

Protein Science

Conversion of trypsin to a functional threonine protease

Teaster T. Baird, Jr, William D. Wright and Charles S. Craik

Protein Sci. 2006 15: 1229-1238; originally published online May 2, 2006;
doi:10.1110/ps.062179006

References

This article cites 26 articles, 9 of which can be accessed free at:
<http://www.proteinscience.org/cgi/content/full/15/6/1229#References>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

Notes

To subscribe to *Protein Science* go to:
<http://www.proteinscience.org/subscriptions/>

Conversion of trypsin to a functional threonine protease

TEASTER T. BAIRD JR.,^{1,2} WILLIAM D. WRIGHT,² AND CHARLES S. CRAIK¹

¹University of California, San Francisco, Department of Pharmaceutical Chemistry, San Francisco, California 94143-2280, USA

²San Francisco State University, Department of Chemistry and Biochemistry, San Francisco, California 94132, USA

(RECEIVED February 21, 2006; FINAL REVISION March 7, 2006; ACCEPTED March 8, 2006)

Abstract

The hydroxyl group of a serine residue at position 195 acts as a nucleophile in the catalytic mechanism of the serine proteases. However, the chemically similar residue, threonine, is rarely used in similar functional context. Our structural modeling suggests that the Ser 195 → Thr trypsin variant is inactive due to negative steric interaction between the methyl group on the β -carbon of Thr 195 and the disulfide bridge formed by cysteines 42 and 58. By simultaneously truncating residues 42 and 58 and substituting Ser 195 with threonine, we have successfully converted the classic serine protease trypsin to a functional threonine protease. Substitution of residue 42 with alanine and residue 58 with alanine or valine in the presence of threonine 195 results in trypsin variants that are 10^2 – 10^4 -fold less active than wild type in k_{cat}/K_M but $>10^6$ -fold more active than the Ser 195 → Thr single variant. The substitutions do not alter the substrate specificity of the enzyme in the P1'–P4' positions. Removal of the disulfide bridge decreases the overall thermostability of the enzyme, but it is partially rescued by the presence of threonine at position 195.

Keywords: serine proteases; trypsin; enzyme mechanisms; protein engineering; substrate specificity

The catalytic machinery that is characteristic of the serine proteases is a classic example of convergent evolution in which nature has selected for a common biological trait through different and unrelated evolutionary pathways. In the case of the serine proteases, the conserved biological trait is the composition and geometry of the essential active site residues. The active site residues that typify the classic serine protease comprise a catalytic triad consisting of a histidine residue, an aspartic acid residue, and the eponymous serine residue. The spatial arrangement of these three residues is nearly identical in serine proteases isolated from both prokaryotic and eukaryotic sources. For example, the crystal structures of subtilisin from

Bacillus subtilis, Kex2 from *Saccharomyces cerevisiae*, trypsin from rats, and thrombin from humans (Bode et al. 1992; Perona et al. 1993a; Steinmetz et al. 1994; Holyoak et al. 2003) show that for each enzyme, the carboxylate of the aspartate side chain accepts a hydrogen bond from the N ϵ 2 of the histidine while N δ 1 of the His residue accepts a hydrogen bond from the hydroxyl group of the serine. The global structures of this family of enzymes are typically classified as either trypsin-fold or subtilisin-fold, although new folds have been identified. Classical examples of the trypsin-fold serine proteases include both digestive proteases such as trypsin and elastase, and regulatory proteases such as Granzyme B and the enzymes of the blood coagulation cascade (Waugh et al. 2000; Davie 2003). Examples of subtilisin-fold proteases include subtilisin BPN', Proteinase K, and furin (Nakayama 1997; Siesen and Leunissen 1997; Betzel et al. 2001). Even though their global three-dimensional structures are unrelated, the identity and local arrangement of

Reprint requests to: Charles S. Craik, University of California, San Francisco, Department of Pharmaceutical Chemistry, 600 16th Street, Box 2280, San Francisco, CA 94143-2280, USA; e-mail: craik@cgl.ucsf.edu; fax: (415) 502-8298.

Article published online ahead of print. Article and publication date are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062179006>.

the three residues required for catalysis in these two families are practically identical. The selection of the same catalytic machinery through genetic selection and evolutionary pressures suggests that this system is one of the most efficient and most fit.

In the catalytic mechanism of the serine proteases, the three residues of the catalytic triad (His 57, Asp 102, and Ser 195, chymotrypsin numbering) act in concert to increase the nucleophilicity of the hydroxyl moiety of Ser 195, the only residue of the three that participates directly in the currently accepted catalytic mechanism (Fig. 1A). The hydroxyl group of Ser 195 attacks the carbonyl carbon of the scissile bond of the substrate and forms a tetrahedral acyl enzyme intermediate that is stabilized by the "oxyanion hole" that is formed by the backbone amide hydrogens of Gly 193 and Ser 195 (chymotrypsinogen numbering) (Ding et al. 1994). The enzyme is deacylated and regenerated when a solvent water molecule acts as the nucleophile and attacks the carbonyl carbon of the acyl enzyme. If any member of the triad is substituted with another amino acid residue, activity of the enzyme is reduced significantly (Carter and Wells 1987; Corey and Craik 1992), underscoring the importance of the relationship between the three residues.

When one considers the role of Ser 195, it is interesting to note that threonine is not used in a similar context. In

fact, only a few examples of proteases in which threonine acts as the nucleophile exist, and none use a Thr-His-Asp triad. One example of a threonine protease is the proteasome. It assembles into a multimeric complex in order to position its substrates, and uses a Thr-Glu/Asp-Lys triad (Brannigan et al. 1995; Lowe et al. 1995; Baumeister and Lupas 1997). Serine and threonine have similar reactivity, and threonine occurs in the genome 1.1% less often than serine. Thus, it is unlikely that the employment of serine over threonine in the traditional triad may be attributed to a difference in reactivity of the hydroxyl group or to the evolutionary availability of serine in the genome (McCaldon and Argos 1988). After elimination of the aforementioned possibilities, the remaining significant difference between the two amino acids within the context of their use or potential use by a protease is one of a structural nature: the presence of a methyl group on the β -carbon of threonine.

