JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT°

ELSEVIER

Crystal Structure of the Bowman–Birk Inhibitor from Barley Seeds in Ternary Complex with Porcine Trypsin

Eun Young Park[1], Jeom-A. Kim[1], Hyung-Wook Kim[1], Young Sil Kim[2] and Hyun Kyu Song[1]*

[1]*Research Institute, National Cancer Center, 809 Madu-dong Ilsan-gu, Goyang-si Gyeonggi-do 411-764 South Korea*

[2]*School of Chemistry, College of Natural Sciences, Seoul National University, Seoul 151-742, South Korea*

*Corresponding author

The structure and function of Bowman–Birk inhibitors (BBIs) from dicotyledonous plants such as soybean have been studied extensively. In contrast, relatively little is known about the BBIs from monocotyledonous plants such as barley, which differ from dicot BBIs in size and tertiary structure. The BBI from barley seeds (BBBI) consists of 125 amino acid residues with two separate inhibitory loops. Previously we determined the high-resolution structure of a 16 kDa BBBI in the free state. The BBBI folds into two compact domains (N and C domain) with tertiary structures that are similar to that of the 8 kDa BBI from dicots. Here we report the structure of a 1 : 2 complex between BBBI and porcine pancreatic trypsin (PPT) at 2.2 Å resolution. This structure confirms that several regions, including the inhibitory loops in the free BBBI structure, show exceptionally low temperature factors and a distorted conformation due to crystalline packing in the lattice. Extensive analysis of the interaction between BBBI and trypsin, and comparison with other known canonical inhibitor–protease complexes, reveals that the mode of interaction between BBBI and PPT is similar to that of known serine protease inhibitors, as expected; however, several unique features are also identified in the primary binding sites near the inhibitory loops as well as in additional binding sites. The carboxy-terminal tail of the inhibitor extends into the interface between the two trypsin molecules and interacts with both of them simultaneously. The longest distance between the two P1 residues (Arg17 and Arg76) in the complex structure is approximately 34 Å, which is shorter than in the free inhibitor, but it is still possible for BBBI to bind and inhibit two trypsin molecules simultaneously and independently.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: anticarcinogenic activity; double-headed inhibitor; gene duplication; monocotyledonous plant; inhibitory loop

Introduction

There are numerous types of serine protease inhibitors in plant seeds that function both in plant metabolism and as protection against animals, fungi, and bacteria. Bowman–Birk protease inhibitors (BBIs) are among the many types of protease inhibitors from plants that have been studied extensively.[1] The first BBI to be isolated was one from leguminous seeds,[2] and its molecular properties were well characterized.[3] BBIs from dicotyledonous seeds are relatively small (typically ~8 kDa) and rich in disulfide bonds. These BBIs are characteristically double-headed;[1] the 8 kDa inhibitor contains two inhibitory loops, each of which specifically inhibits trypsin or chymotrypsin. These BBIs are stable at cooking temperatures and also toward the acidic pH value found in the digestive systems of humans and animals,[4] most probably due to their large number of disulfide bonds (seven bonds out of ~70 amino acid residues) and the polar interactions between subdomains.[5,6] They have attracted much interest due

to their anti-carcinogenic activity, although the exact mechanism of this activity has yet to be established.[7] Human populations consuming large quantities of BBIs in their diet have been shown to exhibit lower rates of colon, breast, prostate, and skin cancers.[4]

In contrast to BBIs from dicots, BBIs from monocots, including cereal grains, have received relatively little attention. The seeds of mono-cotyledonous plants such as barley contain not only 8 kDa single-headed inhibitors but also 16 kDa inhibitors.[8] The existence of only a single inhibitory loop within the 8 kDa inhibitors of BBIs from monocots is likely due to the second inhibitory loop found in dicots becoming non-functional during evolution. Subsequent to this second loop becoming non-functional, 16 kDa inhibitors such as the Bowman–Birk inhibitor from barley (BBBI) may have evolved in monocots by gene duplication.[8] No experimental data on the anti-carcinogenic activity of BBIs from monocots have been reported.

