Engineered Metal Regulation of Trypsin Specificity†

W. Scott Willett,‡ Sarah A. Gillmor,§ John J. Perona,|| Robert J. Fletterick,⊥ and Charles S. Craik*,‡

Departments of Pharmaceutical Chemistry, Biochemistry and Biophysics, and Graduate Group in Biophysics, University of California at San Francisco, San Francisco, California 94143-0446

Received August 15, 1994; Revised Manuscript Received November 30, 1994⊗

ABSTRACT: Histidine substrate specificity has been engineered into trypsin by creating metal binding sites for Ni^{2+} and Zn^{2+} ions. The sites bridge the substrate and enzyme on the leaving-group side of the scissile bond. Application of simple steric and geometric criteria to a crystallographically derived enzyme−substrate model suggested that histidine specificity at the P2′ position might be acheived by a tridentate site involving amino acid residues 143 and 151 of trypsin. Trypsin N143H/E151H hydrolyzes a P2′-His-containing peptide (AGPYAHSS) exclusively in the presence of nickel or zinc with a high level of catalytic efficiency. Since cleavage following the tyrosine residue is normally highly disfavored by trypsin, this result demonstrates that a metal cofactor can be used to modulate specificity in a designed fashion. The same geometric criteria applied in the primary S1 binding pocket suggested that the single-site mutation D189H might effect metal-dependent His specificity in trypsin. However, kinetic and crystallographic analysis of this variant showed that the design was unsuccessful because His189 rotates away from substrate causing a large perturbation in adjacent surface loops. This observation suggests that the reason specificity modification at the trypsin S1 site requires extensive mutagenesis is because the pocket cannot deform locally to accommodate alternate P1 side chains. By taking advantage of the extended subsites, an alternate substrate specificity has been engineered into trypsin.

Metal binding sites in proteins serve a variety of functions, including structural stability, catalysis, folding, and molecular recognition. Nearly one third of all known proteins require metal ions for their structure or function (Ibers & Holm, 1980). This indicates that a thorough understanding of the chemistry of protein/metal interactions will yield increased insight into many physiological processes as well as provide a basis for protein engineering efforts. The wealth of information about metalloproteins has allowed the design of simple metal binding sites for specified functions (Higaki et al., 1992). These sites have been incorporated into recombinant proteins to aid in purification, to enhance structural stability, to facilitate structure determination by X-ray crystallography, and to regulate activity (Tainer et al., 1991). For example, metal binding sites have been incorporated successfully into staphylococcal nuclease (Corey & Schultz, 1989) and trypsin (Higaki et al., 1990) to reversibly inhibit these enzymes. A calcium binding site was introduced into subtilisin to stabilize the enzyme against thermal denaturation (Braxton & Wells, 1992), and a nickel binding site incorporated into glycogen phosphorylase allowed insight into the mechanism of allosteric regulation (Browner et al., 1994). These reports indicate that enzymes can be engineered to bind metal ions at designed sites, providing a powerful new tool to address structure/function relationships and for redesigning enzymatic activity.

Trypsin functions well as a model system to test the design of metal binding sites because it has an established expression/purification system, well-defined substrates for kinetic analysis by continuous and discontinuous assays, and reproducible crystallization conditions for X-ray structure analysis of the modified proteins. Trypsin specificity at the P1 substrate position is exclusively toward Arg and Lys side chains for amide hydrolysis (Evnin et al., 1990), while there is little preference at position P2′ (Schellenberger et al., 1994). Attempts to modify the P1 substrate preference with a small number of amino acid substitutions has met with little success (Graf et al., 1987; Sprang et al., 1988); however, a specificity conversion to large hydrophobic residues was effected by extensive exchange of amino acids for those of chymotrypsin [Hedstrom et al., 1994, Perona et al., 1995; for review, see Perona and Craik (1995)]. As an alternative way to alter substrate specificity of trypsin, the use of transition metal binding sites has been explored. Using this approach, subsites on the enzyme were engineered to bind a metal ion which bridges to a metal-binding amino acid within the substrate. The concept is illustrated in Figure 1 using histidines introduced into trypsin to create metal-dependent P2′-His specificity. This paper explores the structural environments of both the S1 and S2′ sites of trypsin as potentially suitable targets for the introduction of metal-dependent specificity. We describe the successful design, production, and kinetic analysis of a trypsin variant exhibiting metal-dependent histidine specificity at position P2′.

† This work was supported by NSF Grant DMB921806 (to C.S.C.), NIH Grant DK39304 (to R.J.F.), NIH Training Grant T32-GM08388 and University of California Systemwide Biotechnology Reaearch and Educational Program (to W.S.W.), NSF Predoctoral Fellowship Award (to S.A.G), and Postdoctoral NRSA Award GM13818-03 (to J.J.P.)

* Author to whom correspondence should be addressed at the Department of Pharmaceutical Chemistry, University of California at San Francisco, Box 0446, San Francisco, CA 94143-0446. Phone: (415) 476-8146. Fax: (415) 476-0688. E-mail: craik@cgl.ucsf.edu.

‡ Department of Pharmaceutical Chemistry.

§ Graduate Group in Biophysics.

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

⊥ Department of Biochemistry and Biophysics.

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1995.

0006-2960/95/0434-2172$09.00/0

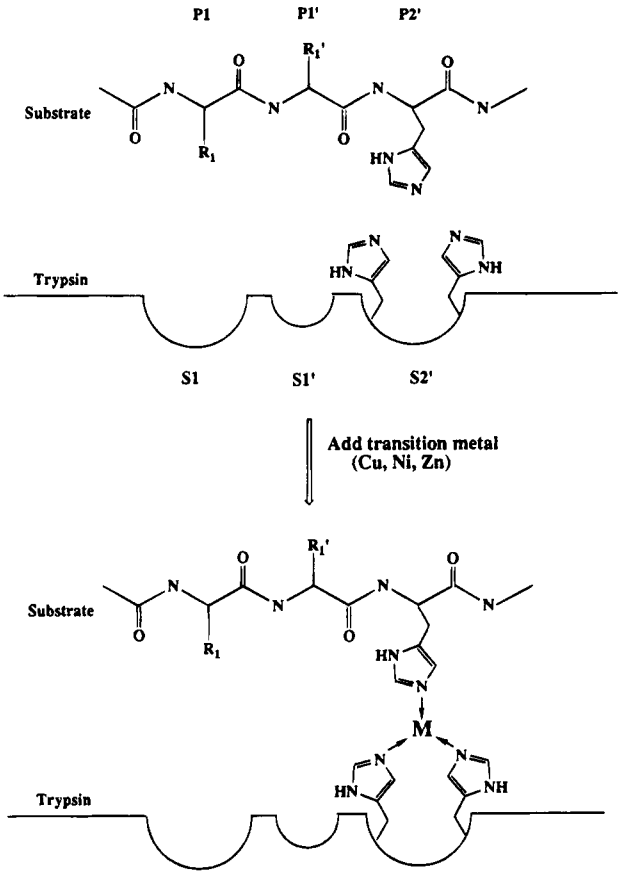

FIGURE 1: Schematic representation of the trypsin-substrate complex showing enzyme subsites (S1-S2′) and substrate residues (P1-P2′) according to the nomenclature of Schechter and Berger (1968). The sissile bond is between residues P1 and P1′, and the primary binding determinant is the P2′ residue. This concept involves an engineered metal binding site on the enzyme allowing a metal ion to bridge enzyme and substrate so that substrate binds and is cleaved only in the presence of metal ion.

