Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

High-resolution structures of three new trypsin–squash-inhibitor complexes: a detailed comparison with other trypsins and their complexes

Ronny Helland,[a] Gunnar I. Berglund,[a] Jacek Otlewski,[c] Wlodek Apostoluk,[c] Ole A. Andersen,[a] Nils P. Willassen[b] and Arne O. Smalås[a]*

[a]Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway, [b]Department of Biotechnology, Faculty of Medicine, University of Tromsø, N-9037 Tromsø, Norway, and [c]Institute of Biochemistry and Molecular Biology, University of Wrocław, Tamka 2, 50-137 Wrocław, Poland

Correspondence e-mail: arne.smalas@chem.uit.no

Received 8 April 1998
Accepted 3 August 1998

PDB References: ST-CMTI, 2sta; ST-CPTI, 2stb; BT-CPTI, 2btc.

An anionic trypsin from Atlantic salmon and bovine trypsin have been complexed with the squash-seed inhibitors, CMTI–I (*Cucurbita maxima* trypsin inhibitor I, P1 Arg) and CPTI–II (*Cucurbita pepo* trypsin inhibitor II, P1 Lys). The crystal structures of three such complexes have been determined to 1.5–1.8 Å resolution and refined to crystallographic *R* factors ranging from 17.6 to 19.3%. The two anionic salmon-trypsin complexes (ST–CPTI and ST–CMTI) and the bovine-trypsin complex (BT–CPTI) have been compared to other trypsin–inhibitor complexes by means of general structure and primary and secondary binding features. In all three new structures, the primary binding residue of the inhibitor binds to trypsin in the classical manner, but with small differences in the primary and secondary binding patterns. Lysine in CPTI–II binds deeper in the specificity pocket of bovine trypsin than lysine in other known lysine–bovine-trypsin complexes, and anionic salmon trypsin lacks some of the secondary binding interactions found in the complexes formed between squash inhibitors and bovine trypsin. The ST–CMTI complex was formed from the reactive-site-cleaved form of the inhibitor. However, well defined electron density was observed for the P1—P1′ peptide bond, together with a hydrogen-bonding pattern virtually identical to those of all serine-protease–protein-inhibitor complexes, indicating a resynthesis of the scissile bond.

1. Introduction

Proteolytic enzymes are found in organisms living in extremely cold to extremely warm environments. Among these, the serine proteinases are one of the most extensively studied groups of enzymes; they have been well characterized for both mammals and cold-adapted fish species. The gene coding for trypsin, a member of this family, is found in at least four different isoforms in Atlantic salmon (*Salmo salar*): three anionic forms and one cationic form (Male *et al.*, 1995; GenBank/EMBL codes SalTRP-I, SalTRP-IA, SalTRP-II and SalTRP-III). The anionic and cationic isoforms differ in both physical behaviour (Outzen *et al.*, 1996) and three-dimensional structure (Schrøder *et al.*, 1998). The anionic form follows the trend of other enzymes from cold-adapted species (*e.g.* Àsgeirsson *et al.*, 1989; Àsgeirsson & Bjarnason, 1991; Kristjansson & Nielsen, 1992; Simpson & Haard, 1984; Osnes & Mohr, 1985; Gildberg & Øverbø, 1990; Feller *et al.*, 1994) where increased catalytic efficiency is at the expense of reduced temperature stability compared with the mammalian counterparts. The cationic form, on the other hand, resembles the mammalian trypsins in these respects. For a specific

substrate, the anionic fish enzyme is almost 40 times more efficient, measured by the k_{cat}/K_M ratio, than the mammalian counterparts (Outzen *et al.*, 1996). K_M (the Michaelis–Menten constant) is reduced by a factor of up to 20, while k_{cat} (the catalytic rate constant) is about twice as high. Thus, the increased catalytic efficiency observed for anionic salmon trypsin is primarily due to a higher binding affinity of the substrate. This suggests differences in the substrate-binding region; in this study the trypsin–squash-inhibitor complexes of both salmon and bovine trypsins serve as models for studying differences in substrate binding. A thorough comparison of bovine and anionic salmon trypsins has been made by Smalås *et al.* (1994) and Heimstad *et al.* (1995), but the structural basis for the differences in physical behaviour are not yet fully understood.

Of particular importance in the determination of the strength of association with trypsin is the P1 residue at position 5 [P1, P1′ *etc*.; notation according to the nomenclature of Schechter & Berger (1967)], which forms almost half of all contacts with the enzyme (Bode *et al.*, 1989). Truncation of this side chain to a Gly residue removes about 70% of the total association energy (O. Buczek, M. Dadlez & J. Otlewski, unpublished results). Interactions of basic side chains in the S1 binding pocket of trypsin thus form a 'hot spot' of enzyme–inhibitor interaction and precise explanation of their structural accommodation seems crucial. The role of the P1 residue is further supported by free-energy calculations, which indicate that the P1 residue plays a dominant role in the binding energy of trypsin complexes (Krystek *et al.*, 1993). In a more intensive study of the importance of the P1 residue in complex formation, we intend to use inhibitors with various P1 amino acids. In this paper, we describe the X-ray structures of three squash-inhibitor complexes with either Arg or Lys side chains at position P1. The squash inhibitors *Cucurbita maxima* trypsin inhibitor I (CMTI-I) and *Cucurbita pepo* trypsin inhibitor II (CPTI–II) are identical, except for two residues. The reactive-site P1 residue which extends into the specificity pocket of trypsin is arginine in CMTI–I and lysine in CPTI–II, and residue 21 is exchanged from a valine in CMTI–I to an isoleucine in CPTI–II. The P1 arginine and P1 lysine side chains make favourable contacts inside the S1 pocket of trypsin. P1 Arg and P1 Lys squash inhibitors recognize bovine and anionic salmon trypsins with association constants in the 10^{11}–10^{12} M^{-1} range. The two enzymes bind the squash inhibitor with P1 lysine with approximately similar association constants (about 8×10^{11} M^{-1}). The association constant is 3 and 6 times lower for binding the inhibitor with P1 arginine to bovine and salmon trypsins, respectively (Otlewski & Zbyryt, 1994; Otlewski, unpublished data). Structure determination and association-energy measurements for P1 side chains should provide detailed information about trypsin specificity and, more generally, about the role of 'hot spots' in the determination of the energy of protein–protein complexes.

The present salmon- and bovine-trypsin complexes are compared with other trypsin complexes with squash inhibitors and the corresponding complex with the slightly larger (58 residues) bovine pancreatic trypsin inhibitor (BPTI; Bode & Huber, 1992). Proteinase inhibitors from squash seeds form an uniform family of small proteins consisting of 27–33 amino-acid residues, cross-linked with three disulfide bridges (for reviews, see Otlewski, 1993; Otlewski & Krowarsch, 1996). Squash inhibitors are highly stable and rigid proteins. They inhibit a number of serine proteinases, including trypsin, plasmin, kallikrein, cathepsin G and blood-clotting factors X_a and XII_a (Otlewski *et al.*, 1990). Squash inhibitors inhibit different serine proteinases *via* the standard mechanism of inhibition (Laskowski & Kato, 1980; Otlewski & Zbyryt, 1994). According to the mechanism, the inhibitors are substrates and their P1–P1′ reactive-site peptide-bond residues can be cleaved by the cognate proteinase. An inhibitor with a cleaved reactive site forms the same stable complex with the enzyme as an inhibitor with the reactive site intact. Thus, complex formation from the cleaved inhibitor and proteinase comprises resynthesis of the reactive site. The character of the complex formed from the cleaved inhibitor, to the best of our knowledge, has never been structurally verified.

