Converting Trypsin to Chymotrypsin: Residue 172 Is a Substrate Specificity Determinant†

Lizbeth Hedstrom,*,‡ John J. Perona,§ and William J. Rutter‖

Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02254, and Department of Pharmaceutical Chemistry, School of Pharmacy, Hormone Research Institute, and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143-0446

Received January 27, 1994; Revised Manuscript Received April 25, 1994®

ABSTRACT: Trypsin and chymotrypsin have very similar tertiary structures, yet very different substrate specificities. Recent site-directed mutagenesis studies have shown that mutation of the residues of the substrate binding pocket of trypsin to the analogous residues of chymotrypsin does not convert trypsin into a protease with chymotrypsin-like specificity. However, chymotrypsin-like substrate specificity is attained when two surface loops are changed to the analogous residues of chymotrypsin, in conjunction with the changes in the S1 binding site [Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249–1253). This mutant enzyme, Tr→Ch[S1+L1+L2], is improved to a protease with 2–15% of the activity of chymotrypsin by the mutation of Tyr172 to Trp. Residue 172 interacts synergistically with the residues of the substrate binding pocket and the loops to determine substrate specificity. Further, these trypsin mutants demonstrate that substrate specificity is determined by the rate of catalytic processing rather than by substrate binding.

The differences in substrate specificity of enzymes with homologous structures provide clues to the structural basis of the exquisite selectivity of enzymatic reactions. The diversity of substrate specificity within a single structural motif is well illustrated by the trypsin family of serine proteases. Trypsin, chymotrypsin, and elastase have very different substrate specificities: trypsin hydrolyzes peptides with arginine or lysine residues at the P1 position [nomenclature from Schechter and Berger (1967)], and chymotrypsin prefers large hydrophobic residues at the P1 position, while elastase prefers small aliphatic residues [for a review, see Polgar (1989)]. These different substrate specificities appear to result from simple structural alterations of the S1 binding pockets of these proteases: Asp189 at the bottom of the S1 binding pocket of trypsin accounts for its specificity for positively charged residues (Kossiakoff, 1987); the analogous residue in chymotrypsin is Ser, which creates a more hydrophobic S1 binding pocket (Blow, 1971). The S1 binding pocket of elastase is occluded by the side chains of Val216 and Thr226 (both residues are Gly in trypsin and chymotrypsin), which accounts for elastase's specificity for small aliphatic residues (Shotton & Watson, 1969). This model predicts that exchanging these few structural elements should be sufficient to exchange substrate specificity (Steitz et al., 1969).

Recent site-directed mutagenesis experiments indicate that this simple model cannot be correct (Graf et al., 1988; Hedstrom et al., 1992). The mutation of Asp189 to Ser in trypsin does not convert trypsin into a protease with chymotrypsin-like specificity; D189S is a poor, nonspecific protease. Further mutations of the S1 binding pocket of trypsin to the analogous residues of chymotrypsin (e.g., Gln192 to Met, Ile138 to Thr, and insert Thr219) do not transfer chymotrypsin-like specificity to trypsin. However, a trypsin mutant containing a chymotrypsin-like S1 binding pocket does acquire chymotrypsin-like specificity when two surface loops are replaced with the analogous loops from chymotrypsin. Loop 1 (residues 185-188) and loop 2 (residues 221–225) connect the walls of the S1 binding pocket, but do not contact the substrate (Hedstrom et al., 1992) (Figure 1). The amino acid sequence of loop 1 is conserved among trypsins and distinct from the equally conserved sequences of loop 1 in chymotrypsin and elastase (Figure 2). Loop 2 has two sequence motifs in trypsin, which again are distinct from the conserved sequences of loop 2 in chymotrypsin and elastase. The mechanism by which these loops influence substrate specificity in the trypsin family is not understood; they may cause small differences in S1 binding pocket structure which are crucial in determining substrate specificity, or they may mediate conformational flexibility which occurs during the hydrolysis reaction.

This mutant, Tr→Ch[S1+L1+L2], has several interesting properties which have important implications for the mechanism of substrate discrimination (Hedstrom et al., 1992) (see Figure 2 for mutant designations). Tr→Ch[S1+L1+L2] hydrolyzes Suc-Ala-Ala-Pro-Phe-AMC[1] 10^3-fold slower than chymotrypsin as measured by k_{cat}/K_m. Tr→Ch[S1+L1+L2], like chymotrypsin, can utilize substrate interactions with the S2, S3, and S4 binding sites to increase the rate of substrate hydrolysis. More importantly, Tr→Ch[S1+L1+L2] is comparable to chymotrypsin in the rate of acylation and deacylation of Suc-Ala-Ala-Pro-Phe-AMC, but is impaired in substrate binding. The following model accounts for these results: Tr→Ch[S1+L1+L2] exists as two conformations in equilibrium;

† This work was supported by the Lucille P. Markey Charitable Trust to Brandeis University and by NIH Postdoctoral Award GM13818-03 (J.J.P.) and NIH Grant DK21344 (W.J.R.).

‡ Graduate Department of Biochemistry, Brandeis University.

§ Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco.

‖ Hormone Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco.

® Abstract published in *Advance ACS Abstracts,* June 15, 1994.

[1] Abbreviations: AMC, 7-amino-4-methylcoumarin; Suc-Ala-Ala-Pro-Phe-AMC, succinylalanylalanylprolylphenylalanyl-AMC; Suc-Ala-Ala-Pro-Phe-SBzl, succinylalanylalanylprolylphenylalanine thiobenzyl ester; Suc-Ala-Ala-Pro-Phe-pNA, succinylalanylalanylprolylphenylalanyl-*p*-nitroanilide; Suc-Ala-Ala-Pro-Tyr-AMC, succinylalanylalanylprolyl-tyrosyl-AMC; Suc-Ala-Ala-Pro-Trp-AMC, succinylalanylalanylprolyl-tryptophanyl-AMC; Suc-Ala-Ala-Pro-Lys-AMC, succinylalanylalanylalanyl-prolyllysyl-AMC; Ac-Phe-pNP, *N*-acetylphenylalanine *p*-nitrophenyl ester; Ac-Phe-NH_2, *N*-acetylphenylalaninamide.

