Protein Science (1999), *8*:253–258. Cambridge University Press. Printed in the USA.
Copyright © 1999 The Protein Society

FOR THE RECORD

Comparison of anionic and cationic trypsinogens: The anionic activation domain is more flexible in solution and differs in its mode of BPTI binding in the crystal structure

ANNETTE PASTERNAK,[1] DAGMAR RINGE,[1,2] AND LIZBETH HEDSTROM[2]

[1]Department of Chemistry, Brandeis University, Waltham, Massachusetts 02454
[2]Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02454

(RECEIVED June 5, 1998; ACCEPTED October 7, 1998)

Abstract: Unlike bovine cationic trypsin, rat anionic trypsin retains activity at high pH. This alkaline stability has been attributed to stabilization of the salt bridge between the N-terminal Ile16 and Asp194 by the surface negative charge (Soman K, Yang A-S, Honig B, Fletterick R., 1989, *Biochemistry 28*:9918–9926). The formation of this salt bridge controls the conformation of the activation domain in trypsin. In this work we probe the structure of rat trypsinogen to determine the effects of the surface negative charge on the activation domain in the absence of the Ile16–Asp194 salt bridge. We determined the crystal structures of the rat trypsin-BPTI complex and the rat trypsinogen-BPTI complex at 1.8 and 2.2 Å, respectively. The BPTI complex of rat trypsinogen resembles that of rat trypsin. Surprisingly, the side chain of Ile16 is found in a similar position in both the rat trypsin and trypsinogen complexes, although it is not the N-terminal residue and cannot form the salt bridge in trypsinogen. The resulting position of the activation peptide alters the conformation of the adjacent autolysis loop (residues 142–153). While bovine trypsinogen and trypsin have similar CD spectra, the CD spectrum of rat trypsinogen has only 60% of the intensity of rat trypsin. This lower intensity most likely results from increased flexibility around two conserved tryptophans, which are adjacent to the activation domain. The NMR spectrum of rat trypsinogen contains high field methyl signals as observed in bovine trypsinogen. It is concluded that the activation domain of rat trypsinogen is more flexible than that of bovine trypsinogen, but does not extend further into the protein core.

Keywords: disorder-order; natively unfolded; serine protease; X-ray crystallography; zymogen

Reprint requests to: Lizbeth Hedstrom, Department of Biochemistry, Brandeis University, Mail Stop 009, Waltham, Massachusetts 02454; e-mail: hedstrom@binah.cc.brandeis.edu.

Abbreviations: R-Tg, rat trypsinogen; R-Tp, rat trypsin; B-Tg, bovine trypsinogen; B-Tp, bovine trypsin; UV, ultraviolet; CD, circular dichroism; NMR, nuclear magnetic resonance; ppm, parts per million; BPTI, bovine pancreatic trypsin inhibitor; PSTI, porcine pancreatic secretory trypsin inhibitor; PDB, Protein Data Bank; RMSD, root-mean-square deviation.

Many important regulatory proteins are disordered in the absence of their effector molecules. This disordered state is believed to be crucial for proper function of these proteins. The manner in which the sequences of these proteins dictates their undefined structures is not understood. One of the first disorder–order transitions described was the conversion of trypsinogen to trypsin. Approximately 15% of trypsinogen is disordered, which renders trypsinogen inactive (Fehlhammer et al., 1977). This disordered region is localized to three loop segments and the N-terminal activation peptide, together termed the activation domain (Fig. 1; Table 1). Trypsinogen is activated upon proteolytic removal of the activation peptide. The new N-terminus, Ile16, folds into a pocket and makes a salt bridge with Asp194. The resulting conformational change orders the activation domain, creating the substrate binding site. Alternatively, the activation domain of trypsinogen can be ordered by binding BPTI, which induces a conformation similar to that of trypsin (Bode et al., 1978). In this complex, the Ile16 pocket contains solvent molecules while the N-terminus is disordered to Gly18.

These seminal structural studies used cationic bovine trypsin (B-Tp) and trypsinogen (B-Tg). Anionic rat trypsin (R-Tp) has a very similar structure and catalytic constants to B-Tp, although their net charges differ by 12.5 units. The residues of the active sites of these two enzymes are conserved. However, the sequences of the activation domains are not conserved despite their crucial structural and regulatory functions (Table 1). Differences in the stability of the activation domains of B-Tp and R-Tp can be inferred from the pH dependence of activity. B-Tp is inactive at high pH (Keil, 1971). This loss of activity is attributed to deprotonation of Ile16, loss of the salt bridge, and the resulting collapse of the activation domain (Fersht, 1972). In contrast, R-Tp retains activity at high pH, which suggests that the pK_a of Ile16 is increased (Craik et al., 1987; Hedstrom et al., 1996). This increased pK_a is believed to result from the surface negative charge of R-Tp, which produces a greater negative electrostatic potential in the Ile16 pocket (Soman et al., 1989).

