

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

Hyun Kyu Song and Se Won Suh*

Department of Chemistry and
Center for Molecular Catalysis
College of Natural Sciences
Seoul National University
Seoul 151-742, Korea

The Kunitz-type trypsin inhibitor from soybean (STI) consists of 181 amino acid residues with two disulfide bridges. Its crystal structures have been determined in complex with porcine pancreatic trypsin in two crystal forms (an orthorhombic form at 1.75 Å resolution and a tetragonal form at 1.9 Å) and in the free state at 2.3 Å resolution. They have been refined to crystallographic *R*-values of 18.9%, 21.6% and 19.8%, respectively. The three models of STI reported here represent a significant improvement over the partial inhibitor structure in the complex, which was previously determined at a nominal resolution of 2.6 Å by the multiple isomorphous replacement method. This study provides the first high-resolution picture of the complex between a Kunitz-type proteinase inhibitor with its cognate proteinase. Many of the external loops of STI show high *B*-factors, both in the free and the complexed states, except the reactive site loop whose *B*-factors are dramatically reduced upon complexation. The reactive site loop of STI adopts a canonical conformation similar to those in other substrate-like inhibitors. The P1 carbonyl group displays no out-of-plane displacement and thus retains a nominal trigonal planar geometry. Modeling studies on the complex between a homologous Kunitz-type trypsin inhibitor DE-3 from *Erythrina caffra* and the human tissue-type plasminogen activator reveal a new insight into the specific interactions which could play a crucial role in their binding.

© 1998 Academic Press Limited

*Corresponding author

Keywords: crystal structure; Kunitz-type proteinase inhibitor; soybean trypsin inhibitor; β-trefoil fold; inhibition of plasminogen activator

Abbreviations used: BPT, bovine pancreatic trypsin; BPTI, bovine pancreatic trypsin inhibitor; ETI, trypsin inhibitor DE-3 from *Erythrina caffra*; ETIa, trypsin inhibitor a from *Erythrina variegata*; ETIb, trypsin inhibitor b from *Erythrina variegata*; Hepes, *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; MCTI, trypsin inhibitor from bitter gourd seeds (*Momordica charantia*); MIR, multiple isomorphous replacement; ortho, orthorhombic crystal; PEG, polyethylene glycol; PPT, porcine pancreatic trypsin; PSTI, pancreatic secretory trypsin inhibitor; r.m.s., root-mean-square; SSI, *Streptomyces subtilisin* inhibitor; STI, Kunitz-type soybean trypsin inhibitor; tetra, tetragonal crystal; tPA, tissue-type plasminogen activator; WCI, Kunitz-type chymotrypsin inhibitor from winged bean seeds.

Introduction

The Kunitz-type soybean trypsin inhibitor (STI), an archetypal member of its family, was characterized and crystallized as early as in 1947 (Kunitz, 1947a,b). Its partial structure in complex with porcine pancreatic trypsin (PPT) was reported in 1974 (Blow *et al.*, 1974; Sweet *et al.*, 1974), and amino acid sequences were reported for the three isoforms of STI, Ti^a, Ti^b, and Ti^c (Kim *et al.*, 1985). They all contain 181 amino acid residues with two disulfide bridges (Cys39-Cys86 and Cys136-Cys145). Between the sequences of Ti^a and Ti^c, there is only one substitution at position 55. On the other hand, eight amino acid substitutions exist between Ti^a and Ti^b (Kim *et al.*, 1985). Isolation of

the cDNA clones encoding the precursors for Ti^a and Ti^b was reported (Song *et al.*, 1991).

The mechanism of inhibition of PPT by STI was proposed (Blow *et al.*, 1974; Baillargeon *et al.*, 1980). A stable complex between the virgin inhibitor and trypsin is rapidly formed and dissociates very slowly into the free enzyme and a mixture of unmodified and modified inhibitors. The modified STI is specifically cleaved at the scissile peptide bond between Arg63 (P1 residue) and Ile64 (P1' residue). In the refined structures of complex between bovine pancreatic trypsin inhibitor (BPTI) and bovine pancreatic trypsin (BPT), the scissile peptide bond remains nearly intact, with a slight out-of-plane deformation of the carbonyl oxygen atom being observed (Bode & Huber, 1992). However, the absence of a large upfield shift in the ¹³C of the P1 residue of STI upon complex formation suggested that the STI:PPT complex is not a covalent, fully tetrahedral adduct (Baillargeon *et al.*, 1980). More recently, no out-of-plane distortion around the inhibitor's scissile peptide bond was observed in the refined structure of the complex between porcine β -trypsin and MCTI-A, a trypsin inhibitor from the squash family (Huang *et al.*, 1993).

Other Kunitz-type proteinase inhibitors from various sources show a high degree of homology to STI. Among them, three-dimensional structures have been reported for trypsin inhibitor DE-3 from *Erythrina caffra* (ETI) at 2.5 Å (Onesti *et al.*, 1991), proteinase K/ α -amylase inhibitor from wheat at 2.5 Å (Zemke *et al.*, 1991), and Kunitz-type chymotrypsin inhibitor from winged bean seeds (WCI) at 2.95 Å (Dattagupta *et al.*, 1996), all in the free state. Other homologous proteins include a trypsin inhibitor from winged bean seeds (Yamamoto *et al.*, 1983), potato cathepsin D inhibitor (Mareš *et al.*, 1989), inhibitors of subtilisin/endogeneous α -amylase from barley and wheat (Leah & Mundy, 1989), and winged bean albumin-1 (Kortt *et al.*, 1989). Kunitz-type proteinase inhibitors are further divided into three groups based on their abilities to inhibit chymotrypsin, trypsin and tissue-type plasminogen activator (tPA). tPA is currently being used clinically as a thrombolytic agent. Group a inhibitors are relatively specific for chymotrypsin but are poor inhibitors of trypsin, and they do not inhibit tPA. Group b inhibitors are more specific towards trypsin than chymotrypsin and they do not inhibit tPA, either. Group c inhibitors inhibit trypsin, chymotrypsin, and tPA. Some proteinase inhibitors of the Kunitz-type in the seeds of *Erythrina* species belong to group c and are of particular interest due to their unique ability to bind to and to inhibit tPA (Heussen *et al.*, 1984). ETI, for example, provides an effective affinity reagent for the one-step purification of both one- and two-chain forms of tPA (Heussen *et al.*, 1984).

The crystal structure of the STI:PPT complex was previously determined at a nominal resolution of 2.6 Å by the MIR method (Blow *et al.*, 1974; Sweet *et al.*, 1974). However, only a very incomplete

model of the inhibitor could be built, since the X-ray data used were relatively incomplete (56%) in the 3.0 to 2.6 Å range and a significant part of the inhibitor showed only poor electron density. Until now, there has been no report of a highly refined structure of any Kunitz-type proteinase inhibitor in complex with its cognate proteinase. Furthermore, the structure of STI in the free state has not yet been reported, precluding a detailed analysis of structural changes associated with complexation with a proteinase.

Therefore, there exists a need for high-resolution crystallographic analyses of STI, in complex with a proteinase as well as in the free state, in order to provide a high-resolution picture of STI and also to gain insight into the interactions between a Kunitz-type proteinase inhibitor with a serine proteinase such as trypsin. In this paper, we report the determination of three structures of STI, that is, the inhibitor alone at 2.3 Å resolution and its complex with PPT in two crystal forms, an orthorhombic crystal form at 1.75 Å resolution and a tetragonal form at 1.9 Å. The structure of STI:PPT complex reported in this study provides a basis for modeling studies to gain insights into the interaction mode between homologous inhibitors and proteinases, including that between ETI and tPA. Modeling experiments were undertaken to understand why ETI, but not STI, is capable of inhibiting tPA.

Results and Discussion

Quality of refined models

Table 1 summarizes the refinement statistics as well as model quality parameters. All three models of STI:PPT complex in two crystal forms and free STI have reasonable crystallographic *R*-factors and excellent stereochemistry. In this paper, residues of PPT and other homologous proteinases are numbered according to the chymotrypsin sequence numbering scheme and a prime (') is appended to the residue number for clarity.

Missing from the inhibitor part of STI:PPT(ortho) model are two segments (residues 125 to 126 and 139 to 142) and the four C-terminal residues (178 to 181). In this model, the electron density for the side-chains of the following residues in the inhibitor, mostly in the loop regions, is either poor or absent: 98, 109, 111 and 112, 124, 127, 138, 143 and 144, 154 and 155, 165 to 167 and 177. The residues 125 and 126 missing in this model are clearly seen in the STI:PPT(tetra) model, whereas the segment 139 to 143 and the four C-terminal residues (178 to 181) are still missing. Missing from the free STI model are two segments (125 and 126; 140 to 142) and the four C-terminal residues (178 to 181), with the electron density for the side-chains of the following residues being either poor or absent: 38, 98, 111 and 112, 124, 127, 138 and 139, 143 and 144, 154, 165 to 167, and 177. Only a single residue, Gln138, in the free STI model is in the disallowed region of Ramachandran plot. Electron density

Table 1. Refinement statistics

	STI:PPT(ortho)	STI:PPT(tetra)	Free STI
Resolution range (Å)	8.0–1.75	8.0–1.90	8.0–2.30
No. of reflections ($F > 2\sigma$)	31,038	23,967	9079
Crystallographic R -factor (%)	18.9	21.6	19.8
Free R -factor (%) ^a	21.4	28.0	26.9
No. of residues	171 (STI), 223 (PPT)	172 (STI), 223 (PPT)	172
No. of non-hydrogen atoms			
Protein	1268 (STI), 1630 (PPT)	1286 (STI), 1630 (PPT)	1,281
Water	142	139	44
Ca ²⁺	1 (PPT)	1 (PPT)	—
Average B -factor (Å ²)			
Main-chain	34.4 (STI), 19.6 (PPT)	35.6 (STI), 19.1 (PPT)	37.1
Side-chain	34.5 (STI), 23.8 (PPT)	34.6 (STI), 20.6 (PPT)	39.4
Water	37.2	34.2	36.0
Ca ²⁺	19.6 (PPT)	33.6 (PPT)	—
R.m.s. deviations from ideal geometry			
Bond lengths (Å)	0.013	0.011	0.012
Bond angles (°)	1.32	1.68	1.40
Ramachandran outliers	0	0	1 (Gln138)

^a The free R -factor was calculated using 10% of the data.

around it is not well defined and its ψ angle lies slightly outside the allowed region. In the previously reported partial model of STI (Sweet *et al.*, 1974), residues 1 to 93 were built and refined with a certain degree of confidence but only a C α tracing is available for the segments 94 to 106 and 130 to 176, while another stretch of density was tentatively assigned to residues 116 to 122. In the present models of STI, a highly flexible loop between 107 and 115 has been unambiguously traced.