Several trypsin inhibitors are found in various tissues of barley.[9–13] A 16 kDa trypsin inhibitor isolated from barley rootlets was characterized as a member of the BBI family. This BBI contains two inhibitory loops, making it capable of inhibiting trypsin in a molar ratio of 1 : 2.[13] A 16 kDa trypsin inhibitor has also been isolated from barley seeds.[14] N-terminal sequencing and the sequence deduced from the 1.9 Å resolution crystal structure of this inhibitor showed that it possesses a sequence that is nearly identical with that of the inhibitor from barley rootlet. It consists of 125 amino acid residues and shows an intramolecular sequence identity of 56% between the amino-terminal half (residues 1–63) and the carboxy-terminal half (residues 64–125). The inhibitory loops are the most conserved regions in the sequence and the P1 residues are Arg17 and Arg76 for the N and C domains, respectively.

The three-dimensional structures of several 8 kDa dicotyledonous BBIs have been characterized in the free form[15–18] as well as in complex with trypsin.[19,20] We previously reported the crystal structure of BBBI in the free state at high resolution and the detailed structural comparison of BBIs in cereal grains (i.e. monocots) with those in legumi-nous plants (i.e. dicots).[14] BBBI consists of two separate domains (the N and C domains) and each domain shares nearly the same overall structure with the 8 kDa BBIs.[14] Although a 1 : 2 complex between BBBI and trypsin has been modeled based on the assumption that BBBI acts as a canonical inhibitor, the accuracy of this analysis is question-able and needs to be confirmed by experiment. Recently, the structure of the complex between wheat germ BBI I-2b and bovine trypsin has been reported, although only to a very limited extent,[21] and the structural data are not available from the RCSB Protein Data Bank. Here we present the crystal structure of BBBI in complex with porcine pancreatic trypsin (PPT) at 2.2 Å resolution. This work markedly expands our previous modeling

study of a 1 : 2 complex between BBBI and trypsin and allows a direct structural comparison between the free and complexed states.[14] We also provide a wealth of structural information on a 16 kDa monocot inhibitor complexed with its target pro-tease that helps to elucidate the critical interactions in the primary binding sites near the inhibitory loops and in additional binding sites such as the carboxy-terminal tail.

Results and Discussion

Structure determination and model quality

The structure of the BBBI–PPT complex was solved by molecular replacement using the model of PPT refined to 1.75 Å (PDB ID: 1AVW). An initial round of MOLREP trials gave two clear molecular replacement solutions among the four PPT molecules in the asymmetric unit. For the next round of the calculation, we fixed the first two solutions and then obtained the solutions of the other two PPT molecules.[22] Although the rotation and translation solutions of the trypsin molecules were readily obtained, calculating those of the inhibitor molecules was not straightforward. How-ever, the electron density for both inhibitory loops of the inhibitors was clearly interpretable with the partial phases from the PPT models. The free BBBI model (PDB ID: 1C2A) was slotted into the electron density between the two trypsin models. The complete model (two 1 : 2 BBBI–PPT complexes) was refined to a free R-factor of 27.4% for 30.0–2.2 Å data with excellent stereochemistry. Table 1 sum-marizes the refinement statistics as well as model quality parameters. Here, residues of PPT are numbered according to the chymotrypsin sequence numbering scheme and a prime (′) is appended to the residue number for clarity.

The model accounts for all 892 (223×4) PPT residues and 224 of 250 (125×2) BBBI residues, as well as four calcium ions and 492 water molecules.

Table 1. Refinement statistics

Resolution range (Å)	30.0–2.2
No. of reflections (no cut-off)	57,380
Completeness (%)	91.6
R-factor/R_{free} (%)	22.3/27.4
No. of BBBI residues	110 (I chain), 114 (J chain)
No. of PPT residues	223 (all A, B, C, D chains)
No. of non-hydrogen atoms	
Protein	1705 (BBBI), 6568 (PPT)
Water/Ca^{2+}	492/4
Average B-factor (Å^2)	
Main-chain	50.3 (BBBI), 34.3 (PPT)
Side-chain	51.9 (BBBI), 35.4 (PPT)
Water/Ca^{2+}	37.9/62.2
r.m.s.deviation from ideal geometry	
Bond lengths (Å)	0.006
Bond angles (°)	1.38
Ramachandran outliers (%)	None