0006-2960/94/0433-8757$04.50/0

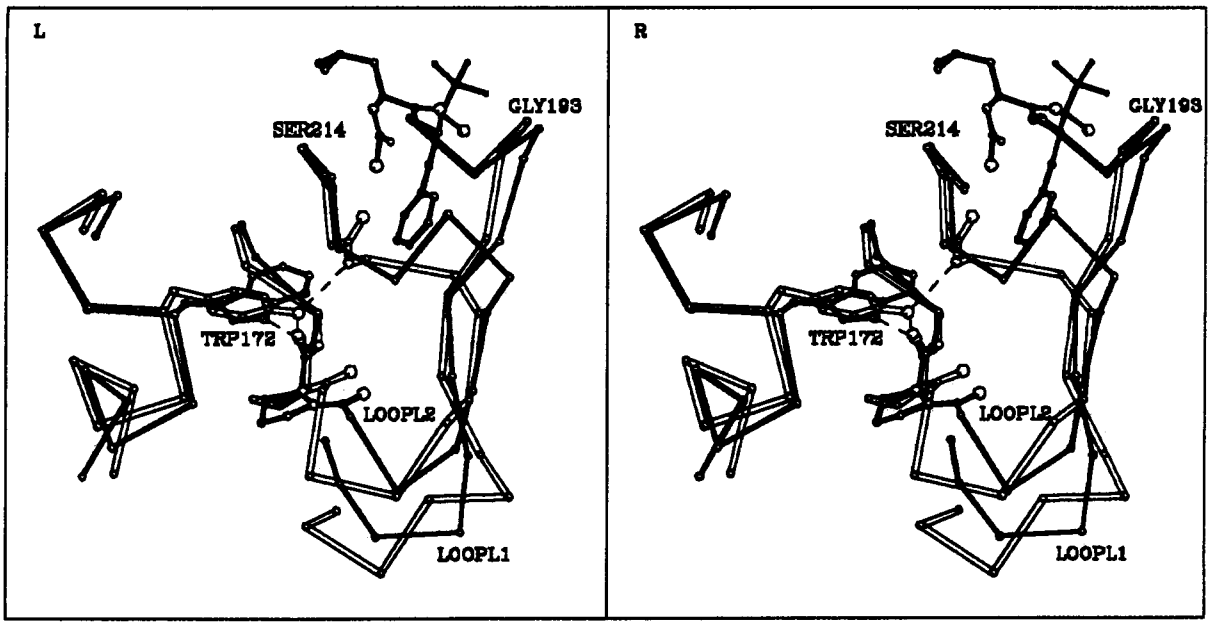

FIGURE 1: Interactions of residue 172 with the S1 binding pockets of trypsin and chymotrypsin. The complex of rat trypsin and benzamidine (open lines) (M. E. McGrath and R. J. Fletterick unpublished results) was superimposed on the complex of γ chymotrypsin and *N*-acetyl-L-leucyl-L-phenylalanine trifluoroketone (solid lines) (Brady et al., 1990) using OVRLAP (Rossman & Argos, 1976). The α carbons of the S1 binding pockets of both enzymes are displayed, as are loops 1 and 2, residues 172 and Pro225, and the trifluoroketone inhibitor. Hydrogen bonds between residue 172 and Pro225 are shown by dashed lines.

	138	172	182	190	200	210	220	230
Cow chymotrypsin A*	t	w	cag-asgv-**ssc**m**gds**ggplvckkngawtlvgiv**swgsstc**s-tst**pgvy**arv
Dog chymotrypsin†	t	.	.....q.d.v...a.....**g**.....s.
Cow trypsin§	i	y	...yle.gk**d**..**q**.....v..--s.k--.q.....-.aqknk....tk.
Rat trypsin-II¶	i	y	.v.fle.gk**d**..**q**.....v..--n.e--.q.....**yg**-.alpdn....tk.
D189S	i	y	.v.fle.gk**S**..**q**.....v..--n.e--.q.....**yg**-.alpdn....tk.
Y172W	i	W	.v.fle.gk**S**..**q**.....v..--n.e--.q.....**yg**-.alpdn....tk.
Tr->Ch[S1+L1+L2]	T	y	.v.-AS.g-**S**..**M**.....v..--n.e--.q.....**SgT**.S-TST....tk.
Tr->Ch[S1+L1+L2+D189]	T	y	.v.-AS.g-**d**..**M**.....v..--n.e--.q.....**SgT**.S-TST....tk.
Tr->Ch[S1+L1*+L2]	T	y	.A.-AS.V-**S**..**M**.....v..--n.e--.q.....**SgT**.S-TST....tk.
Tr->Ch[S1+L1+L2+Y172W]	T	W	.v.-AS.g-**S**..**M**.....v..--n.e--.q.....**SgT**.S-TST....tk.
			------				-----	
			Loop 1				*Loop 2*	

FIGURE 2: Sequences of chymotrypsin, trypsin, and trypsin mutants. Chymotrypsin numbering is used. Periods denote residues which are identical to cow chymotrypsin, bold lettering identifies residues that form the S1 binding pocket, and capital letters denote mutations. Loops 1 and 2 are identified by dashed lines. *Brown and Hartley (1966); †Pinsky et al. (1983); §Titani et al. (1975); ¶(Craik et al., (1984).

the predominant species is in an inactive conformation where the S1 binding pocket is deformed, and the minor species is in an active, chymotrypsin-like conformation which can be trapped by high substrate concentrations. Therefore, substrate binding is impaired relative to that of chymotrypsin because only a small fraction of Tr→Ch[S1+L1+L2] can bind substrate. However, when substrate is bound, it reacts as rapidly as if it were bound to chymotrypsin. This model implies that Tr→Ch[S1+L1+L2] can be improved by designing mutations to stabilize this active conformation.

Here we identify an additional substrate specificity determinant in the trypsin family of serine proteases, residue 172. Mutation of Tyr172 to Trp, the analogous residue in chymotrypsin, improves the activity of Tr→Ch[S1+L1+L2] 20–50-fold, creating an enzyme that has 2–15% of the activity of chymotrypsin.