Average B -factors for the main-chain atoms of STI in both the complexed and the free states are much higher than for PPT in the complex models (Table 1). This indicates a greater degree of inherent flexibility of the inhibitor. An average B -factor for the main-chain atoms of each residue in STI models is plotted as a function of residue number in Figure 1. Besides the C terminus for all three models and also the N terminus for the free inhibitor, regions with high B -factors correspond to long loops connecting β -strands. The electron densities for these regions with high B -factors are weak but the tracing of the polypeptide chain is not ambiguous. In all three models of STI, the region around Asn13 shows B -factors lower than the average (Figure 1). This is due to the involvement of this residue in the network of hydrogen bonds, which stabilizes the conformation of the reactive site loop. The final ($2F_o - F_c$) electron density, calculated using data to 1.75 Å, for the reactive site loop of STI in the STI:PPT complex in the orthorhombic crystal is very well defined, reflecting the low B -factors around this region (Figure 2).

Overall structure of STI

STI is roughly spherically shaped with approximate dimensions of 45 Å × 42 Å × 40 Å. The STI structure consists of twelve antiparallel β -strands, long loops connecting these β -strands, and a 3_{10} -helix (Figure 3). It shows the same overall fold that has been described for the structures of ETI (Onesti

et al., 1991) and WCI (Dattagupta *et al.*, 1996). The reactive site (Arg63-Ile64) is located at one end of the structure on a protruding loop. Six of the strands (A β 1, A β 4, B β 1, B β 4, C β 1, and C β 4) form a short antiparallel β -barrel, with one side of the barrel being closed by a "lid" consisting of the other six strands. This common fold, termed the

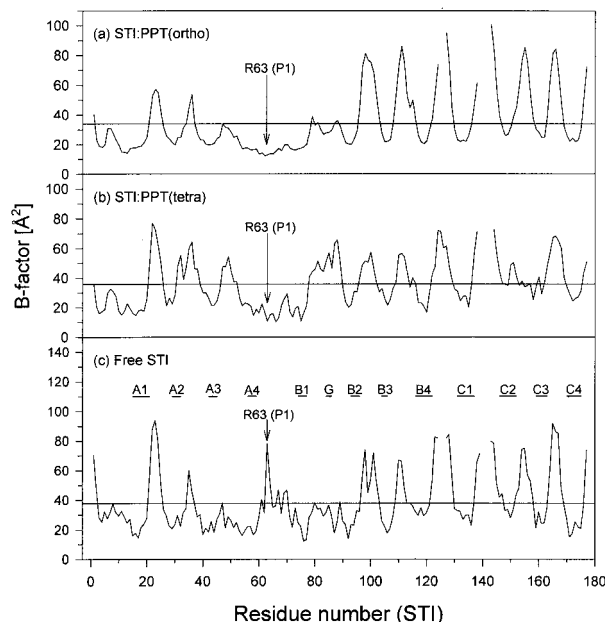


Figure 1. B -factor plot. Average B -factors for the main-chain atoms of each residue are plotted as a function of residue number for STI models (a) in STI:PPT complex in the orthorhombic crystal, (b) in STI:PPT complex in the tetragonal crystal; and (c) in the free state. Overall average values are indicated by horizontal lines. An assignment of secondary structure elements is also given (β -strands of subdomains A, B, and C: A1, residues 15 to 21; A2, 29 to 32; A3, 42 to 45; A4, 56 to 59; B1, 74 to 77; B2, 93 to 96; B3, 104 to 106; B4, 116 to 122; C1, 131 to 137; C2, 146 to 152; C3, 159 to 163; C4, 170 to 175; 3_{10} -helix: G, 84 to 86).

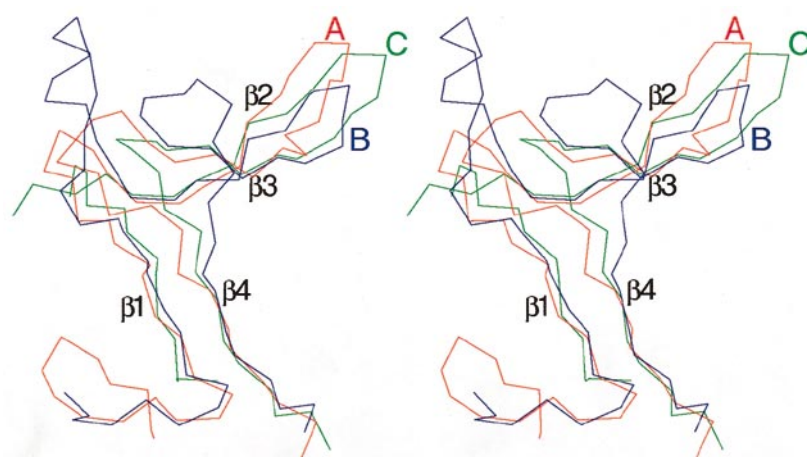


Figure 4. Stereo diagram showing the superposition of C^α atoms among three subdomains (A, B, and C) in STI structure. Red, blue and green lines represent domains A, B and C, respectively. The numbers (1 to 4) indicate the four β -strands in each subdomain.

tetragonal crystal forms and show B -factors lower than the average (13.4 and 16.5 \AA^2 , respectively). The water molecule Wat804 in the active-site pocket occupies a position to form hydrogen bonds with some of Arg63 NH2, Pro61 O, Gln192' OE1, Gly215' O, and Gly219' O atoms. Another water molecule, Wat810, makes hydrogen bonds with Arg63 NH1 and Ser190' OG atoms. Water molecules equivalent to Wat804 and Wat810 of STI:PPT complex are also found in other trypsin-inhibitor complexes: Wat403 and Wat416 in BPTI:BPT (Marquart *et al.*, 1983) and Wat58 and Wat5 in MCTI:PPT (Huang *et al.*, 1993). An interesting integral water molecule is present in the center of a triangular structure made up of β -strands A β 3, B β 3 and C β 3, located almost on the pseudo three-fold axis (Wat34 in Figure 5). An equivalent water molecule is also present in the ETI model, but a water molecule which mediates the interaction among the three β -strands at the bottom of β -barrel in the ETI model (Wat220 in Onesti *et al.*, 1991) is not observed in the STI models.

Comparison of inhibitor structures

When the refined structure of STI in the STI:PP-T(ortho) complex model is superposed with the C^α coordinates of STI in the early structure of the com-

plex (Sweet *et al.*, 1974), which were kindly supplied to us by Drs David Blow and Silvia Onesti during the final stage of our manuscript preparation, an r.m.s. deviation of 1.98 \AA is calculated for 147 C^α atom pairs (residues 1 to 97, 116 to 122, 130 to 138, and 143 to 176), with the maximum C^α difference of 6.02 \AA . Since the early STI model has not been refined well and is rather incomplete, our detailed structural comparison is limited to the refined models of STI in the free state and in its complex with PPT, which are reported in this study. A summary of the superpositions among STI models is given in Table 2, which lists the rms differences for C^α atom pairs. Table 2 also gives the results for comparing the trypsin models as well as for comparing the structures of three homologous inhibitors in their free state. It can be seen from Table 2 that the structural variability is greater for STI than for PPT.

Comparison of the free and trypsin-complexed STI models

All three STI models are superposed in Figure 6(a). The most striking change in STI upon complexation with PPT is a lowering in B -factors of its reactive site loop (residues 60 to 65) and also of the N terminus (Figure 1). This indicates that the

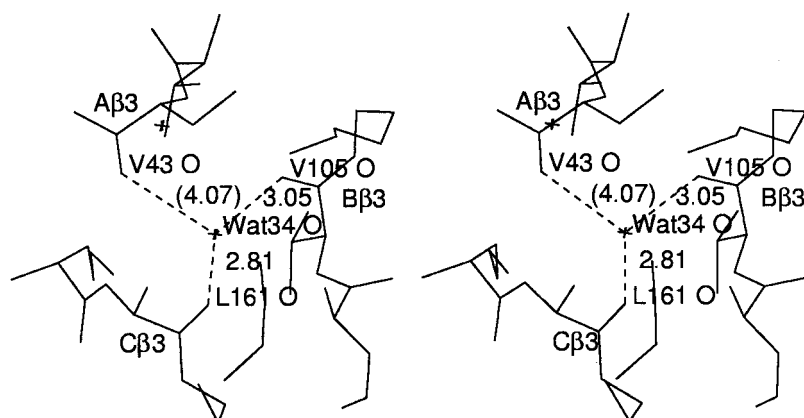


Figure 5. Stereo diagram showing the interactions between the bound water molecule (Wat34) and three β -strands (A β 3, B β 3 and C β 3). The distances between the water and oxygen atoms of the inhibitor are given, with a distance too long for a hydrogen bond in parenthesis.

Table 2. R.m.s. difference for C $^{\alpha}$ superposition among the trypsin and inhibitor models (Å)

PPT	STI	STI family inhibitors
MCT ^a -PPT(ortho), 0.29 (223) ^b	Free STI-STI(ortho), 0.68 (171)	Free STI-ETI, 2.39 (160)
MCT-PPT(tetra), 0.35 (223)	Free STI-STI(tetra), 0.79 (170)	Free STI-WCI, 2.37 (165)
PPT(ortho)-PPT(tetra), 0.30 (223)	STI(ortho)-STI(tetra), 0.50 (170)	WCI-ETI, 1.10 (158)

^a Starting porcine β -trypsin model for molecular replacement.
^b The value in the parenthesis is the number of C $^{\alpha}$ pairs for calculating the r.m.s. difference.

relatively flexible reactive site loop as well as the N terminus of STI becomes rigid upon its binding to trypsin. Similar behaviors have been previously observed for the reactive site loops of other inhibitors (Read & James, 1986).