Only 5% of the data were used for the free R-factor calculation.
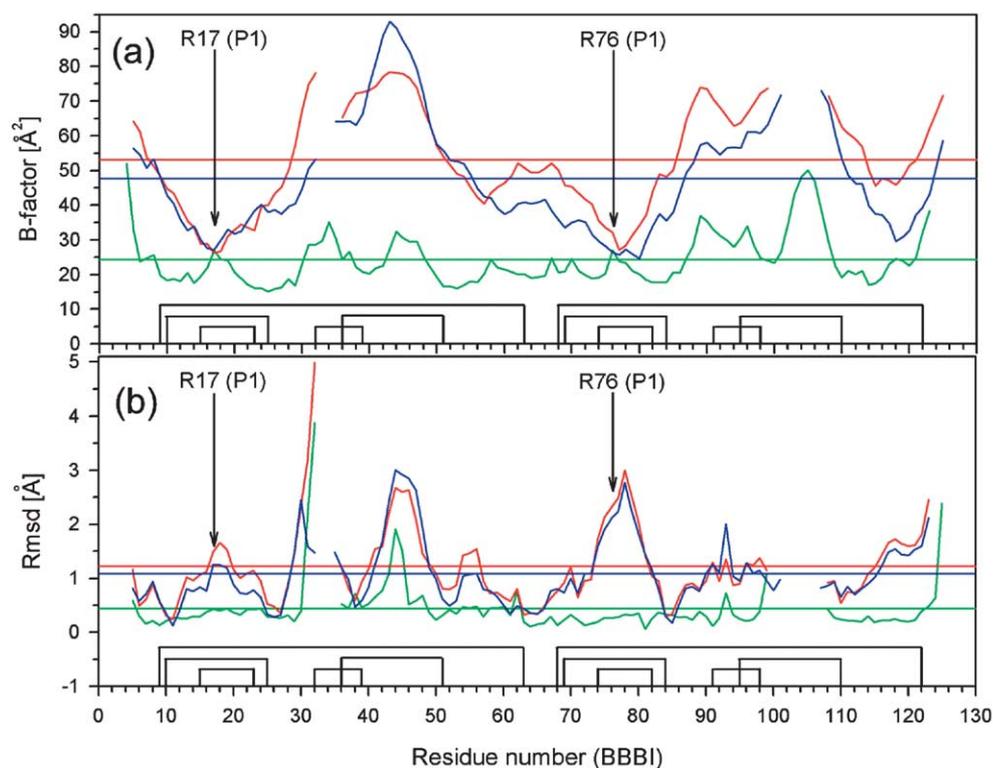
Figure 1. (a) *B*-factor plot for the BBBI models. The average *B*-factors for the main-chain atoms of each residue are plotted as a function of residue number (I chain, red line; J chain, blue line; free, green line). (b) Plot of the difference between free and complexed BBBI models. The r.m.s. differences for the main-chain atoms of each residue are plotted as a function of residue number (free BBBI *versus* I-chain of BBBI–PPT, red line; free BBBI *versus* J chain of BBBI–PPT, blue line; I chain *versus* J chain, green line). P1 residues (Arg17 and Arg76) are indicated as well as the ten disulfide bridges (9:63, 10:25, 15:23, 32:39, and 36:51 in the N domain and 68:122, 69:84, 74:82, 91:98, and 95:110 in the C domain). The horizontal lines indicate the average values of each plot.

Missing from the inhibitor part of BBBI-PPT model are the four N-terminal residues (1–4) and two segments (residues 33–35; 100–107 for the I-chain, residues 33–34; 102–106 for the J-chain) that are otherwise clearly seen in the free BBBI model.[14] When the free BBBI molecule is exposed to trypsin,