Previous studies with the S195T Tn variant demonstrated that this substitution of serine with the chemically similar threonine resulted in a protein that was virtually inactive (Corey and Craik 1992). When the structure of the active site of wild-type trypsin was subsequently modeled with a threonine in position 195 instead of a serine, it suggested that in order for the hydroxyl group of Thr 195 to participate in the essential hydrogen bond

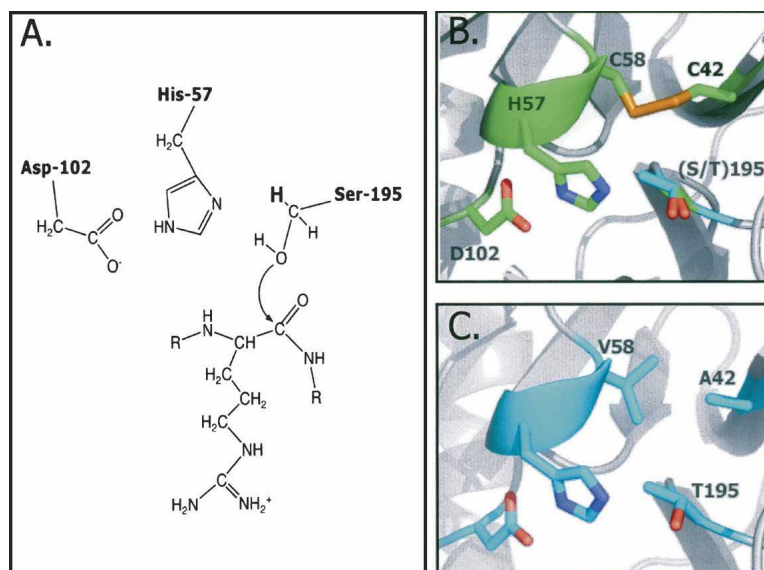


Figure 1. (A) Accepted mechanism of serine proteases. Schematic of the catalytic triad in a serine protease in the presence of a bound Arg-containing substrate. The hydroxyl group of Ser 195 attacks the carbonyl of the Arg residue. The hydrogen atom of the β -carbon closest to the S1' pocket is in bold. (B) Arrangement of the catalytic triad in wild-type trypsin. The arrangement of the residues in the active site of wild-type rat trypsin are shown in stick model. The amino acid side chains of interest are in green for wild type and cyan for the S195T variant. Oxygen atoms are red, nitrogen atoms are blue, and sulfur atoms are orange. This figure was constructed with PYMOL using the coordinates from PDB file 1ANE. (C) Model of optimal arrangement of the catalytic triad in C42A/C58V/S195T trypsin (AVT-Tn) variant. Ala 42, Val 58, and Thr 195 are labeled. The expected arrangement for an active threonine protease variant of trypsin. The hydroxyl group of Thr 195 is in a position comparable to that of Ser 195 in the native enzyme.

network of the catalytic triad in a manner similar to Ser 195 and produce an active enzyme, the methyl group of Thr 195 would have to occupy a spatial position disallowed by the presence of the disulfide bridge formed by cysteines 42 and 58. In this paper, we describe experiments, based on computational analysis of structural properties of threonine and trypsin, designed to determine the structural changes that would be required to convert the classical serine protease, trypsin, into a novel, active threonine protease. We demonstrate that trypsin may indeed be converted into a functional threonine protease that utilizes a catalytic mechanism similar to that of the wild-type enzyme.

Results and Discussion

Mutagenesis and expression of variants

Several trypsin variants were constructed using PCR mutagenesis and expressed as mature enzymes (without the N-terminal propeptide sequence) in *Pichia pastoris*, as described in Materials and Methods. The expression of the variants in the mature form eliminated the need for activation by removal of the propeptide by enteropeptidase, thereby circumventing a potential source of false positives. Expression levels ranged from 1 to 20 mg/L for each variant in the following order: C42A/C58V/S195T > C42A/C58A/S195T > C42A/C58V > C42A/C58A. After hydrophobic chromatography on a phenyl sepharose CL-4B column, the variants were >90% pure, as determined by PAGE. Subsequent DEAE anion exchange chromatography resulted in samples that were >98% pure.

Design of variants

Given the spatial relationship among the residues of the catalytic triad and the chemical similarities between the amino acids serine and threonine, one might expect that a trypsin variant in which Ser 195 was substituted with threonine would exhibit hydrolytic activity. However, previous studies have demonstrated that if Ser 195 is substituted with threonine in native trypsin (S195T Tn), the result is a trypsin variant that is 10^5 -fold less active than wild-type trypsin (Corey and Craik 1992). Molecular modeling studies that were carried out in this investigation suggested that the reason for this large decrease in activity is the inability of the hydroxyl group of the substituting threonine to donate a hydrogen bond to His 57 in a manner similar to Ser 195 in the native enzyme.

In the crystal structure for wild-type trypsin complexed with benzamidine (PDB ID 1ANE), the hydroxyl oxygen of Ser 195 is 3.6 Å from Nε2 of His 57, too far for an efficient hydrogen bond (Fig. 1B, green). However, in the

cocrystal structure of trypsin and bovine pancreatic trypsin inhibitor (PDB ID 1BRB), the distance between the hydroxyl oxygen of Ser 195 and Nε2 of His 57 is reduced to 2.7 Å (Perona et al. 1993b). Similar distances (<3 Å) between these two atoms are observed in cocrystal structures of trypsin and other macromolecular inhibitors. Therefore, one may consider the trypsin–benzamidine cocrystal structure and the increased distance between His 57 and Ser 195 to be representative of the enzyme conformation in the absence of substrate. Altogether, such cocrystal structures suggest that a conformational change that brings the hydroxyl group of Ser 195 within hydrogen bonding distance of His 57 occurs upon substrate binding. This interaction increases the nucleophilicity of Ser 195 and facilitates its use in catalysis. From this observation, it becomes apparent that in order for the hydroxyl group of Thr 195 to be used as a nucleophile in a manner similar to Ser 195, its hydroxyl group would have to be within hydrogen bonding distance the Nε2 of His 57 upon binding a peptide substrate.

