Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 313 (2004) 8–16
BBRC
www.elsevier.com/locate/ybbrc

Crystal structure of trypsin–turkey egg white inhibitor complex☆

B. Syed Ibrahim and Vasantha Pattabhi*

Department of crystallography and Biophysics, University of Madras, Chennai 600 025, India

Received 4 November 2003

Abstract

Crystal structure of the complex between porcine β-trypsin and the second domain of the Kazal-type ovomucoid turkey egg white trypsin inhibitor (OMTKY2) has been determined at 1.9 Å resolution. A peptide fragment from the first domain has been crystallized with the complex. Restrained-refinement of the structure led to an *R*-factor of 0.19 for the 32206 reflections. OMTKY2 exhibits the canonical Kazal-type fold with a central α-helix and a short two-stranded anti-parallel β-sheet. The carbonyl carbon of the reactive site prefers trigonal geometry. The reactive site loop geometry of the inhibitor is complementary to the surface and charge of the binding site in β-trypsin.

Keywords: Serine protease; Turkey egg white inhibitor; Kazal-type inhibitor; Peptide bound inhibitor

The inhibitors of serine proteases have been grouped into 18 different families [1] ranging from low to high molecular weight [2]. Ovomucoids, the major glycoprotein proteinase inhibitor from avian egg white such as chicken, turkey, and golden pheasant have long been known and have been widely studied. Based on their inhibitory properties, ovomucoids from various birds have been classified as single, double, and triple headed inhibitors. Interestingly, inhibitors from closely related birds were found to have different specificities.

Turkey egg white inhibitor is a double-headed inhibitor, where the second domain is specific for trypsin, and the third domain for chymotrypsin and elastase. The connecting peptide between the domains can easily be hydrolysed and the resulting domains are independently active. There is no evidence for noncovalent interactions among these domains [3]. The Kazal-type inhibitors consist of several domains, comprising characteristic disulphide bridges and high structural homology [4]. There are many inhibitors available in nature showing variance in size and shape, but they possess the same specific orientation of the loop at the binding site of the proteases [5]. It has been emphasized that the binding site residue P_1 (nomenclature as described by Schechter and Berger, [6]) plays a vital role in inhibiting the enzyme and is responsible for up to 50% of all contacts at the protein-inhibitor binding interface. The residues located in this region from P_3 to P_4' exhibit conserved conformation and interactions [7–9].

The reactive site of the inhibitor reacts with the active site of the enzyme in a substrate-like manner. In each case contact occurs over a small portion of the enzyme and the inhibitor, but over this portion the fit is excellent and numerous van der Waals interactions, hydrogen bonds, and salt bridges are formed. In each case complex formation occurs with relatively small conformational change and is predominantly a lock and key interaction, although a small conformational adaptation of the inhibitor does take place. In all the studies reported on trypsin-inhibitor complexes, the interesting feature was the geometry of C′ atom of the reactive site residue and in most cases it was concluded that this atom prefers a tetrahedral geometry [3]. The crystal structure of turkey egg white inhibitor is not known but that of the third domain has been reported previously [10]. The crystal structure of the second domain of this ovomucoid inhibitor as trypsin complex (OTPT) has been reported here. The crystal structures of the second

☆ Supplementry data associated with this article can be found, in the online version, at doi: 10.1016/j.bbrc.2003.11.082.

* Corresponding author. Fax: +91-44-223-00-122.
E-mail addresses: vasantha@unom.ac.in, pvasantha@hotmail.com (V. Pattabhi).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2003.11.082

and third domains (OMTKY3) have been compared to understand their specificities.

Materials and methods

Ovomucoid turkey egg white trypsin inhibitor (TEWI), porcine β-trypsin, and the substrate benzoyl arginyl *p*-nitroanilide (BAPNA) were purchased from Sigma Chemicals, USA. For the complex formation, 20 mg trypsin and 40 mg TEWI were dissolved in 500 μl of 10 mM acetate buffer, pH 5.3. The mixed sample was incubated for 1 h at 10 °C. This solution was loaded onto a Sephadex G-200 gel filtration column of size (1.5×50 cm), which had been previously equilibrated with the same buffer, with a flow rate of 1 ml/min. Three milliliters of fractions were collected per tube. The fractions containing complex were pooled after checking with UV (Cary) at 280 nm. The complexed protein was assayed for activity using BAPNA and showed no activity. Protein concentration of 22 mg/ml was used for crystallization.

Crystallization. Crystallization was achieved by the hanging drop vapour diffusion method using 24-well tissue culture plates at 20 °C. The hanging drop was prepared by mixing 4 μl protein and 2 μl reservoir solution. The reservoir solution contained 1.5 M ammonium sulphate, 10 mM zinc acetate, and 50 mM acetate buffer, pH 5.3. The complexed protein crystallized in the $P6_5$ space group. All the solutions used for crystallization were made in acetate buffer, pH 5.3.

X-ray data collection. X-ray data were collected on a single crystal of size ($0.2 \times 0.5 \times 0.9$ mm), mounted in 1 mm capillary tube, using MARXDS image plate, at Indian Institute of Science, Bangalore, India. The data were reduced and scaled using DENZO [11]. The crystal diffracted up to 1.9 Å. The cell parameters and other details are listed in Table 1.