2. Materials and methods

2.1. Crystallization and data collection

Anionic salmon trypsin (ST) was complexed with the squash-seed inhibitors from *Cucurbita maxima* (CMTI–I) and the inhibitor from *Cucurbita pepo* (CPTI–II), and bovine trypsin (BT) was complexed with CPTI–II. For crystallization experiments, CMTI with a cleaved reactive site (P1—P1′ peptide bond) was used, while the peptide bond was intact in CPTI. The reactive-site peptide-bond-hydrolyzed form of the inhibitor was prepared by a 48 h incubation of 70 mg of the inhibitor in 50 m*M* sodium formate and 20 m*M* $CaCl_2$, pH 2.8, with 0.2 mg of bovine trypsin (Worthington Chemical Corp.) in a total volume of 3 ml at 295 K. Trypsin was subsequently removed by chromatography on a Bio-Gel P-10 column equilibrated with 50 m*M* formic acid pH 1.5. The inhibitor fraction was subsequently applied to a SP-Sepharose column equilibrated with 50 m*M* sodium formate pH 3.2. Intact and cleaved inhibitor were resolved by a linear gradient of NaCl (0–0.3 *M*). Finally, the cleaved inhibitor was desalted on a Biogel P-2 column in 50 m*M* formic acid pH 1.5 (Otlewski & Zbyryt, 1994). Anionic trypsin from salmon was purified from pancreatic extract using affinity ion-exchange and gel-filtration chromatography (Outzen *et al.*, 1996), while bovine trypsin was purchased from Sigma Chemical Co. The complexes between squash inhibitor and trypsin were formed by mixing the components in a 2:1 molar ratio, yielding a final protein concentration of 20 mg ml^{-1}. The complexes were crystallized using the hanging-drop vapour-diffusion method. Both salmon-trypsin complexes were crystallized from 1.0 *M* citrate buffer at pH 6.0–6.4 at 277 K. The BT–CPTI complex was crystallized as described by Bode *et al.* (1989). All crystals belong to the orthorhombic space group $P2_12_12_1$, and there is one complex molecule in the asymmetric unit for all complexes. The salmon-trypsin crystals diffracted beyond

Table 1
Data-collection and refinement characteristics of complexes between trypsin and the squash inhibitors CMTI and CPTI.

	BT–CPTI	ST–CMTI	ST–CPTI
Resolution (Å)	1.5	1.8	1.8
Cell dimensions (Å)	59.1 × 55.7 × 74.9	62.3 × 63.7 × 82.3	62.3 × 63.4× 82.5
Number of observations (∞–maximum resolution)	266925	213810	279741
Number of unique reflections	60871	30958	35789
R_{merge} (%)	8.4	8.5	5.0
Completeness (%)	89.8	98.9	99.3
Wilson B (Å^2)	15.1	19.9	21.0
Number of protein atoms in refinement	1820	1796	1797
Number of solvent molecules (including Ca^{2+})	166	139	168
Number of reflections used in refinement	35087	29878	29851
Resolution range in refinement (Å)	6.0–1.5	6.0–1.8	6.0–1.8
R factor (%)	18.3	19.3	17.6
R_{free} (%)	21.2	23.3	20.1
B factor (Å^2)			
All	18.6	22.11	21.86
Protein	17.1	21.32	20.51
Main chain	15.9	19.65	18.97
Side chain	18.7	23.19	22.23
Water	34.5	32.89	37.29
R.m.s. deviations from ideality			
Bond lengths (Å)	0.012	0.013	0.012
Bond angles (°)	2.538	2.625	2.537
R.m.s. coordinate error (Å)			
σ_A	0.18	0.20	0.18
Luzzati	0.10	0.10	0.10

1.8 Å and the bovine complex diffracted to 1.5 Å. All lattice parameters are given in Table 1.

X-ray data were collected on the Swiss–Norwegian Beamline (BM01) at ESRF using a MAR image-plate system (MAR Research). One crystal was used for each complex. Data (Table 1) were integrated using *DENZO* (Otwinowski, 1993; Minor, 1993). Scaling and merging were performed using *SCALA* and *AGROVATA* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994), yielding overall R_{merge} values from 5.0 to 8.4% (Table 1) and completeness between 90 (BT) and 99% (ST).

2.2. Molecular replacement

The orientation of the anionic ST–CMTI complex was determined by molecular replacement using *X-PLOR* (Brünger, 1992). The search model was the bovine-β-trypsin–CMTI complex (Protein Data Bank entry 1PPE; Bernstein *et al.*, 1977; Bode *et al.*, 1989), in which the bovine trypsin was replaced with the coordinates for anionic salmon trypsin (Protein Data Bank entry 2TBS; Smalås & Hordvik, 1993). Reflections with $F > 2\sigma(F)$ were used throughout the molecular replacement. A real-space Patterson rotation search was performed using data in the resolution range 15.0–4.0 Å. One well separated solution 10σ above average and 4σ above the second highest peak was found at Eulerian angles $\theta_1 = 202.15$, $\theta_2 = 66.95$ and $\theta_3 = 61.09°$. The rotated and translated model provided an R value of 42.0% in the resolution range 8.0–2.5 Å.

2.3. Model building and refinement

The rotated and translated model of ST–CMTI served as the starting coordinates for the refinement of both salmon-trypsin–squash-inhibitor complexes. The refinement of BT–CPTI was initiated from the coordinates of the BT–CMTI complex (PDB entry 1PPE). The models were improved first by rigid-body refinement and simulated-annealing methods using *X-PLOR* (Brünger, 1992). Further improvement of the models included alternate cycles of model building using *O* (Jones *et al.*, 1991) followed by conventional positional refinement and individual B-factor refinement using *X-PLOR*. The models were rebuilt on the basis of $2F_o - F_c$ and $F_o - F_c$ electron-density maps. The data used in the refinement were gradually extended to 1.8 and 1.5 Å for the salmon-trypsin and the bovine-trypsin complexes, respectively, and 10% of the data were kept out of refinement for cross validation (Kleywegt & Brünger, 1996). Solvent molecules were added to the model where the difference density exceeded 4σ and within reasonable hydrogen-bonding distances ($\leq$3.4 Å). During refinement, all bonded and non-bonded energy terms on the inhibitor scissile bond were set to zero to allow an unbiased description of its geometry. Non-bonded energy terms for the atoms in the catalytic triad which are involved in proton transfer during catalysis were also set to zero. The final refinement characteristics are given in Table 1.

3. Results

3.1. General description

The quality of all structures is, in general, good. The mean coordinate error estimated from a Luzzati plot (Luzzati, 1952) is about 0.1 Å for all complexes, while the corresponding r.m.s. coordinate errors calculated by the σ_A method suggested by Read (1986) are 0.20 Å for the CMTI complex and 0.18 Å for both CPTI complexes. The Ramachandran plots of φ,ψ angles (Ramachandran & Sasisekharan, 1968) show that about 90% of the non-glycine residues fall within the most favourable regions of the plot for all structures. The remaining 10% lie within the additional allowed regions as defined in *PROCHECK* (Laskowski *et al.*, 1993), except for one residue. As reported for both BT–CMTI (Bode *et al.*, 1989) and PT–MCTI [complex between porcine trypsin and inhibitor from *Momordica charantia* seeds (Huang *et al.*, 1992)], the Φ and Ψ

values for inhibitor residue Ala18I are outside the allowed regions.