MATERIALS AND METHODS

Saccharomyces cerevisiae strain DLM101α [*Mat a, leu 2-3,-112 his2 3-11,-15 can1, ura 3Δ, pep4Δ [cir°]DM23*] was obtained from Dr. L. Hedstrom. *Escherichia coli* strain X90 (F′ *lac I^q, lac ZY, pro AB/Δ-(lac-pro), ara, mal A, argEam, thi, rif*′) was obtained from Dr. A Vershon. All restriction enzymes, T4 DNA ligase, and T4 DNA polymerase were purchased from New England Biolabs, Inc., and used according to manufacturer's recommendations. Enteropeptidase (EK2) was purchased from Biozyme, Inc. SP Sepharose Fast Flow resin was purchased from Pharmacia. Benzamidine agarose was purchased from Pierce, and Affi-Gel 10 was purchased from Bio-Rad. BPTI was purchased from Merck (Trasylol) and coupled to the Affi-Gel beads according to the manufacturer's protocol. *N*α-Benzyloxycarbonyl-L-glycylprolylarginine-7-amino-4-methylcoumarin (Z-GPR-AMC) was purchased from Bachem Biosciences, Inc. The peptide substrates AGPYAHSS and AGPYAASS were purchased from BioServ Labs (San Jose, CA). Oligonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer (Foster City, CA).

Computer Modeling. Modeling was performed using the InsightII program (BioSym Technologies, Inc) displaying the crystal structure of rat anionic trypsin mutant D189S complexed to BPTI on a Silicon Graphics workstation (Perona et al., 1994). Potential metal binding sites utilizing histidines were modeled according to the following criteria: (1) The side chains were solvent-accessible for exogenous metal binding; (2) the distances between the α-carbons of the unsubstituted side chains were less than 13 Å to allow imidazole side chains to achieve optimal distances of about 2 Å for metal chelation; (3) the His side chains possessed one of five favorable rotamers (Ponder & Richards, 1987); (4) all atoms of the substituted His side chains lay farther than 3.0 Å from other atoms of the protein unless a hydrogen bond could be formed. Subroutines from the program InsightII were used to perform the alignments and torsional angle adjustments of the side chains (Higaki et al., 1990).

Modeling in S1 was performed interactively using InsightII. Asp189 and the P1 lysine of BPTI were replaced by histidines and moved until a suitable coordination geometry was found. Second generation variants in S1 were designed using the FRODO (Jones, 1978) and X-PLOR (Brunger et al., 1987) computer packages.

Mutagenesis and Expression. Mutagenesis was performed by the method of Kunkel (1985) on a bluescript-based plasmid (pST) containing the gene for rat anionic trypsinogen fused to the yeast α-factor leader and ADH/GAPDH promoter sequences (Hedstrom et al., 1992). The *Bam*HI/*Sal*I fragment of the resulting trypsin variant was subcloned into the corresponding vector fragment of the yeast expression vector (pYT). Yeast was transformed by electroporation in 2 mm cells using a Bio-Rad Gene Pulser set at 600 V, 200 Ω, and 25 μF. Transformed cells were selected on uracil-deficient plates and grown at 30 °C for 2 days. Colonies were restreaked onto leucine-deficient plates and grown for 2 days at 30 °C to increase plasmid copy number (Erhart & Hollenberg, 1983). Colonies from the Leu$^-$ plates were grown in Leu$^-$ minimal media containing 8% glucose for 2 days (Gardell et al., 1988). Trypsinogen was expressed and secreted into the media by inoculating a 1 L flask containing 1% yeast extract/2% bactopeptone/2% glucose and allowing growth at 30 °C for 48−72 h with shaking.

Purification. Secreted trypsinogen was isolated by centrifugation (10 min at 5000*g*) to remove cells, and the resulting supernatant was adjusted to pH 3.0 by addition of glycine-HCl to 50 mM. This solution was recentrifuged (10 min at 5000*g*) and loaded onto an SP-Sepharose fast flow column, and trypsinogen was eluted with a 0−1.25 M NaCl gradient. Fractions containing trypsinogen as determined by SDS−PAGE were pooled and dialyzed against 10 mM MES/1 mM $CaCl_2$, pH 6, at 4 °C. Enteropeptidase (EK2) dissolved in 10 mM MES, pH 6 at 1 mg/mL was added to the dialyzed trypsinogen pool at a ratio of 1:100 (w/w) EK2 to trypsinogen. Activation was monitored by SDS−PAGE. This activation mixture was incubated at 37 °C for 2−4 h to cleave the propeptide and liberate mature trypsin and then loaded onto an appropriate affinity column equilibrated in 10 mM MES/1 mM $CaCl_2$, pH 6. Variants containing an intact P1 pocket were purified over benzamidine agarose, whereas variants containing the D189H mutation were purified via BPTI-coupled agarose to utilize the extended binding contacts necessary for purification of trypsins containing an altered P1 pocket. The affinity column (either benzamidine or BPTI) was washed with 20 column volumes of equilibration buffer containing 0.5 M NaCl to remove unactivated trypsinogen and trypsin degradation products. Pure trypsin was step-eluted with 0.1 M HCO_2H/0.5 M NaCl/1 mM $CaCl_2$. Fractions containing trypsin were

diafiltered using a Centricon 10 concentrator (Amicon) into 1 mM HCl/1 mM $CaCl_2$ at pH 3 for storage. All purified trypsins were stored under these conditions at 4 °C.

Kinetic Parameter Determination. Kinetic assays of all trypsin variants were performed in 5 mM PIPPS/1 mM Tris/100 mM NaCl/1 mM $CaCl_2$, pH 8. This buffer was chosen for low metal binding capacity, with Tris functioning as a carrier for metal ions (Bai & Martell, 1969). The fluorogenic substrate Z-GPR-AMC was used to test activity against a typical amide trypsin substrate, and Z-Lys-thiobenzyl ester was used to detect low levels of activity.

S2′ Variants. An octapeptide substrate, AGPYAHSS (Bioserv, Inc., San Jose, CA), was used to observe specificity at P2′. This allowed us to measure small kinetic effects at P2′ because of the very slow turnover of P1-Tyr by trypsin. A discontinuous peptide cleavage assay was performed as follows: 2.5 μL of 40 μM enzyme in 1 mM HCl/1 mM$CaCl_2$ was added to 87.5 μL of assay buffer (5 mM PIPPS/1 mM Tris/100 mM NaCl/1 mM $CaCl_2$, pH 8) containing either 228 μM $CuCl_2$, 228 μM $NiCl_2$, 228 μM $ZnCl_2$, or 0.5 mM EDTA as a control. Ten microliters of peptide substrate dissolved in assay buffer was added to the enzyme mixture, and the reaction was allowed to proceed at 37 °C for 2 h. The final enzyme concentration was 1 μM, metal concentration was 200 μM, and substrate concentration ranged from 200 μM to 2 mM. The reaction was stopped by diluting 10 μL of reaction mix into 140 μL of 0.1% TFA containing 0.001% dimethylformamide as an internal standard. Cleavage products were analyzed by RP-HPLC on a C18 guard column (Perkin Elmer) with a 0−30% acetonitrile/0.1% TFA gradient over 2.5 min. Absorbance was measured at 220 nm. Quantitation of product peak areas was performed by the Rainin HPLC Dynamax program. Initial rates were calculated and plotted vs substrate concentration, and the curve was fitted to the Michaelis−Menton equation using Kaleidograph on a Macintosh computer.