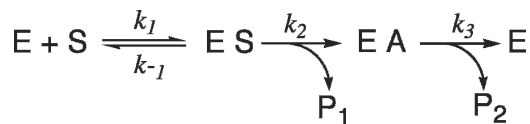
For efficient hydrogen bonding between Thr 195 and His 57 to occur, the methyl and hydroxyl groups of Thr 195 would have to rotate clockwise about its β-carbon. Using the PDB structure 1ANE as a starting point, this interaction was modeled using the programs PyMOL (DeLano 2002) and SwissPDB Viewer (Guex and Peitsch 1997). When modeled using the crystal structure of the trypsin–benzamidine complex, putting the hydroxyl group of Thr 195 in a position similar to that observed with Ser 195 in the native enzyme necessitates that the methyl group of the putative S195T substitution occupy a position that is sterically disallowed. In such a conformation, the methyl group of the substituting threonine would be 2.0 Å from His 57 (Fig. 1B, cyan). In order to interact favorably with the hydroxyl group of Thr 195, the interaction between the imidazole ring of His 57 and the methyl group of Thr 195 would have to be minimized. This may be achieved by a rotation of the imidazole ring of His 57 and/or repositioning of the methyl group of Thr 195. Of particular interest to achieving this local structural modification is the disulfide bridge near the catalytic triad, formed by cysteine residues 42 and 58. Given its location, the disulfide bridge likely plays a role in maintaining the relative position of His 57 in the catalytic triad. Furthermore, its presence precludes the repositioning of the methyl group of Thr 195 to a catalytically viable location; further rotation about the β-carbon would result in negative steric interaction between the methyl group and the disulfide bridge. Thus, one would expect that eliminating the bridge would relax the steric constraints it imposes and allow rotation about the β-carbon of Thr 195 into a catalytically viable location.

With such a rotation, the hydroxyl group of the substituting threonine may then participate in the essential

hydrogen-bond network characteristic of the catalytic triad allowing catalysis to occur. To test this hypothesis, several Tn variants were constructed in which the disulfide bond was removed and Ser 195 was substituted with threonine. The C42A/C58V/S195T-Tn variant (AVT-Tn) was designed to remove the Cys 42–Cys 58 disulfide bond and introduce a hydrophobic character into the modified region. This change was designed to minimize any loss in local structural stability that might occur by elimination of the two cysteines. The Cys 42–Cys 58 disulfide bridge (Fig. 1C) is strictly conserved. In an alignment of 1327 chymotrypsin-like serine proteases using the MEROPS database (Rawlings et al. 2004), the Cys 42–Cys 58 disulfide bridge is found in 97%. A similar triple variant, C42A/C58A/S195T Tn, (AAT-Tn) was constructed and investigated in parallel to determine if decreasing the van der Waals volume in the region formerly occupied by the disulfide bond could result in a variant with greater activity relative to its AVT-Tn counterpart. Finally, two double variants, C42A/C58V (AV-Tn) and C42A/C58A (AA-Tn) were constructed to dissociate the change in nucleophile from the possibly destabilizing effects of removing the disulfide bond.

Activity of the trypsin variants

To examine the effect of the substitutions on specific steps of the catalytic mechanism, the activity of each variant was assayed by measuring its hydrolysis of thiobenzyl (SbzI), para-nitroanilide (pNA) and 7-amino-4-methyl coumarin (AMC) derivatives of benzyloxycarbonyl-glycylprolylarginine (Z-GPR) (Scheme 1). For wild-type trypsin, deacylation (k_3) is rate limiting when measuring hydrolysis of the highly activated thiobenzyl ester substrate, while acylation (k_2) is rate limiting with pNA, AMC, and peptide substrates. In this study, when the activity of the single variant S195T-Tn was assayed, little to no activity was observed with any of the above substrates, suggesting that substitution of serine with a threonine residue in this particular structural context essentially eliminates activity. The S195T single variant only displayed activity toward the thiobenzyl ester substrate when measured at high ($\geq 1 \mu\text{M}$) enzyme concentrations (Fig. 2). Even at such high concentrations, the



Scheme 1. Mechanism of serine proteases. E is the free enzyme, ES is the enzyme–substrate complex, EA is the covalent acyl enzyme intermediate, P1 and P2 are the N- and C-terminal hydrolysis products, respectively.

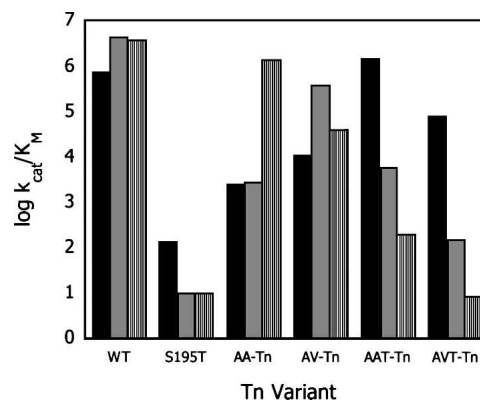


Figure 2. Activity of trypsin variants. The dark bars are activities of the trypsin variants toward Z-GPR-SBzI. The light bars are activities toward Z-GPR-pNA (pNA). The striped bars are activities toward Z-GPR-AMC (AMC). The abbreviations for the trypsin variants are as follows: WT = wild type; S195T = Ser 195 → Thr; AA-Tn = C42A/C58A trypsin; AV-Tn = C42A/C58V trypsin; AAT = C42A/C58A/S195T trypsin; AVT-Tn = C42A/C58V/S195T trypsin. (This result differs slightly from those of previous experiments in which the S195T variant was reported to have activity [Corey and Craik 1992]. This difference may be resolved by considering differences in experimental design. In the previous study, enzyme and substrate assay mixtures were incubated at 25°C until at least one turnover occurred. In the case of the S195T, an assay time >8 h was required to observe one turnover. Results obtained from such a limited data set over such a large time scale are more likely to be susceptible to error. To resolve any potential conflicts, identical experiments with S195T Tn were carried out in parallel for this study.)

catalyzed rate was ≤ 1.5 times that of the background rate, suggesting that the enzyme contributed little to substrate hydrolysis. Considering that little activity was detected with the highly activated SbzI ester substrate, it was not surprising that no activity was observed when the enzyme catalyzed hydrolysis of the less activated pNA, AMC, and peptide substrates were assayed over similar time frames (Fig. 2).