A C $^{\alpha}$ superposition of STI models in the free state and its STI:PPT(ortho) complex gives an r.m.s. difference of 0.68 Å for 171 C $^{\alpha}$ atom pairs (1.09 Å for all 1262 non-hydrogen atoms). A plot of the distance between equivalent C $^{\alpha}$ atoms as a function of residue number is given in Figure 7(a). Residues showing deviations greater than 1.0 Å for C $^{\alpha}$ atoms are 22 and 23, 98, 111, 124, 128, 138, 144, 154 and 155, 166, and 177. A C $^{\alpha}$ superposition of the free STI and STI:PPT(tetra) models gives an r.m.s. difference of 0.79 Å for 170 C $^{\alpha}$ atom pairs (1.23 Å for all 1248 non-hydrogen atoms). A plot of the distance between equivalent C $^{\alpha}$ atoms as a function of residue number is given in Figure 7(b). Residues showing deviations greater than 1.0 Å for

C $^{\alpha}$ atoms are 48 to 49, 124, 144, 155, 166 and 167, and 177. In both cases, the large differences are confined to the external loops and the C-terminal residue Asp177. These regions show high *B*-factors for the main-chain atoms (Figure 1), indicating that the coordinate errors of these regions are larger than the average. Thus large deviations in these regions are partly due to the large coordinate errors but they also reflect the inherent conformational variability of these regions. In other parts of the chain, the magnitude of the deviation is roughly what is expected from the estimated coordinate errors. A superposition of the STI models in the two complex structures gives an r.m.s. difference of 0.50 Å for 170 C $^{\alpha}$ atom pairs (0.75 Å for all 1264 non-hydrogen atoms). A plot of the distance between equivalent C $^{\alpha}$ atoms as a function of residue number is given in Figure 7(c). Residues showing deviations greater than 1.0 Å for C $^{\alpha}$ atoms are 48 and 49, 111, 128, 155, 167, and 177. This large

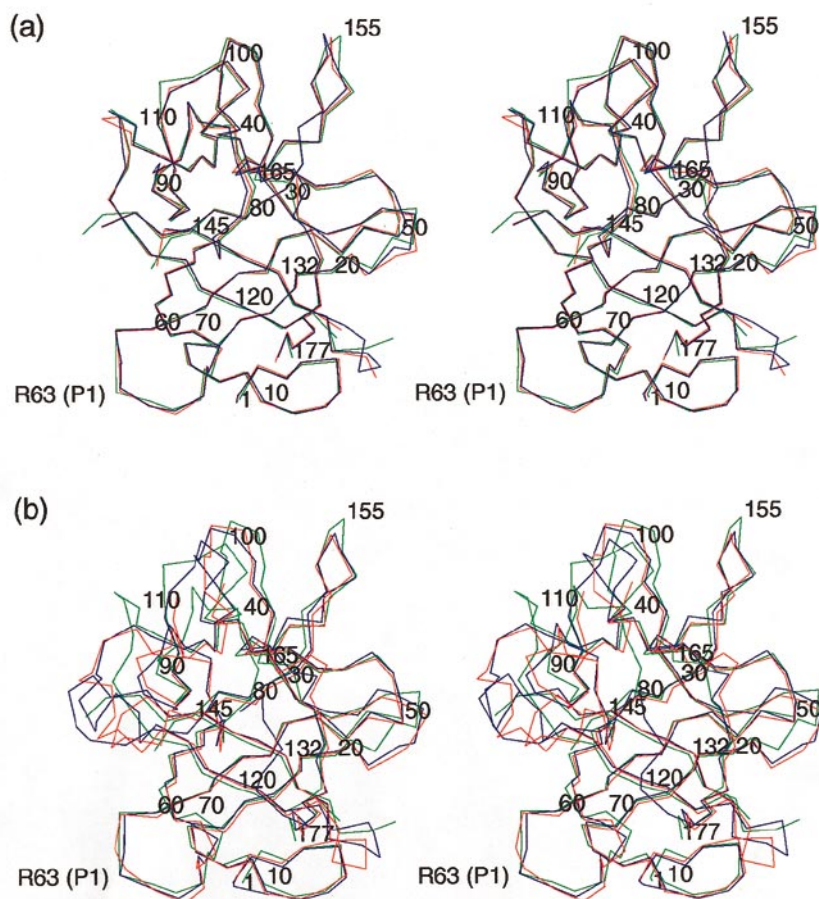


Figure 6. (a) Stereo diagram showing the superposition of C $^{\alpha}$ atoms among the free and complexed STI models. Green lines represent the free STI, while red and blue lines represent the inhibitor in STI:PPT(ortho) and STI:PPT(tetra) models, respectively. (b) Stereo diagram showing the superposition of C $^{\alpha}$ atoms among the STI family inhibitors. Green, red and blue lines represent STI, ETI and WCI, respectively. Every twentieth residue is labeled.

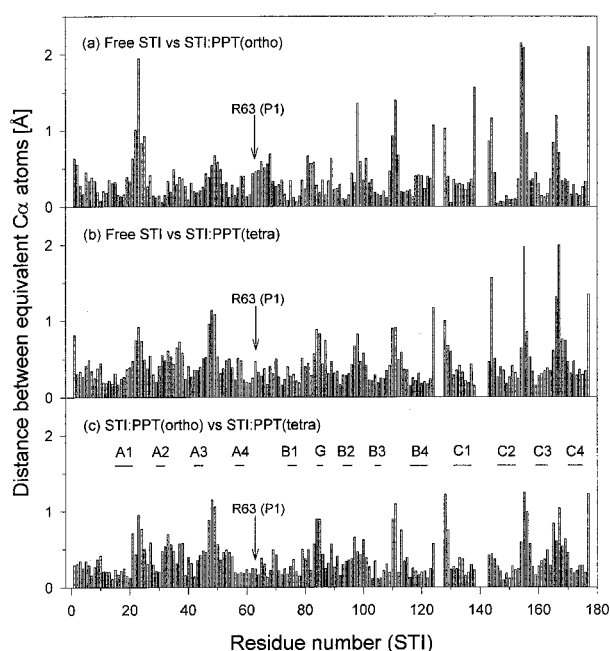


Figure 7. Plot of the distance between equivalent C^α atoms between the free and complexed STI models as a function of residue number. (a) Free STI *versus* STI:PPT(ortho), (b) Free STI *versus* STI:PPT(tetra); and (c) STI:PPT(ortho) *versus* STI:PPT(tetra).

discrepancy has contributions from the large coordinate errors in these regions as well as from the different loop conformations due to different crystalline environments, as in the comparison between the free and complexed STI models. Therefore, the structural difference between the two models of STI in different environments of the orthorhombic and tetragonal crystals of the complex should be regarded as only minor. In particular, an r.m.s. difference for the seven C^α atoms in the reactive site loop (P4–P3') is only 0.17 Å. Thus, the model of STI:PPT(ortho), which has been refined with a more complete data extending to 1.75 Å resolution, will be used for most of the structure descriptions below, unless otherwise stated.

Comparison of free STI with ETI and WCI

The alignment of amino acid sequences of STI, ETI, and WCI based on their three-dimensional structures is shown in Figure 8. The sequence identity between ETI and WCI is 57%, whereas that between ETI and STI sequences is 40% and that between WCI and STI is 45%. This is also manifested in the results of structural comparisons. Figure 6(b) shows a C^α superposition of the models of free STI, ETI, and WCI, refined at 2.3, 2.5, and 2.95 Å, respectively. Between the free STI and ETI, the r.m.s. deviation is 2.39 Å for 160 C^α atom pairs, with the residues 23, 63, 82 to 84, 88, 108, 110 to 116, 138, and 167 of STI showing C^α deviations greater than 3 Å. Between the free

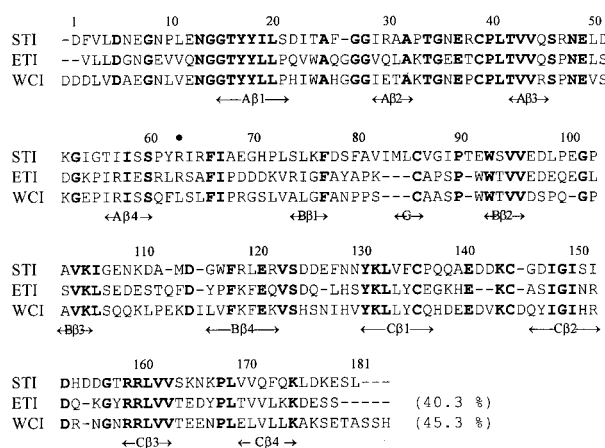


Figure 8. Sequence alignment of STI, ETI and WCI. The numbering is for STI. The secondary structure elements are assigned as in Figure 1. Residues conserved in all three sequences are in bold and the bullet (●) denotes the P1 residue.

STI and WCI, the r.m.s. deviation is 2.37 Å for 165 C^α atom pairs, with the residues 82 to 84, 86, 88 and 89, 98 and 99, 110 to 116, 124, 127, and 143 of STI showing C^α deviations greater than 3 Å. When ETI and WCI models are compared, the r.m.s. deviation is 1.10 Å for 158 C^α atom pairs, only about half as large as in comparing STI with either ETI or WCI (Table 2). A plot of the distance between equivalent C^α atoms as a function of residue number is given in Figure 9. Apparently, the difficulty in map interpretation of free STI (especially, for residues 80 to 88 and 110 to 115) using the phases from the starting ETI model only is due to the large (>6 Å) difference between the structures of ETI and STI (Figure 9). Other discrepant regions are located mostly on the protruding loops of the inhibitors, where insertions or deletions occur.

In the case of ETI, the reactive site loop is displaced with respect to that of free STI, with the distance of 3.15 Å between the C^α atoms of P1 residues. The corresponding distance between ETI and STI:PPT(ortho) is 2.78 Å. The reactive site loop of WCI lies between those of STI and ETI (Figure 6(b)) and its displacement is much smaller, with the distance of 1.02 Å between the C^α atoms of P1 residues in free STI and WCI. Thus the large displacement of the reactive site loop of ETI is largely due to the crystal packing, but is not a consequence of the difference in the complexation state, as speculated by the authors (Onesti *et al.*, 1991). This argument is further supported by an excellent correlation between the C^α distance plot for comparing the complexed STI with ETI and that for comparing the free STI and ETI (Figure 9(a) and (b)). Clearly, most of the discrepancy between the complexed STI and free ETI, including the displacement of the reactive site loop, cannot be ascribed to the difference in

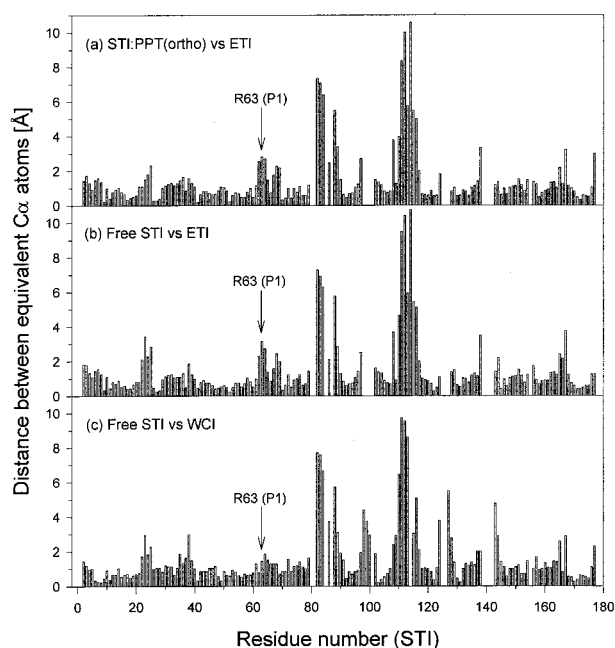


Figure 9. Plot of the distance between equivalent C^α atoms between the STI family inhibitors as a function of residue number. (a) STI:PPT(ortho) versus ETI, (b) Free STI versus ETI, and (c) Free STI versus WCI.

the complexation state. When C^α atoms for P4–P3' residues are superposed, the r.m.s. deviation is 0.71 Å between the free STI and ETI, while it is 0.56 Å between the free STI and WCI. The corresponding r.m.s. deviations are 0.36 Å between the free STI and STI:PPT(ortho) and 0.49 Å between the free STI and STI:PPT(tetra). The local conformations of reactive site loops of STI, ETI, and WCI are, therefore, very similar and are little affected upon complex formation.