S1 Variants. Variants designed to exhibit metal-dependent specificity at P1 were tested against 200 μM suc-AAPH-pNA, suc-AAPD-pNA, and suc-AAPE-pNA in the presence and absence of 100 μM $CuCl_2$, $NiCl_2$, or $ZnCl_2$. Enzyme concentration was 100 nM. These enzymes were also tested against Z-Lys-thiobenzyl ester at an enzyme concentration of 2 μM.

Crystallography. A 1:1 stoichiometric ratio of TnD189H/bovine pancreatic trypsin inhibitor (BPTI) crystallized after 4−7 days at 4 °C in hanging drops (McPherson, 1989) at a final concentration of 10 mg/mL. The well solutions contained 19% PEG 4K, 0.2 M ammonium acetate, and 0.1 M sodium citrate (pH 6.0−6.25); equal volumes (1−3 μL) of the protein and well solutions were mixed. The two crystals used for data collection were both approximately 0.8 × 0.25 × 0.25 mm in size with space group $P3_221$ and cell dimensions $a = b = 92.8$ Å and $c = 62.2$ Å. Diffraction data were recorded to 1.6 Å on an *R*-axis image plate system (Sato et al., 1992). Ninety-six frames of data were collected with a scan angle of 1.3−1.4° per frame. The recording time was 30 min per frame. The data were reduced using the *R*-axis software package (Sato et al., 1992). Phases were determined using the structure of trypsin D189S complexed with BPTI (Perona et al., 1994) as a starting model from which the side chains of Lys15 of BPTI, Ser189, Ile183, Gly158, Val199, Gly183, Tyr228, Ser190, and all waters in the P1 pocket were deleted. The initial *R* factor prior to refinement was 36%. Positional and *B* factor refinement was accomplished with X-PLOR (Brunger et al., 1987) and was alternated with model building into electron density maps computed from coefficients ($2F_o - F_c$) and ($F_o - F_c$) using FRODO (Jones, 1978) running on an Evans and Sutherland PS390 graphics workstation. This completed model has a cumulative R_{cryst} of 18.4% for all data with $I/\sigma(I) > 1$ in the resolution range 6.0−1.8 Å (Table 3). An omit map was also calculated following deletion of residues 189 and 217−226 and subsequent refinement. This map confirmed that this region of the structure had been correctly rebuilt (Figure 7c).

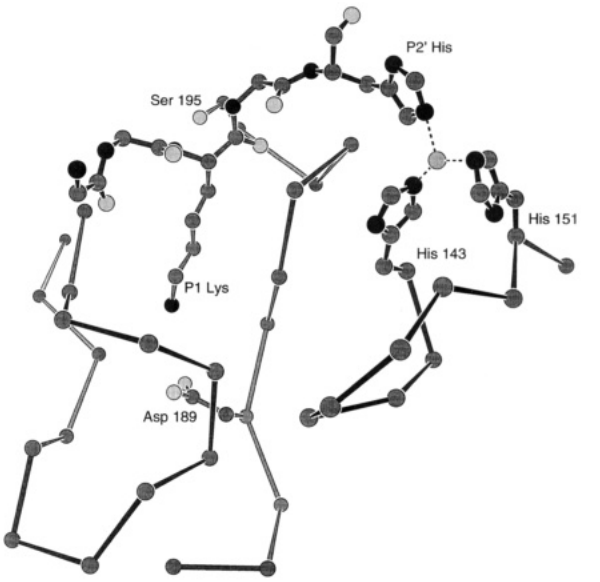

FIGURE 2: Computer model of Tn H143/151 interacting with bound substrate. The P1 lysine of BPTI is shown interacting with Asp189 at the bottom of the P1 pocket of trypsin. The designed metal binding site at S2′ shows histidines at 143 and 151 forming a tridentate site with the P2′ histidine of the substrate and bridging metal ion. Histidine−metal bond lengths are 1.5−2.0 Å, and bond angles are 102−133°. Coordination geometry is roughly tetrahedral.

RESULTS

Histidines were used to create metal binding sites in trypsin because they are known to coordinate transition metals. The variant enzyme designs were based on the X-ray crystal structure of rat anionic trypsin D189S complexed to BPTI (Perona et al., 1994). The gene encoding rat anionic trypsinogen was altered by site-specific mutagenesis, and variant and wild-type trypsins were expressed in yeast and purified to homogeneity by affinity chromatography. Variant enzymes were designed to contain metal binding sites in the S2′ or S1 regions of the enzyme to effect metal-dependent substrate specificity toward substrates containing histidine at either P2′ or P1.

Metal-Dependent P2′ Specificity. Modeling of the trypsin/BPTI structure shows that residues 143 and 151 of the enzyme are candidates for a metal binding site that bridges to a substrate containing a histidine at P2′, forming a tridentate site (Figure 2). The modeling suggests histidine−metal bond lengths of 1.5−2.0 Å and bond angles of 102−133°, showing roughly tetrahedral geometry. On the basis of observed structures of metal binding sites in proteins, transition metals preferring tetrahedral geometry (Ni^{2+} or Zn^{2+}) should bind at this site. Alternatively, each of the

Table 1: Amide Kinetics of S2′ Variants[a]

enzyme	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (min^{-1} μM^{-1})
trypsin	3500	14.1	248
Tn N143H	2670	17.7	151
Tn E151H	4160	18.2	228
Tn H143/151	2970	20.4	145
Tn E151Q	2955	16.2	182

[a] Fluorogenic substrate Z-GPR-AMC was hydrolyzed in 10 mM PIPPS/1 mM Tris/100 mM NaCl/1 mM $CaCl_2$, pH 8, at 25 °C. Enzyme concentration was 2.5 nM. These variants all possess wild-type levels of activity against this substrate.

single mutants, N143H, or E151H might also bind substrate through a metal ion in a bidentate fashion, although presumably with lower affinity than a tridentate site. Although three liganding atoms should bind a metal ion with higher affinity than two, optimal geometry of the site may be the dominant factor in metal binding, leading to a bidentate site with higher affinity. The model prediction of short metal−ligand bond lengths (<2 Å) implies that there would have to be some local structural changes in the protein to allow for tight binding of metal. Trypsins N143H, E151H, and N143H/E151H were constructed, expressed, and purified to homogeneity. Each variant trypsin in this series was tested against Z-GPR-AMC to demonstrate structural integrity. These three variants possess wild-type levels of activity toward this P1-Arg substrate which occupies the P1−P3 sites on the enzyme (Table 1).

A discontinuous RP-HPLC peptide cleavage assay was used to observe specificity at P2′ in the presence and absence of transition metal ions. The substrate octapeptide AGPYAHSS contains a Tyr at P1 to maximize the effect of turnover due to binding at P2′ because very fast hydrolysis rates of a substrate containing Arg or Lys at P1 could mask metal-dependent specificity at P2′. Wild-type trypsin fails to cleave AGPYAHSS at a measurable rate (Figure 3). None of the enzymes cleave this peptide substrate in the absence of metal ion, and neither Tn N143H nor Tn E151H shows activity in the presence of copper, nickel, or zinc. However, the double variant N143H/E151H displays activity in the presence of either nickel or zinc, but not copper (Figures 3−5). The peptide AGPYAASS was used as a control to demonstrate the requirement for histidine at P2′. This substrate is not cleaved by any of the enzymes either in the presence or absence of metal ions. The AGPYAHSS substrate was also incubated with 500 μM nickel and zinc in the absence of enzyme to test for metal-mediated cleavage. No cleavage was observed after 15 h incubation at 37 °C, demonstrating that both enzyme and added metal are required to obtain substrate turnover.