While R-Tp and B-Tp have similar catalytic properties, rat trypsinogen (R-Tg) appears to be less active than B-Tg (Pasternak

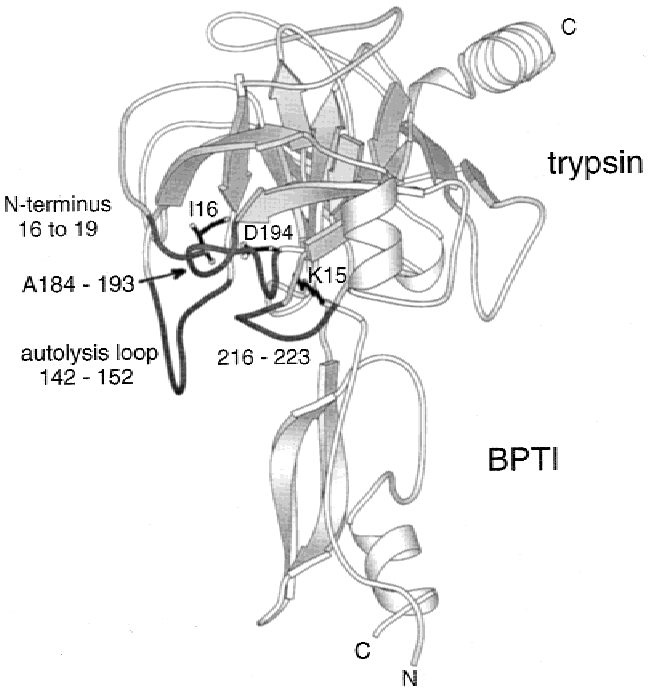

Fig. 1. Wild-type R-Tp complexed with BPTI. The activation domain segments are shaded in dark gray. N and C termini of both proteins are indicated. The side chain of the R-Tp N-terminus, Ile16, is shown in the Ile16 pocket engaged in a salt bridge with Asp194. The Lys15 side chain of BPTI is shown in the primary specificity pocket of R-Tp.

et al., 1998). This observation suggests that the structures of R-Tg and B-Tg may be different. We have determined the crystal structure of the R-Tg·BPTI complex. The structure reveals that BPTI binding induces a trypsin-like conformation in most of the R-Tg activation domain, as observed in the B-Tp·BPTI complex. However, unlike the B-Tp·BPTI complex, Ile16 occupies the Ile16 pocket and residues 142–153 (the autolysis loop) are mostly disordered. CD and NMR spectroscopy indicates that the activation domain of R-Tg is probably more flexible than B-Tg, but that this flexibility does not appear to extend into the protein core. The increased flexibility of R-Tg is consistent with the presumption that negatively charged residues favor disorder.

Table 1. *Sequences of the rat and bovine trypsinogen activation domains*[a]

8–19	Bovine	FPVDDDDK**I**VGG
	Rat	FPVDDDDK**I**VGG
142–152	Bovine	GNTKSSGTSYP
	Rat	GNTLSSGVNEP
183–193	Bovine	GYLEGGKDSCQG
	Rat	GFLEGGKDSCQG
216–223	Bovine	GSGCAQKN
	Rat	GYGCALPD

[a]Underlined residues denote sequence differences between the rat and bovine enzymes. Residue in bold denotes the trypsin N-terminal residue 16.

Results and discussion: *Overall structures of R-Tg·BPTI and R-Tp·BPTI:* We have determined the crystal structures of the BPTI complexes of wild-type R-Tp (PDB accession code 1tgi) and Ser195Ala R-Tg (PDB accession code 1tgj) to 1.8 and 2.2 Å resolution, respectively (Table 2). The Ser195Ala mutant was utilized to prevent autocatalytic conversion of R-Tg to R-Tp. Electron density for the main chain is observed for all residues in R-Tp, and all side chains except for Arg117. In R-Tg, no electron density is observed for residues at the N-terminus to Lys15, or for residues 144–151 of the autolysis loop. As in R-Tp, there is no electron density for the side chain of Arg117. In addition, the side chain of Glu186 in the R-Tg is disordered. There is no interpretable electron density for the C-terminal residues 57–58 of BPTI in either structure.