C42A/C58V/S195T Tn and C42A/C58A/S195T variants

Given the relative positioning of Ser 195 in WT trypsin, one would expect that in order for the hydroxyl group of threonine at position 195 to be used as a nucleophile by the enzyme, it would have to donate a hydrogen bond to His 57 in a manner similar to Ser 195 in the native protein. Modeling studies carried out for this investigation suggested that the reason for inactivity observed with the S195T single variant is that the Cys 42–Cys 58 disulfide bridge precludes this interaction by disallowing rotation about the β -carbon of the substituting threonine, thereby preventing its hydroxyl group from donating a hydrogen bond to His 57. If this were indeed the case, one would expect that if the bridge were not present, Thr 195 would be able to adopt a conformation that would render the enzyme active. To test this hypothesis, the

C42A/C58V/S195T (AVT-Tn) triple variant, which removes the barrier to optimal positioning of the hydroxyl group of Thr 195, was designed and expressed recombinantly. This triple mutant had observable hydrolytic activity on all three substrates. The k_{cat}/K_M value obtained for the AVT-Tn triple variant was ~ 600 -fold greater than that of the S195T single variant against the highly activated SbzI ester substrate (k_{cat} increased by >100 -fold), supporting the assertion that removal of the disulfide bridge formed by Cys 42 and Cys 58 allows the hydroxyl group of threonine at position 195 to substitute and function in a role similar to that of the hydroxyl group of Ser 195 in the native enzyme (see Table 1). More importantly, the AVT-Tn triple variant displayed significant hydrolytic activity toward the less activated substrates Z-GPR-pNA and Z-GPR-AMC (Fig. 2). When the activity of the S195T single variant toward these substrates was measured in similar time frames, hydrolysis above background levels was not observed. Therefore, one may conclude that the activity observed is not an artifact of the experimental design, but is, in fact, due to the activity of the modified enzyme.

The disulfide bridge formed by residues 42 and 58 forms the base of the S1' subsite in the enzyme for binding of the P1' residue of substrates. The C42A and C58V substitutions were initially selected to introduce hydrophobic character in the S1' pocket and minimize any loss of structural stability that might be incurred by removal of the disulfide bridge. However, such a substitution would also decrease the accessible van der Waals volume of the S1' pocket by 19 \AA^3 (Zamyatnin 1972). A decrease in accessible volume in this region would potentially hinder substrate binding and/or recognition by the modified enzyme by impeding association of P1' residues with the S1' subsite. To explore this possibility, a second triple variant in which the cysteines that constituted the disulfide bond were replaced with alanine residues (C42A/C58A/S195T, AAT-Tn) was expressed, characterized, and compared to the AVT-Tn variant. As with AVT-Tn, the AAT-Tn variant exhibited hydrolytic

activity toward the SbzI, pNA, and AMC substrates. The activity, however, was up to 40-fold greater in k_{cat}/K_M for all substrates compared to AVT-Tn (see Table 1). The AAT-Tn variant was 20–40-fold more active in k_{cat}/K_M than the AVT-Tn variant toward the pNA and AMC substrates (see Table 1), with the difference in activity due primarily to an increase in k_{cat} . The K_M values for the AVT-Tn and AAT-Tn variants for the pNA substrate are nearly identical, while the K_M of the AMC substrate for AAT-Tn is ~ 1.5 times that of the AVT-Tn.

The alterations in this region were made to amino acids that are not directly involved in the chemistry. That the K_M values either remain the same or increase suggests that by decreasing the van der Waals volume of the S1' pocket relative to the AVT-Tn, the nucleophilicity of the hydroxyl group of Thr 195 in AAT-Tn is enhanced. This hypothesis is supported further when one considers the relative spatial positioning of His 57, Val/Ala 58, and Thr 195. In addition to forming the base of the S1' pocket, the Cys 42–Cys 58 disulfide bridge may also optimize the positioning of His 57 in the catalytic triad. With an alanine at position 58, the decreased volume in the S1' pocket provides greater accessible volume for rotation of the methyl group of Thr 195. This increase in accessible space for the methyl group would allow its hydroxyl group to hydrogen bond more efficiently with the imidazole ring of His 57 thereby increasing its nucleophilicity. However, if Val is used to replace Cys 58, the accessible volume is decreased by 38 \AA^3 relative to the Ala substitution. This decrease in volume consequently decreases the rotational freedom of the methyl and hydroxyl groups about the β -carbon of Thr 195, resulting in even less optimal interaction between the imidazole ring of His 57 and the hydroxyl group of Thr 195.

C42A/C58V Tn (AV-Tn) and C42A/C58A Tn (AA-Tn) variants

Because of its relative positioning and strict conservation among trypsin-fold serine proteases, the C42–C58 disulfide

Table 1. Activity of wild-type trypsin and variants

Variant	Substrate								
	Z-GPR-SbzI			Z-GPR-pNA			Z-GPR-AMC		
	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($\mu\text{M}^{-1} \text{min}^{-1}$)	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($\mu\text{M}^{-1} \text{min}^{-1}$)	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($\mu\text{M}^{-1} \text{min}^{-1}$)
Wild type	715.5	16.5	43.4	1.2×10^4	46.5	258.06	2646	12.4	220
S195T	0.023	2.85	0.008	—	—	—	0.011	21	0.00052
C42A/C58A	5.0	32.8	0.15	7.4	44.6	0.166	4752	58.0	81.9
C42A/C58V	7.0	10.8	0.65	656.4	29.5	22.2	180	74.7	2.41
C42A/C58A/S195T	121.5	1.4	89	76.	217.7	0.351	0.69	57.4	0.012
C42A/C58V/S195T	3.5	0.73	4.79	2.0	214	0.009	0.021	40.4	0.00052

bridge is thought to have dual roles: (1) to form the S1' binding pocket and (2) to provide structural stability and minimize mobility of His 57 for optimal positioning in the catalytic triad. To examine the contribution of the Cys 42–Cys 58 disulfide bond in maintaining catalytic activity and efficiency, the substitutions for the cysteines used in the triple variants were made while retaining a serine at position 195. Additionally, the substitutions provided a means to discern the catalytic contribution due to the Thr mutation and the contribution of the disulfide removal in the decrease from WT activity.

The removal of the disulfide bond with the retention of serine at position 195 reduced activity up to three orders of magnitude compared to wild type (see Table 1; Fig. 2). In a direct comparison of the activities of C42A/C58V Tn (AV-Tn) to AVT-Tn, one observes that for both the AMC and pNA substrates, the triple variant is less active than the double variant as one might expect from the substitution of Ser 195 with threonine. A similar observation is made when comparing the activities of the C42A/C58A Tn (AA-Tn) and AAT-Tn variants. As seen in the AV and AVT comparison, the activity of the double variant is greater than that of the triple variant when assayed against the AMC substrate. However, the AA-Tn and AAT-Tn variants exhibit similar activity toward Z-GPR-pNA. In fact, the AAT-Tn variant is approximately twofold greater in $k_{\text{cat}}/K_{\text{M}}$ compared to the AA-Tn double variant. This difference may be at least partially explained by considering the kinetic parameters individually.