Mass spectral analysis was used to determine the cleavage site in the peptide substrate. The cleavage products were collected from the HPLC effluent and dried by vacuum centrifugation. Figure 4 illustrates the RP-HPLC assay of the reaction products. The 789.4 Da substrate AGPYAHSS is cleaved to generate the 407.2 Da AGPY peptide and the 401.2 Da AHSS peptide. This result agrees with the masses and retention times predicted by the program MacPro Mass and confirms cleavage following the tyrosine residue, thereby placing the histidine at P2′.

Kinetic analysis shows that TnN143H/E151H displays metal-dependent histidine specificity at P2′, with a k_{cat}/K_m of (82 ± 1.7) × 10^{-3} and (93 ± 0.3) × 10^{-3} μM^{-1} h^{-1} in

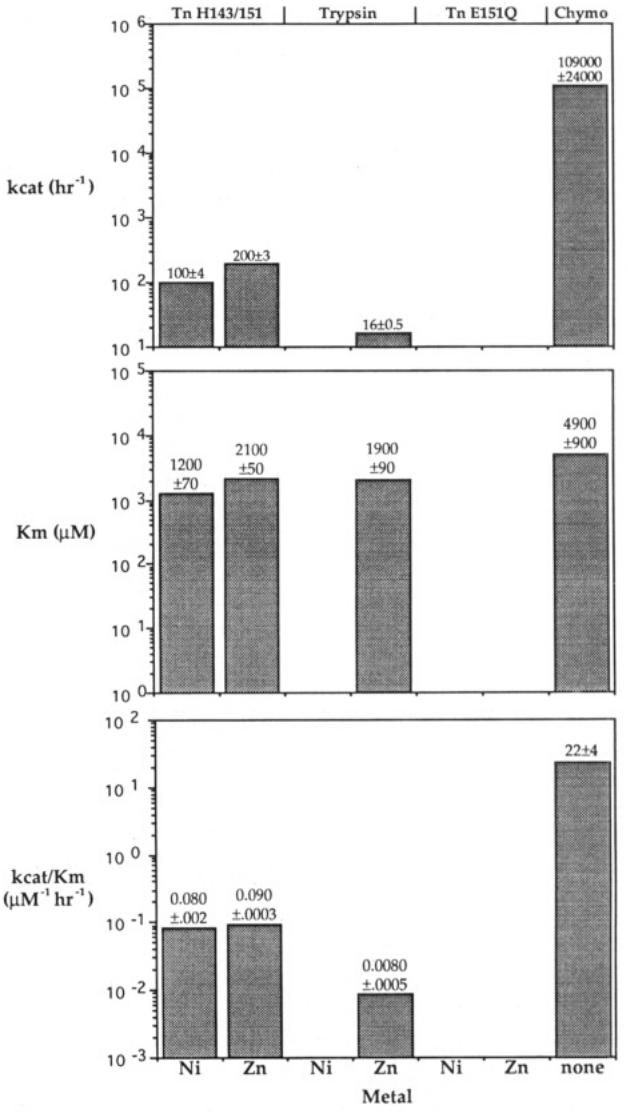

FIGURE 3: Kinetic parameters of AGPYAHSS hydrolysis. Turnover by Tn H143/151 with Zn was only 250-fold lower than chymotrypsin. There was no observable activity toward the peptide substrate by any trypsin molecule in the absence of metal or with copper. The control peptide AGPYAASS lacking a P2′ histidine was not hydrolyzed with or without metal. Substrate incubated alone with metal was not hydrolyzed. The data shown for chymotrypsin were obtained without metal.

the presence of nickel and zinc, respectively. Wild-type trypsin does not hydrolyze AGPYAHSS in the presence of nickel, but does in the presence of zinc with k_{cat}/K_m = (8.3 ±.5) × 10^{-3} μM^{-1} h^{-1}, 10-fold lower than Tn N143H/E151H (Figure 3). To address this unexpected activity of wild-type trypsin, Tn E151Q was made to remove possible metal-binding capability of the carboxylate of Glu151 in the S2′ region. The glutamine side chain has a much lower affinity for metal ions than glutamate and therefore would not be expected to bind the histidine substrate through a bridging metal ion. Trypsin E151Q has activity indistinguishable from that of wild-type toward the Z-GPR-AMC substrate (Table 1) but shows no activity toward the AGPYAHSS substrate under any conditions. This indicates that the carboxylate of Glu151 is responsible for the observed activity of wild-type trypsin in the presence of zinc. Chymotrypsin was used as a positive control to establish a baseline for cleavage following the P1-Tyr; hydrolysis of AGPYAHSS proceeds with k_{cat}/K_m = 22.5 ± 3.8 μM^{-1} h^{-1}. This rate is

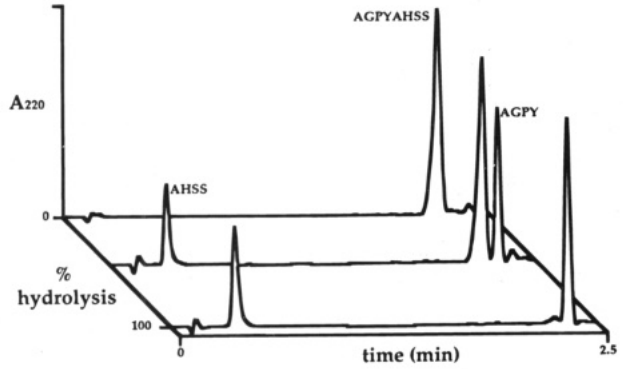

FIGURE 4: HPLC assay of peptide hydrolysis, illustrating progress of reaction from 0 to 100% hydrolysis. Assay was performed by incubation for 2 h at 37 °C of 1 μM enzyme with 200 μM metal ion in 10 mM PIPPS/1 mM Tris/100 mM NaCl/1 mM $CaCl_2$, pH 8. The reaction was stopped by addition of TFA to 0.1%, and reaction products were separated over a C18 cartridge in a 0−30% acetonitrile gradient over 2.5 min. Peak identities were confirmed by mass spectrometry.

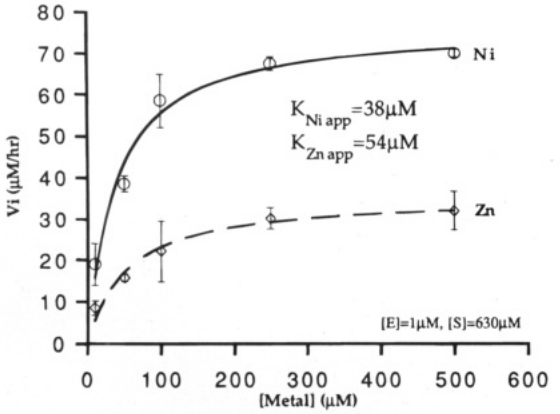

FIGURE 5: Metal titration of Tn H143/151 activity. The peptide hydrolysis assay was performed with increasing concentrations of nickel and zinc. These curves show the saturation of activity by metal ion, yielding apparent metal binding constants of 38 and 54 μM, respectively, when the data were fit to the Michaelis−Menten equation.

275-fold greater than trypsin N143H/E151H in the presence of nickel or zinc.