R-Tp undergoes structural adjustments upon BPTI binding that do not take place in B-Tp: The formation of the complex between BPTI and B-Tp occurs with relatively minor adjustments of inhibitor and enzyme (Huber & Bode, 1978). This perfect complementarity of B-Tp and BPTI yields a high affinity complex ($K_i = 10^{-13}$ M (Vincent & Lazdunski, 1972)). In contrast, R-Tp must rearrange to form the BPTI complex. Two loops, residues 183–188A and 222–224, move from their positions in the R-Tp·benzamidine complex (PDB accession code 1ane). These adjustments may account for the lower affinity of the R-Tp·BPTI complex ($K_i = 10^{-10}$ M (Hedstrom et al., 1996)). Thus, while the benzamidine complexes of R-Tp and B-Tp (PDB accession code 3ptb) are very similar (Sprang et al., 1987), significant differences are found in the BPTI complexes. Residues 183–188A and 222–224 are 0.9 and 0.5 Å, respectively, from the positions in the B-Tp·BPTI complex (PDB accession code 2ptc), while overall the proteases superimpose at the alpha carbons with RMSD of 0.39 Å.

R-Tg has a mostly trypsin-like conformation when bound to BPTI: B-Tg attains a conformation nearly identical to B-Tp upon binding BPTI (Bode et al., 1978). Likewise, the conformation of R-Tg bound to BPTI is similar to that of R-Tp. The RMSD of alpha carbons of the R-Tg·BPTI and R-Tp·BPTI complexes, with the exception of the N-terminus to Gly19 and the autolysis loop, is 0.16 Å. The N-terminus and the autolysis loop have very different conformations in R-Tp and R-Tg as discussed below. In addition, some differences are observed in the catalytic triad residues Ser195, His57, and Asp102, which can be attributed to the substitution of Ala for Ser195 in R-Tg. In the R-Tg complex, His57 shifts 0.3–0.4 Å away from Asp102.

Ile16 of R-Tg binds in the Ile16 pocket close to its position in R-Tp: Although B-Tg attains a trypsin-like conformation in complex with BPTI, the Ile16 pocket contains only solvent and the N-terminus is disordered to Gly18. In contrast, Ile16 is found in the Ile16 pocket of R-Tg·BPTI although it is not the N-terminus (Fig. 2). Electron density is visible in the simulated annealing omit map for the main chain of Lys15 as well as the residues Ile16 and Val17 (Fig. 2). The remainder of the N-terminus is disordered. The average temperature factor for the Ile16 side chain is 30 Å^2 and that for the Val17 side chain is 50 Å^2, compared to the overall average of 18 Å^2. Unlike the N-terminal Ile16 of R-Tp·BPTI, the Ile16 does not interact directly with Asp194 in R-Tg·BPTI. Instead, the N of Ile16 and the carboxylate of Asp194 are linked by

Table 2. *Data collection and refinement*

	Wild-type trypsin-BPTI	Ser195Ala trypsinogen-BPTI
Space group	$P3_221$	$P3_221$
Cell dimensions (Å)	$a = b = 92.72$ $c = 62.19$	$a = b = 92.66$ $c = 62.36$
Resolution (Å)	1.8	2.2
Reflections, observed	97,458	37,016
Reflections, unique	25,741	15,555
Completeness, overall (%)	88.9	97.3
Completeness, highest	80.3	90.5
Resolution shell (%)	(1.80–1.86 Å)	(2.20–2.28 Å)
Signal to noise ratio ($\langle I/\sigma(I)\rangle$)	11	10
R_{sym}[a] (%)	4.9	10.6
R-factor (%)[b]	17.7	19.7
R-free (%)[c]	21.0	24.0
Water molecules	227	106
Sulfate molecules	2	2
Restraints (RMS observed)		
Bond length (Å)	0.006	0.006
Bond angles (°)	1.32	1.34
Dihedral angles (°)	27.6	27.5
Improper angles (°)	0.67	0.69

[a] $R_{sym} = \Sigma|I_i - \langle I\rangle|/\Sigma I_i$, where I_i is the scaled intensity of the ith measurement and $\langle I\rangle$ is the mean intensity for that reflection.
[b] *R*-factor $= \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$ for all reflections.
[c] *R*-free $= \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$ for a control set of 10% of all reflections, randomly chosen.

water molecule 563. Both the visibility of Lys15 and the slightly different position of Ile16 indicate that R-Tg has not been converted to R-Tp (Fig. 2). SDS-PAGE experiments confirm that R-Tg is intact in the crystals (data not shown).