Structure solution and refinement. The structure was solved (Fig. 1) by molecular replacement method using the CNS package [12] and ARP/wARP [13]. Porcine β-trypsin (BPT) was used as the model [14] (PDB1QQU) for molecular replacement. The best solution from rotation search was used for translation search in $P6_5$ space group, in different resolution ranges. The solution obtained was satisfactory and the rigid body refinement gave an *R*-factor of 0.37 and *R*-free 0.34 for 4.0–15.0 Å resolution. Further positional and isotropic temperature factor refinement with CNS brought the *R*-factor down to 0.28 and *R*-free to 0.33, for all the reflections. As the structure of the second domain is unknown, ARP/wARP program version 6.0 was used for building the inhibitor backbone. Every run of ARP/wARP program was followed by refinement using REFMAC5 [15]. Thus, the entire inhibitor structure was built after several cycles. Sixty-two residues of the inhibitor thus built were later mutated with the second domain sequence of common turkey egg white ovomucoid inhibitor (NiceProt View of Swiss-Prot: P01004) using O [16]. The mutated sequence fitted the electron density very well (Fig. 2). Several cycles of model building using $2F_o - F_c$ and $F_o - F_c$ maps were done which was followed by refinement. Apart from the second domain of the inhibitor a small peptide fragment consisting of 7 residues was located adjacent to the N-terminal of the inhibitor. The backbone density was at 1σ level, while the entire side chain appeared at 0.5σ level in the $2F_o - F_c$ map. The sequence of the hepta peptide was assigned by the prediction algorithm in O giving the C^{α} coordinates. The prediction algorithm identified a sequence in the third domain of the inhibitor (OMTKY3). However, several trials were made with many sequences available from both the first and third domains of the inhibitor, which were cleaved during complexation. Thus, 7 residues were fitted in the electron density of the inhibitor by trial and error. The best fitting sequence was identified to be in the N terminal region of the first domain from residue 12 to 18 (Fig. 3).

Water molecules were located from $F_o - F_c$ maps at 3σ level and refined using CNS. Some of the water molecules with high thermal factors were removed in subsequent cycles of refinement. The 1/8th

Table 1
Data collection statistics

Crystal data	
Space group	$P6_5$
Unit-cell parameters (Å)	$a = b = 98.575$, $C = 73.724$ Å $\alpha = \beta = 90.0°$, $\gamma = 120.0°$
Contents of asymmetric unit	1 β-trypsin 1 second domain of the ovomucoid turkey egg white inhibitor 1 hepta peptide from the first domain
Data statistics	
Resolution limits (Å)	80–1.9
No. of observations	189,883
No. of unique reflections	32,206
Completeness (%)	95.2 (92.8)
Multiplicity	3.24
R_{merge} (%)	0.078 (0.461)
Refinement and model statistics	
Resolution limits in refinement (Å)	80.0–1.9
Number of protein non-H atoms	2171
No. of solvent molecules	128
No. of observed reflections ($I/\sigma(I) > 0$)	30831
R_{cryst} (%)	0.19
R_{free} (%)[a]	0.21
Average *B*-factor (Å^2)	
Trypsin and inhibitor	34.1
Hepta peptide	78.5
R.m.s deviations from ideality	
Bond lengths (Å)	0.006
Bond angles (°)	1.236
Dihedral angle (°)	23.2
Ramachandran Map for the structure	
Residues in most favoured regions in (%)	88.7
Residues in additional allowed regions (%)	11.3

Values in parentheses are for the highest resolution shell.
[a] Free *R*-factor was calculated using 5% data.

Omit maps were calculated in order to remove any model bias and the stereochemistry for the final model was checked using PROCHECK [17]. The final *R*-factor was 0.19 and *R*-free 0.21 for all the reflections. There were no residues found in the disallowed region (Table 1). The coordinates have been deposited in the protein data bank (PDBID: 1PWN and RCSBID: RCSB019642). The residues in the additionally allowed regions of Ramachandran map are in general found in the flexible loop regions of the protein.

Results and discussion

Inhibitor structure

Ovomucoid trypsin inhibitor is a 27 kDa glycoprotein [18]; with three homologous domains, each approximately 60 residues long. The third domain is specific for chymotrypsin and elastase, while the second domain is

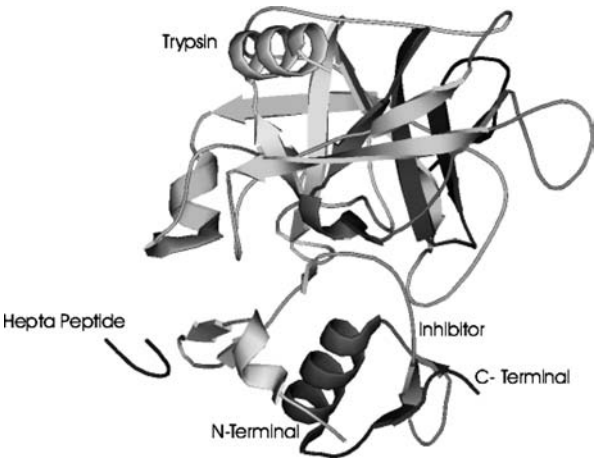

Fig. 1. Crystal structure of OTPT, generated with Pymol (DeLano, W.L.) The hepta peptide is shown closer to the inhibitor structure. The N-terminal of the OMTKY2 has assumed a helical conformation, which is not seen in other Kazal-type inhibitor structures.