The activity of trypsin N143H/E151H is dependent on the concentration of added Ni^{2+} or Zn^{2+} ion. The initial velocity of peptide hydrolysis increases with increasing zinc or nickel concentration until the metal and substrate concentrations are equal. Titration of enzyme activity by metal concentrations below substrate concentration displays Michaelis−Menten kinetics, demonstrating a bireactant system with metal ion acting as an obligate cofactor (Segel, 1975). When the concentration of metal exceeds that of substrate, activity decreases (data not shown). This trend is seen at substrate concentrations ranging from 200 μM to 1 mM. Figure 5 shows nickel and zinc titration of activity at 630 μM substrate yielding an apparent zinc binding constant $K_{Zn\ app} = 54\ \mu M$ and an apparent nickel binding constant $K_{Ni\ app} = 38\ \mu M$. These apparent binding constants are comparable to those measured for bidentate sites, suggesting that only two histidines are bound to the metal ion or that the geometry of the three histidines is poor, resulting in lower affinity.

P1 Specificity. Asp189 was mutated to His to engineer metal-regulated histidine specificity into the trypsin S1 binding pocket. As at position P2′, the design principle relies

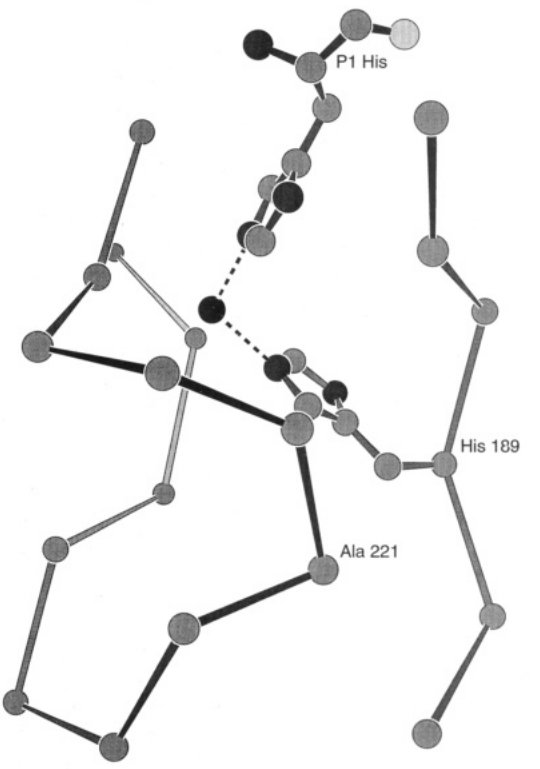

FIGURE 6: Computer model showing designed metal binding site in the trypsin S1 pocket involving His189 of trypsin and the P1 His of substrate. The imidazole−Cu bond lengths are 2 Å, and the Nϵ2−Cu−Nϵ2 bond angle is 100°.

on the affinity of transition metals for histidine side chains, such that a metal ion bridging the enzyme and substrate results in cleavage C-terminal to histidine. Modeling of the primary binding pocket in trypsin showed His189 able to assume the conformation necessary to create a metal binding site in concert with a substrate P1-His (Figure 6). The two histidines coordinate a transition metal ion with bond distances of approximately 2 Å and a N−metal−N angle of 100°. These bond lengths and angles are typically observed in metalloproteins binding transition metals, and both the enzyme and substrate histidines possess allowed rotamers (Ponder & Richards, 1987). In this conformation, however, the His189 side chain makes short steric contacts of 2.5−3.0 Å with main-chain atoms of Ser190. Energy minimization carried out with the program X-PLOR suggested that a small (~0.5 Å) repositioning outward of Ser190 could result in suitable van der Waals interactions with the imidazole ring without introducing other structural changes. It was therefore reasoned that His189 may be accommodated in the modeled conformation because such small rearrangements are common in response to single-site mutations in many proteins.

Trypsin D189H was expressed and purified to homogeneity with a final yield of 12 mg/L of culture. The enzyme possesses no observable activity against the amide substrate Succ-AAPH-pNA either in the presence or absence of 100 μM $CuCl_2$, $NiCl_2$, or $ZnCl_2$. Likewise, there was no observable activity toward Succ-AAPD-pNA or Succ-AAPE-pNA or toward the amide substrate Z-GPR-AMC. However, the enzyme does hydrolyze the more reactive Z-Lys-thiobenzyl ester substrate with a $k_{cat}/K_m = 0.35 \pm 0.01\ min^{-1}\ \mu M^{-1}$ (corrected for background hydrolysis), demonstrating a stably folded protein and integrity of the catalytic triad.

Table 2: Ester Kinetics of S1 Variants[a]

enzyme	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($min^{-1}\ \mu M^{-1}$)
trypsin	6700 ± 150	86 ± 3	78 ± 5
Tn D189H	740 ± 40	2100 ± 170	0.35 ± 0.01
Tn D189H:G226D	760 ± 5	1200 ± 25	0.63 ± 0.01

[a] Kinetic parameters of S1 variants against Z-Lys-SBzl ester. Assays were performed in 10 mM PIPPS/1 mM Tris/100 mM NaCl/1 mM $CaCl_2$, pH 8, at 25 °C with [E] = 2 μM.

Table 3: Trypsin D189H Crystallographic Data

space group	$P3_221$
cell dimensions	$a = b = 92.83$ Å, $c = 62.18$, $\alpha = \beta = 90°$, $\gamma = 120°$
resolution	1.8 Å
%complete	95%
R_{merge}[a]	6.03%
R_{crys}[b]	18.4%
rms bond differences	0.011 Å
rms angle differences	2.655°

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

Wild-type trypsin hydrolyzes this substrate with a k_{cat}/K_m = 78 ± 1 $min^{-1}\ \mu M^{-1}$, 200-fold faster than D189H (Table 2).

A second generation of variants was designed in the P1 site. These additional mutations were introduced in an attempt to stabilize His189 in a suitable orientation for metal binding. The variants D189H/S190G and D189H/G226D were constructed to remove the serine hydroxyl bulk at position 190 and to introduce a hydrogen bonding partner for His189, respectively. The modeled position of Asp226 relative to the His189 and metal ion mimicks the carboxylate−histidine−zinc interaction which forms a stable structure in the active site of metalloproteins (Christianson & Alexander, 1989). In each case, modeling followed by energy minimization showed similar metal ion coordination geometry to that of D189H. Trypsins D189H/S190G, D189H/G226D, and D189H/S190G/G226D were expressed in yeast and purified to homogeneity by BPTI affinity chromatography. Kinetic analysis showed that those variants containing Gly190 exhibited no activity, in the presence or absence of metal cofactors, toward either amide or ester substrates. Trypsin D189H/G226D likewise showed no amidase activity toward Succ-AAPH-pNA under any conditions but did hydrolyze Z-Lys-SBzl ester with k_{cat}/K_m = 0.63 ± 0.01 $min^{-1}\ \mu M^{-1}$, a 2-fold increase in turnover relative to D189H, or 100-fold lower than wild-type trypsin.

Crystal Structure of Trypsin D189H. Trypsin D189H was crystallized in the presence of BPTI and the structure determined at 1.8 Å resolution to understand the structural basis for the failure of the enzyme to exhibit the desired metal-regulated histidine specificity. Crystallographic results are summarized in Table 3. Main chain atoms of the active site residues His57, Asp102, and Ser195 are superimposable with rms deviations of 0.05−0.2 Å when compared to other structures of rat trypsin variants (McGrath et al., 1993; Perona et al., 1993, 1994). The average pairwise rms deviation among the various structures is 0.33−0.40 Å for all backbone Cα atoms. Thus, the D189H mutation does not compromise the overall folding of the enzyme.