The autolysis loop adopts an unusual conformation in R-Tg·BPTI: The autolysis loop makes few contacts with the rest of the trypsinogen and trypsin, which may account for its apparent plasticity (Bennett & Huber, 1984). Disordered in B-Tg, the autolysis loop fails to become well ordered in B-Tg·BPTI (Bode et al., 1978) and adopts yet another conformation in the B-Tg·PSTI structure (Bolognesi et al., 1982). The autolysis loop is also disordered in the R-Tg·BPTI complex. Despite this disorder, it is clear that the autolysis loop is found in a different conformation in R-Tg·BPTI than in R-Tp·BPTI (Fig. 3). The presence of Lys15 prevents Gly142 and Asn143 from occupying a trypsin-like conformation in the wall of the Ile16 pocket in R-Tg. Instead, these residues move away from the N-terminus. Asn143 is not well ordered and the remainder of the loop is disordered with no observable electron density until residue 152. The side chain of Trp141 is in the identical position as in trypsin, but the carbonyl group of Trp141 is rotated, which causes Pro152 and Asp153 to shift. The repositioning of the autolysis loop in the R-Tg·BPTI complex disrupts the

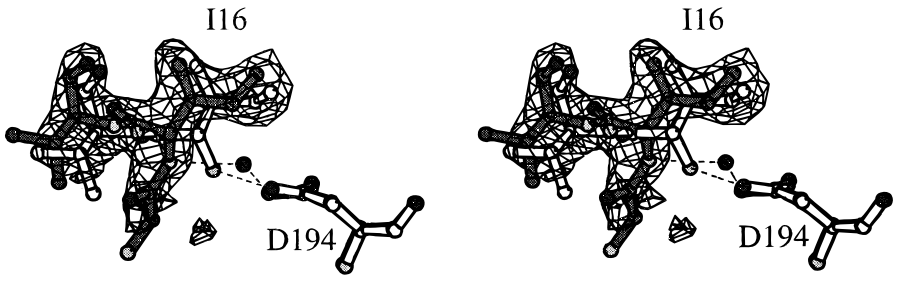

Fig. 2. Stereoview of the electron density of residues 15–17 of Ser195Ala R-Tg (gray bonds) and comparison to wild-type R-Tp (white bonds) in the BPTI complexes. Asp194 for the two complexes superimpose almost exactly. Ile16 is the N-terminal residue in trypsin; the amino group makes a salt bridge with the Asp194 carboxylate oxygen OD2 as represented by the dashed line. Although in R-Tg·BPTI Ile16 is not the N-terminus and the main-chain nitrogen is too far from Asp194 to make a hydrogen bond, the Ile16 occupies a position close to that in trypsin. In R-Tg·BPTI water molecule 563 bridges the main chain N of Ile16 and the 194 carboxylate OD2, indicated by dashed lines. The electron density map for the R-Tg·BPTI structure is a simulated annealing omit map with coefficients $F_o - F_c$ of residues 15–17 contoured at 2.3σ. Electron density for the main chain is seen to continue to residue 15. There is no interpretable electron density for the N-terminus beyond residue 15 or for the Lys15 side chain.

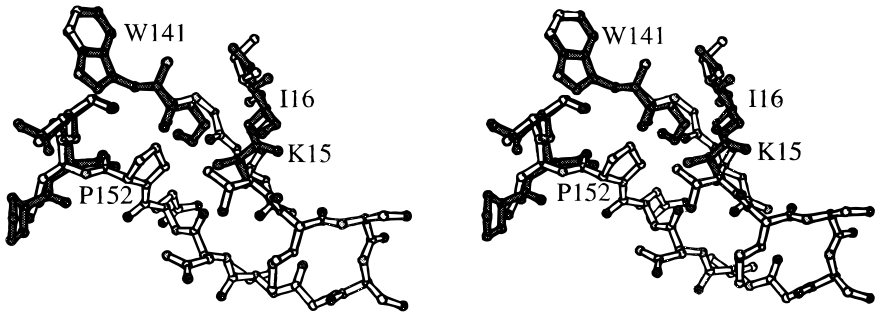

Fig. 3. Stereo comparison of the autolysis loops of R-Tg (gray) and R-Tp (white) in the BPTI complexes. The loop from Trp141–Asp153 as well as the N-terminal Ile16 of the trypsin structure are presented. Most of the residues of the loop are disordered in the trypsinogen structure. Thus, only 141–142 and 152–153 are shown, along with Ile16 and the main chain of Lys15, the side chain of which is disordered. The rest of the N-terminus of trypsinogen is disordered. The position of residue 15 in trypsinogen prevents the autolysis loop from having the same conformation as for trypsin, and the loop is forced to move away such that Pro152 is 7 Å away from its position in trypsin.

interaction between Arg17 of BPTI and Glu151 of R-Tg (Perona et al., 1993a). As a result, the BPTI Arg17 guanidinium group is turned away from its position in R-Tp·BPTI and interacts with the main-chain oxygen of His40.