peptide bonds might be cleaved after one or more of the lysine and arginine residues in the highly mobile loop regions (Figure 1(a)). Indeed, SDS-PAGE analysis of protein from the crystals showed that the polypeptide chain of BBBI was cleaved (data not shown). Among 919 non-glycine and
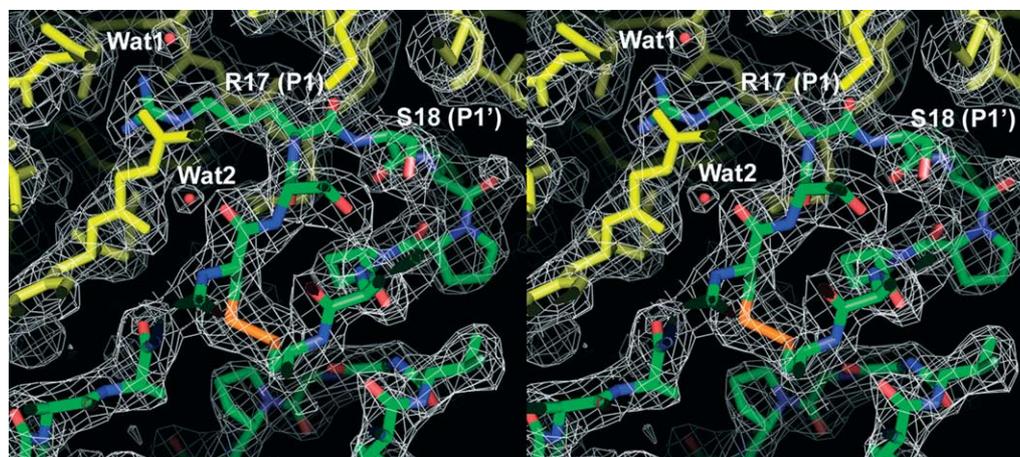
Figure 2. Final $(2F_o - F_c)$ electron density map around the inhibitory loop in the N domain of BBBI. The map was calculated using 30–2.2 Å data and contoured at 1.0σ. Residues in the scissile bond in the inhibitory loop and conserved water molecules are labeled.
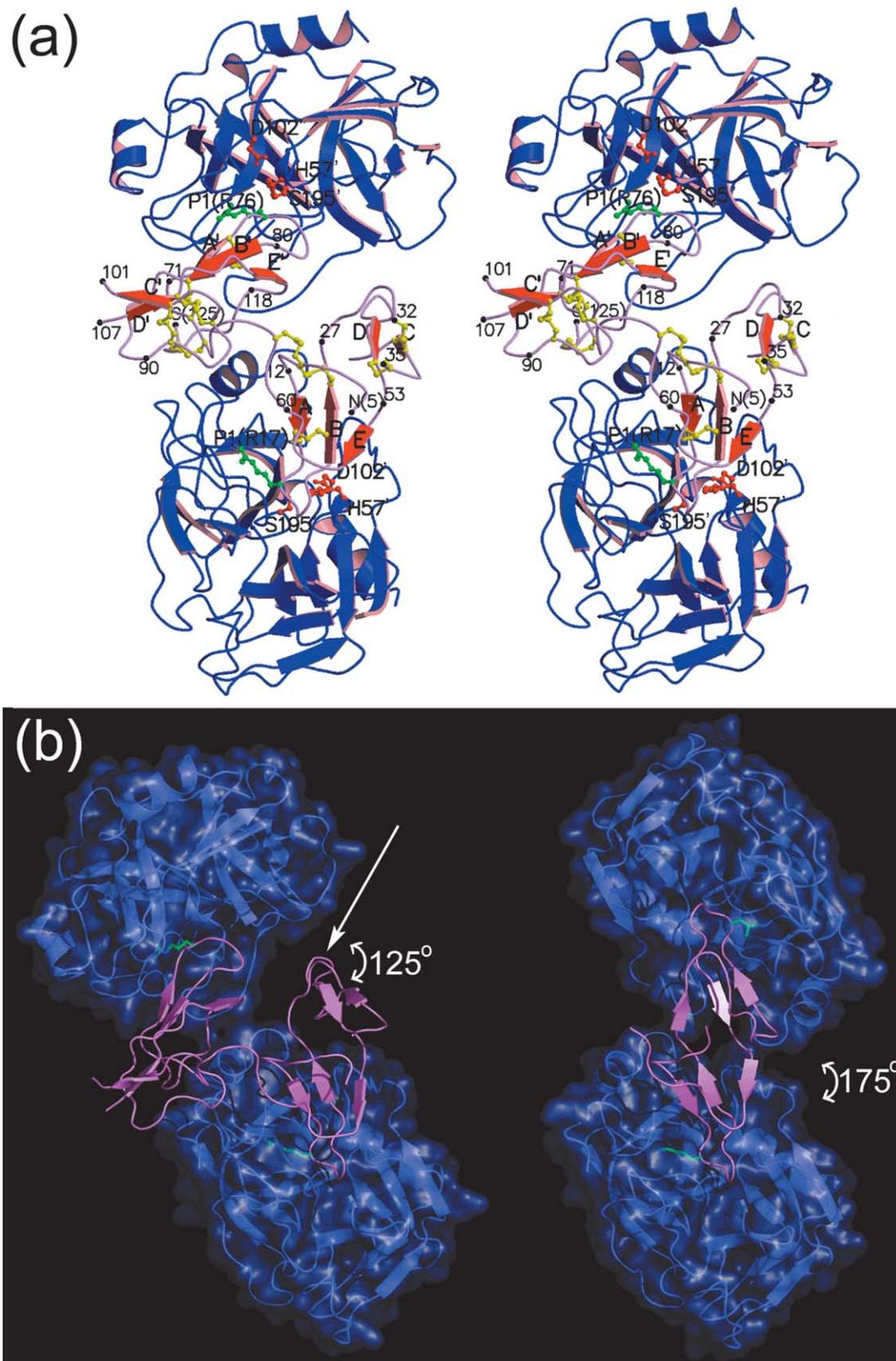
Figure 3. (a) A stereo ribbon diagram of the 1 : 2 complex between BBBI and PPT. Ribbons and arrows represent α-helices and β-strands, respectively. P1 residues (Arg17 and Arg76) are indicated as well as the ten disulfide bridges. The catalytic triad of the trypsin molecule (His57′, Asp102′, Ser195′) is presented and labeled. The secondary structure elements of BBBI are also labeled and approximately every tenth residue is labeled and marked by dots. The secondary structure elements of BBBI were assigned by PROCHECK.[42] This Figure was drawn by MOLSCRIPT[46] and rendered