The His 189 side chain is oriented toward the back of the pocket away from substrate. Its position contrasts strikingly to that of Asp189 in wild-type trypsin, and precludes interaction with a metal ion to form a P1-His binding site (Figures 7a−c). The imidazole ring forms two hydrogen bonds: one with poor geometry between His189 Nε2 and the carbonyl oxygen of Gly187 (N−H−O angle 131°, 2.65 Å) and a second from His189 Nδ1 to water molecule 554 (N−H−O angle 156°, 2.64 Å; Figure 7b). This water bridges to the main-chain carbonyl group of Pro225. Sharp peaks in electron density maps calculated from coefficients ($F_o - F_c$) during refinement indicated that the peptide groups of both Pro225 and Gly187 have been rotated approximately 180° in ψ relative to wild-type trypsin. These backbone adjustments apparently occur to provide the hydrogen-bonding interactions with His189. A further consequence of the main-chain reorientations is a significant perturbation in the overall conformation of the two surface loops bridging the three β-strands of the P1 pocket (loops 1 and 2; Figure 7). While the deviation in position of equivalent α-carbons within loop 1 ranges from 0.2 to 1.0 Å, in loop 2 the deviation is as great as 4 Å relative to wild-type trypsin. An interloop hydrogen bond bridging residues Gly188 and Ala221 is disrupted in this new conformation, although the main-chain hydrogen bond between residues Asp223 and Leu185 is preserved. The water molecules nearest the entrance of the S1 pocket, which interact with the P1-Lys of BPTI, occupy identical positions in this structure relative to wild-type trypsin. Other waters at the base of the pocket are found in novel locations resulting from the presence of His189 and the new conformations of loops 1 and 2.

DISCUSSION

P2′ Designs. Molecular modeling was successful in predicting a trypsin design which introduces metal-dependent histidine specificity at P2′. Trypsin N143H/E151H cleaves a peptide substrate containing a histidine at P2′ in the presence of either nickel or zinc, but not in the presence of copper. The preferred ligand geometries of these three metal ions suggests the observed specificity: distorted octahedral and tetrahedral coordination geometries are observed with nickel and zinc, whereas copper is observed in a square planar geometry. Coordination of metal ions in proteins probably reflects balance of energetic contributions between protein flexibility and metal ion geometry. The S2′ site is formed by the juxtaposition of several surface loops well exposed to solvent. Possibly, the greater potential flexibility of these surface loops, relative to core β-strands critical to the tertiary fold, might be a key to the accommodation of metal binding without a loss of structural integrity. By contrast, the S1 site is a deep cavity which is only partially solvent-accessible and is formed by β-strands which are core secondary structure elements.

The activity of the enzyme toward the peptide substrate was strictly dependent on nickel or zinc concentration, showing saturation kinetics when metal concentrations were lower than substrate concentrations. These data obey Michaelis−Menten kinetics (Figure 5) in a manner consistent with a two-reactant system where the metal ion acts as an obligate cofactor (Segel, 1975). At a substrate concentration of 630 μM, zinc and nickel bind with apparent binding constants $K_{Zn\ app}$ = 54 μM and $K_{Ni\ app}$ = 34 μM. These binding constants are expected to be dependent on substrate concentration as the histidine-containing substrate itself has

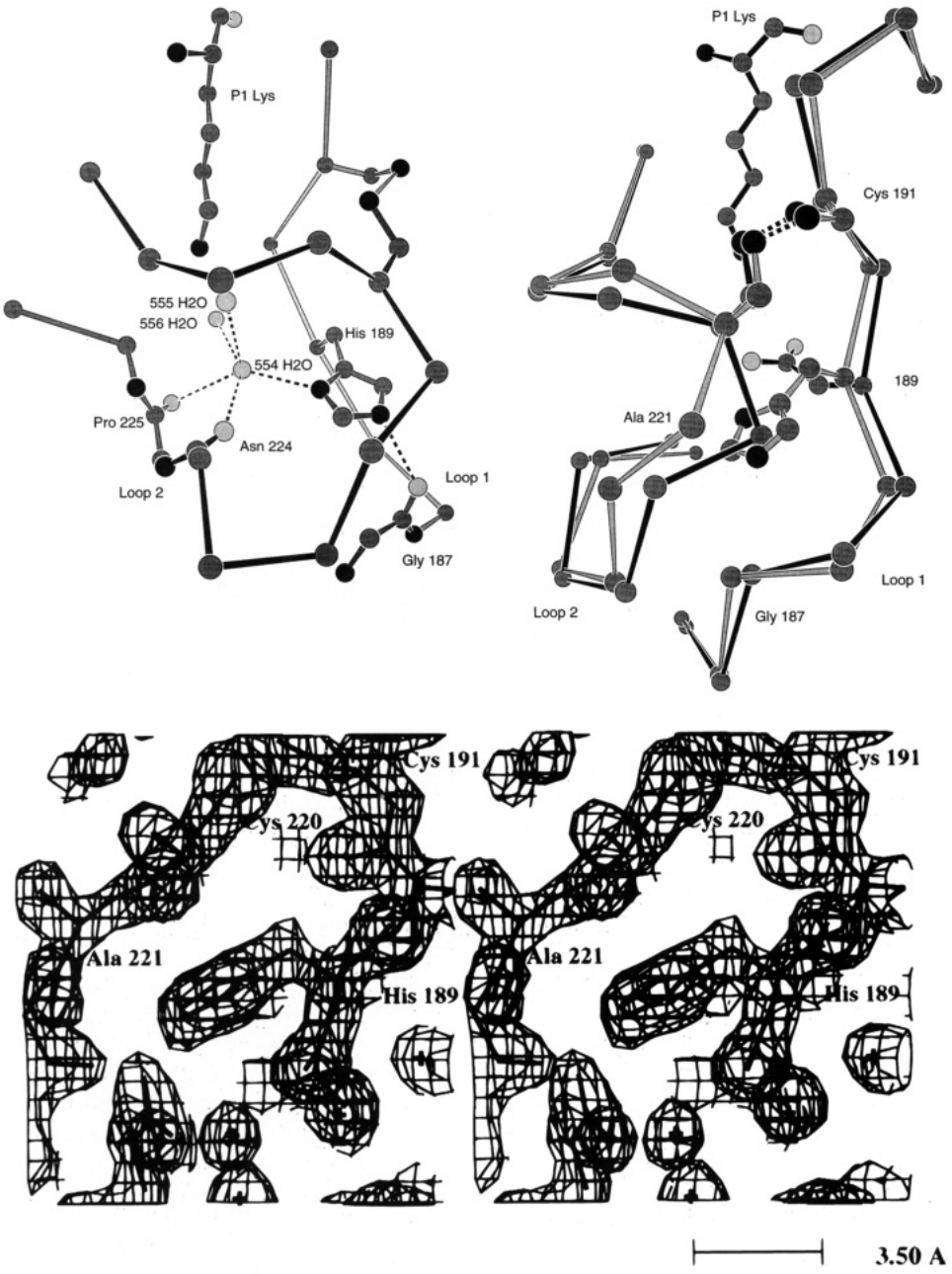

FIGURE 7: (a, top left) Crystal structure of Tn D189H complexed to BPTI showing detail of P1 pocket. Histidine 189 is pointed down and away from the top of the pocket and is unable to participate in metal or substrate binding. Relevant interatomic distances (Å) are as follows: His189Nϵ2 to carbonyl oxygen of Gly187, 2.65; His189Nδ1 to water 554, 2.64; water 554 to water 556, 2.59; water 554 to water 555, 2.89; water 554 to carbonyl oxygen of Pro225, 2.89; and water 554 to carbonyl oxygen of Asn224, 2.71. Loop 1 is residues 185−189, and loop 2 is residues 217−225. The disulfide bond between Cys191 and Cys220 is shown striped. (b, top right) Trypsin D189H structure superimposed upon wild-type trypsin. Wild-type trypsin is shown in black, while trypsin D189H is shown in gray. The His189 side chain is pointing down in the pocket as opposed to the Asp189 of wild-type trypsin. Also visible is the 219−225 loop (loop 2) movement of 1−4 Å. The disulfide bond between C191 and C220 is shown with the dashed line. (c, bottom) Stereo $2F_o - F_c$ omit map contoured at 1σ of trypsin D189H showing the area around His 189 and including the part of the 220s loop which deviated most from wild-type trypsin (loop 2 in panel b). Residues 189 and 217−226 were deleted from the model used to calculate this map. The electron density was displayed using the program CHAIN (Sack, 1988).