Compared to wild-type trypsin, the difference in catalytic efficiency is primarily due to a decrease in k_{cat} for the double variant; the K_{M} values are statistically identical for both the variant and the wild-type enzymes. For trypsin, the K_{M} value may be taken as a binding constant where $k_{-1} \ll k_2$. Therefore, one may conclude that the AA-Tn double variant binds the pNA substrate as well as wild type but the acylation rate has been compromised. One basis for such a compromise could be an increase in entropy of the region formerly occupied by the disulfide bridge. By removing the disulfide bridge, the spatial position of His 57 side chain will be less “fixed” relative to the other members of the catalytic triad potentially increasing the mobility of this residue. The entropic gain incurred by removal the Cys 42–Cys 58 disulfide bridge is ~ 2.23 kcal/mol at 25°C (Thornton 1981). Increased mobility of the imidazole ring would reduce the interaction between His 57 and Ser 195, and consequently decrease the nucleophilicity of the hydroxyl group of Ser 195. If this were indeed the case, one would expect that if the flexibility were decreased, the nucleophilicity of the hydroxyl group and catalytic turnover of the variant would be increased.

By substituting a Val residue at position 58 and an alanine at position 42, the estimated hydrophobic effect

is increased by 1.2 kcal/mol (Karplus 1997) relative to alanines at both positions. One would expect the increased hydrophobicity to provide additional structural stability in the S1' pocket and further stabilize the position of His 57. Consistent with this hypothesis, the k_{cat} value for hydrolysis of the pNA substrate by the AV-Tn double variant is increased 90-fold relative to the AA-Tn double variant, while the K_{M} value decreases slightly. This tighter binding of the pNA substrate may be a consequence of hydrophobic interactions between the phenyl moiety of the pNA leaving group and the valine in the modified S1' site on the protein. The increase in k_{cat} for the AAT-Tn variant may be explained similarly. In this variant, a rotation about the β -carbon of the substituted threonine would place the methyl group of Thr 195 in the region of the S1' site as well. The moderate increase in hydrophobicity of this region by the presence of the methyl group of Thr 195 for the AAT variant may help stabilize this region in a manner similar to the Val substitution of the AV-Tn variant. However, its location would be less buried relative to the AA-Tn variant, and could interfere with substrate binding resulting in the fivefold increase in K_{M} .

Interestingly, the observations noted from comparison of AA-Tn and AAT-Tn are not apparent when comparing AV-Tn and AVT-Tn. In fact, a 10^3 -fold decrease in $k_{\text{cat}}/K_{\text{M}}$ is observed for hydrolysis of Z-GPR-pNA by AVT-Tn relative to hydrolysis by AV-Tn. The decrease in $k_{\text{cat}}/K_{\text{M}}$ results from decreases in both the individual k_{cat} and K_{M} values, suggesting that the substitution of valine at position 58 and alanine at position 42 in the presence of Thr 195 compromises both substrate binding and catalytic turnover. A possible explanation is one that similarly describes the results in comparing the AA-Tn and AAT Tn. With AV-Tn, the valine at position 58 provides additional hydrophobicity that increases catalytic turnover by stabilizing the position of His 57 to optimize its interaction with Ser 195. However, if Ser 195 is substituted with threonine, once rotation about the β -carbon occurs, the methyl group of threonine will be located in this region as well. The bulkier valine in the S1' site and the additional hydrophobic character supplied by the methyl group may stabilize the positioning of His 57 in a similar manner, but consequently introduce greater steric constraints for the binding of putative substrates.

Upon examining the activities of each of the double and triple variants toward each of the substrates used, certain trends appear (Fig. 2). The leaving groups of the tripeptides used in this study occupy the P1' position in the substrate and would occupy the S1' binding site when associated with the protease. The pNA and AMC leaving groups differ in size, with the AMC group being the larger of the two. With the exception of the AA-Tn double variant, the trend that becomes apparent is that the

activity toward substrates with a given leaving group decreases with increasing van der Waals volume of the S1' site. This trend suggests that the amino acid constituency of this region of the protein can influence catalytic efficiency toward certain substrates. When one considers wild-type trypsin, there is no significant difference in k_{cat}/K_M toward any of the substrates that are considered here. This is likely, because the disulfide bridge formed between cysteines 42 and 58 is positioned to stabilize the S1' binding site, and forms a relatively flat surface at the base of the S1' pocket. In the cocrystal structure of trypsin with the macromolecular inhibitor ecotin, methionine 85 of ecotin occupies the equivalent of the P1' position of peptide substrates and interacts with this region. Once the disulfide bridge is removed, the resulting increase in entropy will render the S1' binding site less defined. In the case of AA-Tn, the larger hydrophobic groups in the P1' position of the substrate may stabilize the S1' site by hydrophobic interaction. This stabilization could account for the increased k_{cat}/K_M for this variant and the AMC substrates. However, if the S1' binding site is occupied by larger and bulkier hydrophobic amino acid side chains, as would be the case with the AV-Tn variant, stabilization might be increased, but the increased van der Waals volume in the P1' the S1' site could impede enzyme–substrate interaction. This impedence may lead to decreased catalytic efficiency of the enzyme by compromising the positioning of His 57, making it less susceptible to nucleophilic attack by the hydroxyl group of Ser 195. Substitution of Ser 195 with threonine may increase this negative steric interaction more. By rotating about the β -carbon, the methyl group of threonine would increase the van der Waals volume in the S1' region by 20 Å³ relative to the AV-Tn variant alone. This increased volume would likely further impede the association of amino acids into this subsite, accounting for the lower activity observed with all variants in which Cys 58 was substituted with Val.

Substrate specificity profile and thermal stability

Considering that multiple substitutions were made simultaneously for the variants, there was a possibility that the lower activity observed with the variants relative to wild type may have been the result of altered substrate specificity, i.e., the substrates typically used for wild-type trypsin may not have been the best substrates for the trypsin variants. To ascertain whether the substitutions that were introduced affected the substrate specificity profile of the variants, a positional scanning synthetic combinatorial library (Harris et al. 2000) was used to determine the amino acid preferences for positions P1–P4. The results from these experiments illustrated that the

specificity of each of the variants for residues in the P1–P4 positions mirrored that of wild-type trypsin (Fig. 3). Each of the variants exhibited stronger preference for Arg than Lys at the P1 position. No single amino acid was greatly preferred at the P2 position. However, the variants and wild-type trypsin discriminate against the presence of basic, branched, and aromatic amino acids at P2 in the presence of Arg at P1. The only amino acid discriminated against at P3 is proline, and at P4, no particular amino acid is preferred. Therefore, the decrease in activity relative to wild-type Tn may not be attributed to altered substrate preference for amino acids in the P1–P4 positions.