MATERIALS AND METHODS

Materials. Suc-Ala-Ala-Pro-Phe-AMC was purchased from Enzyme Systems Products (Livermore, CA). Chymotrypsin (TLCK treated), Suc-Ala-Ala-Pro-Phe-SBzl, Suc-Ala-Ala-Pro-Phe-pNA, 4,4′-dithiodipyridine, and Ac-Phe-pNP were purchased from Sigma Chemical Co. Ac-Phe-NH_2 was purchased from Research Plus (Bayonne, NJ). Suc-Ala-Ala-Pro-Tyr-AMC, Suc-Ala-Ala-Pro-Trp-AMC, and Suc-Ala-Ala-Pro-Lys-AMC were the generous gift of Dr. Laszlo Graf (Eotvos Lorand University, Budapest).

Construction of Mutant Trypsins. Site-directed mutagenesis was performed by the method of Kunkel as described previously (Kunkel, 1985; Hedstrom et al., 1992). All mutants were completely sequenced to insure that only the desired mutation was introduced. The following oligonucleotides were used (mismatched bases are underlined):

Tr→Ch[S1+L1*+L2], ATG-GTC-TGT-GCA-GGC-GCC-TCG-GGA-GTC-TCT-TCC-TGC; Y172W and Tr→Ch-[S1+L1+L2+Y172W], GCC-TCC-TGG-CCT-GGA-AAG-ATC; Tr→Ch[S1(D189)+L1+L2], TCG-GGA-GGC-GAT-TCC-TGC-ATG.

Expression and Isolation of Mutant Trypsins. Recombinant rat trypsinogen II was produced in a *Saccharomyces cerevisiae* expression system from the pYT plasmid, which contained the ADH/GAPDH promoter and α factor leader sequence fused to the trypsinogen coding sequences (Hedstrom et al., 1992). Culture medium (2 L) containing trypsinogen was

Table 1: Specificity of Chymotrypsin, Trypsin, and Trypsin Mutants[a]

enzyme	Suc-Ala-Ala-Pro-X-AMC k_{cat}/K_m (M^{-1} s^{-1}), X = Phe	Tyr	Trp	Lys	Phe:Tyr:Trp:Lys
chymotrypsin	$(4.9 \pm 0.5) \times 10^5$	$(8.6 \pm 0.2) \times 10^5$	$(1.3 \pm 0.1) \times 10^6$	69 ± 1	$1:2:3:10^{-4}$
trypsin	6.0 ± 0.2	8.9 ± 0.3	1.48 ± 0.01	$(1.9 \pm 0.1) \times 10^5$	$1:1.5:0.3:10^4$
D189S	38 ± 1	38 ± 8	4.8 ± 3	7.3 ± 0.5	1:1:0.1:0.2
Tr→Ch[S1+L1+L2]	800 ± 20	$(6.1 \pm 0.6) \times 10^3$	$(1.0 \pm 0.1) \times 10^3$	10 ± 1	1:7:1:0.01
Tr→Ch[S1+L1+L2+D189]	1.24 ± 0.01	nd	nd	15.3 ± 0.2	1:nd:nd:13
Tr→Ch[S1 + L1* + L2]	240 ± 10	nd	nd	nd	nd
Tr→Ch[S1 + L1 + L2 + Y172W]	$(9.3 \pm 0.7) \times 10^3$	$(6.1 \pm 0.2) \times 10^4$	$(3.2 \pm 0.3) \times 10^4$	9.0 ± 0.7	$1:6:3:10^{-3}$
Y172W	5.4 ± 0.2	10.1 ± 0.4	≤ 2.7	$(2.5 \pm 0.1) \times 10^3$	1:2:0.5:500

[a] Conditions: 50 mM Hepes, pH 8.0, 100 mM NaCl, and 10 mM $CaCl_2$, 37 °C. Assays were performed at [S] $\ll K_m$, as described in Materials and Methods. Note that the values for chymotrypsin and Tr→Ch[S1 + L1 + L2] are slightly lower (~3-fold) than our previous reports. This discrepancy most likely results from an improvement in instrumentation, resulting in better sensitivity and thermostating. The relative activities of chymotrypsin and Tr→Ch[S1 + L1 + L2] are unchanged. The values for D189S and wild-type trypsin agree with previous reports from the Graf laboratory (Graf et al., 1988). nd, not determined.

absorbed on a 30-mL column of Toyopearl 650M (Supelco) and eluted with a gradient of 100 mM AcOH/2 mM NaOAc to 100 mM Tris-HCl, pH 8.0 (120 mL total). The trypsinogen-containing fractions were pooled and concentrated to ~15 mL with a Centriprep (Amicon). The trypsinogen was activated by addition of enterokinase (Biozyme). Trypsin was isolated on a soybean trypsin inhibitor resin (Sigma Chemical Co.). Trypsin, D189S, and Tr→Ch[S1+L1+L2+Y172W] bind to soybean trypsin inhibitor and elute in 0.1 N formic acid, while the other trypsin mutants elute in 0.5 M NaCl. The trypsin-containing fractions were pooled, dialyzed against 20 mM NaAc, pH 5.2, and chromatographed on a MonoS column (Pharmacia) eluted with a linear gradient of 0–1 M NaCl. Trypsin is stored at 4 °C in 1 mM HCl and 10 mM $CaCl_2$. Enzyme concentration was determined by titration with *p*-nitrophenyl *p'*-guanidinobenzoate.