some intrinsic binding affinity for the metal ions. This is manifested as a reduction in V_{max} when metal ion concentrations exceeds substrate concentrations. The decrease in activity at high metal ion concentrations is consistent with either an interaction between the metal and the substrate, rendering the substrate unavailable to the enzyme, or between metal and enzyme, causing direct enzyme inhibition.

Metal ions are used in naturally occurring proteins for a variety of purposes. In proteases they are used in catalysis and structural stability, but they are yet to be observed in protease−substrate binding interactions. Other types of enzymes are known to use metal ions to bind substrates. Data for the restriction enzyme *Eco*RV show that Mg^{2+} is required for proper substrate specificity and show relaxed substrate discrimination when Mg^{2+} is replaced by Mn^{2+} (Vermote & Halford, 1992). Additionally, studies of xylose isomerase have shown that the bound Mg^{2+} ions are important for proper substrate binding and subsequent catalysis (Jenkins et al., 1992). These results indicate that metal ions are widely used for substrate recognition in enzymes other than proteases. Following this observation is the question of whether proteases also use this mechanism of substrate binding even

though it has not been observed. Possibly trypsin N143H/E151H functions in a manner similar to undiscovered naturally occurring proteases employing metal-dependent molecular recognition.

The observation that zinc activates wild-type trypsin, but nickel or copper does not, raises interesting questions regarding the nature of metal ion specificity for coordination to His−Glu pairs. The selectivity may be geometrically based, requiring proper spatial orientation of ligands. Also, it is notable that copper ion does not activate any of these enzymes at any of the concentrations used in this study. As copper is expected to bind tighter to histidines than either nickel or zinc (Martell & Smith, 1974), one explanation for the failure of copper to activate the enzyme is that it binds too tightly to the enzyme/substrate complex to allow turnover. In this context, the designed metal binding site may function to inhibit the enzyme from hydrolyzing P2′ histidine-containing substrates in the presence of copper. It has been observed that Cu^{2+} inhibits wild-type trypsin with a $K_i = 225\ \mu M$ (W.S.W., unpublished result), indicating that copper interacts with the enzyme in a deleterious fashion. The copper ion may bind between the catalytic His57 and Asp102 causing the observed inhibition as silver ion was shown to do (Chambers et al., 1974). Alternatively, the square planar geometry preferred by copper may not be possible within the structural constraints imposed at the S2′ site. Crystal structures of trypsin N143H/E151H complexed with a P2′-His-containing inhibitor in the presence of different metal ions should provide further insight into this question.

Trypsin D189H Crystal Structure. The crystal structure of trypsin D189H reveals conformational changes in two surface loops (loop 1 and loop 2; Figure 7b) adjacent to the S1 pocket, resulting in a new network of intramolecular interactions. The structural changes result from the interaction of His189 with backbone amides of both loops, which leads to the ψ rotation of 180° of two peptide bonds and consequent adoption of the new conformation. It is of interest that the interaction of a His imidazole with adjacent main-chain groups is not uncommon in proteins (Baker & Hubbard, 1984). For example, in trypsin His40 makes a similar interaction with the carbonyl oxygen atom of Gly193, and His53 in the serine protease inhibitor ecotin (McGrath et al., 1994) interacts with the carbonyl oxygen atom of Trp51 in a similar manner. The geometry in the latter case is very similar to that observed in the His189−Gly187 interaction. In both instances the hydrogen-bond-accepting backbone carbonyl group lies in an adjacent loop. Possibly, a favorable interaction of the His189 imidazole ring with the adjacent Gly187 main chain provides part of the stabilization energy of the new loop conformations. These observations might profitably be used to reorient surface loops in protein engineering designs.

It is, however, difficult to estimate the subtle forces that favor a large rearrangement of the pocket to the new conformation rather than the modest relaxation of the Ser190 backbone apparently required to permit orientation of His189 toward substrate. The additional substitutions introduced into D189H were ineffective in stabilizing the modeled orientation of the imidazole ring. Since the Gly190 variants exhibit no activity at all, it is possible that these enzymes are misfolded and do not adopt stable tertiary structures. Trypsin D189H/G226D manifests a weak esterase activity very similar to that of D189H. Thus, it is seems likely that His189 in this enzyme adopts a D189H-like orientation. None of the changes therefore appear to be sufficient to overcome the combination of forces stabilizing the new conformations of loops 1 and 2.

The rearrangement of the S1 pocket in trypsin D189H provides insight into the requirements for substrate specificity modification at this position. In contrast to the serine proteases subtilisin and α-lytic protease (Bone et al., 1989; Estell et al., 1986), mutagenesis of trypsin S1-site residues directly in contact with substrate fails to improve catalytic efficiencies toward alternate P1 amino acids (Graf et al., 1987, 1988; J.J.P. and C.S.C., unpublished observations). Instead, exchange of the surface loops 1 and 2, together with replacement of four S1-site amino acids, will endow trypsin with a chymotrypsin-like substrate preference (Hedstrom et al., 1992; Perona et al., 1995).

In this context it is of interest that the introduction of His189 causes a large structural change in loops 1 and 2 rather than a small shifting of the main chain directly in the P1 side-chain binding site at Ser190. In α-lytic protease, greatly broadened substrate specificity profiles appear to arise from structural plasticity of the S1 site, which includes small movements (0.5−1.0 Å) of the main chain in an area which directly contacts substrate (Bone et al., 1991). Thus, the enzyme can deform to accommodate alternate P1-substrate side chains in a manner which does not compromise positioning of the scissile bond relative to the catalytic machinery. Structural analysis of a number of S1-site trypsin variants, however, shows some adjustment in second-shell residues but none in amino acids which directly contact substrate (Perona et al., 1993; Wilke et al., 1991). Similarly, we observe here that in trypsin D189H an apparently small deformation of Ser190 to accommodate His189 does not occur. Thus, the His189 imidazole may be oriented away from substrate because there is no capacity for local deformability. The absence of even limited flexibility to permit orientation of His189 toward substrate supports the hypothesis that the requirement for extensive mutagenesis to obtain specificity modification in trypsin is due to a high degree of rigidity of the S1 site (Perona et al., 1995).