The activation domain of unliganded R-Tg is more flexible: We measured near UV CD spectra to probe the unliganded structures of R-Tg and R-Tp (Fig. 4). Ellipticity values of near UV CD spectra reflect the environment around tryptophan and, to a lesser extent, tyrosine residues. Chromophores that are in a flexible part of a protein experience a uniform symmetrical environment and do not contribute to the CD signal. The CD spectra of B-Tg and B-Tp are nearly identical (Kerr et al., 1975), which suggests (1) the aromatic chromophores of the activation domain do not contribute significantly to the CD spectrum; and (2) the environments of the remaining chromophores are similar in B-Tp and B-Tg. The CD spectrum of R-Tp closely resembles the spectrum of B-Tp, with a wide trough from about 270–290 nm of ellipticity -8.4×10^{-3} deg cm^2 dmol^{-1} (Fig. 4) (Kerr et al., 1975), as expected given that all four tryptophan residues of B-Tp are also found in R-Tp. This observation confirms the structural similarity of R-Tp and B-Tp. In contrast, the CD intensity of the R-Tg spectrum is only about 60% of R-Tp (Fig. 4). This observation suggests that, in contrast to B-Tg and B-Tp, the environment of the tryptophan residues is different in R-Tg and R-Tp. Two tryptophan residues, Trp141 and Trp215, are adjacent to the activation domain (Fehlhammer et al., 1977). The less negative ellipticity of R-Tg suggests that the environment around Trp141 and/or Trp215 is more flexible. This conclusion is supported by fluorescence spectroscopy. The fluorescence of tryptophan is quenched in hydrophilic environments. The fluorescence spectrum of R-Tg is about 50% less than that of R-Tp (data not shown), which further suggests that one or more tryptophan residues are more exposed to solvent in R-Tg. Thus both the CD and fluorescence spectra suggest that Trp141 and/or Trp215 are more exposed to solvent. These observations suggest that the activation domain of R-Tg is more flexible than in B-Tg, and may even extend farther into the protein core.

The hydrophobic core of R-Tg is intact: The close proximity of an aromatic ring shifts the ^{1}H-NMR signal of a methyl group to high fields. Ring current shifts are extremely sensitive to the position of the aromatic ring and the methyl protons. Perkins and Wuthrich assigned the two highest field signals, -0.98 and -0.80 ppm, of the B-Tp ^{1}H-NMR spectrum to Ile73 and Val227 (Perkins & Wuthrich, 1980). These upfield signals arise from the proximity of Ile73 and Val227 to Trp141 and Trp215, respectively. Similar high field signals are also observed in the spectrum of B-Tg, as expected, since the structures of B-Tp and B-Tg are similar in this region. If Trp141 or Trp215 are disordered in R-Tg, then these residues would not be closely packed with Ile73 and Val227, and the upfield signals would be lost. However, upfield signals are clearly seen in the R-Tg spectrum (Fig. 5), which implies that Ile73 and Val227 are in close proximity to Trp141 and Trp215. This observation suggests that the hydrophobic core of R-Tg is intact.

These results suggest that the activation domain of R-Tg is more flexible than that of B-Tg, but that this flexibility does not extend farther into the protein core. This greater flexibility provides a

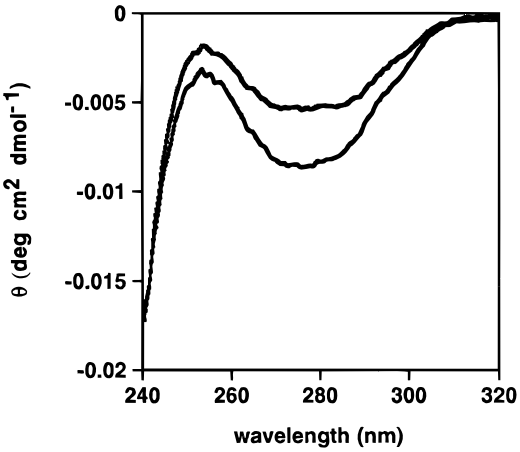

Fig. 4. Near UV circular dichroism spectra of R-Tg and R-Tp. The upper spectrum is that of Ser195Ala R-Tg, and the lower spectrum is Ser195Ala R-Tp. Each sample contained 0.3 mg/mL enzyme. The Ser195Ala R-Tp spectrum agrees well with the wild-type R-Tp spectrum (S. Hung & L. Hedstrom, unpubl. obs.).