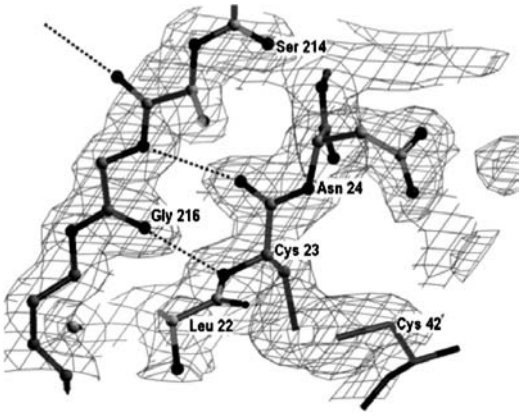

Fig. 2. $F_o - F_c$ Electron density map showing the interface between trypsin and inhibitor. Trypsin residues from 204 to 216 and the inhibitor residues from 22 to 24 form a β-sheet at the interface. The map is shown at 1σ level.

specific for trypsin [3]. The following is the complete sequence of common turkey egg white inhibitor (Nice-Prot View of Swiss-Prot: P01004). The sequence of the second domain, which binds to trypsin, is shown in bold. The other two domains were not found in the electron density suggesting cleavage by trypsin during crystallization. The hepta peptide sequence from the first domain that was located in the electron density has also been highlighted.

VEVDCSRFPN **TTNEEGKD**VL VCTEDLRPIC
GTDGVTHSEC LLCAYNIEYG TNISKEHDGE
CREA**VPMDCS** **RYPNTTSEEG** **KVMILCNKAL**
NPVCGTDGVT **YDNECVLCAH** **NLEQGTSVGK**
KHDGECRKEL AAVSVDCSEY PKPACTLEYR
PLCGSDNKTY GNKCNFCNAV VESNGTLTLS
HFGKC

The sequences of the active site region from P_4 to P_3' for the second and third domains show that only the P_3 position is conserved between the domains.

P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	
L	C	N	K	A	L	N	— Second domain
A	C	T	L	E	Y	R	— Third domain

The sequence alignment between these domains using ClustalW shows 31% identity. Both the domains have three-disulphide bridges in the same region of its structure and the reactive site residues of the domains are buried into the binding pocket of the respective enzyme. Out of the three-disulphide bridges two of them viz. (Cys 6**I**–Cys 45**I** and Cys 23**I**–Cys 42**I**) are located in the central α-helix region of the OMTKY2 inhibitor. As the peptide is cleaved at the N-terminal region of the inhibitor, this region exhibits relatively high thermal factors due to which the first two residues of the inhibitor OMTKY2 have only partial electron density. The third

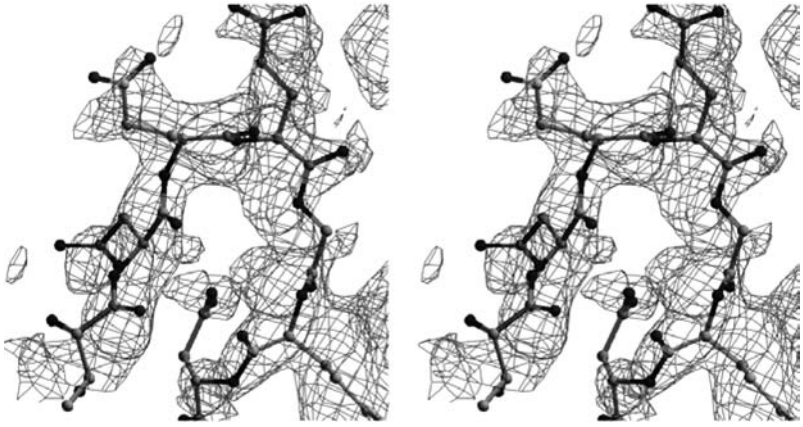

Fig. 3. Stereo view of the hepta peptide with the fitted electron density ($2F_o - F_c$ map). The electron density displayed covers all atoms including side chain at 0.5σ level.

disulphide bridge is formed between the C-terminal residue Cys 63**I** and Cys 31**I**. All the disulphide bridges are of length 2.02 Å (±0.01).

The structure of the second domain reported here is superimposable on the third domain, with an r.m.s deviation of 0.37 Å, except for a 10 residue long extra loop from 11I to 20I (Fig. 4), having an anti-parallel β-sheet structure. The overall secondary structure of inhibitor comprises two helical segments a central α-helix from Glu 40**I** to Gln 52**I**, and in the N-terminal region a small α-helix from Asp 5I to Tyr 9I; an anti-parallel β-sheet in the extra loop between Asn 11**I**–Thr 13**I** and Val19**I**–Ile 21**I**, and a β-strand between Leu 22**I**—Cys 23**I**. Apart from this an anti-parallel β-sheet forms between Thr 37**I**–Tyr 38**I**, Val 30**I**–Gly 32**I**, and Lys 57**I**–His 59**I**. These secondary structural elements interact appreciably to stabilize the intact structure of the OMTKY2 and the interactions are listed in Table 2.

