetic constants K_s and k_2. In most enzymes studied, a precise description of structural elements playing critical roles in the separate binding and catalytic steps has been elusive, in part because the steady-state Michaelis parameters k_{cat} and K_m often reflect combinations of more fundamental mechanistic constants involving both binding affinity and catalytic turnover. Here the mechanistic kinetic parameters K_s, k_2, and k_3, which represent true measures of binding affinity, acylation rate, and deacylation rate, are calculated from the Michaelis constants k_{cat} and K_m by measurement of hydrolysis rates toward structurally similar amide and ester substrates, based on the assumption that deacylation is rate-limiting for ester substrates (Table 1, legend; Zerner & Bender, 1964; Hedstrom et al., 1992, 1994a). The five enzymes trypsin, D189S, Tr⇒Ch[S1+L1+L2], Tr⇒Ch[S1+L1+L2+Y172W], and chymotrypsin possess K_s and k_2 constants which vary over a wide range for hydrolysis of peptidyl chymotrypsin-like substrates. The crystal structures of each of these enzymes are herein compared, resulting in two specific correlations: high acylation rates k_2 depend upon a specific Gly216 conformation, and tight binding K_s depends upon the enzyme and solvent structure at the base of the S1 pocket. Approaches to understanding structure−function relationships which correlate differences in structure with the second-order rate constant k_{cat}/K_m necessarily give more limited insight, because no distinction is made among the individual mechanistic steps (Bone et al., 1991). The approach described here should be of general value to the analysis of specificity in many enzyme families. In some cases, kinetic parameters may be calculated from the Michaelis constants, as is possible for the serine protease mechanism. Alternatively, they can be measured directly through pre-steady-state kinetic analysis.

A General Role for Residue 216 in the Trypsin-like Proteases. A third pancreatic serine protease, elastase, cleaves amide bonds adjacent to small hydrophobic amino acids. Superposition of seven crystal structures of mammalian serine proteases (Figure 9) shows a striking correlation between the P1 site specificity of the enzyme and the conformation of the polypeptide backbone at residue 216. The backbone torsional angles for this amino acid fall into three groups which match the trypsin-like (trypsin, kallikrein, and α-thrombin), chymotrypsin-like (chymotrypsin and mast cell protease), and elastase-like (pancreatic and neutrophil elastases) P1 site specificities (Figure 9, legend). Position 216 is Val in both pancreatic and human neutrophil elastases; in these enzymes this amino acid fulfills a dual role both by providing a hydrophobic platform for the P1 residue and by forming the main-chain hydrogen bonds at position P3 (Navia et al., 1989; Watson et al., 1970). Occupancy of the extended peptide binding site in both pancreatic and neutrophil elastases also increases the rate of acylation in preference to binding affinity or deacylation rate (Thompson & Blout, 1970, 1973; Stein et al., 1987). In the case of neutrophil elastase, it is also known that the interactions with the extended peptide binding site increase k_2 for specific but not for nonspecific substrates (Stein et al., 1987). Therefore, we suggest that hydrogen bonds formed by the P3 residue