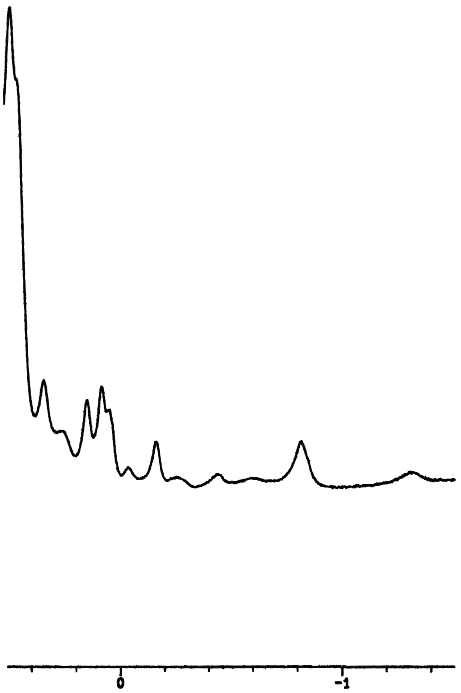

Fig. 5. High-field region of the one-dimensional ^{1}H-NMR spectrum (500 MHz) of R-Tg. The sample contained 1 mM Ser195Ala R-Tg in 10 mM $CaCl_2$ with 10% D_2O at pH 8.0. The spectrum is shown in the range from 0.5 to −1.5 ppm. Chemical shifts of the prominent upfield peaks are −1.31, −0.82, and −0.44 ppm.

possible explanation for the lower activity of R-Tg: more energy will be required to form an ordered conformation (Pasternak et al., 1998). This increased flexibility might also be expected to destabilize the ordered activation domain structure of R-Tp. However, the increased negative charge of R-Tp also appears to stabilize the Ile16–Asp194 salt-bridge (Soman et al., 1989), thus compensating for the greater intrinsic flexibility of the activation domain.

Why is the activation domain of R-Tg more flexible than that of B-Tg?: There are accumulating reports of "natively unfolded" proteins (Weinreb et al., 1996), i.e., proteins that exist naturally in a disordered state until ligand binding induces structure. Natively unfolded proteins are typically highly charged, and intriguingly, most are anionic (see Table 1 of Weinreb et al., 1996). Thus, negative charges may favor disordered structure. Interestingly, a survey of water distributions around amino acids in 16 high-resolution protein structures shows that Asp and Glu residues have a greater tendency to interact with water than Lys residues (Thanki et al., 1988). Similarly, Wolfenden used hydration potentials to demonstrate that Asp and Glu side chains interact more favorably with water than the Lys side chain (Wolfenden et al., 1979). The cationic B-Tg and anionic R-Tg differ in charge by 12.5 units; the activation domain of R-Tg is 4 units more negative than in B-Tg (Table 1). This charge difference may be the source of the increased flexibility of the activation domain of R-Tg.

Methods: *Construction and purification of trypsinogen:* The Ser195Ala mutant of R-Tg was used in all experiments to prevent autoactivation. The mutant was constructed and purified as described previously (Hedstrom et al., 1992; Pasternak et al, 1998).

Circular dichroism spectra of trypsin and trypsinogen mutants: All spectra were measured using a Jovin-Yvon Mark V autodichrograph. Samples contained 0.3 mg/mL enzyme in 50 mM Tris, 10 mM $CaCl_2$, pH 8.0. Near UV spectra were measured in one cycle from 320 to 240 nm in 0.2 nm steps in a 1 cm cell with a sensitivity of 5×10^{-6} and response time of 2 s.

NMR spectroscopy: A one-dimensional ^{1}H-NMR spectrum (512 scans) of Ser195Ala trypsinogen was measured on a Bruker AMX-500 spectrometer at 25 °C. A presaturation pulse of 700 ms was used for water suppression. Water was used as the reference at 4.78 ppm. The sample contained 1 mM R-Tg, 10 mM $CaCl_2$, and 10% D_2O at pH 8.0.