The reactive site loop: its conformation and changes upon complexation

STI belongs to the family of substrate-like inhibitors, which possess an exposed reactive site loop of

a characteristic canonical conformation. The reactive site loop of STI is not constrained by the secondary structure elements or disulfide bridges that could limit its conformational freedom, unlike many other proteinase inhibitors. In STI, the side-chain of Asn13 plays an important role in stabilizing the reactive loop conformation through a network of hydrogen bonds (Figure 10). This results in the low B -factors for the region around Asn13 (Figure 1). Similar stabilizations are also observed for ETI and WCI, although the details of hydrogen bonding pattern are not identical. For example, a water molecule resides in the reactive site loop of WCI, forming hydrogen bonds with the carbonyl O atoms of Phe64 and Ser62 (Dattagupta *et al.*, 1996). An equivalent water molecule is not observed in either the STI or the ETI structure. The backbone conformation of the reactive site loop of STI does not change significantly upon forming a complex with PPT, as indicated by similar main-chain dihedral angles for P4–P3' residues (Table 3). This is similarly found in BPTI and its complex with BPT (Table 3). Table 3 shows that there is a large deviation of ϕ and ψ angles at several positions between our refined models of STI and the early model of STI in its complex with PPT (Sweet *et al.*, 1974). Table 3 also lists the conformational angles of several other inhibitors as well as their amino acid sequences around the scissile peptide bond. Between STI and BPTI, there is a large deviation in the ϕ angle of P4 and ϕ and ψ angles of P2'. Despite a great difference in the main-chain conformational angles at P2' position (arginine) for STI and BPTI, both guanidinium groups point in the same direction as pointed out earlier (Sweet *et al.*, 1974). Between STI and MCTI, a large deviation is observed for ϕ and ψ angles of P3 and the ψ angle of P2', whereas ETI and WCI show similar conformations as STI at P2'. Notwithstanding some differences in the main-chain conformations at some positions, the side-chains, which interact with the proteinase, are quite similarly oriented in several substrate-like inhibitors.

The extension of the reactive site loop of STI family (measured as the distance between C^α

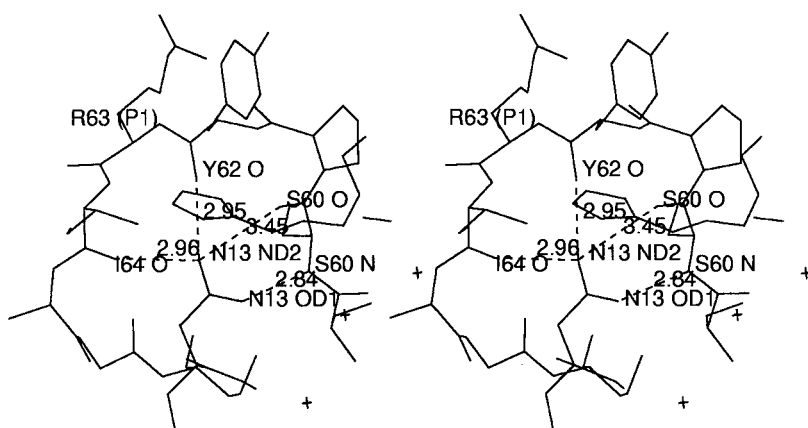


Figure 10. Stereo diagram showing the role of Asn13 in stabilizing the reactive site loop conformation of STI. The conserved residue Asn13 is involved in extensive hydrogen bondings with several residues in the reactive site loop.

Table 3. Comparison of dihedral angles (ϕ/ψ) of reactive site loop residues

	P4	P3	P2	P1	P1'	P2'	P3'
STI	S	P	Y	R	I	R	F
ETI	S	R	L	R	S	A	F
WCI	S	Q	F	L	S	L	F
BPTI	G	P	C	K	A	R	I
MCTI	I	C	P	R	I	W	M
Free STI	-112/149	-70/-28	-62/158	-84/21	-60/148	-90/-22	-127/126
STI:PPT(ortho)	-114/144	-58/-34	-56/139	-89/38	-83/148	-67/-38	-118/155
STI:PPT(tetra)	-126/149	-67/-23	-65/121	-70/10	-59/165	-88/-29	-115/170
STI:PPT(Blow)	-58/127	-40/-39	-65/141	-93/47	-93/123	-165/-178	8/49
ETI	-88/171	-66/-30	-112/150	-66/134	-169/151	-74/-46	-106/163
WCI	-125/-177	-85/-21	-80/173	-111/76	-122/158	-89/-19	-130/157
BPTI	86/176	-86/-6	-81/165	-104/9	-76/172	-127/76	-105/121
BPTI:BPT	78/174	-77/-29	-70/155	-116/39	-87/164	-112/79	-98/124
MCTI	-58/117	-137/112	-54/147	-91/34	-82/146	-84/125	-152/127

STI:PPT(Blow), Sweet *et al.* (1974); ETI, Onesti *et al.* (1991); WCI, Dattagupta *et al.* (1996); BPTI, Wlodawer *et al.* (1987); BPTI:BPT, Marquart *et al.* (1983); MCTI, Huang *et al.* (1993).

atoms of residues P4 and P3') is among the smallest (9.0 Å for STI, 9.2 Å for ETI, and 8.8 Å for WCI compared with 11.3 Å for BPTI, 15.0 Å for MCTI (squash family), 15.6 Å for PSTI (Kazal family), and 15.9 Å for eglin c (PI-1 family)). The reactive site loops of BPTI (Kunitz), Kazal, squash, and PI-1 families are constrained by disulfide bridges that could limit their conformational freedom. No disulfide bridge adjacent to the reactive site is present in the STI and eglin c (PI-1 family). Whereas the reactive site loop of STI is held in a favorable conformation by hydrogen bonds (Figure 10), that of eglin c is stabilized by electrostatic interactions between charged residues on both sides of the reactive site loop.

The magnitude of overall structural changes in proteinase inhibitors upon binding the cognate proteinases ranges from very small to extremely large. For the best-characterized BPTI, the r.m.s. difference between the free and complexed structures is 0.26 to 0.27 Å for 58 C α atom pairs (Perona *et al.*, 1993). The r.m.s. difference between the free *Streptomyces subtilisin inhibitor* (SSI) and the inhibitor in SSI:subtilisin BPN' complex is 0.63 Å for 107 C α atom pairs (Takeuchi *et al.*, 1991). At the other extreme, the serpin family shows the most dramatic conformational change. The structure of cleaved α_1 -antitrypsin showed an unexpected separation of its reactive center peptide (Met358-Ser359; P1-P1') by about 70 Å with incorporation of the reactive loop into the β -sheet structure (Loebermann *et al.*, 1984). A superposition of 54 C α atoms in the central β -sheet of the cleaved and uncleaved α_1 -antitrypsin gave an r.m.s. difference of 2.4 Å (Song *et al.*, 1995). For STI, the r.m.s. difference between the free and complexed inhibitors is 0.68 to 0.79 Å for 171 C α atom pairs. The deviation for C α atoms belonging to P4-P3' residues in the reactive site loop is smaller than the overall change: 0.40 Å between the free STI and STI:PPT(ortho) and 0.26 Å between the free STI and STI:PPT(tetra).

Geometry of the reactive site carbonyl group

The geometry of the carbonyl group at P1 residue is of great importance in understanding the interaction between the inhibitor and the proteinase during the catalytic mechanism. In the earlier and lower resolution structure of STI:PPT complex (Sweet *et al.*, 1974), a tetrahedral intermediate was proposed. This proposal was not, however, validated by the subsequent high-resolution structures of other complexes, because only slight (Marquart *et al.*, 1983; Bode & Huber, 1992) or negligible (Read *et al.*, 1983; Huang *et al.*, 1993) out-of-plane deformations of the P1 carbonyl group were observed. Additionally, the carbonyl carbon atom was found to be within the "sub-van-der-Waals" distance (typically around 2.7 Å) from the Ser195' OG atom (Bode & Huber, 1992). In our STI:PPT(ortho) and STI:PPT(tetra) models, the distance is 2.76 and 2.71 Å, respectively. In order to obtain an unbiased geometry of the P1 carbonyl group, refinements were carried out either with or without applying the restraints to the carbonyl carbon atom. With the usual restraints, the out-of-plane distance of Arg63 C is 0.0084 and 0.0156 Å for STI:PPT(ortho) and STI:PPT(tetra), respectively (the positive value corresponding to a distortion toward the Ser195' OG of trypsin). Without the restraints, it increases to 0.024 and 0.067 Å for STI:PPT(ortho) and STI:PPT(tetra), respectively. As a control experiment, refinements were also performed after removing the restraints for the carbonyl carbon atom of P2' residue (Arg65). The out-of-plane distance of Arg65 C shows a considerable increase from 0.019 to 0.082 Å for STI:PPT(ortho) or from 0.026 to 0.093 Å for STI:PPT(tetra), respectively, when the restraints on the carbonyl carbon atom of Arg65 are removed. Therefore, it is concluded that the P1 carbonyl group of STI displays no significant out-of-plane displacement and thus it retains a nominal trigonal planar when complexed with PPT. This is in agreement with the

result for the MCTI:PPT complex (Huang *et al.*, 1993) and is consistent with STI belonging to a typical substrate-like inhibitor family, whose reactive site loop takes the canonical conformation in both the free and the complexed states. It is worth mentioning that the out-of-plane distance of Lys15 (P1) C ranges between 0.12 and 0.14 Å for BPTI complexes (PDB ID codes: 2PTC, 1TPA, and 1TGP).

Mode of interaction between STI and PPT

The pattern of interaction between STI and PPT is depicted in Figure 11(a) and is summarized in Table 4. There are some minor differences between the interaction pattern between the orthorhombic and tetragonal crystal structures of the complex. Twelve amino acid residues out of the 181 in STI make contact with PPT in the orthorhombic crystal structure. They are: Asp1, Phe2, Asn13, Pro61 (P3), Tyr62 (P2), Arg63 (P1), Ile64 (P1'), Arg65 (P2'), His71, Pro72, Trp117 and Arg119. In the tetragonal crystal structure, the three residues His71, Pro72 and Arg119 do not interact with PPT. However, the pattern for the hydrogen bonding interaction involving the reactive site loop residues from Pro61 (P3) to Arg65 (P2') is well conserved between the two crystal forms, except that a hydrogen bond between Tyr62 and Gly96' is not present in the tetragonal crystal structure (Table 4).