The electron density is, in general, well defined for the trypsin part of all complexes. Regions where the density is weak are generally in loop regions exposed to solvent. The anionic salmon-trypsin main chain is well defined in both complexes, except for the three terminal residues at the C-terminus, a feature similar to that of the benzamidine-inhibited salmon trypsin (Smalås & Hordvik, 1993). Only the side chains of residues Tyr97 and Ser148 are completely missing in both salmon-trypsin complexes. In addition, the side chain of Arg62 is also completely missing in ST–CMTI. Residues 24 and 28 in anionic salmon trypsin were identified as prolines (Fig. 1). These residues are alanines in previously solved structures of the fish trypsin. Although some trypsin side chains are poorly defined in BT–CPTI, all of them can be interpreted. The electron density of the inhibitors in the salmon-trypsin complexes is, in general, not as well defined as for the trypsin molecules or the inhibitor part of the mammalian-trypsin complex. The orientation of the main chain of Arg1I is not clear and residue Glu24I is completely missing in both salmon-trypsin complexes. Only the segment from Val2I to Leu7I, in contact with the enzyme, is fully defined in the two complexes. Regions of the salmon-trypsin complexes where the squash inhibitor has poor or missing electron density include the side chain of Arg1I, the loop from residus 9I to 13I, residue 19I and the loop from 23I to 27I. The inhibitor part of BT–CPTI is involved in several packing interactions with symmetry-related molecules, resulting in an electron density which is considerably better compared with those of the salmon-trypsin complexes. These interactions probably have a stabilizing effect on the loop 9I–13I in the bovine-trypsin complexes, and the presence of the Arg1I side chain in the bovine-trypsin complex stabilizes the 23I–27I loop. The three internal waters in the squash inhibitors reported by Bode *et al.* (1989) and Huang *et al.* (1993) are also found in the three new structures.

The refined average B factor of all atoms is about 22 Å^2 for both salmon-trypsin complexes (Table 1) and 18.6 Å^2 for the BT–CPTI complex. All values correspond reasonably well with the Wilson B factor from scaling of the data sets.

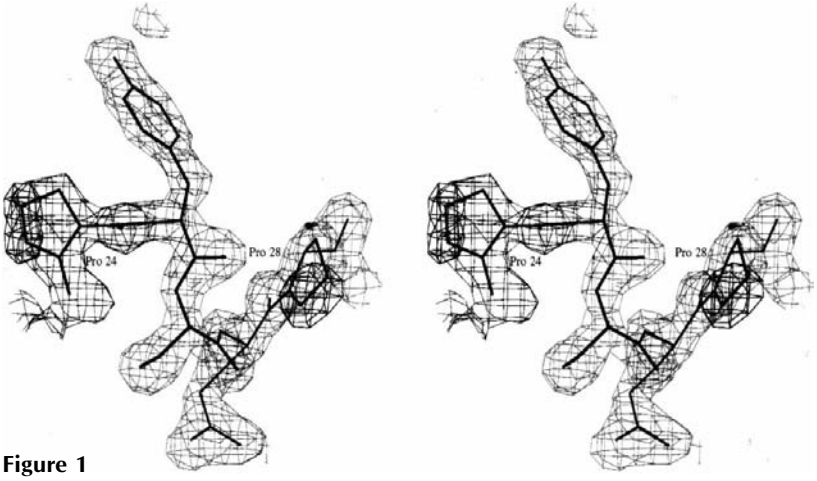

Figure 1
Electron density for residues at the N-terminus of anionic salmon trypsin in complex with the squash-seed inhibitor CPTI-II. The maps are from simulated annealing where residues 24 and 28 were refined as alanines. The $2F_o - F_c$ map (grey) is contoured at 1.5σ and the $F_o - F_c$ map (black) is contoured at 3σ. Both maps clearly identify residues 24 and 28 as prolines. The coordinates of the final refined model are superimposed on the simulated-annealing maps. The figure was produced using *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997).

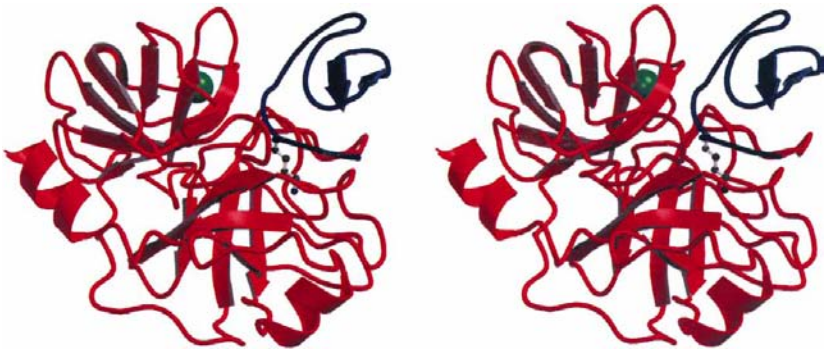

Figure 2
Ribbon-style diagram of the complex between salmon trypsin (red) and CPTI (blue). The P1 lysine residue of CPTI penetrating the active site of trypsin is indicated along with the structurally bound calcium ion (green) of trypsin.

3.2. Root-mean-square comparison of trypsins and their complexes

The three new structures are compared to the structures formed between bovine trypsin and CMTI and porcine trypsin (PT) in complex with MCTI (inhibitor from *Momordica charantia* seeds) (for illustration, see Fig. 2). The complex formed between bovine trypsin and BPTI is added for comparison of lysine binding. In Table 2, all main-chain atoms of the trypsin part of the trypsin-inhibitor complexes are compared with the benzamidine-inhibited bovine (Bode & Schwager, 1975) and anionic salmon (Smalås & Hordvik, 1993) trypsins. The inhibitor part of the complexes is compared to CPTI of the BT–CPTI complex. Differences between trypsins from different species are in the same range as those between other trypsin structures (Smalås *et al.*, 1994) and the mean differences between different crystal structures of the salmon or bovine trypsins (0.25 and 0.34 Å, respectively) are within the range expected from uncertainties in coordinate positions and effects from different packing environments. Regions that appear with peaks higher than 0.5 Å in the main-chain r.m.s. plots (not shown) between various trypsin structures are, in most cases, external loops that are known to be flexible in trypsin (Heimstad, 1996): the Nβ1–Nβ2 loop (residues 37–40), the calcium-binding loop (68–80), the Nβ5–Nβ6 loop (91–103) and the 'autolysis' loop (141–155). The

Table 2
R.m.s. differences (Å) between main-chain atoms in selected trypsin–protein-inhibitor complexes.

	Trypsin part of ST–BENZ	Trypsin part of BT–BENZ	Inhibitor part of BT–CPTI†
ST–BENZ	—	1.467	—
BT–BPTI	1.393	0.338	—
BT–CMTI	1.403	0.340	0.153
BT–CPTI	1.392	0.337	—
ST–CMTI	0.253	1.416	0.314
ST–CPTI	0.251	1.410	0.306
PT–MCTI	1.360	0.505	1.510

† All main-chain atoms of the inhibitors are superimposed on the fold of the inhibitor in the bovine-trypsin–CPTI complex.

main-chain atoms of residue 191 and 192 at the entrance to the specificity pocket also deviate from the benzamidine-inhibited trypsin by more than 0.5 Å for both salmon-trypsin complexes. These differences are most likely to be caused by inhibitor binding. The corresponding deviations for the bovine-trypsin complexes are somewhat lower.