Asn 24**I** which occupies P_2 position has a strong interaction with OD1of Asn 40**I** resulting in a compact structure for the inhibitor. The residues from P_4 to P_3' of both the inhibitor domains (i.e., OMTKY2 and OMTKY3) have similar interactions with the respective enzymes, excepting for the reactive site residue (P_1) interactions of the inhibitors. The recognition site of the second domain consists of three water molecules while in the third domain there are five water molecules in this region. In the case of chymotrypsin the binding pocket is not completely occupied by the reactive site residue side chain and hence more water molecules are required to stabilize this pocket.

Protein–protein interaction

The extra loop in the present structure interacts with Gln 175, LYS 224, TYR 217, TYR 20, and Lys 159 of trypsin and stabilizes it. Gln 175 of trypsin has slightly changed its orientation as compared to free trypsin (PDB1QQU) to stabilize the position of the extra arm of the inhibitor by this interaction. 16 out of 62 residues of the inhibitor make 128 contacts (2.6–4.0 Å) with trypsin. Residues Lys 60, Ser 149, and Tyr 217 of trypsin found in loop regions interact with the inhibitor through both main chain and side chain interactions. Lys 25I (P_1) and Ala 26I (P_1') have numerous contacts with trypsin in the binding pocket and the active site of trypsin. Almost 50% of the interactions are through reactive site residue as seen in other inhibitor complexes. The side chain of Lys 25I (P_1) is oriented in its expected position and occupies the specificity pocket by replacing water molecules in the functional water network found in free trypsin [19]. The NZ atom of Lys 25I interacts with OD1 and OD2 of Asp189 (Fig. 5) of trypsin and one of the interactions is mediated through conserved water (W37) and is responsible for orienting the side chain of the reactive site residue in the binding pocket. The water molecule W37 in OTPT is well ordered and is hydrogen bonded with the carbonyl oxygen of Gly 219. All the other three-conserved water molecules (W33, W9, and W2) found in this region have strong interactions with neighbouring residues (Table 2). The non bonded distance between C′ of Lys 25I and Ser 195 OG is 2.70 Å, which falls in the short contact range and is comparable to those observed in other trypsin-inhibitor complexes.

Fig. 4. Super position of C^{α} trace of the second domain OMTKY2 (black) and third domain OMTKY3 (grey) using INSIGHT II. The extra loop missing in the third domain is in anti-parallel β-sheet structure.

Cys 23**I** occupies the P_3 position in the inhibitor. Bittergourd trypsin inhibitor (1MCT) [20], OMTKY3, and OMTKY2 are the only inhibitors having the same residue at this site. P_3 interactions in these inhibitor structures are similar and the carbonyl oxygen is hydrogen bonded to NH of Gly 216 of trypsin. OH···O type of hydrogen bond was found between the hydroxyl group of Tyr 217 of trypsin and Glu 50**I**. Some of the important interactions are listed in Table 2. The tight binding of the inhibitor with the enzyme explains its complementary nature to the binding cleft of trypsin. Several interactions in the interface of enzyme and inhibitor could explain the strong binding of inhibitor. Either side of the reactive site residue is involved in protein–protein interactions, e.g., Leu 22I–Cys 23I and Gln 204–Gly 216 of trypsin form a β-sheet at the interface, while on the other side of reactive site loop NH···O and CH···O type of hydrogen bonds have been observed. Some of the interactions in the interface are listed in Table 2. The charge surface calculation was done using SPDBV viewer [21]. The active site and specificity pocket of trypsin are negatively charged and the surface of the inhibitor is covered with positively charged residues such as Asn 24I, Lys 25I, Asn 28I and hydrophobic residues Ala 26I, Leu 27I.

Reactive site geometry

It has been observed that all serine protease inhibitors share similar conformation at the binding loop (P_4 to P_3'), even though they have different inter scaffolding [7].