Activity of Mutant Trypsins. Assay mix contained 50 mM Hepes, pH 8.0, or 50 mM Mops, pH 6.5, 10 mM $CaCl_2$, and 0.1 M NaCl. Stock solutions of substrates were prepared in dimethylformamide. The final concentration of the dimethylformamide in the assay solution was less than 4%. Hydrolysis of the AMC substrates was monitored fluorometrically, with excitation at 380 nm and emission at 460 nm (Zimmerman et al., 1977). Assays were performed using 2.0 mL of assay mix containing substrate (generally 1.2 μM to 1 mM) in a stirred cell at 37 °C with a Hitachi F2000 spectrofluorimeter. Concentrations of fluorometric substrates were determined by total hydrolysis with chymotrypsin or trypsin. Suc-Ala-Ala-Pro-Phe-pNA (125 μM to 5 mM) hydrolysis was followed spectrophotometrically (ϵ_{410} = 8.48 mM^{-1} cm^{-1}). Suc-Ala-Ala-Pro-Phe-SBzl (2.5 μM to 50 μM) hydrolysis was monitored spectrophotometrically in the presence of 4,4'-dithiodipyridine (25 μM) (ϵ_{324} = 19.8 mM^{-1} cm^{-1}). Ac-Phe-pNP (10–100 μM) hydrolysis was measured spectrophotometrically (ϵ_{400} = 5.4 mM^{-1} cm^{-1} at pH 6.5). Ac-Phe-NH_2 (1.2–40 mM) hydrolysis was measured using ninhydrin to monitor ammonia release (Brady & Abeles, 1990; Hedstrom et al., 1992). Ac-Phe-NH_2 assay mixes contained 5.5–11 μM Tr→Ch[S1+L1+L2+Y172W] in 250 μL. Aliquots (50 μL) were removed at 4 or 5 intervals over 150 min. Spectrophotometric measurements were performed on a Beckman DU650, Perkin-Elmer 552, or Hitachi U2000 spectrophotometer. Data was analyzed using KinetAsyst software, and reported values are the average of at least two experiments.

RESULTS AND DISCUSSION

Residue 189 Interacts Synergistically with Loops 1 and 2 To Determine Substrate Specificity. The specificity of trypsin for Lys- and Arg-containing peptides results from the interaction of the positive charge of the substrate with Asp189. Mutation of Asp189 to Ser, the analogous residue in chymotrypsin, reduces the activity on Lys- and Arg-containing substrates 10^5-fold, without significantly increasing activity on hydrophobic substrates (Graf et al., 1988). This result suggests that residue 189 interacts with other regions of trypsin to determine substrate specificity. This postulate is substantiated by the observation that introduction of Asp189 into Tr→Ch[S1+L1+L2] does not confer activity on Lys- and Arg-containing substrates (Table 1). This mutant, Tr→Ch-[S1+L1+L2+D189], has 10^3-fold less activity on Suc-Ala-Ala-Pro-Phe-AMC than Tr→Ch[S1+L1+L2] as measured by k_{cat}/K_m [this parameter is used to assess substrate specificity (Fersht, 1985)]; however, activity on Suc-Ala-Ala-Pro-Lys-AMC is not restored by this mutation. These results suggest that loops 1 and 2 direct the function of residue 189 in determining substrate specificity.

Improvements of Tr→Ch[S1+L1+L2]. Two trypsin mutants were designed to improve the chymotrypsin-like activity of Tr→ Ch[S1+L1+L2] (Figure 1). The first, Tr→Ch-[S1+L1*+L2], contains additional changes in the region of loop 1: Val183 and Gly188 are changed to the analogous residues of chymotrypsin (Ala and Val, respectively) (Figure 2). These changes were expected to stabilize loop 1 and, by extension, the S1 binding pocket. However, Tr→Ch[S1-+L1*+L2] displayed activity on Suc-Ala-Ala-Pro-Phe-AMC similar to that of Tr→Ch[S1+L1+L2] (Table I), indicating that these additional mutations did not improve chymotrypsin-like activity.

The interactions between the S1 binding pocket and the remainder of the trypsin and the chymotrypsin were analyzed in order to identify differences that could account for the apparent instability of the S1 binding pocket of Tr→Ch-[S1+L1+L2] (J. J. Perona, L. Hedsrom, W. J. Rutter, and R. J. Fletterick, manuscript in preparation). One difference was identified at residue 172, which is a Tyr in trypsin and a Trp in chymotrypsin. Tyr172 forms hydrogen bonds with the carbonyl of Pro225 and the amide proton of Tyr217 in trypsin, whereas Trp172 makes a single hydrogen bond to Pro225 (Figure 1). The loss of one hydrogen bond might alter the conformational equilibrium of the S1 binding pocket of Tr→Ch[S1+L1+L2], resulting in a more chymotrypsin-like structure. Indeed, addition of the Tyr172 to Trp mutation to Tr→Ch[S1+L1+L2] produced an enzyme with 10–30-fold more activity than Tr→Ch[S1+L1+L2] on Suc-Ala-Ala-Pro-Phe-AMC, Suc-Ala-Ala-Pro-Tyr-AMC, and Suc-Ala-Ala-Pro-Trp-AMC (Table 1). No improvement in the ability to hydrolyze Suc-Ala-Ala-Pro-Lys-AMC was observed. Interestingly, addition of the Tyr172 to Trp mutation to

Table 2: Steady-State Kinetic Parameters for the Hydrolysis of Substrates by Chymotrypsin, Trypsin, and Trypsin Mutants[a]