Our orthorhombic crystal of the STI:PPT complex, although grown under a condition different

from that of Sweet *et al.* (1974), has the same space group with nearly identical cell parameters. However, the interaction pattern observed in our orthorhombic structure is somewhat different from that described by Sweet *et al.* (1974). We believe that this difference is mainly due to the fact that our model is more complete and is better refined at higher resolution. The protein sequencing data (Kim *et al.*, 1985) and the nucleotide sequence of cDNA (Song *et al.*, 1991) indicate that the amino acid residues at P4 and P3 positions in the early model of the complex, Pro60 (P4) and Ser61 (P3), are erroneous. The corrected sequence, Ser60 (P4) and Pro61 (P3), was built into our model and is confirmed by the excellent electron density (Figure 2). The residues Pro60 (P4) and Phe66 (P3') in the early model, suggested to form the region of interaction with PPT, are not involved in the interaction in both of our models. The two residues Trp117 and Arg119 belong to the segment 116 to 122, which was tentatively assigned in the early model of the complex and were not included in the residues participating in the interaction with trypsin (Sweet *et al.*, 1974). A detailed description of more differences is given below.

Most of the contacts between STI and PPT involve the five residues in the reactive site loop (P3-P2'). In our orthorhombic crystal structure, eleven of the 14 hydrogen bonds between STI and PPT are accounted for by the above residues, while in the tetragonal structure, ten of the eleven are accounted for. The P1 residue Arg63 makes the

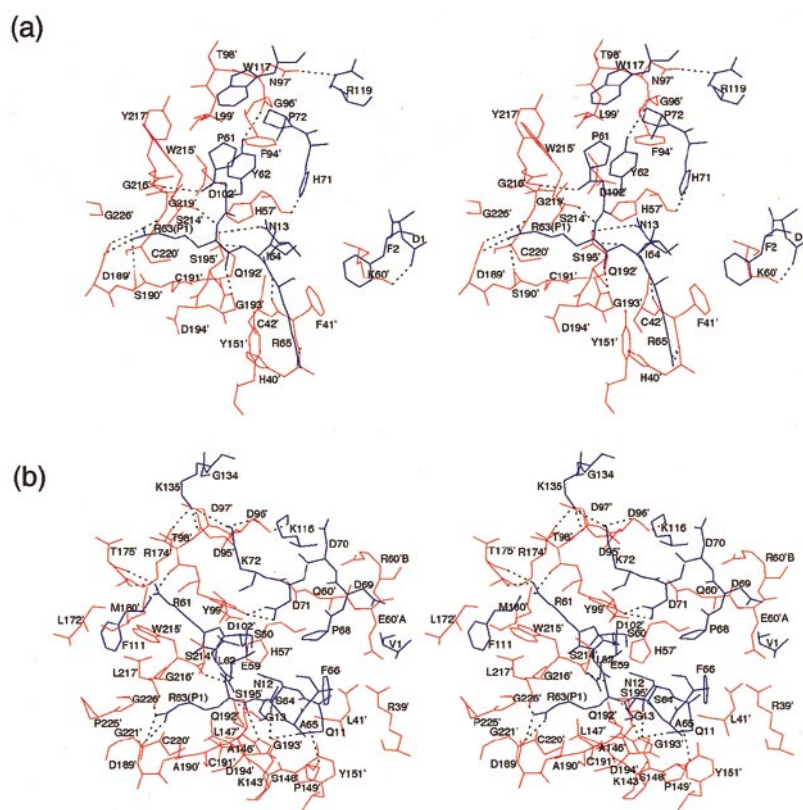


Figure 11. Stereo diagram showing the interaction; (a) between STI and PPT; and (b) between ETI and tPA. Blue and red lines represent the inhibitor and the proteinase, respectively. All residues involved in the interaction are labeled. Dotted lines denote hydrogen bonds.

Table 4. Total interactions between STI and PPT

STI	PPT	^a No. of interactions		^b Note	
		STI:PPT(ortho)	STI:PPT(tetra)		
Asp1	Lys60'	2, —	—, 1 (OD2–NZ; 2.80 Å)	ion	
Phe2	Phe41'	3, —	5, —	a–a	
	Lys60'	3, —	3, —		
Asn13	Gln192'	4, 1 (ND2–OE1; 3.46 Å)	5, —		
Pro61	Trp215'	3, —	3, —		
	Gly216'	1, 1 (O–N; 3.12 Å)	4, 1 (O–N; 3.14 Å)		
Tyr62	His57'	16, —	5, —		
	Phe94'	2, —	—, —		
	Gly96'	—, 1 (OH–O; 2.97 Å)	—, —		
	Leu99'	5, —	8, —		
	Asp102'	1, —	—, —		
	Gln192'	1, —	—, —		
	Ser214'	1, —	1, —		
	Trp215'	1, —	1, —		
	Arg63	Asp189'	4, 2 (NH1–OD1; 2.70 Å) (NH2–OD2; 3.09 Å)	4, 2 (NH1–OD1; 2.68 Å) (NH2–OD2; 2.95 Å)	ion ion
		Ser190'	4, 1 (NH1–OG; 3.11 Å)	3, 1 (NH1–OG; 3.43 Å)	
Ile64	Cys191'	4, —	5, —		
	Gln192'	4, —	4, —		
	Gly193'	2, 1 (O–N; 2.61 Å)	2, 1 (O–N; 2.54 Å)		
	Asp194'	1, —	—, —		
	Ser195'	8, —	5, —		
	Ser214'	1, 1 (N–O; 3.03 Å)	—, 1 (N–O; 3.38 Å)		
	Trp215'	2, —	1, —		
	Gly216'	1, —	—, —		
	Gly219'	—, 1 (NH2–O; 2.79 Å)	—, 1 (NH2–O; 2.82 Å)		
	Cys220'	—, —	3, —		
	Gly226'	1, —	—, —		
	Phe41'	—, —	1, —		
	Cys42'	2, —	1, —		
	His57'	1, —	3, —		
	Gln192'	1, —	2, —		
Gly193'	1, —	1, —			
Arg65	Ser195'	2, 1 (N–OG; 3.12 Å)	1, 1 (N–OG; 2.89 Å)		
	His40'	1, 1 (NH2–O; 3.07 Å)	1, 1 (NH2–O; 2.99 Å)		
	Phe41'	—, 1 (N–O; 3.34 Å)	—, 1 (N–O; 3.41 Å)		
	Tyr151'	3, —	6, —	a–n	
His71	Gly193'	—, —	2, —		
	His57'	1, 1 (NE2–O; 3.08 Å)	—, —		
Pro72	Asn97'	1, —	—, —		
	Leu99'	1, —	—, —		
Trp117	Asn97'	2, —	—, —		
	Tyr217'	—, —	5, —		
Arg119	Asn97'	1, 1 (NE–OD1; 3.21 Å)	—, —		

^a An interatomic contact is counted for every pair of atoms which is within 0.5 Å of theoretical van der Waals contact distance. The table indicates the number of interatomic contacts and hydrogen bonds (in *italics*) between amino acids of STI and those of PPT.

^b Ion: possible ionic interaction, a–a: aromatic–aromatic edge–face interaction, a–n: aromatic–NH group interaction.

most extensive hydrogen bonds with PPT, forming six hydrogen bonds in total. The side-chain of Arg63 occupies its expected position in the primary binding pocket of PPT. The guanidinium group of Arg63 makes an ionic interaction with the carboxylate group of Asp189' in PPT. Hydrogen bonds are possible between the nitrogen atoms of the side-chain of Arg63 and the side-chain oxygen atom (OG) of Ser190' and the carbonyl O atom of Gly219'. This differs from the hydrogen bonding pattern reported previously (Sweet *et al.*, 1974). The phenolic side-chain of Tyr62 (P2) is positioned between the side-chains of Leu99' and His57', lying parallel to the imidazole ring of the latter. Its hydroxyl group forms a hydrogen bond with the carbonyl O atom of Gly96'. It was suggested to form a hydrogen bond with the side-chain of

Asn97' in the previous report (Sweet *et al.*, 1974). Even though the residue Pro61 (P3) was erroneously assigned as Ser in the early model, the main-chain dihedral angles at this position (Table 3) and the hydrogen bonding interaction between its carbonyl O atom and the main-chain N atom of Gly216' are similar. The long side-chain of Arg65 (P2') interacts with Tyr151' in PPT, possibly forming a hydrogen bond between its nitrogen NE atom and the aromatic ring (Levitt & Perutz, 1988).

Gln192' of PPT belongs to a stretch of eight residues strictly conserved between PPT and tPA (Figure 12) and interacts with as many as four residues of STI (Asn13, Tyr62, Arg63 and Ile64). Therefore, its conformation is of great importance in the prediction of ETI:tPA complex structure by homology modeling. Its unique conformation could not

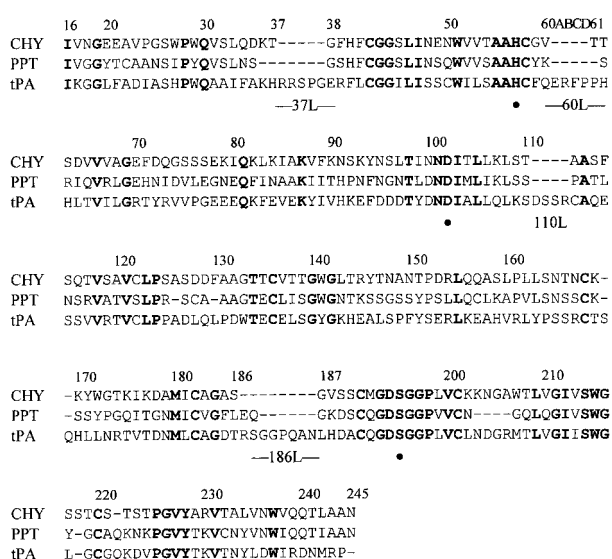


Figure 12. Sequence alignment of serine proteinases based on their three-dimensional structures. The abbreviations are as follows. Bovine chymotrypsin (CHY), porcine pancreatic trypsin (PPT), and human tissue-type plasminogen activation (tPA). The numbering scheme is according to the sequence number of CHY. The bullets (●) denote the positions of catalytic residues His57, Asp102 and Ser195. Several loop regions (L) of tPA are indicated.

be assigned in the early model of Sweet *et al.* (1974), because the electron density for the side-chain was not long enough and there was little room for the side-chain in the complex. In contrast, the side-chain of Gln192' in PPT is well defined by the electron density in our models. This residue was suggested to form a hydrogen bond with the P4 residue of STI (Sweet *et al.*, 1974), but this interaction is absent in both of our complex models.

A close contact between the aromatic rings of Phe2 in STI and Phe41' in PPT was noted earlier (Sweet *et al.*, 1974), with 13 interatomic contacts. In our structure, the two aromatic rings make an edge-face interaction, resulting in much fewer interatomic contacts. This kind of aromatic edge-face interaction is energetically favorable and is the most common among the observed aromatic-aromatic interactions (Burley & Petsko, 1985). A similar interaction was found in the crystal structure of chymotrypsin complexed with 6-benzyl-3-chloro-2-pyrone (Ringe *et al.*, 1985). The ionic interaction between Asp1 and Lys60' was also noted earlier (Sweet *et al.*, 1974) but the side-chains were drawn unreasonably close together by the real-space refinement procedure. In our models, they are much better defined and there is no extensive van der Waals contact between them.