Table 2
Some of the interactions between trypsin and inhibitor

10 B PRO	CB	....	175	A GLN	OE1						3.29
20 B MET	SD	....	224	A LYS	NZ						3.17
20 B MET	CE	....	224	A LYS	NZ						3.38
21 B ILE	O	....	217	A TYR	CD2						3.54
50 B GLU	CD	....	217	A TYR	OH						3.48
50 B GLU	OE2	....	217	A TYR	CZ						3.58
50 B GLU	OE2	....	217	A TYR	OH						2.68
15 B GLU	OE2	....	159	A LYS	CD	*	5	0	0	0	3.39
15 B GLU	O	....	159	A LYS	CD	*	5	0	0	0	2.98
28 B ASN	OD1	....	60	A LYS	CE						3.38
28 B ASN	OD1	....	60	A LYS	NZ						3.24
39 B ASP	OD1	....	149	A SER	OG						3.32
25 B LYS	NZ	....	190	A SER	O						3.01
25 B LYS	NZ	....	189	A ASP	OD1						3.44
25 B LYS	NZ	....	190	A SER	OG						2.91
25 B LYS	O	....	195	A SER	OG						3.05
25 B LYS	O	....	193	A GLY	N						2.65
25 B LYS	O	....	195	A SER	N						3.04
26 B ALA	N	....	195	A SER	OG						3.05
28 B ASN	OD1	....	60	A LYS	NZ						3.24
37 B THR	N	....	93	A ASN	OD1	*	4	0	-1	-1	2.92
39 B ASP	OD1	....	149	A SER	OG						3.32
39 B ASP	OD1	....	151	A TYR	OH						2.67
Water interactions stabilizing the calcium binding loop:											
70 A GLU	OE1	....	W36	HOH	O						2.63
72 A ASN	O	....	W36	HOH	O						2.84
75 A VAL	O	....	W36	HOH	O						2.65
80 A GLU	OE1	....	W36	HOH	O						2.61
Water interactions at the active site:											
189 A ASP	OD2	....	W2	HOH	O						3.36
224 A LYS	O	....	W2	HOH	O						2.64
22laA ALA	N	....	W2	HOH	O						3.17
221 A GLN	N	....	W2	HOH	O						2.74
217 A TYR	O	....	W2	HOH	O						2.79
227 A VAL	O	....	W9	HOH	O						2.85
215 A TRP	O	....	W9	HOH	O						3.04
190 A SER	OG	....	W9	HOH	O						3.15
25 B LYS	NZ	....	W9	HOH	O						2.85
216 A GLY	N	....	W33	HOH	O						3.14
219 A GLY	O	....	W33	HOH	O						2.87
192 A GLN	OE1	....	W33	HOH	O						2.95
23 B CYS	O	....	W33	HOH	O						3.16
25 A LYS	NZ	....	W37	HOH	O						2.98
189 A ASP	OD2	....	W37	HOH	O						2.55
219 A GLY	O	....	W37	HOH	O						2.95
25 B LYS	CE	....	W37	HOH	O						3.47
Strong interactions at the interface of trypsin and inhibitor (0MTKY2):											
23 B CYS	N	....	216	A GLY	O						3.08
23 B CYS	O	....	216	A GLY	N						3.18
24 B ASN	O	....	192	A GLN	OE1						3.16
25 B LYS	N	....	195	A SER	OG						2.87
25 B LYS	N	....	214	A SER	O						3.09
25 B LYS	O	....	194	A ASP	N						3.34
27 B LEU	N	....	41	A PHE	O						2.93
27 B LEU	O	....	41	A PHE	CB						3.20
Interactions stabilizing the inhibitor:											
40 B ASN	OD1	....	24	B ASN	O						3.35
40 B ASN	OD1	....	24	B ASN	O						3.35
11 B ASN	CA	....	20	B MET	O						3.09
11 B ASN	CB	....	20	B MET	O						3.53
11 B ASN	C	....	20	B MET	O						3.38
12 B THR	N	....	20	B MET	O						2.78

Table 2(*continued*)

12 B THR	O	....	19	B VAL	CA	3.50
12 B THR	O	....	19	B VAL	C	3.55
12 B THR	O	....	20	B MET	N	2.69
30 B VAL	N	....	38	B TYR	O	2.97
30 B VAL	O	....	37	B THR	CA	3.43
30 B VAL	O	....	38	B TYR	N	2.92
31 B CYS	N	....	59	B HIS	O	2.91
31 B CYS	O	....	58	B LYS	N	2.70
31 B CYS	O	....	58	B LYS	CA	3.44
31 B CYS	O	....	58	B LYS	CB	3.28
31 B CYS	O	....	59	B HIS	N	3.42
31 B CYS	CB	....	59	B HIS	O	3.67
31 B CYS	CB	....	59	B HIS	NE2	3.37

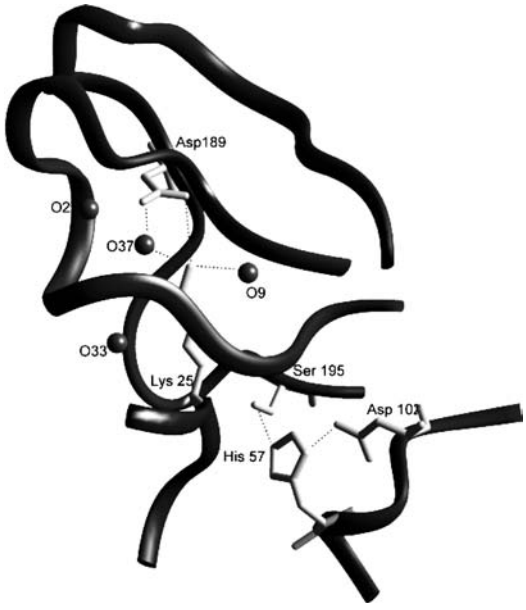

Fig. 5. Ribbon diagram of binding pocket of trypsin and reactive site of the inhibitor in OTPT complex. The recognition amino acid Asp 189, active site residues His 57, Asp 102, Ser 195, and Lys 25 of the inhibitor are shown in stick model. The water molecules in the binding pocket of OTPT shown as balls. Three water molecules forming a water network from Ser195 OG to Asp189 in free trypsin have been replaced by the side chain of Lys 25I.