As expected, the CPTI complex superimposes well on the CMTI complex for both the salmon- and bovine-trypsin complexes, with mean main-chain r.m.s. differences of 0.073 and 0.167 Å for the trypsin part, and 0.114 and 0.153 Å for the inhibitor part, respectively. The trypsin part of the complexes has a fold which is very similar to those found for the benzamidine-inhibited complexes, with an r.m.s difference of 0.25 Å for the salmon-trypsin complexes. The r.m.s. values are slightly higher for the corresponding bovine-trypsin complexes (0.34 Å). R.m.s. values were also calculated for selected segments or residues of the polypeptide chains: all main-chain atoms of the specificity pocket (residues 189–193 and 214–220), all atoms of the catalytic triad (His57, Asp102, Ser195), all main-chain atoms of the inhibitor-binding loop (P3–P2′) and, finally, all atoms of the reactive P1 residue (data not shown). The general impression is that the tertiary fold is slightly more conserved upon complex binding for the salmon trypsin than for the bovine trypsin. This is seen by the lower r.m.s. values for salmon trypsin both when the squash complexes are compared to each other and when they are compared to the benzamidine-inhibited complex. The r.m.s. differences for the catalytic triad atoms are of the same order for the squash-inhibitor complexes when compared to the benzamidine complexes (∼0.44–0.48 Å). The orientation of the catalytic triad is less well conserved for the two bovine-trypsin–squash-inhibitor complexes (BT–CMTI *versus* BT–CPTI, 0.122 Å; ST–CMTI *versus* ST–CPTI, 0.095 Å). The His57 side chain in BT–CMTI is rotated about 10° (χ_1 = 78°) relative to the BT–CPTI complex and the salmon-trypsin–squash-inhibitor complexes. The higher r.m.s. value for the bovine P1 residues, which is 0.225 Å compared with 0.171 Å for the salmon-trypsin–squash-inhibitor complexes, is also probably caused by differences in the χ_1 values (BT–CMTI, −64°; BT–CPTI, −73°; ST–CMTI, −75°; ST–CPTI, −75°). The BT–BPTI complex is generally, and not surprisingly, more similar to BT–CPTI than ST–CPTI. However, the P1 residue seems to have a conformation more similar to those of the salmon-trypsin complexes. The large r.m.s. values for the binding loop are primarily caused by the difference in amino-acid sequence at positions P2 and P3. In general, the tertiary fold of the trypsin–squash-inhibitor complex seems to be better conserved when lysine is bound than when arginine is bound. The orientation of the reactive residue is practically identical for the P1-lysine complexes, and the orientation of the catalytic triad is better conserved for the CPTI complexes. The difference in the specificity pocket and the inhibitor-binding loop seems to be of the same order for P1 arginine and P1 lysine.

3.3. Enzyme–inhibitor contacts

The binding-loop fold of the inhibitor is well conserved in all trypsin–squash-inhibitor complexes. The orientation of the P1 residue is also reasonably well conserved, regardless of whether it is lysine or arginine, and the N^{ζ} atoms of the lysines occupy almost the same position as the C^{ζ} atoms of the arginine residues. Only short segments of the 29-residue-long squash inhibitors (*cf.* Fig. 3) are in direct contact with the trypsin molecule in the complex. In total, 17 different trypsin residues are in contact with the inhibitor in each of the six complexes compared. Of these, only four amino acids differ between the three trypsins. Residue 39 is serine in the porcine trypsin and tyrosine in the bovine and salmon trypsins. Residue 151 is tyrosine in the mammalian trypsins and aspartic acid in salmon trypsin. Residue 175 is methionine in salmon trypsin and glutamine in the two mammalian trypsins. Finally, residue 217 is serine in the bovine trypsin and tyrosine in the two other enzymes.

The P4–P2′ segment of the squash-inhibitor complexes with P1 arginine forms about 130 enzyme–inhibitor contacts shorter than 4.0 Å (Table 3). For the squash-inhibitor P1-lysine complexes the number is about 120. An additional 11–14 such contacts are formed between the inhibitor Arg1I residue and the bovine and porcine trypsins. Well defined electron density is not observed for the P5 residue of the salmon-trypsin complexes. A maximum of 11 contacts are formed by residues other than the P5–P2′ binding loop. The total number of enzyme–inhibitor contacts shorter than 4.0 Å is 131 in the BT–BPTI complex,

	P4	P3	P2	P1	P1′	P2′	P3′	P4′																				
	1	2	3	4	5*	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
ST-CMTI	R	V	C	P	R	I	L	M	E	C	K	K	D	S	D	C	L	A	E	C	V	C	L	E	H	G	Y	C	G
ST-CPTI	R	V	C	P	K	I	L	M	E	C	K	K	D	S	D	C	L	A	E	C	I	C	L	E	H	G	Y	C	G
BT-CMTI	R	V	C	P	R	I	L	M	E	C	K	K	D	S	D	C	L	A	E	C	V	C	L	E	H	G	Y	C	G
BT-CPTI	R	V	C	P	K	I	L	M	E	C	K	K	D	S	D	C	L	A	E	C	I	C	L	E	H	G	Y	C	G
PT-MCTI	R	I	C	P	R	I	Y	M	E	C	T	R	D	S	D	C	M	A	K	C	I	C	V	A	-	G	H	C	G

	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
BT-BPTI	T	G	P	C	K	A	R	I	I	R	Y	F	Y	N	A	K	A	G	L	C	Q	T	F	V	Y	G	G	C	R

Figure 3
Complete amino-acid sequences of the squash inhibitors and of the binding loop of BPTI. Residues which form contacts shorter than 4.0 Å to trypsin in the enzyme–inhibitor complex are boxed.

Table 3
Summary of enzyme–inhibitor contacts (Å) shorter than 4.0 Å.

	BT–BPTI	BT–CPTI	ST–CPTI	ST–CMTI	BT–CMTI	PT–MCTI
P5	2	12	—	—	11	14
P4–P2′	107	105	111	123	120	125
Other	22	11	8	4	7	5

where 22 are formed by residues other than the P5–P2′ loop. 12 of those are formed by the BPTI secondary binding loop [residues 36I–39I, (Scheidig *et al.*, 1997)] and the remaining ten are formed by P3′ (2) and P4′ (8). The five residues from P3 to P2′ (Cys3I, Pro4I, Arg/Lys5I, Ile6I and Leu/Tyr7I) form contacts shorter than 3.4 Å to the enzyme in all complexes. The equivalent residues of BPTI form contacts with the enzyme in the bovine-trypsin BPTI complex. Such contacts are also formed by the P4 residue (Val/Ile2I) in all squash-inhibitor complexes having arginine as the P1 residue, and by the P5 residue (Arg1I) in the four mammalian squash-inhibitor complexes. Other contacts shorter than 3.4 Å are formed by Cys28I in ST–CMTI, Tyr27I in BT–CPTI, Met-17I and Lys19I in PT–MCTI, and Ile19I and Arg39I in BT–BPTI.

3.3.1. Primary binding contacts. As the primary binding residue (P1) is changed from a lysine in BT–BPTI, BT–CPTI and ST–CPTI to an arginine in BT–CMTI, ST–CMTI and PT–MCTI, the binding pattern at the bottom of the primary binding pocket differs between the two groups of trypsin complexes. Intermolecular contacts between the P1 amino acid of each inhibitor and the trypsin molecules for the six complexes compared are listed in Table 4, while additional enzyme–inhibitor contacts are listed in Table 5.