Crystallization: Ser195Ala R-Tg and wild-type R-Tp stored in 1 mM HCl were concentrated to 7 mg/mL and mixed with a slight molar excess of BPTI with $CaCl_2$ such that the Ca^{2+} concentration of the complex solution was 10 mM. Crystals of both the Ser195Ala trypsinogen-BPTI and the wild-type trypsin-BPTI complexes were obtained via hanging drop vapor diffusion at room temperature. In the case of trypsin, well solution contained 30% PEG 4000, 0.2 M $LiSO_4$, and 0.1 M Tris pH 8.5, and produced a crystal with dimensions 1.2 mm × 0.2 mm × 0.2 mm. A trypsinogen crystal, which looked like a faceted sphere of about 0.3 mm diameter, was grown from 24% PEG 4000, 0.2 M $LiSO_4$, and 0.1 M Tris pH 8.0.

Structure determination: Data on a single Ser195Ala trypsinogen-BPTI crystal were collected at 4° with a scan width of 1° per frame and an exposure time of 60 min per frame for 60° on an R-AXIS IIC image plate system mounted on a Rigaku RU-200B X-ray generator running at 45 kV and 120 mA. Data on a single wild-type trypsin-BPTI crystal were collected similarly but with a scan width of 0.5° and a 30 min exposure per frame for 36.5°. Frames were integrated and the data were scaled and merged together using the HKL package (DENZO and SCALEPACK) from HKL Research, Inc. (Charlottesville, Virginia) (Table 2).

Crystals of Ser195Ala trypsinogen-BPTI and wild-type trypsin-BPTI were isomorphous with a previous structure of R-Tp mutant D189G/G226D complexed with BPTI (Perona et al., 1993a). The initial model for the Ser195Ala trypsinogen-BPTI complex was the D189G/G226D trypsin-BPTI complex (accession code 1brb) (Perona et al., 1993a), and the initial model for the wild-type trypsin-BPTI was the Ser195Ala complex. The structures were refined using the X-PLOR package (Brünger et al., 1987) with 10% of the data set aside to compute R_{free} (Brünger, 1992). Initially an overall temperature factor of 20 Å^2 was used. Rigid body refinement was carried out treating trypsin or trypsinogen and BPTI as separate rigid bodies. Positional refinement and *B*-factor refinement were monitored by the R_{free} to prevent overrefinement (Brünger, 1992; Kleywegt & Jones, 1997). Water molecules were added in three rounds using the waterpick script in X-PLOR and keeping waters for which electron density was seen in the difference Fourier electron density map with coefficients $F_o - F_c$, and the difference Fourier electron density map with coefficients $2F_o - F_c$. In addition, after refinement the waters were required to be within hydrogen bond distance (3.4 Å) from hydrogen bond

donors or acceptors, and waters with temperature factors greater than 60 Å^2 were deleted from the model. In the trypsinogen structure disordered residues were omitted from the structure factor calculation. Simulated annealing omit maps (1,000 to 300 K) (Hodel et al., 1992) were calculated in attempts to locate the autolysis loop and the N-terminus, and were also used in positioning the residues at the boundaries of these regions for which electron density was observed (residues 16–17, 142, and 152–153). These residues were included in structure factor calculations for the subsequent refinement. No interpretable electron density for additional residues was observed in the difference Fourier map when the model was otherwise complete. Examination of and manual adjustments to the structure were performed between rounds of refinement using the program O (Jones et al., 1991) on a Silicon Graphics workstation. Ramachandran plots generated by PROCHECK (Laskowski et al., 1993) on the final coordinates indicated that all residues in both structures fell into the "most favored" or "allowed" regions. Least-squares alpha carbon alignments were done using the program LSQMAN (Kleywegt & Jones, 1994) limiting the alignment to residues 20–141 and 154–245 when one of the molecules being aligned was trypsinogen. Figures 1–3 were created using the program MOLSCRIPT (Kraulis, 1991).

Acknowledgments: This work was supported by NIH HL50366 (L.H.) and in part by a grant from the Lucille P. Markey Charitable Trust to Brandeis University. L.H. is a Searle Scholar and a Beckman Young Investigator. We thank Huaping Mo for the NMR spectrum, and David Harrison, Ezra Peisach and other members of the Ringe lab for help with crystallographic computing.

References

Bennett WS, Huber R. 1984. Structural and functional aspects of domain motions in proteins. *CRC Crit Rev Biochem 15*:291–384.