non-proline residues, the number of residues lying in the most favored, additionally allowed, generously allowed, and disallowed regions in the Ramachandran plot are 785 (85.4%), 134 (14.6%), 0, and 0, respectively.

The average *B*-factor for the main-chain atoms of each residue in the BBBI models is plotted as a function of residue number in Figure 1(a). The average *B*-factors for the main-chain atoms of BBBI in the complexed state are much higher than in the free state. Indeed, the free BBBI exhibited an exceptionally low overall *B*-factor (26.6 Å^2) compared with other protease inhibitors,[14] indicating that the inherent flexibility of the inhibitor might be reduced by crystalline contacts in the free structure, as noted previously.[14] Besides the inhibitory loop, which forms a tight complex with trypsin, most of the other regions of the complexed BBBI show high *B*-factors for the main-chain atoms in both chains (Figure 1(a)). The electron densities for several loop regions are weak but the tracing of the polypeptide chain is not ambiguous and, as mentioned earlier, some of these regions are proteolytically susceptible. The final $(2F_o - F_c)$ electron density for the inhibitory loop of BBBI in the BBBI–PPT complex, calculated using data to 2.2 Å, is very well-defined (Figure 2), reflecting the low *B*-factors around this region (Figure 1(a)). This is due to the involvement of these residues in extensive contacts with the trypsin molecule, which stabilizes the conformation of the inhibitory loop.

Overall structure

The stereo ribbon diagram of the overall ternary complex is shown in Figure 3(a). It has an elongated shape with approximate dimensions of 95 Å × 45 Å × 44 Å. The BBBI structure in the BBBI–PPT complex shows essentially the same secondary structural elements and folding topology as those previously reported for free BBBI.[14] The buried surface area of the inhibitor molecule upon complex formation is approximately 1080 Å^2 and 820 Å^2 for the N and C domains of BBBI, respectively (average value for two complex molecules). Therefore, the N domain of BBBI shows more extensive interaction with trypsin than does the C domain. The total buried surface for the complex between N domain of BBBI and trypsin (~2160 Å^2) is larger than that of other protease–inhibitor interfaces, while the total buried surface for the complex between C domain of BBBI and trypsin (~1640 Å^2) is within the commonly observed range of 1600(±400) Å^2.[23] These values are much larger than those of BBI