Loops 1 and 2 also form part of the "activation domains" of trypsin and chymotrypsin, which become ordered only upon transition from the zymogen forms to the mature enzymes (Huber & Bode, 1978). The segments involved in this conformational transition include the N-terminal residues 16−19, the adjacent loop from residues 142 to 152, and loops 1 and 2 together with adjacent residues at positions 216 and 189−193. Kinetic and structural analyses of trypsin mutants possessing partial chymotrypsin-like specificities suggests that the stability of the S1 site in each native enzyme is highly dependent upon the integrity of the adjacent structure (Hedstrom et al., 1994). Both the conformations of loops 1 and 2 as well as their interactions with surrounding portions of the structure are quite different in trypsin and chymotrypsin. Alternate strategies for stabilizing the S1 site in trypsin and chymotrypsin thus appear to have evolved together with the divergence in substrate specificity.

In this context it is noteworthy that in trypsin D189H the entire "activation domain" (Huber & Bode, 1978), including loops 1 and 2, is very well-ordered. The structure shows that the side chain of His189 occupies a central position between loops 1 and 2, where it disrupts existing wild-type

interactions and solvent structure in addition to producing the large conformational change in loop 2. Therefore, the networks of intramolecular interactions present within the activation domains of trypsin and chymotrypsin do not represent exclusive solutions to the problem of producing a stable structure. Trypsin D189H is produced in the zymogen form and activated by enteropeptidase. This shows that the mutation does not block the normal processing pathway, although it may destabilize the partially folded zymogen state. The observation that D189H possesses an alternate stable conformation in the activation domain, rather than destabilizing the structure as observed in D189S (Hedstrom et al., 1994), may be of significance to understanding the mechanism of zymogen activation in these enzymes.

The crystal structure of Tn D189H provides the structural basis for the observed kinetic result. Kinetic data can be used in conjunction with X-ray structural data as the basis for iterative design, making trypsin a good model system to test protein engineering hypotheses. These experiments showed, however, that the primary specificity pocket is not easily reconfigured. The obvious "improvements" suggested by the structure do not lead to the desired kinetic result of metal-dependent specificity. This reflects the current partial understanding of protein structure/function relationships yet allows a significant addition to the database of protein engineering results.

This report demonstrates that an engineered metal binding site may be used to effect histidine specificity in trypsin. Kinetic analysis was used to assess the success of the design, and X-ray crystallographic analysis was used to elucidate structural details of the mutant enzyme. Iterative use of kinetic and structural analysis may improve current levels of activity so that these design principles can be used to create highly specific, metal-dependent proteases.

ACKNOWLEDGMENT

We thank C. Tsu for technical advise and permission to cite unpublished kinetic data, members of the Craik lab for critical reading of the manuscript, and J. M. Parks for inspiration. Figures 2, 6, and 7 were drawn using the program MAXIMAGE written by M. Rould. Mass spectrometry was performed at the University of California San Francisco Mass Spectrometry Facility (A. L. Burlingame, Director), which is supported by the National Institute of Environmental Health Schiences Grant ES 04705.

REFERENCES

Bai, K. S., & Martell, A. E. (1969) *J. Inorg. Nucl. Chem. 31*, 1697−1707.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97−179.

Bone, R., Silen, J. L., & Agard, D. A. (1989) *Nature 339*, 191−195.

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

Braxton, S., & Wells, J. A. (1992) *Biochemistry 31*, 7796−7801.

Browner, M. F., Hackos, D., & Fletterick, R. J. (1994) *Nature Struct. Biol. 1*, 327−333.

Brunger, A. T., Huber, R., & Karplus, M. (1987) *Biochemistry 26*, 5153−5162.

Chambers, J. L., Christoph, G. G., Krieger, M., Kay, L., & Stroud, R. M. (1974) *Biochem. Biophys. Res. Commun. 59*, 70−74.

Christianson, D. W., & Alexander, R. S. (1989) *J. Am. Chem. Soc. 111*, 6412−6419.

Corey, D. R., & Schultz, P. G. (1989) *J. Biol. Chem. 264*, 3666−3669.

Erhart, E., & Hollenberg, C. P. (1983) *J. Bacteriol. 156*, 625−635.

Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., & Wells, J. A. (1986) *Science 233*, 659−663.

Evnin, L. B., Vasquez, J. R., & Craik, C. S. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6659−6663.

Gardell, S. J., Craik, C. S., Clauser, E., Goldsmith, E. J., Stewart, C. B., Graf, M., & Rutter, W. J. (1988) *J. Biol. Chem. 263*, 17828−17836.

Graf, L., Craik, C. S., Patthy, A., Roczniak, S., Fletterick, R. J., & Rutter, W. J. (1987) *Biochemistry 26*, 2616−2623.

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.

Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249−1253.

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

Higaki, J. N., Haymore, B. L., Chen, S., Fletterick, R. J., & Craik, C. S. (1990) *Biochemistry 29*, 8582−8586.

Higaki, J. N., Fletterick, R. J., & Craik, C. S. (1992) *Trends Biochem. Sci. 17*, 100−104.

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

Ibers, J. A., & Holm, R. H. (1980) *Science 209*, 223−235.

Jenkins, J., Janin, J., Rey, F., Chiadmi, M., van Tilbeurgh, H., Lasters, I., De Maeyer, M., Van Belle, D., Wodak, S. J., Lauwereys, M., et al. (1992) *Biochemistry 31*, 5449−5458.

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

Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.

Martell, A. D., & Smith, R. M. (1974) *Critical Stability Constants*, Plenum Press, New York.

McGrath, M. E., Haymore, B. L., Summers, N. L., Craik, C. S., & Fletterick, R. J. (1993) *Biochemistry 32*, 1914−1919.

McGrath, M. E., Erpel, T., Bystroff, C., & Fletterick, R. J. (1994) *EMBO J. 13*, 1502−1507.

McPherson, A. (1989) *Preparation and Analysis of Protein Crystals*, Robert E. Kreiger Publishing Co., Malabar, FL.

Perona, J. J., Tsu, C. A., Craik, C. S., & Fletterick, R. J. (1993) *J. Mol. Biol. 230*, 934−939.

Perona, J. J., Hedstrom, L., Wagner, R., Rutter, W. J., Craik, C. S., & Fletterick, R. J. (1994) *Biochemistry 33*, 3252−3259.

Perona, J. J., Hedstrom, L., Rutter, W. J., & Fletterick, R. J. (1995) *Biochemistry* (in press).

Perona, J. J., & Craik, C. S. (1995) *Protein Sci.* (in press).

Ponder, J. W., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 775−791.

Sack, J. S. (1988) *J. Mol. Graphics 6*, 224−225.

Sato, M., Yamamoto, M., Imada, K., & Katsuba, Y. (1992) *J. Appl. Crystallogr. 25*, 348−357.

Schechter, I., & Berger, A. (1968) *Biochem. Biophys. Res. Commun. 27*, 157.

Schellenberger, V., Turck, C. W., & Rutter, W. J. (1994) *Biochemistry 33*, 4251−4257.

Segel, I. H. (1975) *Enzyme Kinetics*, John Wiley & Sons, Inc., New York.

Sprang, S. R., Fletterick, R. J., Graf, L., Rutter, W. J., & Craik, C. S. (1988) *Crit. Rev. Biotechnol. 8*, 225−236.

Tainer, J. A., Roberts, V. A., & Getzoff, E. D. (1991) *Curr. Opin. Biotechnol. 2*, 582−591.

Vermote, C. L., & Halford, S. E. (1992) *Biochemistry 31*, 6082−6089.

Wilke, M. E., Higaki, J. N., Craik, C. S., & Fletterick, R. J. (1991) *J. Mol. Biol. 219*, 525−532.

BI941885S