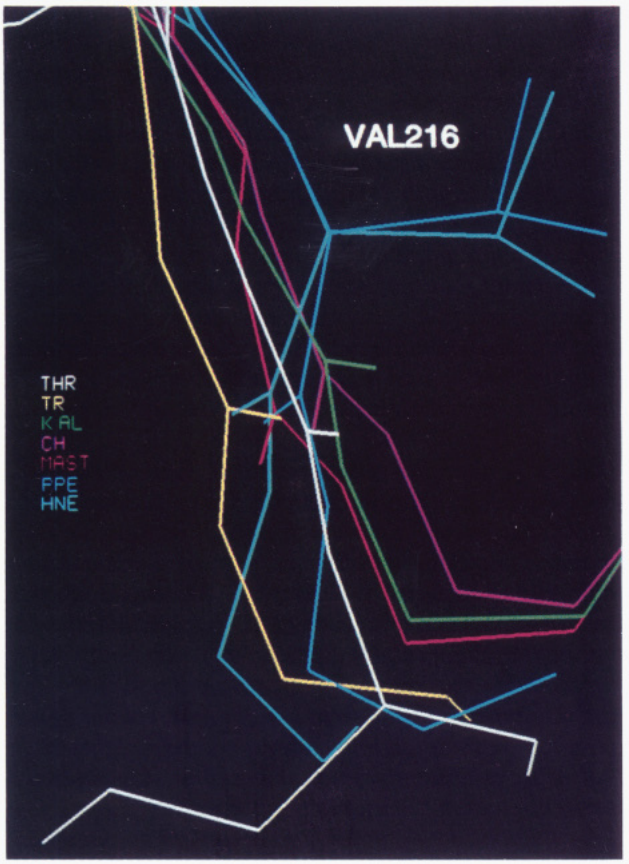


FIGURE 9: Superpositions of the structures of bovine trypsin (TR, yellow), porcine kallikrein (KAL, green), human α-thrombin (THR, white), bovine chymotrypsin (red), rat mast cell protease (MAST, pink), porcine pancreatic elastase (PPE, dark blue), and human neutrophil elastase (HNE, light blue) at residue 216. The backbone conformational angles at this position are TR, $\Phi = 163.5°$, $\Psi = 169.9°$ (PDB 3PTN; Marquardt et al., 1983); KAL, $\Phi = -175.9°$, $\Psi = 170.1°$ (PDB 2PKA; Bode et al., 1983); THR, $\Phi = 173.0°$, $\Psi = 170.3°$ (PDB 1ABJ; Bode et al., 1989); CH, $\Phi = 174.5°$, $\Psi = -133.5°$ (PDB 7GCH; Brady et al., 1990); MAST, $\Phi = -157.8°$, $\Psi = -148.1°$ (PDB 3RP2; Remington et al., 1988); PPE, $\Phi = -133.8°$, $\Psi = 176.7°$ (PDB 3EST; Meyer et al., 1988); and HNE, $\Phi = -118.2°$, $\Psi = 166.8°$ (PDB 1HNE; Navia et al., 1989). For Tr⟹Ch[S1+L1+L2] and Tr⟹Ch[S1+L1+L2+Y172W], these angles are $\Phi = 180.0°$, $\Psi = -155.5°$ and Φ 176.5°, $\Psi = -157.5°$, respectively. All of the structures are determined at resolutions equal to or better than 2.0 Å, with the exception of α-thrombin determined at 2.4 Å. Examination of all other available structures of these enzymes substantiates the division into three groupings.

of elastase substrates with Val216 are similarly critical to accurate substrate positioning. Conversion of either trypsin or chymotrypsin to an elastase-like protease is thus predicted to require mutations sufficient to reorient the main chain of position 216 to an elastase-like conformation. Further, introduction of the S1 site residues of human neutrophil elastase into trypsin fails to introduce specificity toward elastase-specific substrates (J. J. Perona and C. S. Craik, unpublished observations). This suggests that, as for chymotrypsin, specificity determinants for cleavage by elastase also include amino acids not directly in contact with substrate.

The correlation between the conformation of the backbone at position 216 and the P1 site specificity holds for all available structures of mammalian trypsin-like proteases [coordinates for the recently determined factor Xa (Padmanabhan et al., 1993) and complement factor D (Narajana et al., 1994) enzymes are not yet available; the structure of rat tonin (Fujinaga & James, 1987) was determined in an inactive conformation]. It is also true for the microbial *S griseus* trypsin (Read & James, 1988) which despite its origin is nonetheless highly homologous to the mammalian enzymes. However, the correlation breaks down for the other microbial enzymes *Streptomyces griseus* proteases A and B (Delbaere et al., 1979; Moult et al., 1985) and *L. enzymogenes* α-lytic protease [Brayer et al., 1979; Bone et al., 1987; coordinates for the recently determined structure of *S. griseus* protease E (Nienaber et al., 1993) are not yet available]. These enzymes retain the trypsin fold but are smaller in size and possess substantial structural differences in surface loops including loops 1 and 2. Other structural mechanisms for determining specificity may thus be extant in these proteases (see below).