substrate/enzyme	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
Ac-Phe-NH_2			
chymotrypsin[b]	0.43 ± 0.08	(2.3 ± 0.08) × 10^4	18 ± 3
trypsin[b]	(2.2 ± 0.6) × 10^{-4}	1000 ± 200	0.2 ± 0.1
D189S[b]	0.018 ± 0.006	(1.7 ± 0.1) × 10^5	0.10 ± 0.01
Tr→Ch[S1 + L1 + L2][b]	0.018 ± 0.003	(6.5 ± 0.2) × 10^4	0.32 ± 0.08
Tr→Ch[S1 + L1 + L2 + Y172W][c]	≥0.08	≥1 × 10^5	0.75 ± 0.15
Suc-Ala-Ala-Pro-Phe-pNA			
chymotrypsin[b]	49 ± 2	86 ± 5	(5.6 ± 0.5) × 10^5
trypsin[d]	≥0.2	≥2.5 × 10^4	9 ± 1
D189S[b]	0.29 ± 0.01	(1.5 ± 0.2) × 10^4	21 ± 1
Tr→Ch[S1+L1+L2][b]	13.0 ± 0.5	7300 ± 300	(1.7 ± 0.1) × 10^3
Tr→Ch[S1 + L1 + L2 + Y172W]	24.7 ± 0.7	300 ± 40	(8.3 ± 0.8) × 10^4
Suc-Ala-Ala-Pro-Phe-AMC			
chymotrypsin	10.4 ± 0.8	21 ± 4	(4.9 ± 0.5) × 10^5
trypsin	(7.0 ± 0.1) × 10^{-3}	1100 ± 200	6.0 ± 0.2
D189S	0.10 ± 0.01	2800 ± 500	38 ± 1
Tr→Ch[S1 + L1 L2][e]	≥4	≥6000	800 ± 20
Tr→Ch[S1 + L1 + L2 + Y172W][f]	≥30	≥3000	(9.3 ± 0.7) × 10^3
Ac-Phe-pNP, pH 6.5			
chymotrypsin[b]	30 ± 3	50 ± 15	(6 ± 1) × 10^5
trypsin[b]	29.8 ± 0.7	650 ± 80	(4.6 ± 0.1) × 10^4
D189S[b]	40 ± 10	70 ± 20	(6.0 ± 0.6) × 10^5
Tr→Ch[S1 + L1 + L2][b]	33 ± 3	27 ± 8	(1.2 ± 0.1) × 10^6
Tr→Ch[S1 + L1 + L2 + Y172W]	38 ± 3	39 ± 9	(1.1 ± 0.2) × 10^6
Suc-Ala-Ala-Pro-Phe-SBzl			
chymotrypsin[b]	52 ± 2	16 ± 1	(3.1 ± 0.3) × 10^6
trypsin[b]	36 ± 1	470 ± 35	(7.0 ± 0.2) × 10^4
D189S[b]	33 ± 5	27 ± 3	(1.23 ± 0.03) × 10^6
Tr→Ch[S1 + L1 + L2][b]	36 ± 5	21 ± 3	(1.7 ± 0.1) × 10^6
Tr→Ch[S1 + L1 + L2 + Y172W]	63 ± 4	10 ± 2	(6.0 ± 0.7) × 10^6

[a] Conditions: 50 mM Hepes, pH 8.0, or 50 mM Mops, pH 6.5, 100 mM NaCl, and 10 mM $CaCl_2$, 37 °C. [b] Hedstrom et al (1992). [c] No saturation is observed at 40 mM substrate. [d] No saturation is observed at 5.0 mM substrate. [e] No saturation is observed at 1.0 mM substrate. [f] No saturation is observed at 0.45 mM substrate.

Tr→Ch[S1+L1+L2] has the greatest effect on the hydrolysis of Suc-Ala-Ala-Pro-Trp-AMC; like chymotrypsin, Tr→Ch-[S1+L1+L2+Y172W] prefers Trp-containing substrates over Phe-containing substrates.

The S1 Binding Pocket and Loops 1 and 2 Interact Synergistically with Residue 172. Interestingly, the mutation of Tyr172 to Trp does not improve the ability of trypsin to hydrolyze chymotrypsin substrates in the absence of loops 1 and 2 and the other alterations of the S1 binding pocket (Table 1). The hydrolysis of Suc-Ala-Ala-Pro-Lys-AMC by Y172W is 100-fold slower than that by trypsin. Thus Y172W has lost specificity for trypsin substrates without acquiring specificity for chymotrypsin substrates. These results indicate that position 172 interacts synergistically with the S1 binding pocket at loops 1 and 2 to determine specificity in the trypsin family of serine proteases. Residue 172 does not contact the substrate, although it does interact with one wall of the S1 binding pocket. Therefore, position 172 is another example of a structural determinant of substrate specificity which does not contact the substrate.

Characterization of Amide Hydrolysis by Tr→Ch-[S1+L1+L2+Y172W]. The hydrolysis of amide and ester substrates by Tr→Ch[S1+L1+L2+Y172W] was characterized in order to assess how the Tyr172 to Trp mutation improved the function of Tr→Ch[S1+L1+L2]. The steady-state kinetic parameters for the hydrolysis of Ac-Phe-NH_2, Suc-Ala-Ala-Pro-Phe-AMC, Suc-Ala-Ala-Pro-Phe-pNA, Ac-Phe-pNP, and Suc-Ala-Ala-Pro-Phe-SBzl by chymotrypsin, trypsin, D189S, Tr→Ch[S1+L1+L2], and Tr→Ch[S1-+L1+L2+Y172W] are shown in Table 2. Tr→Ch[S1+L1-+L2+Y172W] hydrolyzed the amide substrates 2–50-fold faster than Tr→Ch[S1+L1+L2] as measured by k_{cat}/K_m. This improvement in activity results from an increase in k_{cat} for Ac-Phe-NH_2. The increase in the hydrolysis of Suc-Ala-Ala-Pro-Phe-pNA resulted primarily from a decrease in K_m. Unfortunately, no saturation was observed for Suc-Ala-Ala-Pro-Phe-AMC for Tr→Ch[S1+L1+L2] and Tr→Ch-[S1+L1+L2+Y172W]; therefore, the k_{cat} and K_m values reported in Table 2 are lower limits. The K_m's for the hydrolysis of amide substrates by Tr→Ch[S1+L1+L2-+Y172W] are 3–100-fold higher than those of chymotrypsin, while the k_{cat}'s for amide hydrolysis by Tr→Ch[S1+L1-+L2+Y172W] and chymotrypsin are comparable.

All of the enzymes hydrolyze Suc-Ala-Ala-Pro-Phe-AMC and Suc-Ala-Ala-Pro-Phe-pNA faster than Ac-Phe-NH_2. This behavior cannot be attributed to the reactivity of the different leaving groups; the k_{cat} for the hydrolysis of Bz-Arg-NH_2 by trypsin is equivalent to that for the hydrolysis of Bz-Arg-pNA, and the k_{cat}/K_m's for hydrolysis of these two compounds differ only 10-fold (data not shown). More extensive studies of the leaving group dependence of amide hydrolysis by chymotrypsin have shown that the Bronsted β value for these reactions is low (Fastrez & Fersht, 1973; Inward & Jencks, 1965). Therefore, the difference in the rates of hydrolysis of Ac-Phe-NH_2 and the oligopeptide substrates most likely is a product of interactions between the P2–P4 substrate residues and the S2–S4 enzyme binding sites. These extended binding interactions determine the specificity of amide hydrolysis as evidenced by the increase in k_{cat}/K_m. This phenomenon is best illustrated by comparing the mechanistic rate constants for amide hydrolysis (see below).