As shown in Table 4, the main chain of the P1 residue is bound to the enzyme in the classical manner described by Bode & Huber (1992) for all six complexes (for illustration, see Fig. 4). Since the lysine side chain is about one C—C bond shorter than the arginine side chain, substrates and substrate analogues with P1 lysine are known to bind to the substrate-binding Asp189 *via* a water molecule (Marquart *et al.*, 1983). This is also observed for the ST–CPTI and BT–CPTI complexes. In addition, the BT–CPTI complex forms a hydrogen bond directly to Asp189 $O^{\delta 2}$ (3.19 Å). This hydrogen bond is also observed to some extent in the ST–CPTI complex, where the distance is 3.41 Å. This direct hydrogen bond is also observed in the complex formed by bovine trypsin with an engineered subtilisin inhibitor (Takeuchi *et al.*, 1992) and bovine trypsinogen complexed with pancreatic secretory trypsin inhibitor (Bolognesi *et al.*, 1982). The distance is 3.63 Å in the BT–BPTI complex. Four hydrogen-bonding-type interactions between the main-chain P1 residue of the inhibitor and trypsin are common to the squash-inhibitor complexes and BT–BPTI. The peptide carbonyl O atom of P1 is bound in the oxyanion hole (193 N and 195 N) with similar hydrogen-bonding distances for the squash inhibitors. Both the amide N atom and the carbonyl O atom of P1 participate in hydrogen-bonding contacts with the hydroxyl O atom of Ser195. In the squash-inhibitor complexes, the P1 amide N atom also forms a relatively strong hydrogen bond to the carbonyl O atoms of Ser214. In BPTI, the cysteine and proline preceding the P1 residue have switched places compared with the squash inhibitors, and the main chain takes a completely different orientation from that of the P3 residue; hence, the distance between the P1 amide N atom and 214 O of trypsin is longer in the BT–BPTI complex (3.56 Å). Whether the longer distances in the P1–S1 and P3–S3 binding sites of the BPTI complex are solely caused by the difference in the amino-acid sequence of the binding loop is not yet quite clear. The dihedral ψ angles of all squash-inhibitor P1 residues are about -90°, compared with -116° in the BT–BPTI complex. The cysteine in the P2 position of BPTI could be responsible for forcing the P1 ψ angle into a different conformation.

The width of the specificity pocket is approximately determined by the distance between the peptide planes of residues 215 and 191 on opposite sides of the pocket. This distance is ~7.90 Å in both benzamidine-inhibited trypsin complexes. The distance between the carbonyl C atoms of these residues for the six trypsin–protein inhibitor complexes (Table 4) shows that there is a significant enlargement of the pocket upon complex formation with the protein inhibitors compared with the benzamidine-inhibitor structures. The enlargement is more pronounced for the squash P1-Lys complexes, and less for the mammalian P1-Arg complexes. BT–BPTI and ST–CMTI fall in between.

3.3.2. Secondary binding contacts. The binding-loop residues from P5 to P2′ of the squash inhibitors and residues from P3 to P4′ and residue 39I of BPTI form close contacts with four to seven residues of trypsin in addition to those formed by the P1 residue (Table 5; Figs. 3 and 4). In addition to P4–P2′, MCTI also forms contacts to trypsin with residues 17I and 19I in the PT–MCTI complex. MCTI differs from the other two squash inhibitors by the deletion of residue 25I and longer side chains for residues 17I and 19I, and hence gives rise to additional contacts with trypsin. The main difference in binding properties between BPTI and the smaller inhibitors of the squash family is that the residues on the C-terminal side of the P1 residue are more heavily involved in contacts in the former (see Bode & Huber, 1992). BPTI forms seven such hydrogen bonds, while all five squash inhibitors only form the one between P2′ and 41 O, with almost identical lengths. Owing to its larger and more extended shape, BPTI forms additional contacts with two loops in trypsin: residues 39–40 and 97. The number of hydrogen bonds formed between the inhibitor and the enzyme on the N-terminal side of the P1 residue varies among the complexes. The P5–S5 (P5 N217–O^{γ}) interaction is only found in the two bovine-trypsin complexes owing to the difference in amino-acid composition. Only the hydrogen bond between P3 N and 216 O is conserved in all squash-inhibitor complexes and has a similar distance, except for BT–CMTI where it is about 0.2 Å longer than in the other complexes. This bond is not formed in BT–BPTI owing to the amino-acid difference in the inhibitor-binding loop. The

Table 4
Selected inter- and intramolecular distances (Å) within the primary binding pocket of trypsin complexes.

	BT–BPTI†		BT–CPTI		ST–CPTI		ST–CMTI		BT–CMTI		PT–MCTI	
	H-bond‡	Other	H-bond	Other	H-bond	Other	H-bond	Other	H-bond	Other	H-bond	Other
P1 N–195 O^{γ}	3.15	—	2.95	—	2.90	—	2.96	—	2.98	—	3.02	—
P1 N–214 O	3.56	—	3.06	—	2.99	—	2.99	—	3.07	—	3.02	—
P1 C^{α}–195 O^{γ}	—	3.10	—	3.01	—	2.95	—	2.96	—	2.95	—	3.04
P1 C–195 O^{γ}	—	2.68	—	2.68	—	2.65	—	2.61	—	2.69	—	2.73
P1 C–195 C^{β}	—	3.22	—	3.38	—	3.37	—	3.31	—	3.41	—	3.35
P1 O–193 N	2.76	—	2.63	—	2.70	—	2.73	—	2.78	—	2.74	—
P1 O–194 N	3.06	—	3.35	—	3.40	—	3.33	—	3.41	—	3.35	—
P1 O–195 N	2.81	—	2.99	—	3.02	—	2.94	—	2.89	—	2.92	—
P1 O–195 O^{γ}	3.04	—	2.99	—	2.98	—	2.89	—	2.88	—	2.91	—
P1 O–195 C^{β}	—	3.15	—	3.17	—	3.19	—	3.07	—	3.17	—	3.08
P1 C^{β}–195 O^{γ}	—	3.15	—	3.22	—	3.18	—	3.14	—	3.28	—	3.22
P1 C^{ε}–190 O	—	3.33	—	3.39	—	3.41	—	—	—	—	—	—
P1 N^{ζ}–189 $O^{\delta1}$	3.63	—	3.19	—	3.41	—	—	—	—	—	—	—
P1 N^{ζ}–190 O	3.01	—	3.01	—	3.05	—	—	—	—	—	—	—
P1 N^{ζ}–190 O^{γ}	3.13	—	3.09	—	3.23	—	—	—	—	—	—	—
P1 C^{ζ}–190 O	—	—	—	—	—	—	—	3.37	—	3.52	—	3.36
P1 $N^{\eta}1$–189 $O^{\delta1}$	—	—	—	—	—	—	3.04	—	3.03	—	2.85	—
P1 $N^{\eta}1$–190 O	—	—	—	—	—	—	3.20	—	3.31	—	3.22	—
P1 $N^{\eta}1$–190 O^{γ}	—	—	—	—	—	—	2.87	—	2.94	—	2.94	—
P1 $N^{\eta}1$–189 C^{γ}	—	—	—	—	—	—	—	3.54	—	3.60	—	3.40
P1 $N^{\eta1}$–226 C^{α}	—	—	—	—	—	—	—	3.45	—	3.41	—	3.31
P1 $N^{\eta2}$–189 $O^{\delta2}$	—	—	—	—	—	—	2.78	—	2.89	—	2.79	—
P1 $N^{\eta2}$–219 O	—	—	—	—	—	—	3.16	—	3.09	—	2.78	—
P1 N^{ζ}–OW1	2.79	—	2.80	—	2.81	—	—	—	—	—	—	—
P1 $N^{\zeta}/N^{\eta1}$–OW2	2.84	—	2.93	—	3.00	—	3.27	—	3.12	—	3.15	—
P1 N^{ε}–OW3	—	—	—	—	—	—	3.12	—	2.96	—	3.01	—
215 C–191 C	—	8.18	—	8.29	—	8.26	—	8.18	—	8.13	—	8.13

† 189 $O^{\delta1}$ and $O^{\delta2}$ are switched in BPTI relative to the squash inhibitors. ‡ H-bond, hydrogen bond.