A structural basis of tPA inhibition by ETI

There is a strong clinical interest in human tPA (Collen & Lijnen, 1995). Figure 12 compares the

amino acid sequences of chymotrypsin, PPT and tPA, as aligned on the basis of their three-dimensional structures. In this paper, residue numbers of all serine proteinases follow the chymotrypsin numbering for consistency. Human tPA shows a sequence identity of 41% against PPT. The recent crystal structure of recombinant two-chain human tPA showed that the S1 pocket of tPA is almost identical to that of trypsin, whereas the S2 site is considerably reduced in size (Lamba *et al.*, 1996; PDB ID code, 1RTF).

Although the three inhibitors STI, ETI and WCI show a high degree of sequence identity (Figure 8), only ETI is capable of inhibiting tPA. In order to gain a better insight into this specificity, docking studies were undertaken using the model of ETI, after manual correction of the scissile peptide bond towards the canonical conformation (Lamba *et al.*, 1996). This kind of docking studies will benefit significantly from a high-resolution structure of a Kunitz-type trypsin inhibitor in its complex with a cognate proteinase. Our highly refined model of STI:PPT complex provides an excellent template to build a homology model of ETI:tPA complex. The hypothetical models of both ETI:tPA and STI:tPA complexes were subject to energy minimization, after the following changes were made to STI to mimic ETI: removal of Asp1 and replacement of Phe2 by Val and Arg65 by Ala. The results of our modeling studies largely confirm those of the previous docking studies (Lamba *et al.*, 1996), which indicated that the exposed reactive site loop of ETI should slot into tPA's active-site cleft from the P3 to the P2' residue. Furthermore, potentially significant ionic and hydrogen bonding interactions involving the residues outside the P3-P3' segment have also been indicated by our studies (Figure 11(b)), in addition to the interactions between the reactive site loop of ETI from the P3 to P3' residues and tPA described previously (Lamba *et al.*, 1996).

N-terminal Val1 ammonium group of ETI was suggested to be located close to the distal polar groups of Gln60' and Glu60'A with the Val1 side-chain approaching the Arg39' side-chain of tPA (Lamba *et al.*, 1996). Our modeling studies indicate that the Val1 side-chain is located above the ionically interacting pair of side-chains of Arg39' and Glu60'A, but the side-chain of Gln60' points away from the Val1 side-chain (Figure 11(b)). This interaction with Glu60'A explains the much weaker binding of the mutant ETI with an extra aspartic acid at the N terminus (Teixera *et al.*, 1994). This interaction is also partly responsible for STI and WCI being unable to inhibit tPA. STI has an extra aspartic acid residue at the N terminus and WCI two extra aspartic acid residues (Table 5). These extra residues of STI and WCI at the N terminus protrude from the body of the molecule, when compared with the ETI model. The protruding N termini of both STI and WCI may restrict their access to the active site of tPA by making a close contact with the characteristic insertion loop

Table 5. Residues of ETI involved in its interaction with tPA

			P3	P2	P1	P1'	P2'	P3'							
(ETI No.)			1	61	62	63	64	65	66	69	70	71	72	116	135
ETI			<u>V</u>	R	<u>L</u>	<u>R</u>	S	<u>A</u>	<u>F</u>	D	D	D	K	<u>K</u>	K
ETIa			<u>V</u>	R	<u>L</u>	<u>R</u>	<u>S</u>	<u>A</u>	<u>F</u>	D	D	D	K	<u>K</u>	K
ETIb			<u>E</u>	P	<u>F</u>	<u>R</u>	S	<u>Y</u>	<u>F</u>	K	G	<u>S</u>	<u>L</u>	<u>K</u>	<u>E</u>
(STI No.)	-1	1	2	61	62	63	64	65	66	69	70	71	72	119	139
STI		D	F	P	Y	R	I	R	F	E	G	H	P	R	Q
WCI	D	D	D	Q	F	L	S	L	F	R	G	S	L	K	D

The underlined residues are predicted to interact with tPA. The residues in boldface are critical for specific binding of tPA by ETI (and ETIa). Other residues are also listed to facilitate the discussion.

(60Loop) as well as by making a repulsive ionic interaction. ETIa from *Erythrina variegata*, belonging to group c, has the same valine residue at the N terminus as ETI and shows the inhibitory activity against tPA. In contrast, the N-terminal residue of ETIb is a glutamic acid and it does not inhibit tPA, even without extra negatively charged residues at the N terminus (Table 5). This is readily explained by the ionic interaction of the N-terminal Val1 of ETI with Glu60'A in tPA.

There is a pronounced difference in the surface features of STI and ETI, despite similarity in their overall three-dimensional structures (Figure 13). Another region of tPA important for inhibitor binding is 97Loop. This β -hairpin loop protrudes considerably from the active site of tPA. Together with Glu93', the side-chains of Asp95', Asp96' and

Asp97' form a highly acidic surface patch in the tPA structure (Figure 13). The three residues Asp95', Asp96' and Asp97' interact with the positively-charged side-chains of Lys72, Lys116 and Lys135 of ETI (Figure 13). In STI, WCI and ETIb, which do not show inhibitory activity toward tPA, two of these three lysines (Lys72 and Lys135) are replaced by neutral or negatively charged residues, while a positive charge is retained at the position corresponding to Lys116 of ETI (Table 5). The side-chain nitrogen atom of Lys72 makes salt bridges with side-chain oxygen atoms of Asp96' and Asp97' in 97Loop of tPA. The side-chain nitrogen atom of Lys116 in ETI also interacts with the side-chain of Asp96' in tPA. However, this interaction may not be important in the selective inhibition of tPA by ETI, since the corresponding residue

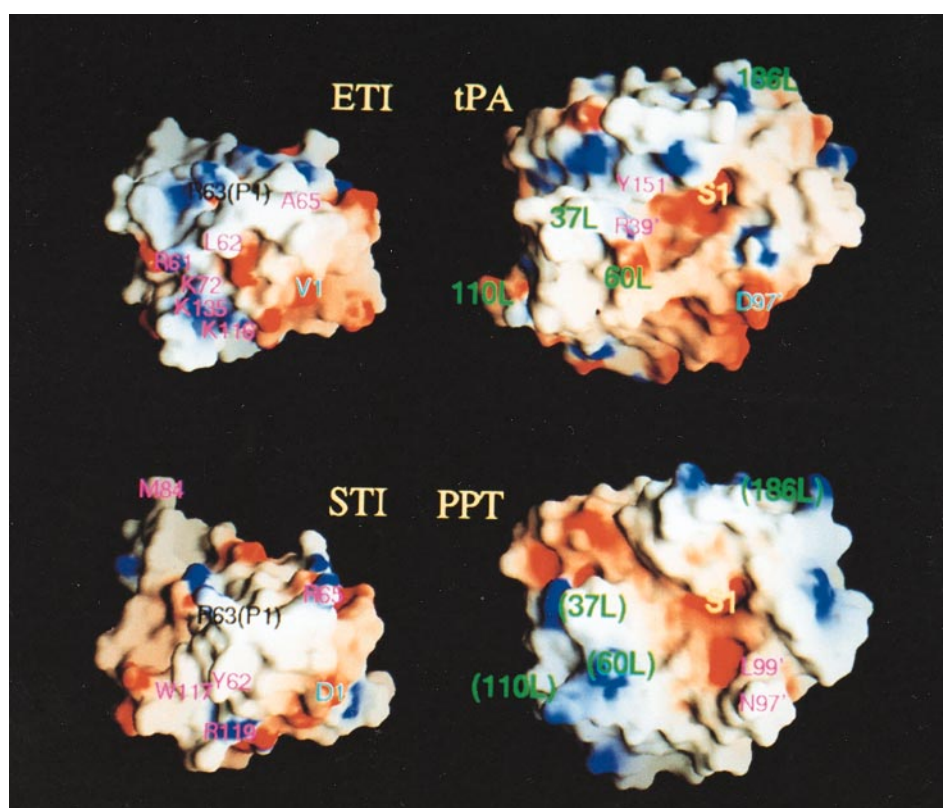


Figure 13. A comparison of molecular surfaces. Positively charged regions are blue and negatively charged regions red. The P1 residue, S1 pocket and some basic and acidic patches are labeled. This figure was generated using GRASP (Nicholls, 1992).

Arg119 in STI is predicted to make a similar interaction with Asp96' of tPA. The side-chain of Lys135 in ETI makes at least three hydrogen bonds with the side-chains of Asp95' and Thr98' in addition to the ionic interaction with Asp95' (Figure 11(b)).

A mutagenesis study on ETIa from *E. variegata* showed that the tPA inhibitory activity of the single mutants R61P (P3) and L62F (P2) is significantly reduced and, furthermore, the double mutant R61P/L62F (P3/P2) lacks the tPA inhibitory activity, despite retaining the trypsin inhibitory activity (Kouzuma *et al.*, 1997). These mutants were designed on the basis of the sequence difference between ETIa and ETIb (Table 5). The guanidyl group of Arg61 in ETI makes hydrogen bonds with the main-chain oxygen atoms of Arg174' and Thr175' in tPA (Figure 11(b)). The P3 residues of STI and WCI are proline and glutamine, respectively. Both crystal structures of STI:PPT(ortho) and STI:PPT(tetra) show that the side-chain of Tyr62 (P2) in STI makes a van der Waals contact with Leu99' in PPT. Interestingly, the interacting side-chains are switched in the ETI:tPA complex and a comparable interaction is possible between the side-chains of Leu62 (P2) in ETI and Tyr99' in tPA. A unique feature of ETI (and ETIa also), as compared with the inhibitors which do not inhibit tPA, is the presence of three consecutive acidic residues, Asp69, Asp70 and Asp71 (Table 5). However, its significance is not clear, since only Asp71 makes hydrogen bonds with the side-chain of Tyr99'. A mutant A65Y (P2') of ETIa from *E. variegata* exhibited the same level of tPA inhibitory activity as the wild-type ETIa. (Kouzuma *et al.*, 1997). Modeling indicates that an aromatic ring of tyrosine at P2' position can interact favorably with that of Tyr151' in tPA through an edge-face interaction, the most common mode of aromatic-aromatic interaction (Burley & Petsko, 1985). Other residues of ETI which are possibly involved in the interaction with tPA are Gln11, Asn12, Gly13 and Phe111. Table 5 lists the residues of ETI which play an important role in its interaction with tPA.