The phi/psi values for this region have been tabulated (Table 3) for comparison. Table 3 indicates that the reactive site peptides prefer conserved orientation, though the sequence is not conserved. The inhibitors compared in all the cases have strong interactions with their respective enzyme in a lock and key fashion. Apart from these interactions the conserved orientation is also one of the reasons for inhibition of the enzyme, as mutations in this region lead either to destabilization of the binding loop conformation [25] or lower affinity than that of the wild-type protein [26]. All these facts indicate strongly that the specific binding controls the inhibition of an enzyme.

As mentioned earlier the reactive site of OMTKY2 has Lys 25I (P_1) and Ala 26I ($P_{1'}$) in OMTKY2. The geometry of the P_1 residue is most important in understanding the interaction between inhibitor and the enzyme during catalysis. It has been reported that the C′ atom of P1 residue assumes a tetrahedral geometry which was negated by Song et al. [24]. In the present structure the results agree with those reported by Song et al. In order to assess the effect of restraint refinement on this geometry, unrestrained refinement was carried out. The distortion of C' atom of reactive site residue (Lys 25**I**) was calculated through plane calculation, which shows a deviation of 0.007 Å from the plane and increases to 0.052 Å in the unrestrained model. As a control experiment the same was checked for Leu 21I residue, which had a deviation to 0.008 Å in the restrained refinement and 0.024 Å in the unrestrained structure. As the deviations from planarity are insignificant it can be concluded that C′ atom is in trigonal geometry in this structure.

Peptide fragment

A small peptide fragment consisting of 7 residues was identified near the N-terminal region of the OMTKY2. It has strong interactions with OMTKY2 as well as trypsin. The sequence of this hepta peptide was arrived at, based on the cleavage sites of the inhibitor and the electron density. The hepta peptide assumes a β-turn structure with 3rd and 4th residues at corners, but $(i + 1)$ to $(i + 3)$ hydrogen bond is absent. This hepta peptide is held strongly by Asn 11**I** of OMTKY2 and is situated near the surface of calcium binding loop of trypsin. Trypsin residues namely His 71, Asn 72, Asp 74, Ser 153, and Leu 154 (Table 4) interact strongly with the hepta peptide. The loop starting from 71 to 80 (calcium binding loop) of trypsin deviates significantly from the native trypsin (r.m.s. deviation 1.07 Å) conformation to facilitate interactions with the hepta peptide and thus stabilizing it. The phi/psi values of this hepta peptide are listed in Table 5.

Table 3
Comparison of dihedral angles (ϕ/Ψ) of the reactive site loop residues (P_4–$P_{3'}$) of inhibitors in trypsin complexes and in free state

	P_4	P_3	P_2	P_1	$P_{1'}$	$P_{2'}$	$P_{3'}$
Sequence							
OMTKY2[a]	L	C	N	K	A	L	N
BPTI	G	P	C	K	A	R	I
MCTI	I	C	P	R	I	W	M
STI	S	P	Y	R	I	R	F
ETI	S	R	L	R	S	A	F
Conformation(ϕ/Ψ)							
Complex structures							
OMTKY2[a]	−94/123	−114/144	−70/161	−105/35	−87/151	−101/94	−130/74
BPTI: BPT	78/174	−77/−29	−70/155	−116/39	−87/164	−112/79	−98/124
STI: PPT	−114/144	−58/−34	−56/139	−89/38	−83/148	−67/−38	−118/155
MCTI	−58/117	−137/112	−54/147	−91/34	−82/146	−84/125	−152/127
Free inhibitor structures							
BPTI	86/176	−86/−6	−81/165	−104/9	−76/172	−127/76	−105/121
ETI	−88/171	−66/−30	−112/150	−66/134	−169/151	−74/−46	−106/163
Free STI	−112/149	−70/−28	−62/158	−84/21	−60/148	−90/−22	−127/126

Torsion angles at P_2, P_1, and $P_{1'}$ are highly conserved in the complex structures.
[a] OMTKY2 present study; BPTI, [22]; ETI, [23]; and STI:PPT (orthorhombic), [24].

Table 4
Some of the interactions of the hepta peptide with trypsin and inhibitor

6 C LYS	N	....	121	HOH	O						3.14
11 B ASN	CG	....	6	C LYS	NZ						3.42
11 B ASN	OD1	....	6	C LYS	NZ						2.74
11 B ASN	ND2	....	6	C LYS	NZ						3.67
3 C GLU	OE1	....	71	A HIS	NE2	*	4	0	−1	0	3.42
3 C GLU	OE2	....	71	A HIS	NE2	*	4	0	−1	0	3.09
2 C ASN	ND2	....	72	A ASN	OD1	*	4	0	−1	0	2.98
2 C ASN	ND2	....	72	A ASN	ND2	*	4	0	−1	0	3.36
2 C ASN	ND2	....	74	A ASP	CB	*	4	0	−1	0	2.95
2 C ASN	ND2	....	74	A ASP	CG	*	4	0	−1	0	3.43
2 C ASN	ND2	....	74	A ASP	OD2	*	4	0	−1	0	3.03
2 C ASN	OD1	....	153	A SER	OG	*	4	0	−1	0	2.78
4 C GLU	**OE2**	....	**154**	**ALEU**	**CD1**	*	**4**	**0**	**−1**	**0**	**2.85**