Alternate Mechanisms by Which Global Protein Structure Influences Specificity. Since the distal elements loops and 1 and 2 and Trp172 specify the conformation of crucial amino acids directly contacting substrate, they must be considered as specificity determinants. We distinguish then between two classes of such determinants: (1) amino acids such as Gly216 and Asp189 which directly contact substrate and may be viewed as primary; (2) the distal elements, which play an active role beyond providing some stability to the primary site. The existence of the secondary distal determinants suggests that substrate specificity in trypsin and chymotrypsin may be better viewed as a distributed property of the protein fold. Since Tr⟹Ch[S1+L1+L2+Y172W] remains 6−50-fold less efficient than chymotrypsin, it appears likely that additional distal determinants remain to be identified. The high value of K_s toward peptidyl P1-Phe substrates exhibited by this enzyme is almost certainly a consequence of the still incompletely formed S1 site (Figures 6 and 7) and suggests further amino acids as good targets for continued mutagenesis.

Although determining the conformation of proximal amino acids may be a common strategy by which distal elements influence specificity, an alternative mechanism could be by modulating the degree of flexibility present in the substrate binding site. Good evidence for this second possibility is provided by the properties of two Met⟹Ala mutants in the S1 site of α-lytic protease, which greatly broaden specificity from an elastase-like profile to include as well the large aromatic side chains preferred by chymotrypsin (Bone et al., 1989). Crystallographic analysis suggests that the ability of the enzyme to shift to a broad specificity profile is due to backbone flexibility of the S1 site, particularly in the region of residue 216 and subsequent amino acids, which shift position significantly in different enzyme−substrate analog structures (Bone et al., 1991). In contrast, the crystal structures of trypsins D189S, D189G/G226D, and G226A show that there is no effect on the conformation of Gly216 as a result of substitution inside the S1 pocket (Perona et al., 1993b, 1994; Wilke et al., 1991).

These observations suggest that trypsin and α-lytic protease may differ in the degree of rigidity present at residue 216 in the S1 site. The more rigid trypsin pocket would likely lead to a very strong dependence of catalytic efficiency on retention of the full native structure. By this model, D189S does not acylate chymotryptic substrates at a high rate because the necessary flexibility required for conversion of the Gly216 backbone to a chymotrypsin-like conformation is not present. Exchange of the distal surface loops, which