Characterization of Ester Hydrolysis by Tr→Ch-[S1+L1+L2+Y172W]. In contrast to amide hydrolysis, ester hydrolysis by Tr→Ch[S1+L1+L2+Y172W] is virtually identical to that by chymotrypsin, D189S, and Tr→Ch-[S1+L1+L2]. This observation is not surprising because trypsin hydrolyzes Ac-Phe-pNP and Suc-Ala-Ala-Pro-Phe-SBzl 40-fold less efficiently than chymotrypsin. As shown

Table 3: Mechanistic Kinetic Parameters for Amide Hydrolysis by Chymotrypsin, Trypsin, and Trypsin Mutants[a]

enzyme/substrate	K_s (M)	k_2 (s^{-1})	k_3 (s^{-1})
chymotrypsin			
Ac-Phe-NH_2[b]	0.023	0.43	60
Suc-Ala-Ala-Pro-Phe-pNA[b]	1.5×10^{-3}	850	52
Suc-Ala-Ala-Pro-Phe-AMC	2.5×10^{-4}	12	52
trypsin			
Ac-Phe-NH_2[b]	1.0×10^{-3}	2.2×10^{-4}	30
Suc-Ala-Ala-Pro-Phe-pNA	≥0.25	≥0.2	36
Suc-Ala-Ala-Pro-Phe-AMC	1.1×10^{-3}	0.007	36
D189S			
Ac-Phe-NH_2[b]	0.17	0.018	39
Suc-Ala-Ala-Pro-Phe-pNA[b]	0.015	0.29	33
Suc-Ala-Ala-Pro-Phe-AMC	2.8×10^{-3}	0.10	33
Tr→Ch[S1 + L1 + L2]			
Ac-Phe-NH_2[b]	0.065	0.018	33
Suc-Ala-Ala-Pro-Phe-pNA[b]	0.011	20	37
Suc-Ala-Ala-Pro-Phe-AMC	≥0.006	≥4	37
Tr→Ch[S1 + L1 + L2 + Y172W]			
Ac-Phe-NH_2	≥0.1	0.08	38
Suc-Ala-Ala-Pro-Phe-pNA	5×10^{-4}	41	63
Suc-Ala-Ala-Pro-Phe-AMC	≥0.006	≥60	63

[a] The mechanistic kinetic parameters for the hydrolysis of Ac-Phe-NH_2, Suc-Ala-Ala-Pro-Phe-pNA, and Suc-Ala-Ala-Pro-Phe-AMC are derived from the steady-state kinetic constants of Table 2. Assuming that deacylation is completely rate determining for the hydrolysis of Ac-Phe-pNP and Suc-Ala-Ala-Pro-Phe-SBzl, then K_s, k_2, and k_3 for amide hydrolysis can be determined from the following equations: (a) $k_{cat,ester} = k_3$; (b) $K_m = K_s[k_{cat,ester}/(k_2 + k_{cat,ester})]$; (c) $k_{cat} = k_2 k_{cat,ester}/(k_2 + k_{cat,ester})$. [b] Hedstrom et al. (1992).

previously, the D189S mutation is sufficient to convert trypsin into an esterase with chymotrypsin-like specificity (Hedstrom et al., 1992). All of the enzymes characterized here are efficient, although nonspecific, esterases. In addition, these enzymes catalyze the hydrolysis of Suc-Ala-Ala-Pro-Phe-SBzl and Ac-Phe-pNP with equal efficiency, indicating that interactions between the P2–P4 substrate residues and the S2–S4 enzyme binding sites do not influence ester hydrolysis.

Mechanistic Rate Constants for Amide Hydrolysis by Tr→Ch[S1+L1+L2+Y172W]. Serine proteases hydrolyze substrates via a three-step mechanism involving substrate binding (K_s), acylation of the active site serine (k_2), and hydrolysis of the acylenzyme intermediate (deacylation, k_3) (Polgar, 1989):

$$E + AcX \underset{}{\overset{K_s}{\rightleftharpoons}} E{\cdot}S \xrightarrow{k_2} E\text{-}Ac \xrightarrow[HOH]{k_3} E + P \qquad (X^- \text{ released at } k_2)$$

Assuming that deacylation is completely rate determining for ester substrates, the mechanistic parameters for the hydrolysis of amide substrates can be derived from the data of Table 2 (Zerner et al., 1964). These parameters are shown in Table 3.

Deacylation Does Not Determine Substrate Specificity. As noted previously, k_3 is similar for all of the enzymes, and all of the amide substrates, studied (Hedstrom et al., 1992). This similarity is a consequence of the similarity of ester hydrolysis by these enzymes. These observations demonstrate that interactions between the P2–P4 substrate residues and the S2–S4 enzyme binding sites are not important for the deacylation reaction and that deacylation does not determine substrate specificity. The covalent enzyme–substrate interactions are sufficient to position the acylenzyme intermediate for optimal interaction with the enzyme catalytic groups in the absence of favorable noncovalent interactions with the S1 site. This lack of specificity in the deacylation step makes biological sense; discrimination in the deacylation step would inactivate a serine protease.

Influence of Substrate Binding on Substrate Specificity. The binding affinity of substrates containing Phe at P1 does not correlate with the ability to hydrolyze such substrates. Chymotrypsin has the highest affinity and activity for Suc-Ala-Ala-Pro-Phe-AMC (Table 3). However, trypsin has the highest affinity for Ac-Phe-NH_2. D189S, Tr→Ch[S1+L1+L2], and Tr→Ch[S1+L1+L2+Y172W] have comparable affinities for Suc-Ala-Ala-Pro-Phe-pNA, although their activities vary greatly. Thus the affinity of substrate binding appears to be less important in determining substrate specificity than the ability to hydrolyze substrates.