Materials and Methods

Purification and crystallization

Kunitz-type soybean trypsin inhibitor (isoform Ti^a, T-9003) and porcine pancreatic trypsin (T-7418) were purchased from Sigma. For the complex formation, STI (138 mg) and PPT (102 mg) were dissolved in 100 mM sodium chloride solution buffered at pH 8.0 with 100 mM Tris-HCl and were left for 2 h at 4°C. This solution was loaded on a Sephacryl S-200 HR gel filtration column (2.2 cm × 76.0 cm), which had been previously equilibrated with the same buffer. The fractions containing the STI:PPT complex were collected and concentrated to ~110 mg/ml concentration (1.6 ml volume) by ultrafiltration using a YM 10 membrane (Amicon) for crystallization. The protein concentration was estimated by measuring the absorbance at 280 nm, assuming a correspondence of 1.0 mg/ml concentration to the unit

absorbance at 280 nm for the 1.0 cm path length. Crystallization was achieved by the hanging drop vapor diffusion method using 24-well tissue culture plates (Flow Laboratories) at room temperature (22 ± 1°C). The hanging drop was prepared by mixing equal volumes of the protein solution and the reservoir solution. The STI:PPT complex have been crystallized in two crystal forms. For the orthogonal form, the reservoir solution is 30% (v/v) PEG 400 and 100 mM lithium sulfate in 100 mM Hepes at a final pH of 7.67. Crystals of typical dimensions of 0.5 mm × 0.6 mm × 1.5 mm grew within a week. For the tetragonal form, the reservoir solution is 1.70 M ammonium sulfate and 10% (v/v) dioxane in 100 mM Mes at a final pH of 5.68. Crystals of typical dimensions of 2.0 mm × 2.0 mm × 0.4 mm grew within a week. The detailed crystallization condition for the free STI has been described previously (Lee *et al.*, 1993). Crystals of dimensions of 0.5 mm × 0.4 mm × 0.2 mm grew within about 25 days.

X-ray diffraction studies

For all X-ray data collection, a crystal was mounted in a thin-walled glass capillary and both ends of the capillary were filled with the mother liquor and then sealed with wax. Three sets of data were collected at three places using different types of detectors. X-ray diffraction studies on the free STI have been described previously (Lee *et al.*, 1993). A summary of data collection is given in Table 6.

Data from an orthorhombic crystal of STI:PPT complex were collected on a MAR image plate X-ray detector at the beamline X12C of the National Synchrotron Light Source, Brookhaven National Laboratory (Sweet, 1989). All data were reduced and scaled using DENZO (Gewirth, 1993) and SCALEPACK (Otwinowski, 1992). The crystal diffracted to beyond 1.7 Å resolution using synchrotron X-rays. Despite much difference in crystallization conditions, our orthorhombic crystal of the complex is similar to the crystal previously grown by ethanol (Sweet *et al.*, 1974).

Data from a tetragonal crystal of STI:PPT complex were collected at 14°C using a Weissenberg camera for macromolecular crystallography at the BL-6A2 experimental station of Photon Factory, Japan (Sakabe, 1991). A Fuji image plate (20 × 40 cm) was placed at a distance of 429.7 mm from the crystal. The oscillation range per image plate was 3.0°, with a speed of 2.0°/sec and a coupling constant of 1.0°/mm. An overlap of 0.5° was allowed between two contiguous image plates. Image plates were digitized by a Fuji BA100 scanner. Raw data were processed using the program WEIS (Higashi, 1989). The crystal belongs to the tetragonal space group $P4_12_12$ or $P4_32_12$. The former space group was later found to be correct on the basis of translation function calculations.

Molecular replacement

All three structures were solved by the molecular replacement method (Crowther & Blow, 1967) using the programs X-PLOR (Brünger, 1992) and AMoRe (Navaza, 1994). Reflections with $F_o > 2\sigma_F$ were used throughout the molecular replacement and subsequent refinement calculations. First, a molecular replacement solution of the free STI structure was attempted, using the previously reported model of ETI as the probe structure (Onesti *et al.*, 1991; PDB ID code, 1TIE). A consistent

Table 6. Data collection statistics

	STI:PPT(ortho)	STI:PPT(tetra)	STI
X-rays	1.000 Å (NLSL ^a)	1.000 Å (PF ^b)	CuK α (SNU ^c)
No. of crystals	1	1	1
Resolution limit (Å)	1.75	1.80	2.1
Space group	$P2_12_12_1$	$P4_12_12$	$P2_12_12_1$
Cell axes (<i>a</i> , <i>b</i> , <i>c</i> in Å)	58.91, 62.33, 151.46	62.45, 62.45, 229.11	39.42, 58.08, 96.35
Asymmetric unit	1 (STI:PPT)	1 (STI:PPT)	1 (STI)
V_M (Å ³ /Da)	3.16	2.56	2.75
Solvent content (%)	61	52	55
No. of total reflections	419,433	118,553	27,868
No. of unique reflections	52,687	26,956	10,238
R_{merge} (%)	7.8	7.9	4.5
Data completeness (%)	91.8 (50.0–1.75 Å) 68.6 (1.80–1.75 Å)	62.0 (60.0–1.80 Å) 37.2 (1.90–1.80 Å)	75.9 (33.0–2.10 Å) 26.7 (2.30–2.10 Å)

^a National Synchrotron Light Source, Brookhaven National Laboratory, USA.

^b Photon Factory, Tsukuba, Japan

^c Seoul National University, Seoul, Korea.

rotation and translation solution was obtained for different resolution ranges. And the refinement of the free STI model converged slowly to an *R*-factor of about 25% for 8.0 to 2.3 Å data. At this time, X-ray data from an orthorhombic crystal of STI:PPT complex became available and its structure was solved by molecular replacement using the model of porcine β -trypsin refined at 1.6 Å (Huang *et al.*, 1993; PDB ID code, 1MCT). After orienting and positioning the PPT model alone according to the molecular replacement solution, an *R*-factor of 42.4% and a correlation coefficient of 0.50 were obtained for 10.0 to 4.0 Å data. A model for the complex was then generated by positioning the partially refined model of free STI through a careful inspection of the electron density and also the interaction between trypsin and the inhibitor. Electron density for the STI part in the complex improved during the successive refinement steps and the complex structure was refined to an *R*-factor of 18.9% for 8.0 to 1.75 Å data. Now, this highly refined model of STI in the complex was used in turn for a structure re-determination of free STI. It gave a clear rotation solution ($\alpha = 88.10^\circ$, $\beta = 31.45^\circ$, $\gamma = 10.49^\circ$) and a translation solution ($x = 0.2719$, $y = 0.2867$, $z = 0.3261$ in fractional coordinates). After the STI model was oriented and positioned according to the above solution, it gave an *R*-factor of 32.6% and a correlation coefficient of 0.73 for 12.0 to 4.0 Å data.

The starting model for STI:PPT(tetra) was the refined model of STI:PPT(ortho). The search for the molecular replacement solution was performed similarly. The correct space group was determined to be $P4_12_12$ by translation function calculations. For 10.0 to 4.0 Å data, an *R*-factor of 35.6% and a correlation coefficient of 0.51 were obtained for $P4_12_12$, compared with an *R*-factor of 46.0% and a correlation coefficient of 0.26 for $P4_32_12$.

Refinement

The STI:PPT(ortho) model was subject to refinement by the program X-PLOR version 3.1 (Brünger, 1992). Initially, a rigid-body refinement was carried out with 8.0 to 6.0 Å data to further improve the positional and orientational parameters. The high-resolution limit of the diffraction data was increased stepwise from 6.0 to 3.5 Å. The *R*-factor at this stage was 36.9% for 8.0 to 3.5 Å data. Next, atomic positions were refined by the conventional conjugate gradient minimization, with

higher resolution data being added in steps. During this stage, the *R*-factor dropped to 27.5% for 8.0 to 2.5 Å data. This model was then subject to a simulated annealing refinement, employing the standard slow cooling protocol: from 3000 K to 300 K (time-step 0.5 fs; decrement of temperature 25 K; number of steps at each temperature 50). Solvent molecules, modeled as water, were located as high peaks in the ($F_o - F_c$) maps and were included in the subsequent rounds of refinement. A water molecule was removed when its temperature factor (*B*-factor) exceeded 60 Å². Higher resolution reflections up to 1.75 Å were added in steps. Further simulated annealing refinement and the refinement of isotropic *B*-factors for individual atoms gave the final *R*-factor of 18.9% for 31,038 unique reflections with $F_o > 2\sigma_F$ in the range 8.0 to 1.75 Å. The final free *R*-factor is 21.4%. The refinements of STI:PPT(tetra) and free STI models were performed similarly. A summary of the refinement statistics is given in Table 1.

Model building and structure analysis

The program CHAIN version 5.4 running on a Silicon Graphics Indigo² XZ workstation was used for model rebuilding (Sack, 1988) and the model stereochemistry was assessed by PROCHECK (Laskowski *et al.*, 1993). Structural comparisons were made using the program LSQKAB in the CCP4 program package (SERC Daresbury Laboratory, 1994). The orthorhombic crystal structure of STI:PPT complex was used as a template to build a homology models of tPA complexes with inhibitors. In homology modeling of ETI:tPA complex, the structure of recombinant two-chain human tPA (Lamba *et al.*, 1996; PDB ID code, 1RTF) was superposed with PPT (with an r.m.s. deviation of 0.92 Å for 181 C α atom pairs). The ETI structure (Onesti *et al.*, 1991; PDB ID code, 1TIE) was also superposed with STI. The resulting models of the complexes between the inhibitors and tPA were then subject to energy minimization by the CHARMM dynamics in the program package QUANTA (Molecular Simulations, Inc.). Programs CHAIN version 5.4 (Sack, 1988), GRASP (Nicholls, 1992), and MOLSCRIPT version 2.1.4 (Kraulis, 1991) were used on a Silicon Graphics Indigo² XZ workstation for model display. The atomic coordinates have been deposited with the Protein Data Bank (ID codes: 1AVU for the free STI, 1AVW for the STI:PPT(ortho), and 1AVX for the STI:PPT(tetra)).

Acknowledgments

This work was supported by the Center for Molecular Catalysis, Seoul National University, S.N.U. Dae Woo Research Fund (95-05-2057), Korea Science and Engineering Foundation (95-0501-07-01-3), and the Korea Ministry of Education, Basic Sciences Research Institute. We thank Dr Robert M. Sweet of the beamline X12C at National Synchrotron Light Source, Brookhaven National Laboratory, USA, and Professor N. Sakabe and Dr M. Suzuki of the beamline BL-6A2, Photon Factory, KEK, Japan for their assistance with data collection. The Inter-University Center for Natural Science Research Facilities, Seoul National University is also thanked for providing the X-ray equipment, which is partially supported by the Specialization Fund from KOSEF.