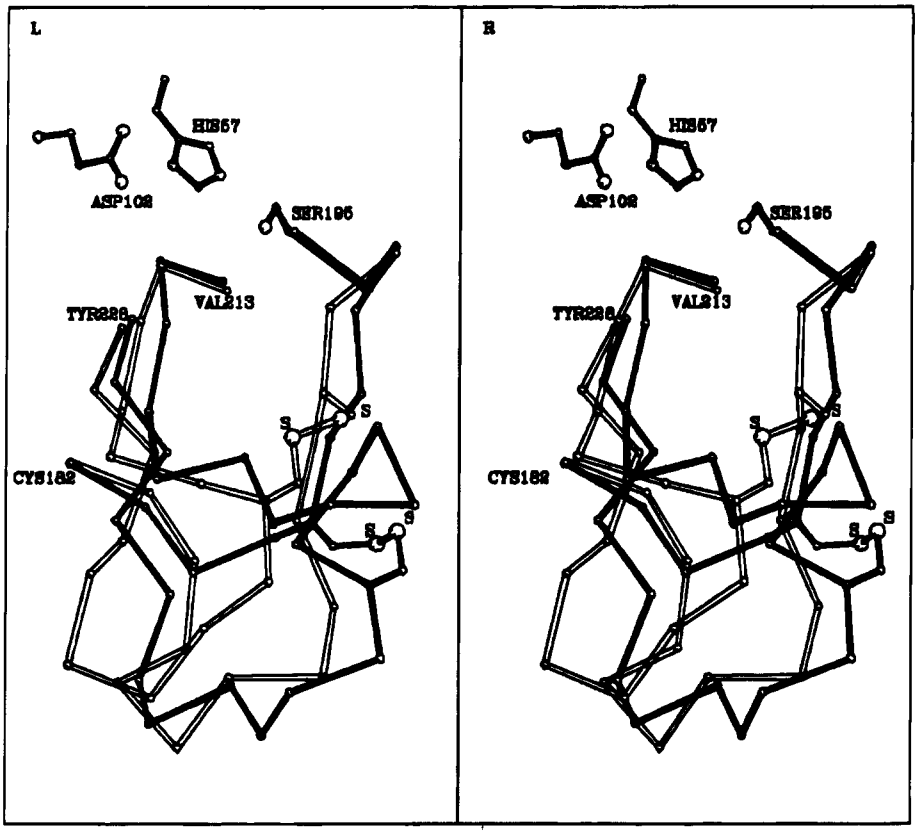


FIGURE 10: Stereoview of a superposition of the primary S1 binding pockets and adjacent surface loops of wild-type rat trypsin (open lines; M. E. McGrath and R. J. Fletterick, unpublished) and α-lytic protease (filled lines). α-Carbon atoms of the enzymes were optimally superimposed using OVRLAP (Rossman & Argos, 1976). The ends of the chains are labeled with the amino acids present in trypsin. The locations of the disulfide bonds are marked with the label S at the sulfur atoms; the divergence in connecting loop structure results in their positioning approximately 7 Å apart in the two enzymes. Loops 1 and 2 of trypsin are indicated at the bottom of the figure.

maintain the Gly216 conformation, is therefore necessary to achieve specificity modification. Of particular interest now is an understanding of the differing ways in which the second-shell scaffolds of trypsin and α-lytic protease may function to create rigid and flexible S1 sites, respectively. These scaffolds differ substantially in structure, since in α-lytic protease loop 1 is missing entirely and loop 2 is greatly enlarged (Figure 10). Further mutational analysis, targeting the specific interactions which tether the S1 sites of each enzyme to these and other surrounding portions of the structure, may offer an excellent opportunity to decipher the essential design features which specify a flexible or rigid active site.

ACKNOWLEDGMENT

We thank C. Collins and R. Wagner for assistance with crystal growth and analysis and C. Tsu and S. Gillmor for critical reading of the manuscript.

REFERENCES

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237−282.

Bode, W., Mayr, I., Baumann, U., Huber, R., Stone, S. R., & Hofsteenge, J. (1989) *EMBO J. 8*, 3467−3475.

Bone, R., Shenvi, A. B., Kettner, C. A., & Agard, D. A. (1987) *Biochemistry 26*, 7609−7614.

Bone, R., Silen, J. L., & Agard, D. A. (1989) *Nature 339*, 191.

Bone, R., Fujishige, A., Kettner, C. A., & Agard, D. A. (1991) *Biochemistry 30*, 10388−10398.

Brady, K., Wei, A., Ringe, D., & Abeles, R. H. (1990) *Biochemistry 29*, 7600.

Brayer, G. D., Delbaere, L. T. J., & James, M. N. G. (1979) *J. Mol. Biol. 131*, 743−775.

Brunger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458.

Brunger, A. T., Krukowski, A., & Erickson, J. W. (1990) *Acta Crystallogr. A46*, 585−593.

Dayringer, H., Tramontano, A., Sprang, S., & Fletterick, R. J. (1986) *J. Mol. Graphics 4*, 82.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature 314*, 255.

Fierke, C. A., & Jencks, W. P. (1986) *J. Biol. Chem. 261*, 7603.

Fujinaga, M., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 373−396.

Graf, L., Jancso, A., Szilagyi, L., Hegyi, G., Pinter, K., Naray-Szabo, G., Hepp, J., Medzihradszky, K., & Rutter, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4961−4965.

Grobelny, D., Goli, U. B., & Galardy, R. E. (1989) *Biochemistry 28*, 4948.

Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249.

Hedstrom, L., Perona J. J., & Rutter, W. J. (1994a) *Biochemistry 33*, 8757−8763.

Hedstrom, L., Farr-Jones, S., Kettner, C. A., & Rutter, W. J. (1994b) *Biochemistry 33*, 8764−8769.