One role that disulfide bonds play in proteins is to provide structural stability by reducing the conformational entropy of the denatured state (Betz 1993). The Cys 42–Cys 58 disulfide bridge is one of six found in wild-type trypsin. Studies in which the C191–C220 bridge near the P1 pocket was removed determined that the bond was not essential for the function, structure, or stability of trypsin (Wang et al. 1997). However, given the proximity of the C42–C58 bridge to the active site, and that this disulfide bridge is nearly absolutely conserved among trypsin-fold serine proteases (Rawlings et al. 2004), one would expect its removal to have a more dramatic effect, particularly in stability and structure of the active site. To investigate the effect that removing the Cys 42–Cys 58 disulfide bond had on the structural stability of trypsin, the temperature dependence of hydrolysis of Z-GPR-pNA catalyzed by WT, AV-Tn, and AVT-Tn variants was measured as described in Materials and Methods. For wild-type trypsin, the activity of the enzyme at each time point and temperature was statistically identical. However, as might be expected, the removal of the disulfide bridge in the presence of Ser 195 resulted in a protein with decreased thermostability. When the activity of the AV-Tn variant enzyme was measured in 10-min intervals over 40 min, a linear decrease in activity was observed with the slope increasing with increasing temperature (Fig. 4). One would expect that if the β -methyl group of Thr 195 in the AVT-Tn variant interacts favorably with Val 58, the hydrophobic interaction might increase the stability of the protein relative to AV-Tn. When the AVT-Tn was assayed, its behavior was more similar to that of the wild-type enzyme than the AV-Tn variant, suggesting that the Val 58/Thr 195 combination sufficiently compensates for the loss of the disulfide bridge. The activities of the AVT-Tn measured at different time points and temperatures were statistically identical, suggesting that the presence of the methyl group of Thr 195 significantly stabilizes the enzyme relative to its Ser 195 counterpart, likely by hydrophobic interaction with Val 58. Furthermore, the fact that the activity does not decrease suggests that, at least over this time period, the AVT-Tn variant is as stable as wild type.

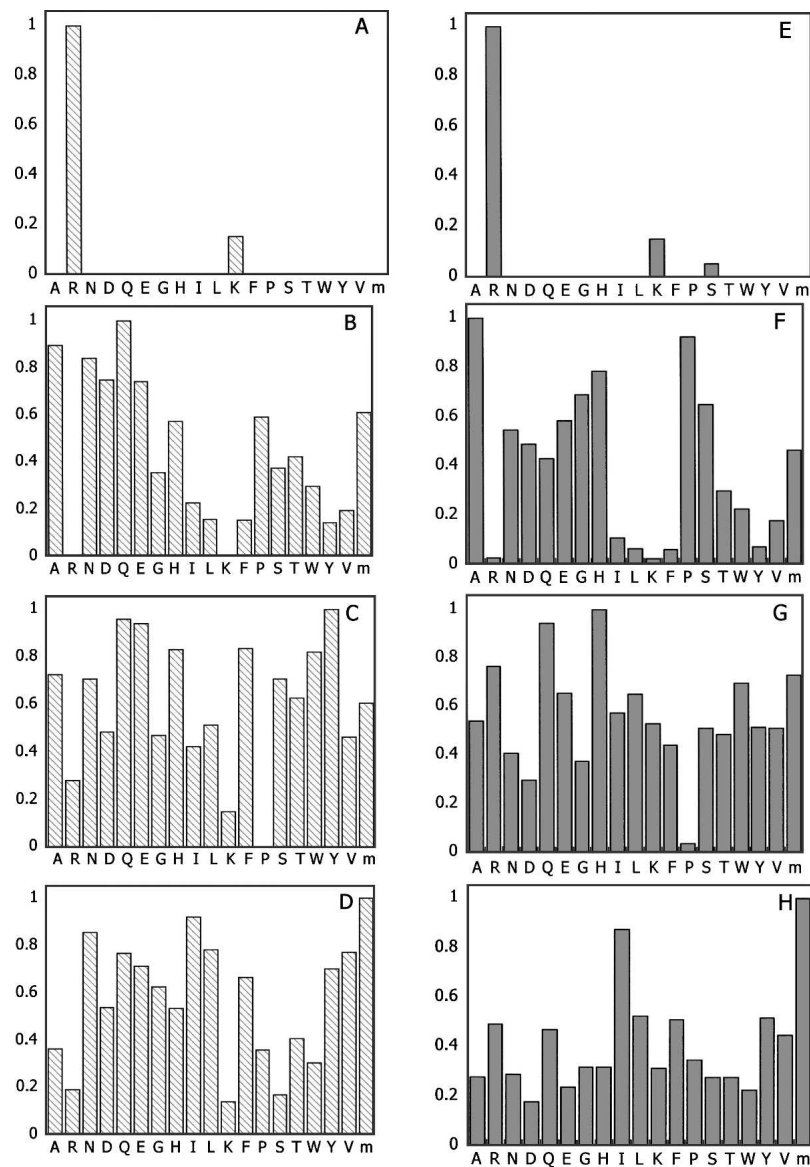


Figure 3. Substrate specificity profiles. The amino acid preferences for substrate positions P1–P4 were determined for each of the variants and wild-type trypsin as described in Materials and Methods. Panels A–D correspond to preferences for positions P1–P4, respectively, for wild-type trypsin. Panels E–H correspond to preferences for positions P1–P4, respectively, for the AVT-Tn triple variant. The AA-Tn, AV-Tn, and AAT-Tn variants yielded similar results.

We have demonstrated that the serine protease trypsin can be converted to a functional threonine protease by removal of a single disulfide bond. The activity of the newly fashioned threonine protease is significantly lower than that of wild-type trypsin, but its activity is comparable to serine proteases that play functional roles in biology such as KSHV protease (Pray et al. 1999). Furthermore, the activity exhibited by the engineered trypsin threonine protease is comparable to that of a naturally occurring threonine protease, the proteasome. The proteasome and the serine proteases share no structural similarities or active

site residues outside of the nucleophile. Nonetheless, it is interesting to note that the activities of engineered and naturally occurring proteases that use the hydroxyl group of threonine as the nucleophile in peptide cleavage exhibit similar activities.