from soybean (SBBI), which is smaller in size, due to the additional contacts observed in the BBBI–PPT complex (Figure 3). The buried surface area of sub-domain 1 (the N-terminal half) in SBBI is also larger than that in sub-domain 2 (C-terminal half), as in the case of BBBI.[24] Over 25% of the total surface area of BBBI is covered by two trypsin molecules (~1890 Å^2 out of 7550 Å^2). In the case of SBBI, the total covered surface is nearly 35% of the molecular surface (~1500 Å^2 out of 4240 Å^2). In the structure of the 1 : 2 complex, bound trypsin molecules come very close to one another, but the interaction between them is minimal. The main-chain oxygen atom of Ser170′ makes a water-mediated interaction with that of Ser147′ in the adjacent trypsin molecule. However, the electron density of the water molecule involved in this interaction is weak and the interaction occurs in only one of the two 1 : 2 complexes in the asymmetric unit, making the significance of the water-mediated interaction unclear.

There is an extensive hydrogen-bonding network between the two domains (Figure 4(a)). Detailed analysis of the interaction between the two domains reveals that the side-chain atoms of Arg64 are very important for stabilization of the relative orientation of these two domains (Figure 4(a)). The side-chain oxygen atoms of Asp11 in the N domain form hydrogen bonds with the main-chain nitrogen atom of Arg64 and the side-chain hydroxyl group of Tyr114 in the C domain, as reported. The two reactive sites (Arg17-Ser18 and Arg76-Ser77; P1–P1′) are located at opposite sides of the inhibitor structure on protruding loops between strands A and B in the N domain (or A′ and B′ in the C domain) (Figure 3(a)). Interestingly, the longest distance between the two P1 residues is approximately 34 Å, which is slightly shorter than in the free BBBI (40 Å) and 8 kDa double-headed BBIs from dicotyledonous seeds (36–40 Å).[5,14,18] This separation still allows the two inhibitory loops to bind to and inhibit two trypsin molecules simultaneously and independently. However, the relative spatial orientations of the two protease molecules when bound to double-headed BBIs from dicotyledonous and monocotyledonous seeds are markedly different (Figure 3(b)). This difference arises because the two inhibitory loops in the 8 kDa double-headed dicotyledonous BBIs are related by a pseudo-2-fold (approximately 175°) axis relating the two sub-domains, whereas those in the 16 kDa double-headed BBBI are related by approximately 125° in the polar angle (Figure 3(b)). Intriguingly, the loops related by the pseudo-2-fold symmetry in

with Raster3D.[47] (b) A comparison of trypsin's orientation in the BBBI–PPT and SBBI–trypsin complexes. This view is designed to show the pseudo-2-fold symmetry of the SBBI–trypsin complex, and the bottom trypsin molecules were superimposed for viewing the BBBI–PPT complex. Trypsin molecules are presented as transparent surfaces and inhibitor molecules as ribbons. The side-chain atoms of P1 residues are drawn and colored green. The white arrow indicates the position of the non-functional inhibitory loop of BBBI. This Figure was drawn with PyMOL (http://pymol.sourceforge.net/).
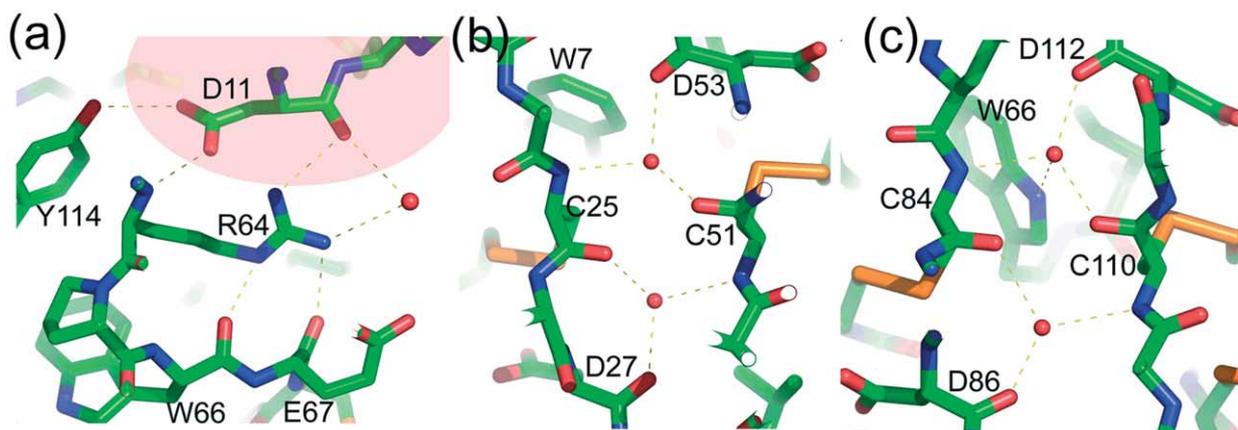
Figure 4. (a) A diagram showing the interface between the N and C domains of BBBI. The hydrogen bonding interactions are indicated by dots. The area belonging to the N domain of BBBI is shaded by transparent pink. A diagram showing the interactions between the two buried water molecules and surrounding backbone atoms: (b) N domain and, (c) C domain. The hydrogen bonding interactions are indicated by dots and the intriguing tryptophan residues (Trp7 and Trp66) are also drawn. This Figure was drawn with PyMOL (http://pymol.sourceforge.net/).