Not surprisingly, increasing the length of the oligopeptide substrate usually decreased K_s. Chymotrypsin binds Ac-Phe-NH_2 10–100-fold less tightly than Suc-Ala-Ala-Pro-Phe-AMC and Suc-Ala-Ala-Pro-Phe-pNA, while D189S and Tr→Ch[S1+L1+L2+Y172W] bind Ac-Phe-NH_2 10–100-fold less tightly than the two oligopeptide substrates. The difference in binding the monomeric substrate and the oligopeptide substrates was only 5–10-fold for Tr→Ch[S1+L1+L2]. Interestingly, oligopeptide substrates bind to trypsin with affinity comparable to or lower than that of Ac-Phe-NH_2. Thus no correlation exists between P1–S1 subsite specificity and the influence of P2–P4 and S2–S4 interactions on binding.

Influence of the Acylation Reaction on Substrate Specificity. The acylation reaction determines substrate specificity. The acylation rate constants for chymotrypsin catalyzed oligopeptide amide hydrolysis are the greatest, followed by those of Tr→Ch[S1+L1+L2+Y172W] and Tr→Ch[S1+L1+L2]. The values of k_2 for trypsin- and D189S-catalyzed hydrolysis of oligopeptide amides are much less. A similar trend is observed for hydrolysis of Ac-Phe-NH_2, although the range of variation is much less (10^3-fold versus 10^5-fold for oligopeptides). Again chymotrypsin has the greatest value for k_2, followed by Tr→Ch[S1+L1+L2+Y172W], D189S, Tr→Ch[S1+L1+L2], and trypsin. Thus the increase in specificity for substrates containing Phe parallels the increase in the acylation rate, not substrate binding. The discrimination in the acylation step can be circumvented by very reactive substrates, which accounts for the low specificity of ester hydrolysis.

The length of the amide substrate also influences the rate of acylation. Previous results suggested that interactions between the P2–P4 residues and the S2–S4 sites increased k_2 for oligopeptides containing the "correct" P1 residue, but not for oligopeptides containing the "incorrect" P1 residue (Hedstrom et al., 1992). However, the more complete data set presented here indicates that this conclusion is not accurate. The values of k_2 for hydrolysis of Suc-Ala-Ala-Pro-Phe-AMC and Suc-Ala-Ala-Pro-Phe-pNA by chymotrypsin are 30–10^3-fold greater than the k_2 for hydrolysis of Ac-Phe-NH_2. Similar increases in k_2 are observed for the trypsin catalyzed hydrolysis of oligopeptides containing Arg or Lys at P1 relative to the analogous monomeric substrates (L. Hedstrom, unpublished data). However, the values of k_2 for the hydrolysis of Suc-Ala-Ala-Pro-Phe-AMC and Suc-Ala-Ala-Pro-Phe-pNA by trypsin are also 30–10^3-fold greater than that for Ac-Phe-NH_2 hydrolysis (Table 3). Thus trypsin can utilize extended binding interactions to increase k_2 for both specific substrates (i.e., P1 = Lys) and nonspecific substrates (P1 = Phe). Both Tr→Ch[S1+L1+L2] and Tr→Ch[S1+L1+L2+Y172W] use interactions between P2–P4 and S2–S4 to produce large increases in k_2, while k_2 for oligopeptide hydrolysis by D189S

is only 10-fold greater than k_2 for Ac-Phe-NH_2 hydrolysis. Thus the ability to utilize binding energy from these distal interactions to increase k_2 seems to be a characteristic of efficient proteases, and not dependent on the correct P1–S1 interaction.

These results contrast with similar work on elastase, where interactions with the extended binding site increase k_2 for specific substrates (P1 = Ala or Val), but not for nonspecific substrates (P1 = Phe) (Stein et al., 1987). Such contradictory observations suggest that the extended binding sites may influence substrate specificity by a variety of mechanisms. It seems reasonable to expect that the influence of interactions between the S1 site and the extended binding sites is context dependent. Strong interactions between an oligopeptide and the extended binding sites may cause an oligopeptide to bind in the active site despite unfavorable interactions between the P1 residue and the S1 site. Alternatively, strong interactions between the P1 residue and the S1 site may cause anomalous binding at the extended binding sites. This binding must be productive for the reaction to occur; i.e., the substrate must be in register with the catalytic residues. Obviously, interactions between an enzyme and a good substrate will be optimal at the catalytic residues and both the S1 and the extended binding sites; these interactions will be perturbed for a poor substrate. However, the nature of these perturbations will depend on the particular substrate.

The Y172W Mutation Stabilizes the S1 Binding Pocket. We initially proposed a model for the action of Tr→Ch-[S1+L1+L2], where this enzyme exists as an equilibrium between an active, chymotrypsin-like conformation and an inactive conformation where the S1 binding pocket is deformed. This model accounts for the defect in substrate binding observed for Tr→Ch[S1+L1+L2] while explaining the efficient rates of hydrolysis, and it predicted that Tr→Ch-[S1+L1+L2] could be improved by stabilizing this active conformation. Such a stabilizing mutation is naively predicted to lower K_s by improving substrate binding with little effect on k_2. However, although the binding of Suc-Ala-Ala-Pro-Phe-pNA to Tr→Ch[S1+L1+L2+Y172W] is improved relative to Tr→Ch[S1+L1+L2] binding, no improvement was observed in the binding of Ac-Phe-NH_2. In contrast, k_2 is increased 15-fold for this substrate. This observation appears to conflict with our proposed model.

However, an increase in k_2 can also be consistent with stabilization of the S1 binding pocket. Jencks has noted that exact positioning of a substrate in an enzyme active site will accelerate a reaction. However, such exact positioning has a large entropic cost, which results from ordering of both the enzyme and the substrate. This entropic cost results in the generally low substrate binding affinities of enzymes when compared to the binding affinities of antigen–antibody complexes (Jencks, 1987). Clearly, a rigid substrate binding site will require less ordering than a deformed binding site. This entropic savings will be realized by an increase in k_2. This effect will be most pronounced for a substrate such as Ac-Phe-NH_2 which makes few interactions with the enzyme and consequently has little binding energy available to stabilize a flexible binding site. In contrast, Suc-Ala-Ala-Pro-Phe-pNA can use additional interactions to stabilize a flexible S1 binding pocket; the advantage of a rigid S1 binding site is therefore less.