Table 5
(ϕ/Ψ) angles of hepta peptide

Res. Name	Phi (ϕ)	Psi (Ψ)
Thr	79.1	−67.5
Asn	−33.8	144.0
Glu	−127.4	−127.1
Glu	−119.8	58.1
Gly	−159.8	−120.0
Lys	−139.3	155.2
Asp	159.1	34.1

Trypsin in the free and complexed state

Superposition of trypsin in the complexed and free state (BPT) was done using INSIGHTII [27] and LSQ_MOL of O [16]. The r.m.s deviation including side chain atoms was 2.92 Å while the deviation for main chain atoms only was 0.56 Å. However, the deviations of main chain atoms in the loop regions were significantly higher (2.98 Å). Main differences were observed only at the active site region. The residues involved in inhibitor interactions have moved from the native trypsin position to facilitate anchoring of the reactive site residues.

The hydrogen bond geometry at the active site is in agreement with other complex structures and the hydrogen bond strength is more as compared to the native trypsin (Table 6). The shorter hydrogen bond distances in the present complex (OTPT) indicate rigid packing at the active site. The widening of hydrogen bond angle OG of Ser195$\cdots$ N of His57 as compared to native trypsin is due to Ser 195 OG atom moving away from its original position during complexation. The salt bridges found in the native trypsin have been retained in the complex with minor variations and 4 additional salt bridges were found within the inhibitor. There are a total of 482 hydrogen bonds calculated using HbPlus program [28] in the present complex structure, including all the water molecules.

Table 6
H-bond geometry at the active site region of trypsin

	OTPT (Å)	BPT (Å)	AVW (Å)	MCT (Å)
His 57 ND1–Asp 102 OD2	2.61	2.73	2.68	2.75
His 57 ND1–Asp 102 OD1	3.36	3.57	3.27	3.43
His 57 N–Asp 102 OD1	2.94	2.90	2.91	2.93
His 57 NE2–Ser 195 OG	2.67	3.07	2.52	2.80
Asp 189 OD2–Ala 221 N	3.00	2.91	2.95	2.77
Angle				
Ser 195 CB–OG–His 57NE2	96.63°	92.03°	103.95°	97.3°

AVW [26], MCT [20]

Ca^{2+} ion binds and stabilizes trypsin [29] by reducing the autolysis of trypsin [30]. Most of the structures so far solved have Ca^{2+} ion in the loop region Lys 60–Glu 80 of trypsin while there is no Ca^{2+} ion in the present study, as Ca^{2+} was not used in the preparation or crystallization of the complex. A water molecule W36 has replaced Ca^{2+} ion and it interacts with Glu 80, Glu 70, Asn 72, and Val 75. Almost all the residues, excepting Glu 77, which was interacting with Ca^{2+} in BPT, are involved in Ca^{2+} interactions. The hepta peptide found in the structure offers additional stability to this loop and vice versa.

Solvent interactions

Solvent molecules at the active site were compared with native trypsin and other complex structures. A network of water molecules at the active site of trypsin has been reported to play an important role in the activity of trypsin [19]. Three water molecules from this network are replaced in most of the trypsin-inhibitor complexes by the inhibitor, which is in agreement with the present study. The water molecules (W77, W43, and W86) forming the water network in α-trypsin (1AKS) are replaced by the side chain of Lys 25**I** in OTPT. Only one water molecule (W37) from this network is not expelled if the specificity pocket is occupied by lysine rather than arginine. W9 is the invariant water molecule in OTPT found in almost all the complexes and free trypsin. It interacts with both Asp189 of trypsin and Lys 25**I**, in such a way as to stabilize the side chain of Lys 25**I**. Some of the water molecules are involved in the water-mediated interaction between trypsin and inhibitor. Cys 23**I** interacts with Gln 192 and Gly 219 through a water molecule W33. A similar network was found between Leu 27**I** and Ser 39 through the water molecule W46.

Acknowledgments

We thank Department of Science and Technology, India, Center for Scientific and Industrial research (S.I), for financial assistance. We acknowledge the use of National facility for Macromolecular data collection at IISc., Bangalore, India.