second P3–S3 bond (P3 O–216 N) is conserved among all complexes having lysine as P1. This interaction is only found in the salmon-trypsin complex when P1 is arginine. The distance is similar for all four complexes. The P2–S2 interaction from P2 O to 192 $N^{\varepsilon2}$ is also conserved among the P1-lysine complexes, but with differing distances. This interaction is also observed for PT–MCTI. The side chain of the N-terminal arginine (I1) forms a strong ion-pair with the squash-inhibitor C-terminal carboxyl group in both bovine-trypsin–squash-inhibitor complexes. Furthermore, the positively charged N-terminus forms a strong hydrogen bond with the hydroxyl group of Ser217 of trypsin. The larger tyrosine at position 217 of salmon trypsin cannot participate in a corresponding bond in the squash-inhibitor complexes, as Tyr217 in ST–CPTI and ST–CMTI partly occupies the same space as the N-terminal of the inhibitor in the complexes with bovine trypsin, thus forcing the N-terminal arginine to a different position and probably disrupting possible salt-bridge formation with the C-terminus. This salt bridge is probably an important stabilizing factor of the squash inhibitors, and lack of the salt-bridge could be a reason for the generally less well defined inhibitor structure seen for the salmon complexes. Residue 217 is also a tyrosine in porcine trypsin, and should not participate in a hydrogen bond with the N-terminal of the inhibitor in the PT–MCTI complex. However, a salt bridge between the N-terminal arginine and the C-terminus is reported for this complex (Huang *et al.*, 1993), albeit with very high mean temperature factor and hence low stability. Accomodation of the P5 arginine side chain in the porcine complex is probably possible because of a slightly more flexible Tyr217 side chain. The mobility of this side chain in the salmon-trypsin complexes is restricted because of the formation of hydrogen bonds between Glu221 $O^{\varepsilon1}$ and Asn224 $O^{\delta1}$ and a proline at position 222 of extension loop 2 (nomenclature according to Hedstrom *et al.*, 1992). The positioning of the Arg1I side chain between the inhibitor and residue 217 probably prevents the formation of a hydrogen bond between P3 O and 216 N in BT–CMTI and PT–MCTI, as the enzyme and inhibitor are forced slightly apart. The length of the P3–S3 bonds in the BT–CPTI complex are of the same order as for the salmon complexes. The water molecule (OW1) which bridges P1 N^{ε} and Asp189 $O^{\delta2}$ for the P1-lysine complexes causes a shift in the position of residue 216 relative to the P1 arginine complexes. The water–216 Cα distance is within close van der Waals contacts (3.37 Å).

3.3.3. Waters in contact region. All inhibitor complexes have two water molecules bound at the bottom of the binding pocket. One of the molecules has approximately the same position in all complexes, regardless of whether P1 is lysine or arginine (OW2; Fig. 4). In addition to the water molecules buried in the active-site pocket (OW1, OW2 and OW3; Fig. 4), the P1-lysine and P1-arginine squash-inhibitor complexes have three to four (BT–CMTI only) water molecules serving as bridges between trypsin and the inhibitor. Two of those are common for all six squash-inhibitor complexes (OW4 and OW6). OW5 is also present at approximately the same position in all squash-inhibitor complexes, but a hydrogen bond to residue 151 in the salmon-trypsin complexes is impossible

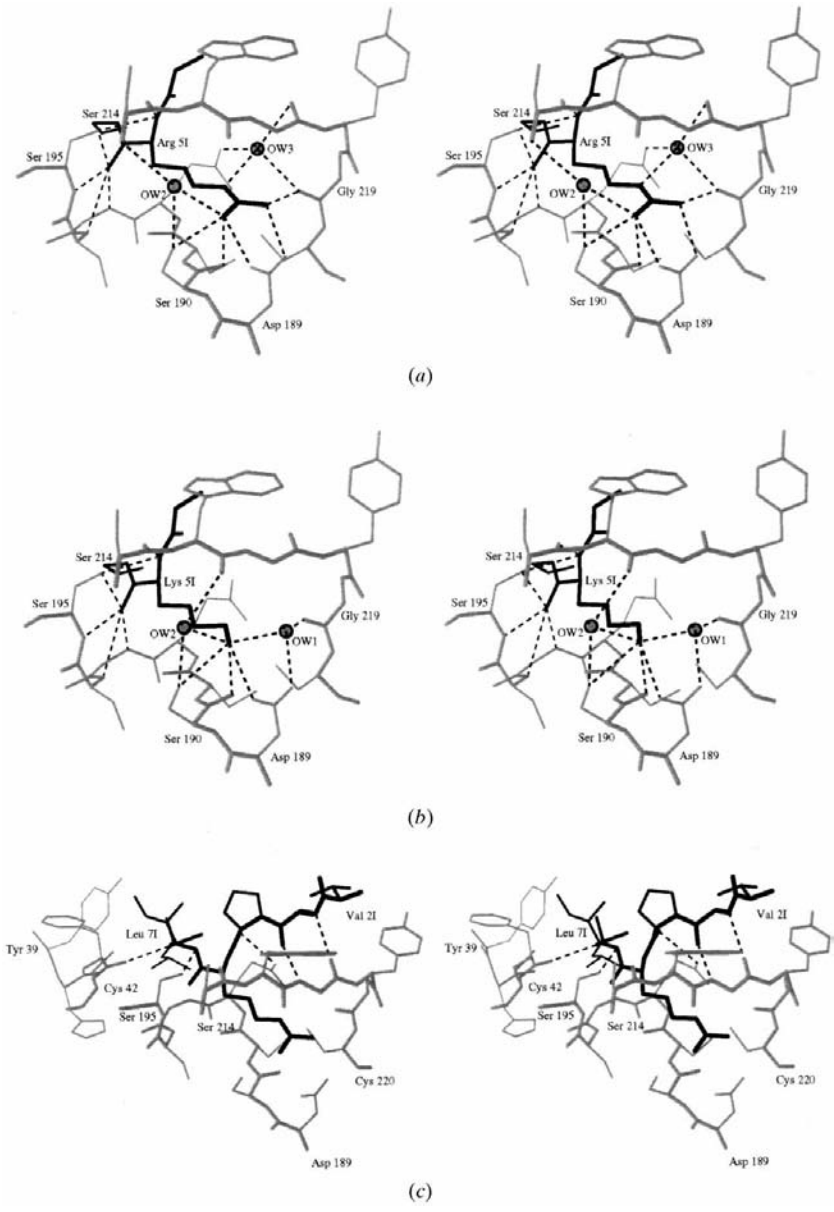

Figure 4
Stereoview of the primary and secondary binding of the squash-seed inhibitors to anionic salmon trypsin. (*a*) The two amine groups of the guanidino group of P1 arginine substrates bind to both O atoms of the Asp189 carboxy group. (*b*)The N^{ζ} atom of P1 lysine contacts the carboxy group of Asp189 through one direct hydrogen bond to $O^{\delta1}$ and through a water molecule to $O^{\delta2}$. (*c*) Of the five hydrogen bonds found in the secondary binding site of ST–CPTI, only the two on the C-terminal side of the scissile bond are conserved for the six trypsin-inhibitor complexes studied. Two of the hydrogen bonds on the N-terminal side are conserved for all complexes with lysine as the reactive residue. The figure was produced using *BOBSCRIPT* (Kraulis, 1991; Esnouf, 1997).