It is more difficult to predict the effect of stabilizing the S1 binding pocket on K_s. As discussed previously, stabilization of the S1 binding pocket will increase the fraction of enzyme that can bind substrate, which should decrease K_s. However, a deformed binding site may also allow the substrate to bind in many similar orientations (these orientations can also include variations in rotational and vibrational motion of the substrate). These very small differences in orientation can cause large decreases in catalytic rates. A rigid binding site will no longer permit these less productive orientations. The loss of the less productive binding modes will increase k_2 but will also increase K_s. Thus the observed K_s could decrease or increase, depending on which mechanism predominates. It seems reasonable to expect a substrate that has many contacts with the enzyme (e.g., Suc-Ala-Ala-Pro-Phe-pNA) to bind in fewer orientations than a substrate with minimal enzyme contacts (e.g., Ac-Phe-NH_2). This reasoning suggests that K_s will be lowered for Suc-Ala-Ala-Pro-Phe-pNA, but not necessarily for Ac-Phe-NH_2.

Alternatively, the improvement in Suc-Ala-Ala-Pro-Phe-pNA binding may result from changes in the interaction of the pNA group with the enzyme. Alterations in the S1 site may modify the position of the leaving group in the S1′ site, thus changing binding affinity. Indeed, the specificity of the S1′ site of Tr→Ch[S1+L1+L2] is slightly different from that of trypsin (Schellenberger et al., 1993).

Further evidence that the Y172W mutation stabilizes the S1 binding pocket is presented in the following paper (Hedstrom et al., 1994). Proflavin binding is a direct probe of the integrity of the S1 binding pocket of trypsin and chymotrypsin (Fersht, 1972). Tr→Ch[S1+L1+L2+Y172W] binds proflavin, while no proflavin binding to Tr→Ch-[S1+L1+L2] can be observed. This observation indicates that the S1 binding pocket of Tr→Ch[S1+L1+L2] is deformed relative to those of trypsin and chymotrypsin and that the Tyr172 to Trp mutation stabilizes the S1 binding pocket. X-ray structural studies also support this conclusion (J. J. Perona, L. Hedstrom, W. J. Rutter, and R. J. Fletterick, manuscript in preparation). The structures of both Tr→Ch-[S1+L1+L2] and Tr→Ch[S1+L1+L2+Y172W] complexed with Suc-Ala-Ala-Pro-Phe chloromethyl ketone have been determined. The shape of the S1 binding pocket of both of these enzymes resembles that of the S1 binding pocket of chymotrypsin. The residues of the S1 binding pocket of Tr→Ch[S1+L1+L2] have high thermal factors, and some of the side chains are disordered. The S1 binding pocket of Tr→Ch[S1+L1+L2+Y172W] is more structured; Gly187 of loop 1 is visible in Tr→Ch[S1+L1+L2+Y172W] but not in Tr→Ch[S1+L1+L2]. Thus the Y172W mutation stabilizes the S1 binding pocket.

In conclusion, residue 172 is a substrate specificity determinant in the trypsin family of serine proteases. Additional substrate specificity determinants remain to be identified.

ACKNOWLEDGMENT

We thank Dorrie Chang for technical assistance.

REFERENCES

Blow, D. M. (1971) In *The Enzymes* (Boyer, P. D., Ed.) Vol. 3, pp 213–248, Academic Press, Boca Raton, FL.
Brady, K., & Abeles, R. H. (1990) *Biochemistry 29*, 7608–7617.
Brown, J. R., & Hartley, B. S. (1966) *Biochemistry J. 101*, 214.
Craik, C. S., Choo, Q.-L., Swift, G. H., Quinto, C., MacDonald, R. J., & Rutter, W. J. (1984) *J. Biol. Chem. 259*, 14255–14264.
Evnin, L. B., Vasquez, J. R., & Craik, C. S. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6659–6663.
Fastrez, J., & Fersht, A. R. (1973) *Biochemistry 12*, 1067–1074.
Fersht, A. R. (1972) *J. Mol. Biol. 64*, 497–509.

Fersht, A. R. (1985) *Enzyme Structure and Mechanism,* W. H.Freeman and Co., New York.
Graf, L., Jancso, A., Szilagyi, L., Hegyi, G., Pinter, K., Naray-Szabo, G., Hepp, J., Medzihradsky, K., & Rutter, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85,* 4961–4965.
Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255,* 1249–1253.
Hedstrom, L., Farr-Jones, S., Kettner, C., & Rutter, W. J. (1994) *Biochemistry* (Following paper in this issue).
Inward, P. W., & Jencks, W. P. (1965) *J. Biol. Chem. 240,* 1986.
Jencks, W. P. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52,* 65–73.
Kossiakoff, A. A. (1987) In *Biological Macromolecules and Assemblies* (Jurnak, F. A., & McPherson, A., Eds.) pp 369–412, John Wiley and Sons, New York.
Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82,* 488–492.
Pinsky, S. D., LaForge, K. S., Luc, V., & Scheele, G. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80,* 7486.
Polgar, L. (1989) *Mechanisms of Protease Action,* CRC Press, Inc., Boca Raton, FL.
Rossman, M. G., & Argos, P. (1976) *J. Mol. Biol. 105,* 75–95.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27,* 157–162.
Schellenberger, V., Turck, C. W., Hedstrom, L., & Rutter, W. J. (1993) *Biochemistry 32,* 4349–4353.
Shotton, D. M., & Watson, H. C. (1969) *Nature 25,* 811–816.
Stein, R. L., Strimpler, A., Hori, H., & Powers, J. C. (1987) *Biochemistry 26,* 1301–1305.
Steitz, T. A., Henderson, R., & Blow, D. M. (1969) *J. Mol. Biol. 46,* 337–348.
Titani, K., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1975) *Biochemistry 14,* 1358.
Zerner, B., Bond, R. P. M., & Bender, M. L. (1964) *J. Am. Chem. Soc. 86,* 3674–3679.
Zimmerman, M., Ashe, B., Yurewciz, E. C., & Patel, G. (1977) *Anal. Biochem. 78,* 47–51.