Noting that threonine proteases and serine proteases that exhibit activities similar to that of the engineered threonine protease are found in nature and they serve important roles in their respective environment, the question arises as to why threonine proteases are not observed as frequently in nature. At least in the case of

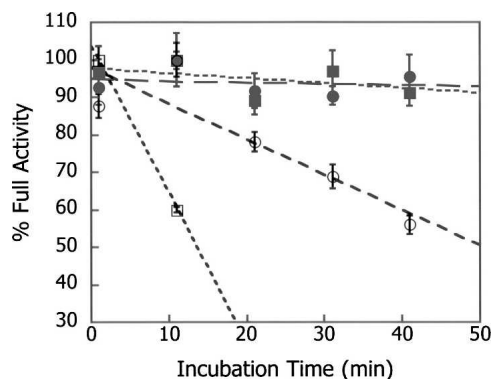


Figure 4. Temperature dependence of hydrolytic activity. The thermostability of AVT-Tn and AV-Tn variants was determined by a modified kinetic assay as described in Materials and Methods. The rate of Z-GPR-pNA hydrolysis catalyzed by WT at 37°C (◆), C42A/C58V/S195T at 25°C (●) and 37°C (■), and C42A/C58V at 25°C (○) and 37°C (□).

the trypsin-fold, it seems that the answer may primarily be a structural one. Trypsin is considered to be a digestive protease with broad specificity. Its primary requirement for a viable substrate is a basic side chain at the P1 position. Serving in such a capacity, any higher degree of substrate selectivity would decrease its biological function. The disulfide bridge formed by Cys 42 and Cys 58 is positioned such that it both (1) optimizes the positioning of the essential His of the catalytic triad and (2) forms a relatively flat surface for the S1' pocket. The disulfide bond helps to position the His 57 side chain such that it accepts a hydrogen bond from Ser 195 and increases its nucleophilicity. The flat surface that it forms in the S1' pocket allows any residue to occupy the P1' position. If the disulfide under consideration is removed, both of these structural features are altered, resulting in a less active enzyme with less digestive (more selective) character.

Materials and methods

Mutagenesis, expression, and purification of trypsin variants

Trypsin variants C42A/C58A, C42A/C58V, C42A/C58V/S195T, and C42A/C58V/S195T were constructed using PCR mutagenesis and the gene for rat anionic trypsin. Oligonucleotides used to introduce the desired mutations were synthesized on a PerSeptive Biosystems Expedite 8909 DNA synthesizer. The nucleic acid sequence of each oligonucleotide was complementary to the sense strand of the trypsin gene except at the point of mutagenesis. The altered gene was then subcloned into the expression vector pPICZ α -A (Invitrogen). To verify that only the desired substitutions were introduced, each mutant was sequenced by the UCSF DNA sequencing facility. The resultant plasmids were transformed into X-33 *P. pastoris* cells for expression. Single colonies were isolated and cultured overnight at 30°C in 10 mL of YPD/Zeocin media (1% yeast extract, 2%

peptone, 2% dextrose, 100 μ g/mL Zeocin). The 10-mL overnight culture was then used to inoculate 1 L of BMGY (1% yeast extract, 2% peptone, 1% glycerol) buffered at pH 6.0 with 0.1 M K_2HPO_4/KH_2PO_4 and incubated at 30°C while shaking for 24–36 h. Expression was induced by the addition of methanol to 0.5% every 24 h for an additional 72 h. The cells were removed by centrifugation, and the supernatant was brought to 3 M NaCl, 20 mM MES (pH 6.0) by the addition of crystalline NaCl and 0.5 M MES (pH 6). The high salt supernatant was loaded onto a Phenyl-Sepharose CL-4B column (Pharmacia) that was pre-equilibrated with 3 M NaCl, 20 mM MES (pH 6.0). The column was then washed with the same buffer until the flow-through was colorless. The trypsin variants were eluted using a negative NaCl step gradient in 0.5 M increments in 20 mM MES (pH 6.0), collecting 50–100 mL fractions at each step. Fractions that contained the trypsin variant were identified by Coomassie staining of PAGE gels. The positive fractions were pooled and concentrated to ≤ 10 mL via diafiltration using an Amicon diaflow apparatus (Millipore). The concentrated sample was then dialyzed into 10 mM Tris-Cl (pH 7.4) at 4°C. The dialyzed sample was then loaded onto a DEAE column that had been pre-equilibrated in the same buffer. The trypsin variant was eluted with a 0.0–0.5 M linear gradient of NaCl in 10 mM Tris-Cl (pH 7.4). Fractions containing trypsin were identified by either PAGE or Western blot, concentrated to ≤ 10 mL, dialyzed into 10 mM citrate (pH 3.0), and stored at 4°C until further use.

Kinetic characterization of variants

The activity of the trypsin variants was assayed spectrophotometrically by monitoring cleavage of the tripeptide N- α -benzyloxycarbonyl-L-glycylprolylarginine with thiobenzyl ester (Z-GPR-SBzl) and *p*-nitroanilide (Z-GPR-pNA) leaving groups at 324 and 405 nM, respectively. Fluoremetric assays were conducted using the same tripeptide sequence with a 7-amido-4-methylcoumarin leaving group (Z-GPR-AMC, $\lambda_{ex} = 380$, $\lambda_{em} = 460$). All assays were conducted in 50 mM HEPES, 100 mM NaCl, 20 mM $CaCl_2$ (pH 8.0). Spectrophotometric experiments were carried out at 25°C using 0.01–1.3 μ M trypsin variant, 5–50 μ M Z-GPR-Sbzl (Synpep), and 250 μ M 4,4-dithiodipyridine (Sigma-Aldrich) or 10–600 μ M Z-GPR-pNA (Sigma-Aldrich) for 5–30 min in a final volume of 500 μ L. Data was collected on a Bio-Tek Uvikon 923 double beam UV/Vis spectrophotometer. Fluoremetric assays were carried out similarly at 25°C using 0.1–1.0 μ M trypsin variant and 10–200 μ M Z-GPR-AMC (Sigma-Aldrich) in a final volume of 1.0 mL. Data was collected on an ISA SPEX Fluorolog-3 fluorimeter. The kinetic parameters k_{cat} , K_M , and k_{cat}/K_M were determined by fitting the data to the Michaelis-Menten equation using the curve-fitting program Kaleidagraph (Synergy Software).