3.4. The geometry of the reactive site

Omission of constraints on the inhibitor scissile bond during refinement yielded a reactive-site geometry for all three complexes which is very similar to those of the already known bovine-trypsin complexes (Marquart *et al.*, 1983). The distance between the reactive hydroxyl O atom of Ser195 and the carbonyl C atom of the P1 residue is about 2.6 Å, which is not significantly different to that found for other comparable complexes (Table 6), and is intermediate between the van der Waals distance and a covalent bond, as discussed by Bode & Huber (1992). All other inter- and intramolecular distances at the reactive site, including the oxyanion-hole interactions, are also similar to those of previously determined trypsin complexes. The scissile peptide bond is intact for all three complexes, indicating that a resynthesis of the scissile bond must have taken place in the initially cleaved CMTI inhibitor. For both salmon-trypsin complexes the scissile bond is shorter (1.23 and 1.25 Å) than a mean peptide bond (1.33 Å) and the scissile carbonyl double bond is longer (1.34 and 1.35 Å relative to 1.23 Å), indicating a higher degree of delocalization of the double bond. In all the mammalian complexes the C—N bond is typically 1.33 ± 0.1 Å, except for the BPTI complex where both the scissile C—N and C—O distances are 1.26 Å. The C—O distance for the BT–CPTI complex is longer than for the other complexes with mammalian trypsins, but shorter than for the salmon-trypsin complexes. As in the ovomucoid-inhibitor complexes (Bode & Huber, 1992), no out-of-plane distortion of the carbonyl C atom was observed for the three new complexes in the present study. However, the importance of these observations should not be overestimated, as the differences, after all, are less than the crystallographic coordinate errors.

4. Discussion

The crystal structures of the salmon-trypsin complexes reveal a new form of salmon trypsin. The electron density clearly shows prolines at positions 24 and 28. From the genetic studies, gene loci coding for four different trypsins were identified and sequenced (Male *et al.*, 1995). One of these has a proline in position 28 (SSTRYPII), as do the bovine and porcine trypsins, but all known anionic salmon trypsins have alanine at position 24. However, additional forms of trypsin from Atlantic salmon are not surprising, since salmoid fish species are well known for their complex isoenzyme patterns (see, for

owing to a different fold of the bovine-trypsin autolysis loop. Tyr151 occupies the same position in all mammalian complexes, but MCTI has a deletion at position 25I, thus forming a tighter loop which disables formation of the hydrogen bond from OW5 to Tyr151 O^{η}. OW8 is only found in BT–CMTI and links the inhibitor to the autolysis loop. OW7 in the salmon-trypsin complexes occupies the position close to the guanido group of Arg1I in the bovine-trypsin complexes.

Table 5
Secondary binding contacts (Å).

	BT–BPTI	BT–CPTI	ST–CPTI	ST– CMTI	BT–CMTI	PT–MCTI
P5 N–217 O^{γ}†	—	2.71	—	—	2.78	—
P3 N–216 O‡	4.84	3.12	3.13	3.05	3.30	3.11
P3 O–216 N	3.22	3.22	3.25	3.18	3.44	3.42
P2 O–192 $N^{\varepsilon 2}$§	2.89	3.19	3.33	3.45	3.41	3.27
P1′ N–195 O^{γ}	3.08	3.05	3.05	3.06	3.09	3.10
P2′ N–41 O	2.83	3.03	3.03	3.03	2.97	3.14
P2′ N^{ε}–40 O	3.34	—	—	—	—	—
P2′ $N^{\eta 2}$–40 O	2.81	—	—	—	—	—
P4′ N–39 O^{η}	2.98	—	—	—	—	—
19I N^{ζ}–217 O^{η}	—	—	—	—	—	2.84
39I N^{ε}–97 O	2.88	—	—	—	—	—
39I $N^{\eta 2}$–97 O	2.75	—	—	—	—	—

† Residue 217 is serine in bovine trypsin and tyrosine in salmon and porcine trypsin. ‡ Difference in amino-acid sequence. Cys in the squash inhibitors and Pro in BPTI. § $O^{\delta 1}$ in PT–MCTI.

Table 6
Close intramolecular distances (Å) involving the catalytic residues.

	ST–BENZ	BT–BENZ	BT–BPTI	BT–CPTI	ST–CPTI	ST–CMTI	BT–CMTI	PT–MCTI
57 $N^{\varepsilon 2}$–195 O^{γ}	3.17	2.96	2.62	2.56	2.53	2.65	2.65	2.80
57 $N^{\delta 1}$–102 $O^{\delta 2}$	2.73	2.67	2.62	2.64	2.70	2.55	2.68	2.75
102 $O^{\delta 1}$–56 N	2.92	2.90	2.72	2.81	2.93	3.02	2.74	2.79
102 $O^{\delta 1}$–57 N	2.75	2.76	3.11	2.95	2.93	3.00	2.97	2.92
102 $O^{\delta 2}$–214 O^{γ}	2.51	2.83	2.69	2.69	2.59	2.64	2.59	2.77
195 O^{γ}–P1 C	—	—	2.68	2.68	2.65	2.61	2.69	2.73
195 O^{γ}–P1 N	—	—	3.15	2.95	2.90	2.96	2.98	3.02
195 O^{γ}–P1 O	—	—	3.04	2.99	2.98	2.89	2.88	2.91
195 O^{γ}–P1′ N	—	—	3.08	3.05	3.05	3.06	3.09	3.10

example, Male *et al.*, 1995). The substitutions to proline do not seem to influence the structures, as the polypeptide chain exhibits almost exactly the same conformation in this region. However, the substitutions to proline reduce the flexibility of this loop considerably, as judged from the temperature factors, and could also restrict the relative mobility of the two domains of trypsin, as the segment from residue 23–29 participates in linking the two domains together.

The crystal structures presented above show that the two squash inhibitors, one with P1 arginine (CMTI) and one with P1 lysine (CPTI), bind to the active sites of both salmon and bovine trypsins in a classical manner in most respects. The approximate sixfold and threefold (salmon and bovine trypsins, respectively) higher association of CPTI (P1 lysine inhibitor) compared to CMTI (P1 arginine inhibitor) is, however, not straightforwardly interpretable from these results. The present study reveals that lysine at P1 is able to form direct hydrogen-bond interactions with the trypsin-specific Asp189 at the bottom of the binding pocket, and not only through a water molecule as previously assumed. The relatively close interaction between the P1-lysine N^{ζ} atom and the carboxylate O atom of Asp189 (3.2 and 3.4 Å for BT–CPTI and ST–CPTI, respectively) is not found in the BT–BPTI complex. The shorter distance for the CPTI complexes seems to arise from the substitution of a cysteine in position P2 of BPTI by a proline in CPTI. This introduces a different dihedral ψ angle on the P2 peptide, allowing the P1 lysine to bind slightly deeper in the binding pocket. Additionally, the side chain of Asp189 is slightly tilted in the trypsin–CPTI complexes compared with BT–BPTI, which is probably also a result of a stronger interaction of the former. According to Evnin *et al.* (1990), lysine/arginine substrate discrimination in trypsin depends on the volume at the bottom of the specificity pocket and the distribution of the negative charge of Asp189 and Ser190. The tight binding of both CMTI and CPTI to bovine trypsin could be caused by a more flexible specificity pocket which adjusts itself according to the P1 residue. Hedstrom *et al.* (1992) showed that the two loops from residues 184–188 and 221–225 are important for substrate affinity. Accommodation of the bulkier tyrosine at position 217, the hydrogen bond formed by the side chains of residues 221 and 224 and the proline at position 222 in salmon trypsin all contribute to a rigidification of the bottom of the specificity pocket and could be a reason for the higher substrate discrimination for salmon trypsin, as both trypsins prefer lysine at the primary binding site. Although the number of hydrogen bonds between the P1 side chain and trypsin is higher for arginine than for lysine and should, therefore, favour arginine binding, the larger size of arginine also results in closer intermolecular contacts that might be unfavourable (Fig. 4 and Table 4). Arginine to lysine substitution in the complex between chicken lysozyme and the Fab-fragment of a monoclonal antibody (Chacko *et al.*, 1995), shows a 100-fold preference for the arginine side chain. However, interactions between mutants of BPTI, STI (soybean trypsin inhibitor) or squash inhibitors with bovine and salmon trypsins, show only a moderate preference for lysine or arginine. Comparing the antibody–antigen and proteinase inhibitor complexes in terms of the hydrogen-bonding network does not easily account for the observed thermodynamic differences.