Huber, R., & Bode, W. (1978) *Acc. Chem. Res. 11*, 114−122.

Jencks, W. P. (1969) *Catalysis in Chemistry and Enzymology*, Dover, New York.

Jencks, W. P. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 65.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157.

Kettner, C. A., Bone, R., Agard, D. A., & Bachovchin, W. W. (1988) *Biochemistry 27*, 7682.

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

Marquardt, M., Walter, J., Deisenhofer, J., Bode, W., & Huber, R. (1983) *Acta Crystallogr. B39*, 480.
Matthews, B. W., Sigler, P. B., Henderson, R., & Blow, D. M. (1967) *Nature 214*, 652.
Meyer, E., Cole, G., Radhakrishnan, R., & Epp, O. (1988) *Acta Crystallogr. B44*, 26.
Moult, J., Sussman, F., & James, M. N. G. (1985) *J. Mol. Biol. 182*, 555–566.
Narajana, S. V., Carson, M., el-Kabbiani, O., Kilpatrick, J. M., Moore, D., Chen, X., Bugg, C. E., & DeLucas, L. J. (1994) *J. Mol. Biol. 235*, 695–708.
Navia, M. A., McKeever, B. M., Springer, J. P., Lin, T.-Y., Williams, H. R., Fluder, E. M., Dorn, C. P., & Hoogsteen, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7–11.
Nienaber, V. L., Breddam, K., & Birktoft, J. J. (1993) *Biochemistry 32*, 11469–11475.
Padmanabhan, K., Padmanabhan, K. P., Tulinsky, A., Park, C. H., Bode, W., Huber, R., Blankenship, D. T., Cardin, A. D., & Kisiel, W. (1993) *J. Mol. Biol. 232*, 947–966.
Perona, J. J., Tsu, C. A., Craik, C. S., & Fletterick, R. J. (1993a) *J. Mol. Biol. 230*, 919–933.
Perona, J. J., Tsu, C. A., McGrath, M. E., Craik, C, S., & Fletterick, R. J. (1993b) *J. Mol. Biol. 230*, 934–949.
Perona, J. J., Hedstrom, L., Wagner, R., Rutter, W. J., Craik, C. S., & Fletterick, R. J. (1994) *Biochemistry 33*, 3252–3259.
Polgar, L. (1989) *Mechanisms of Protease Action* Chapter 3, CRC Press, Boca Raton, FL.
Poulos, T. L., Alden, R. A., Freer, S. T., Birktoft, J. J., & Kraut, J. (1976) *J. Biol. Chem. 251*, 1097–1103.
Powers, J. C., & Harper, J. W. (1986) in *Proteinase Inhibitors* (Barrett, A. J., & Salvesen, G., Eds) p 95, Elsevier, Amsterdam and New York.
Read, R. J., & James, M. N. G. (1988) *J. Mol. Biol. 200*, 523–551.
Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., & Neurath, H. (1988) *Biochemistry 27*, 8097.
Rossmann, M. G., & Argos, P. (1976) *J. Mol. Biol. 105*, 75–95.
Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417.
Schechter, I., & Berger, A. (1968) *Biochem. Biophys. Res. Commun. 27*, 157.
Stein, R. L., Strimpler, A. M., Hori, H., & Powers, J. C. (1987) *Biochemistry 26*, 1301.
Steitz, T. A., Henderson, R., & Blow, D. M. (1969) *J. Mol. Biol. 46*, 337.
Stroud, R. M. (1974) *Sci. Am. 231* (1) (July), 74–88.
Thompson, R. C., & Blout, E. R. (1970) *Proc. Natl. Acad. Sci. U.S.A. 67*, 1734.
Thompson, R. C., & Blout, E. R. (1973) *Biochemistry 12*, 57.
Watson, H. C., Shotton, D. M., Cox, J. C., & Muirhead, H. (1970) *Nature 225*, 806.
Wilke, M. E., Higaki, J. N., Craik, C. S., & Fletterick, R. J. (1991) *J. Mol. Biol. 219*, 525–532.
Zerner, B., & Bender, M. L. (1964) *J. Am. Chem. Soc. 86*, 3669.

BI9414101