Thermostability measurements

A modified activity kinetic assay was employed in order to monitor the rate at which trypsin variants lost activity as a function of temperature. Enzyme solution was diluted with 2 \times incubation buffer, so that the diluted solution contained 10 mM glycine (pH 3), 20 mM $CaCl_2$, and NaCl to bring ionic strength of the incubation buffer equal to that of the assay buffer. Sufficient enzyme solution was used such that five activity measurements could be made from the contents, one for each time point. Once mixed with incubation buffer, the enzyme

solutions were placed in a water bath (for 25°C and 37°C assays) or in ice water (for 4°C assay). This was taken to be the zero time point. The solutions were allowed to come to temperature for ~1 min before the first activity measurement was made. For each run, a 5 μ L aliquot of enzyme solution was taken and added to the wells of a 96-well plate containing assay buffer (243 μ L) and 144 μ M Z-GPR-pNA for a final volume of 250 μ L/well. The plate reader, in kinetic mode, then monitored absorbance/time at 410 nm for 90 sec. This procedure was repeated every 10 min for a total of 40 min. This, in turn, was done for each variant, and for each temperature. Data were normalized to percent relative activity with the highest activity for a given run taken to be 100%.

Substrate specificity profiles

The activity of the trypsin variants was assayed using positional scanning synthetic combinatorial substrate library in a 96-well plate. Each well contained 0.01–1.0 μ M Tn variant and 361 substrates per well at a concentration of 250 nM/substrate in a final volume of 100 μ L. For determination of P1 selectivity, each well had its P1 position fixed with one of 19 amino acids (all standard amino acids were used except cysteine and methionine; norleucine was added to approximate methionine) while all other positions were randomized. For determination of selectivity for positions P2–P4, each well had substrates with two positions fixed; P1 was arginine in all samples while P2, P3, or P4 were fixed in one well with one of 19 amino acids. Substrate cleavage was measured by monitoring the release of the fluorometric AAC group ($\lambda_{\text{ex}} = 380$, $\lambda_{\text{em}} = 460$) at 25°C for 90 min in a Molecular Devices SpectrMax Genini microtiter plate reader.

Acknowledgments

This work was supported by funds from grants by the National Cancer Institute (NIH CA 72006) and the National Center on Minority Health and Health Disparities (5P20-MD000262), and a fellowship from the National Science Foundation. We thank Mikael Perakyla and acknowledge the late Professor Peter Kollman for initial discussions regarding the design of the variants. We also thank Sandra W. Ruggles for critically reviewing the manuscript.

References

Baumeister, W. and Lupas, A. 1997. The proteasome. *Curr. Opin. Struct. Biol.* **7**: 273–278.
 Betz, S.F. 1993. Disulfide bonds and the stability of globular proteins. *Protein Sci.* **2**: 1551–1558.
 Betzel, C., Gourinath, S., Kumar, P., Kaur, P., Perbandt, M., Eschenburg, S., and Singh, T.P. 2001. Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution. *Biochemistry* **40**: 3080–3088.
 Bode, W., Turk, D., and Karshikov, A. 1992. The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human

α -thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure–function relationships. *Protein Sci.* **1**: 426–471.
 Brannigan, J.A., Dodson, G., Duggleby, H.J., Moody, P.C., Smith, J.L., Tomchick, D.R., and Murzin, A.G. 1995. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**: 416–419.
 Carter, P. and Wells, J.A. 1987. Engineering enzyme specificity by “substrate-assisted catalysis.” *Science* **237**: 394–399.
 Corey, D.R. and Craik, C.S. 1992. An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc.* **114**: 1784–1790.
 Davie, E.W. 2003. A brief historical review of the waterfall/cascade of blood coagulation. *J. Biol. Chem.* **278**: 50819–50832.
 DeLano, W.L. 2002. The PyMOL molecular graphics system. San Carlos, CA.
 Ding, X., Rasmussen, B.F., Petsko, G.A., and Ringe, D. 1994. Direct structural observation of an acyl-enzyme intermediate in the hydrolysis of an ester substrate by elastase. *Biochemistry* **33**: 9285–9293.
 Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.
 Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A., and Craik, C.S. 2000. Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc. Natl. Acad. Sci.* **97**: 7754–7759.
 Holyoak, T., Wilson, M.A., Fenn, T.D., Kettner, C.A., Petsko, G.A., Fuller, R.S., and Ringe, D. 2003. 2.4 Å resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg boronic acid inhibitor. *Biochemistry* **42**: 6709–6718.
 Karplus, P.A. 1997. Hydrophobicity regained. *Protein Sci.* **6**: 1302–1307.
 Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. 1995. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* **268**: 533–539.
 McCaldon, P. and Argos, P. 1988. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **4**: 99–122.
 Nakayama, K. 1997. Furin: A mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem. J.* **327**: 625–635.
 Perona, J.J., Tsu, C.A., McGrath, M.E., Craik, C.S., and Fletterick, R.J. 1993a. Relocating a negative charge in the binding pocket of trypsin. *J. Mol. Biol.* **230**: 934–949.
 Perona, J.J., Tsu, C.A., Craik, C.S., and Fletterick, R.J. 1993b. Crystal structures of rat anionic trypsin complexed with the protein inhibitors APPI and BPTI. *J. Mol. Biol.* **230**: 919–933.
 Pray, T.R., Nomura, A.M., Pennington, M.W., and Craik, C.S. 1999. Auto-inactivation by cleavage within the dimer interface of Kaposi’s sarcoma-associated herpesvirus protease. *J. Mol. Biol.* **289**: 197–203.
 Rawlings, N.D., Tolle, D.P., and Barrett, A.J. 2004. MEROPS: The peptidase database. *Nucleic Acids Res.* **32**: D160–D164.
 Siezen, R.J. and Leunissen, J.A. 1997. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Sci.* **6**: 501–523.
 Steinmetz, A.C., Demuth, H.U., and Ringe, D. 1994. Inactivation of subtilisin Carlsberg by N-((tert-butoxycarbonyl)alanylprolylphenylalanyl)-O-benzoylhydroxyl-amine: Formation of a covalent enzyme-inhibitor linkage in the form of a carbamate derivative. *Biochemistry* **33**: 10535–10544.
 Thornton, J.M. 1981. Disulphide bridges in globular proteins. *J. Mol. Biol.* **151**: 261–287.
 Wang, E.C., Hung, S.H., Cahoon, M., and Hedstrom, L. 1997. The role of the Cys191–Cys220 disulfide bond in trypsin: New targets for engineering substrate specificity. *Protein Eng.* **10**: 405–411.
 Waugh, S.M., Harris, J.L., Fletterick, R., and Craik, C.S. 2000. The structure of the pro-apoptotic protease granzyme B reveals the molecular determinants of its specificity. *Nat. Struct. Biol.* **7**: 762–765.
 Zamyatnin, A.A. 1972. Protein volume in solution. *Prog. Biophys. Mol. Biol.* **24**: 107–123.