Studies of association between trypsin and small synthetic inhibitors which form interactions exclusively with the P1 binding pocket (Os & Otlewski, unpublished results), show that anionic salmon trypsin binds both lysine and arginine analogues up to 20 times more strongly than the bovine counterpart. Also, the Michaelis–Menten constant (K_M) for the binding of small synthetic substrates indicates that salmon trypsin forms interactions within the primary-binding pocket stronger than those formed by porcine and bovine trypsins (Outzen *et al.*, 1996). However, association constants for salmon and bovine trypsins are very similar for both squash

inhibitors. From these results it is likely that the increased primary binding observed for salmon trypsin is compensated for by weaker secondary interactions, which results in similar values of the association constant K_a for the binding of squash inhibitors to bovine and salmon trypsins.

The ST–CMTI complex was crystallized with the P1–P1′ peptide bond of CMTI in a hydrolyzed state. The crystal structure of the complex, however, shows that the peptide bond is intact and that there are no differences between this bond and the corresponding bond in the other complexes where the peptide bond was initially intact. A resynthesis of the cleaved bond has thus taken place during the formation of the complex. This is probably accomplished from the strong binding to trypsin, forcing the terminals of the cleaved peptide into a position where resynthesis can occur. The crystallographic refinements of all three complexes were carried out with no energy constraints on the P1 peptide, thus allowing an unbiased judgment of the state of the reactive site. There is no out-of-plane distortion for the reactive peptide in any of the three complexes, and no indication of partly cleaved inhibitors. On the other hand, as reported for other proteinase–inhibitor complexes (Bode & Huber, 1992), the distance between the active serine hydroxyl group of trypsin and the P1 peptide carbonyl C atom is shorter than expected from their van der Waals radii. The O^{γ}–C distances of about 2.6 Å indicate that the interaction is intermediate between a van der Waals interaction and a covalent bond. It should also be stressed that clear resynthesis of the reactive-site peptide bond and also all other interactions developed in the ST–CMTI complex indicate that the complex is exactly the same as that formed from the reactive-site-intact inhibitor.

We would like to thank the organizers at the Swiss–Norwegian Beamline (SNBL) at the European Synchrotron Radiation Facility (ESRF) for kindly supplying us with beamtime, and Hanna-Kirsti Schrøder for help during data collection.

References

Àsgeirsson, B. & Bjarnason, J. B. (1991). *Comput. Biochem. Physiol. B*, **99**, 327–335.

Àsgeirsson, B., Fox, J. W. & Bjarnason, J. B. (1989). *Eur. J. Biochem.* **180**, 85–94.

Bernstein, F. C., Koetzle, T. F., Williams, G. J, Meyer, E. E. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Bode, W., Greyling, H. J., Huber, R., Otlewski, J. & Wilusz, T. (1989). *FEBS Lett.* **242**, 285–292.

Bode, W. & Huber, R. (1992). *Eur. J. Biochem.* **204**, 433–451.

Bode, W. & Schwager, P. (1975). *J. Mol. Biol.* **98**, 693–717.

Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E. & Huber, R. (1982). *J. Mol. Biol.* **162**, 839–868.

Brünger, A. T. (1992). *X-PLOR. A system for X-ray crystallography and NMR.* New Haven: Yale University Press.

Chacko, S., Silverton, E., Kam-Morgan, L., Smith-Gill, S., Cohen, G. & Davies, D. (1995). *J. Mol. Biol.* **245**, 261–274.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 133–138.

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

Feller, G., Payan, F., Theys, F., Qian, M., Haser, R. & Gerday, C. (1994). *Eur. J. Biochem.* **222**, 441–447.

Gildberg, A. & Øverbø, K. (1990). *Comput. Biochem. Physiol. B*, **97**, 775–782.

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

Heimstad, E. S. (1996). Doctoral thesis, University of Tromsø, Norway.

Heimstad, E. S., Hansen, L. K. & Smalås, A. O. (1995). *Protein Eng.* **8**, 379–388.

Huang, Q., Liu, S. & Tang, Y. (1993). *J. Mol. Biol.* **229**, 1022–1036.

Huang, Q., Liu, S., Tang, Y., Zeng, F. & Qian, R. (1992). *FEBS Lett.* **297**, 143–146.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kleywegt, G. J. & Brünger, A. T. (1996). *Structure*, **4**, 897–904.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Kristjansson, M. M. & Nielsen, H. H. (1992). *Comput. Biochem. Physiol. B*, **101**, 247–253.

Krystek, S., Stouch, T. & Novotny, J. (1993). *J. Mol. Biol.* **234**, 661–679.

Laskowski, M. Jr & Kato, I. (1980). *Annu. Rev. Biochem.* **49**, 593–626.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

Male, R., Lorens, J. B., Smalås, A. O. & Torrissen, K. R. (1995). *Eur. J. Biochem.* **232**, 677–685.

Marquart, M., Walter, J., Diesenhofer, J., Bode, W. & Huber, R. (1983). *Acta Cryst.* B**39**, 480–490.

Minor, W. (1993). *DENZO*. Purdue University, West Lafayette, IN, USA.

Osnes, K. K. & Mohr, V. (1985). *Comput. Biochem. Physiol. B*, **82**, 607–619.

Otlewski, J. (1993). *Innovations in Proteases and Their Inhibitors*, edited by F. X. Aviles, pp. 369–388. New York, Berlin: Walter de Gruyter.

Otlewski, J. & Krowarsch, D. (1996). *Acta Biochim. Pol.* **43**, 431–444.

Otlewski, J. & Zbyryt, T. (1994). *Biochemistry*, **33**, 200–207.

Otlewski, J., Zbyryt, T., Krokoszynska, I. & Wilusz, T. (1990). *Biol. Chem. Hoppe-Seyler*, **371**, 589–594.

Otwinowski, Z. (1993). *DENZO*. Yale University, New Haven, CT, USA.

Outzen, H., Berglund, G. I., Smalås, A. O. & Willassen, N. P. (1996). *Comput. Biochem. Biophys. B*, **115**, 33–45.

Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–437.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.

Scheidig, A. J., Hynes, T. R., Pelletier, L. A., Wells, J. A. & Kossiakoff, A. A. (1997). *Protein Sci.* **6**, 1806–1824.

Schrøder, H. K., Willassen, N. P. & Smalås, A. O. (1998). *Acta Cryst.* D**54**, 780–798.

Simpson, B. K. & Haard, N. F. (1984). *Can. J. Biochem. Cell Biol.* **62**, 894–900.

Smalås, A. O., Heimstad, E. S., Hordvik, A., Willassen, N. P. & Male, R. (1994). *Proteins*, **20**, 149–166.

Smalås, A. O. & Hordvik, A. (1993). *Acta Cryst.* D**49**, 318–330.

Takeuchi, Y., Nonaka, T., Nakamura, K. T., Kojima, S., Miura, K.-I. & Mitsui, Y. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 4407–4411.