References

- Baillargeon, M. W., Laskowski, M., Neves, D. E., Porubcan, M. A., Santini, R. E. & Markley, J. L. (1980). Soybean trypsin inhibitor (Kunitz) and its complex with trypsin. Carbon-13 nuclear magnetic resonance studies of the reactive site arginine. *Biochemistry*, **19**, 5703–5710.
- Blow, D. M., Janin, J. & Sweet, R. M. (1974). Mode of action of soybean trypsin inhibitor (Kunitz) as a model for specific protein–protein interactions. *Nature*, **249**, 54–57.
- Bode, W. & Huber, R. (1992). Natural protein proteinase inhibitors and their interactions with proteinases. *Eur. J. Biochem.* **204**, 433–451.
- Brünger, A. T. (1992). *X-PLOR Manual, Version 3.1: A System for Crystallography and NMR*, Yale University Press, New Haven.
- Burley, S. K. & Petsko, G. A. (1985). Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.
- Collen, D. & Lijnen, H. R. (1995). Molecular basis of fibrinolysis, as relevant for thrombolytic therapy. *Thromb. Haemostasis*, **74**, 167–171.
- Crowther, R. A. & Blow, D. M. (1967). A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallog.* **23**, 544–548.
- Dattagupta, J. K., Podder, A., Chakrabarti, C., Sen, U., Dutta, S. K. & Singh, M. (1996). Structure of a Kunitz-type chymotrypsin inhibitor from winged bean seeds at 2.95 Å resolution. *Acta Crystallog. sect. D*, **52**, 521–528.
- Eriksson, A. E., Causens, L. S., Weaver, L. H. & Matthews, B. W. (1991). Three-dimensional structure of human basic fibroblast growth factor. *Proc. Natl Acad. Sci. USA*, **88**, 3441–3445.
- Finzel, B. C., Clancy, L. L., Holland, D. R., Muchmore, S. W., Watenpugh, K. D. & Einspahr, H. M. (1989). Crystal structure of recombinant human interleukin-1 β at 2.0 Å resolution. *J. Mol. Biol.* **209**, 779–791.
- Gewirth, D. (1993). *The Denzo Manual: An Oscillation Data Processing Program for Macromolecular Crystallography by Zbyszek Otwinoski*, Yale University, New Haven, CT.
- Graves, B. J., Hatada, M. H., Hendrickson, W. A., Miller, J. K., Madison, V. S. & Satow, Y. (1990). Structure of interleukin-1 α at 2.7 Å resolution. *Biochemistry*, **29**, 2679–2684.
- Habazetti, J., Gondol, D., Witscheck, R., Otlewski, J., Schleicher, M. & Holak, T. A. (1992). Structure of hisactophilin is similar to interleukin-1 β and fibroblast growth factor. *Nature*, **359**, 855–858.
- Heussen, C., Joubert, F. & Dowdle, E. B. (1984). Purification of human tissue plasminogen activator with *Erythrina* trypsin inhibitor. *J. Biol. Chem.* **259**, 11635–11638.
- Higashi, T. (1989). The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. *J. Appl. Crystallog.* **22**, 9–18.
- Huang, Q., Liu, S. & Tang, Y. (1993). Refined 1.6 Å resolution structure of the complex formed between porcine β -trypsin and MCTI-A, a trypsin inhibitor of the squash family. *J. Mol. Biol.* **229**, 1022–1036.
- Kim, S.-H., Hara, S., Hase, S., Ikenaka, T., Toda, H., Kitamura, K. & Kaizuma, N. (1985). Comparative study on amino acid sequences of Kunitz-type soybean trypsin inhibitors, Ti^a, Ti^b, and Ti^c. *J. Biochem.* **98**, 435–448.
- Kortt, A. A., Strike, P. M. & De Jersey, J. (1989). Amino acid sequence of a crystalline seed albumin (winged bean albumin-1) from *Psophocarpus tetragonolobus* (L.) DC. *Eur. J. Biochem.* **181**, 403–408.
- Kouzuma, Y., Yamasaki, N. & Kimura, M. (1997). The tissue-type plasminogen activator inhibitor ETIa from *Erythrina variegata*: structural basis for the inhibitory activity by cloning, expression, and mutagenesis of the cDNA encoding ETIa. *J. Biochem.* **121**, 456–463.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
- Kunitz, M. (1947a). Crystalline soybean trypsin inhibitor. *J. Gen. Physiol.* **30**, 291–310.
- Kunitz, M. (1947b). Isolation of a crystalline protein compound of trypsin and soybean trypsin inhibitor. *J. Gen. Physiol.* **30**, 311–320.
- Lamba, D., Bauer, M., Huber, R., Fischer, S., Rudolph, R., Kohnert, U. & Bode, W. (1996). The 2.3 Å crystal structure of the catalytic domain of recombinant two-chain human tissue-type plasminogen activator. *J. Mol. Biol.* **258**, 117–135.
- Laskowski, M., Jr., Kato, I., Leary, T. R., Schrode, J. & Sealock, R. W. (1974). Evolution of specific protein proteinase inhibitors. In *Proteinase Inhibitors: Bayer Symposium V* (Fritz, H., Tschesche, H., Greene, L. J. & Truscheit, E., eds), pp. 597–611, Springer, Berlin.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
- Leah, R. & Mundy, J. (1989). The bifunctional α -amylase/subtilisin inhibitor of barley: nucleotide sequence and patterns of seed-specific expression. *Plant Mol. Biol.* **12**, 673–682.
- Lee, J., Song, H. K., Hwang, K. Y., Kim, K. K. & Suh, S. W. (1993). Crystallization and preliminary X-ray crystallographic study of Kunitz-type soybean trypsin inhibitor. *Mol. Cells*, **3**, 335–337.
- Levitt, M. & Perutz, M. F. (1988). Aromatic rings act as hydrogen bond acceptors. *J. Mol. Biol.* **201**, 751–754.
- Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984). Human α_1 -proteinase inhibitor. *J. Mol. Biol.* **177**, 531–556.
- Mareš, M., Meloun, B., Pavlik, M., Kostka, V. & Baudyš, M. (1989). Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. *FEBS Letters*, **251**, 94–98.

- Marquart, M., Walter, J., Deisenhofer, J., Bode, W. & Huber, R. (1983). The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallog. sect. B*, **39**, 480–490.
- McLachlan, A. D. (1979). 3-fold structural pattern in the soybean trypsin inhibitor (Kunitz). *J. Mol. Biol.* **133**, 557–563.
- Murzin, A. G., Lesk, A. M. & Chothia, C. (1992). β -trefoil fold. *J. Mol. Biol.* **223**, 531–543.
- Navaza, J. (1994). AMoRe: an automatic package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157–163.
- Nicholls, A. (1992). *GRASP: Graphical Representation and Analysis of Surface Properties*, Columbia University, New York.
- Onesti, S., Brick, P. & Blow, D. M. (1991). Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds. *J. Mol. Biol.* **217**, 153–176.
- Otwinoski, Z. (1992). *The Scalepack Manual*, Yale University, New Haven, CT.
- Perona, J. J., Tsu, C. A., Craik, C. S. & Fletterick, R. J. (1993). Crystal structures of rat anionic trypsin complexed with the protein inhibitors APPI and BPTI. *J. Mol. Biol.* **230**, 919–933.
- Priestle, J. P., Schär, H.-P. & Grütter, M. (1988). Crystal structure of the cytokine interleukin-1 β . *EMBO J.* **7**, 339–343.
- Read, R. J. & James, M. N. G. (1986). Introduction to protein inhibitors: X-ray crystallography. In *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G., eds), pp. 301–335, Elsevier, Amsterdam.
- Read, R. J., Fujinaga, M., Sielecki, A. R. & James, M. N. G. (1983). Structure of the complex of *Streptomyces griseus* protease B and the third domain of the turkey ovomucoid inhibitor at 1.8 Å resolution. *Biochemistry*, **22**, 4420–4433.
- Ringe, D., Seaton, B. A., Gelb, M. H. & Abeles, R. H. (1985). Inactivation of chymotrypsin by 5-benzyl-6-chloro-2-pyrone: ^{13}C NMR and X-ray diffraction analyses of the inactivator-enzyme complex. *Biochemistry*, **24**, 64–68.
- Rutenber, E. & Robertus, J. D. (1991). Structure of ricin B-chain at 2.5 Å resolution. *Proteins: Struct. Funct. Genet.* **10**, 260–269.
- Sack, J. S. (1988). CHAIN: a crystallographic modeling program. *J. Mol. Graphics.* **6**, 244–245.
- Sakabe, N. (1991). X-ray diffraction data collection system for modern protein crystallography with a Weissenberg camera and an imaging plate using synchrotron radiation. *Nucl. Instrum. Methods A*, **303**, 448–463.
- SERC Daresbury Laboratory (1994). The CCP4 suite: Programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
- Song, S. I., Kim, C. H., Baek, S. J. & Choi, Y. D. (1991). Molecular cloning and nucleotide sequencing of cDNA encoding the precursor for soybean trypsin inhibitor (Kunitz type). *Mol. Cells*, **1**, 317–324.
- Song, H. K., Lee, K. N., Kwon, K.-S., Yu, M.-H. & Suh, S. W. (1995). Crystal structure of an uncleaved α_1 -antitrypsin reveals the conformation of its inhibitory reactive loop. *FEBS Letters*, **377**, 150–154.
- Sweet, R. M. (1989). Facilities at the National Synchrotron Light Source at Brookhaven National Laboratory. In *Synchrotron Radiation in Structural Biology* (Sweet, R. M. & Woodhead, A. D., eds), pp. 63–66, Plenum Press, New York.
- Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H. & Blow, D. M. (1974). Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6 Å resolution. *Biochemistry*, **13**, 1599–1608.
- Takeuchi, Y., Satow, Y., Nakamura, K. T. & Mitsui, Y. (1991). Refined crystal structure of the complex of subtilisin BPN' and *Streptomyces* subtilisin inhibitor at 1.8 Å resolution. *J. Mol. Biol.* **221**, 309–325.
- Teixeira, A. V., Dowdle, E. B. D. & Botes, D. P. (1994). Site-directed mutagenesis of the synthetic *Erythrina* trypsin/tissue plasminogen activator (t-PA) inhibitor encoding-gene to compare the interaction of *Erythrina caffra* and soybean trypsin inhibitor with tPA. *Biochim. Biophys. Acta*, **1217**, 23–28.
- Wlodawer, A., Deisenhofer, J. & Huber, R. (1987). Comparison of two highly refined structures of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **193**, 145–156.
- Yamamoto, M., Hara, S. & Ikenaka, T. (1983). Amino acid sequence of two trypsin inhibitors from winged bean seeds (*Psophocarpus tetragonalobus* (L.) DC). *J. Biochem.* **94**, 849–863.
- Zemke, K. J., Müller-Fahrnow, A., Jany, K.-D., Pal, G. P. & Saenger, W. (1991). The three-dimensional structure of the bifunctional proteinase K/ α -amylase inhibitor from wheat (PKI3) at 2.5 Å resolution. *FEBS Letters*, **279**, 240–247.
- Zhang, J., Cousens, L. S., Barr, P. J. & Sprang, S. R. (1991). Three-dimensional structure of human basic fibroblast growth factor, a structural homolog of interleukin-1 β . *Proc. Natl Acad. Sci. USA*, **88**, 3446–3450.
- Zhu, M., Komiya, H., Chirino, A., Faham, S., Fox, G. M., Arakawa, T., Hsu, B. T. & Rees, D. C. (1991). Three-dimensional structures of acidic and basic fibroblast growth factors. *Science*, **251**, 90–93.

Edited by R. Huber

(Received 30 June 1997; received in revised form 7 October 1997; accepted 9 October 1997)