References

[1] M. Laskowski Jr., M.A. Qasim, What can the structures of enzyme–inhibitor complexes tell us about the structures of enzyme–substrate complexes?, Biochim. Biophys. Acta 1477 (2000) 324–337.
[2] W. Bode, R. Huber, Natural protein proteinase inhibitors and their interaction with proteinases, Eur. J. Biochem. 204 (1992) 433–451.
[3] M. Laskowski Jr., I. Kato, Protein inhibitors of proteinases, Ann. Rev. Biochem. 49 (1980) 593–626.
[4] B. Schlott, J. Wohnert, C. Icke, M. Hartmann, R. Ramachandran, K.H. Guhrs, E. Glusa, J. Flemming, M. Gorlach, F. Grobe, O. Ohlenschlager, Interaction of Kazal-type inhibitor domains with serine proteinases: biochemical and structural studies, J. Mol. Biol. 318 (2002) 533–546.
[5] W. Apostoluk, J. Otlewski, Variability of the canonical loop conformations in serine proteinases inhibitors and other proteins, Proteins: Struct. Funct. Genet. 32 (1998) 459–474.
[6] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain, Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[7] K. Huang, W. Lu, S. Anderson, M. Laskowski Jr., M.N. James, Water molecules participate in proteinase–inhibitor interactions: crystal structures of Leu18, Ala18, and Gly18 variants of turkey ovomucoid inhibitor third domain complexed with *Streptomyces griseus* proteinase B, Protein Sci. 4 (1995) 1985–1997.
[8] W. Lu, I. Apostol, M.A. Qasim, N. Warne, R. Wynn, W.L. Zhang, S. Anderson, W.Y. Chiang, E. Ogin, I. Rothberg, K. Ryan, M. Laskowski Jr., Binding of amino acid side-chains to S1 cavities of serine proteinases, J. Mol. Biol. 266 (1997) 441–461.
[9] D. Krowarsch, M. Dadlez, O. Buczek, I. Krokoszynska, A.O. Smalas, J. Otlewski, Interscaffolding additivity: binding of P_1 variants of bovine pancreatic trypsin inhibitor to four serine proteases, J. Mol. Biol. 289 (1999) 175–186.
[10] M. Fujinaga, A.R. Sielecki, R.J. Read, W. Ardelt, M. Laskowski Jr., M.N.G. James, Crystal and molecular structures of the complex of α-chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution, J. Mol. Biol. 195 (1987) 397–418.
[11] Z. Otwinowski, Oscillation Data Reduction Program, in: L. Sawyer, N. Isaacs, S. Bailey (Eds.), Proceeding of the CCP4 Study Weekend: "Data Collection and Processing", SERC Daresbury Laboratory, England, 1993, pp. 56–62.
[12] A.T. Brunger, P.D. Adams, G. Marius Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J. Jiang-Sheng, J. Kuszewski, M. Nilger, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography and NMR system: a new software suite for macromolecular structure determination, Acta Cryst. D 54 (1998) 905–921.
[13] R.J. Morris, A. Perrakis, V.S. Lamzin, ARP/wARP's model-building algorithms. I. The main chain, Acta Cryst. D 58 (2002) 968–975.

[14] A. Johnson, N. Gautham, V. Pattabhi, Crystal structure at 1.63 Å resolution of the native form of porcine β-trypsin: revealing an acetate ion binding site and functional water netwotk, Biochim. Biophys. Acta V 1435 (1999) 7–21.
[15] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Cryst. D 53 (1997) 204–255.
[16] T.A. Jones, J-Y. Zou, S.W. Cowan, M. Kjeldgard, Improved methods for building protein models in electron density maps and the location of errors in these models, Acta Cryst. A 47 (1991) 110–119.
[17] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thronton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crtystallogr. 26 (1993) 283–291.
[18] H. Lineweaver, C.W. Murray, Identification of the trypsin inhibitor of egg white with ovomucoid, J. Biol. Chem. 171 (1947) 565.
[19] A. Johnson, L. Damodharan, V. Pattabhi, J. Biochem. Mol. Biol. Biophys. V 5 (2001) 559–565.
[20] Q. Huang, S. Liu, Y. Tang, Refined 1.6 Å resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the squash family. Detailed comparison with bovine beta-trypsin and its complex, J. Mol. Biol. V 229 (1993) 1022.
[21] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-Pdb Viewer (1997).
[22] A. Wlodawer, J. Deisenhofer, R. Huber, Comparison of two highly refined structures of bovine pancreatic inhibitor, J. Mol. Biol. 193 (1987) 145–156.
[23] S. Onesti, P. Brick, D.M. Blow, Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra seeds*, J. Mol. Biol 217 (1991) 153–176.
[24] A. Grzesiak, R. Helland, A.O. Smalas, D. Krowarsch, M. Dadlez, J. Otlewski, Substitutions at the P1′ position in BPTI strongly affect the association energy with serine proteinases, J. Mol. Biol. 301 (2000) 205–217.
[25] H. Czapinska, J. Otlewski, S. Krzywda, G.M. Sheldrick, M. Jaskolski, High-resolution structure of bovine pancreatic trypsin inhibitor with altered binding loop sequence, J. Mol. Biol. 295 (2000) 1237–1249.
[26] H.K. Song, S.W. Suh, Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator, J. Mol. Biol. 275 (1998) 347–363.
[27] Biosym/MSI Release, San Diego, CA, USA. 95.0 92121–3752.
[28] I.K. McDonald, J.M. Thronton, Satisfying hydrogen bonding potential in proteins, J. Mol. Biol 238 (1994) 777–793.
[29] R.B. Martin, In metal ions, in: H. Siger (Ed.), Biologicl Systems, V17, Marcel Dekker, New York, 1984, pp. 1–49.
[30] D. Gabel, V. Kasche, Autolysis of β-trypsin Influence of calcium ions and heat, Acta Chem. Scandinavica. V 27 (1973) 1971–1981.