Bode W, Schwager P, Huber R. 1978. The transition of bovine trypsinogen to a trypsin-like state upon strong ligand binding. The refined crystal structures of the bovine trypsinogen-pancreatic trypsin inhibitor complex and of its ternary complex with Ile-Val at 1.9 Å resolution. *J Mol Biol 118*:99–112.

Bolognesi M, Gatti G, Menegatti E, Guarneri M, Marquart M, Papamokos E, Huber R. 1982. Three dimensional structure of the complex between pancreatic secretory trypsin inhibitor (Kazal type) and trypsinogen at 1.8 Å resolution. Structure solution, crystallographic refinement and preliminary structural interpretation. *J Mol Biol 162*:839–868.

Brünger AT. 1992. Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472–475.

Brünger AT, Kuriyan J, Karplus M. 1987. Crystallographic *R*-factor refinement by molecular dynamics. *Science 235*:458–460.

Craik CS, Roczniak S, Largman C, Rutter WJ. 1987. The catalytic role of the active site aspartic acid in serine proteases. *Science 237*:909–913.

Fehlhammer H, Bode W, Huber R. 1977. Crystal structure of bovine trypsinogen at 1.8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. *J Mol Biol 111*:415–438.

Fersht AR. 1972. Conformational equilibria in α- and δ-chymotrypsin. The energetics and importance of the salt bridge. *J Mol Biol 64*:497–509.

Hedstrom L, Lin T-Y, Fast W. 1996. Hydrophobic interactions control zymogen activation in the trypsin family of serine proteases. *Biochemistry 35*:4515–4523.

Hedstrom L, Szilagyi L, Rutter WJ. 1992. Converting trypsin to chymotrypsin: The role of surface loops. *Science 255*:1249–1253.

Hodel A, Kim S-H, Brünger AT. 1992. Model bias in macromolecular crystal structures. *Acta Cryst A 48*:851–858.

Huber R, Bode W. 1978. Structural basis of the activation and action of trypsin. *Acct Chem Res 11*:114–122.

Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. 1991. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Cryst A 47*:110–119.

Keil B. 1971. Trypsin. *The enzymes 3*:250–275.

Kerr MA, Walsh KA, Neurath H. 1975. Catalysis by serine proteases and their zymogens. A study of acyl intermediates by circular dichroism. *Biochemistry 14*:5088–5094.

Kleywegt GJ, Jones TA. 1994. A super position. *CCP4/ESF-EACBM newsletter on protein crystallography 31*:9–14.

Kleywegt GJ, Jones TA. 1997. Model-building and refinement practice. *Methods Enzymol 277*:208–230.

Kraulis PJ. 1991. MolScript: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst 24*:946–950.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst 26*:283–291.

Pasternak A, Liu X, Lin TY, Hedstrom L. 1998. Activating a zymogen without proteolytic processing: Mutation of Lys15 and Asp194 active trypsinogen. *Biochemistry 37*:16201–16210.

Perkins SJ, Wuthrich K. 1980. Conformational transition from trypsinogen to trypsin. ^{1}H nuclear magnetic resonance at 360 MHz and ring current calculations. *J Mol Biol 138*:43–64.

Perona JJ, Hsu CA, Craik CS, Fletterick FJ. 1993a. Crystal structures of rat anionic trypsin complexed with the protein inhibitors APPI and BPTI. *J Mol Biol 230*:919–933.

Perona JJ, Hsu CA, Craik CS, Fletterick RJ. 1993b. Relocating the negative charge in the binding pocket of trypsin. *J Mol Biol 230*:934–949.

Soman K, Yang A-S, Honig B, Fletterick R. 1989. Electrical potentials in trypsin isozymes. *Biochemistry 28*:9918–9926.

Sprang S, Standing T, Fletterick RJ, Stroud RM, Finer-Moore J, Xuong N-H, Hamlin R, Rutter WJ, Craik CS. 1987. The three dimensional structure of Asn102 mutant of trypsin: Role of Asp102 in serine protease catalysis. *Science 237*:905–909.

Thanki N, Thornton JM, Goodfellow JM. 1988. Distributions of water around amino acid residues in proteins. *J Mol Biol 202*:637–657.

Vincent J-P, Lazdunski M. 1972. Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and role of disulfide bridges. *Biochemistry 11*:2967–2977.

Weinreb PH, Zhen W, Poon AW, Conway KA, Lansbury PT Jr. 1996. NACP: A protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry 35*:13709–13715.

Wolfenden RV, Cullis PM, Southgate CCF. 1979. Water, protein folding, and the genetic code. *Science 206*:575–577.