each domain of BBBI (nonfunctional loop in monocot) are relatively accessible to trypsin molecules. Indeed, the corresponding loop in the N domain of BBBI shows the highest *B*-factors (Figure 1(a)), reflecting the highest flexibility, and the corresponding loop in the C domain of BBBI might act as a substrate.

Solvent structure

Bound solvent molecules, mostly water, are an integral part of protein structure. In many proteins, bound water molecules play structurally and/or functionally important roles. In particular, a detailed analysis of a protein–protein interaction requires a set of well-defined water molecules. However, in order to locate water molecules in the crystal with confidence, the protein structure has to be refined to high resolution, typically better than about 2.5 Å. In this study, the BBBI–PPT complex structure has been refined to high enough resolution to locate ordered water molecules with confidence. The refined model of BBBI–PPT includes a total of 492 ordered solvent molecules, all modeled as water. Eight water molecules (two for each BBBI–PPT interface) are involved in mediating the interaction between BBBI and PPT (Figure 2). They occupy nearly equivalent positions in all four trypsin–inhibitor interfaces and show *B*-factors that are less than the average (22.9 Å^2 and 31.3 Å^2, respectively; average value for four equivalent water molecules). The water molecule Wat1 in the active-site pocket occupies a position that favors hydrogen bonds with Arg17 (Arg76; P1) NH2, Trp215′ O, Val227′ O, and Ser190′ OG atoms. Another water molecule, Wat2, makes hydrogen bonds with Arg17 (Arg76; P1) NE, Cys15 (Cys74; P3) O, Gln192′ OE1, Gly216′ O, and Gly219′ O atoms. Similar water molecules are found in other trypsin–inhibitor complexes: Wat804 and Wat810 of

soybean Kunitz-type trypsin inhibitor (STI)–PPT; Wat403 and Wat416 in bovine pancreatic trypsin inhibitor (BPTI)–bovine pancreatic trypsin (BPT); and Wat58 and Wat5 in trypsin inhibitor from bitter gourd seeds: PPT.[25–27] The positions of these water molecules differ within the trypsin–inhibitor complexes by 0.1 Å to 0.6 Å. In the free BBBI structure, a water molecule (Wat161) forms a hydrogen bond with the side-chain oxygen atom (OG) of Ser77 (P1′) in the reactive site loop of the C domain.[14] In the BBBI–PPT complex, this water molecule also forms a hydrogen bond with the side-chain oxygen atom of Thr81 (P5′). Although no equivalent water molecule is present in the N domains of either the BBBI or dicot BBI structures, this water molecule likely plays a significant role in the canonical conformation of the inhibitory loop in the C domain.

Buried water molecules are present in the domain interface and are an integral part of each domain (Figure 4). A water molecule forms a hydrogen bond with the main-chain oxygen atom of Asp11 in the N domain and the side-chain nitrogen of Arg64 in the C domain (Figure 4(a)). The contribution of this water towards stabilizing the interaction between the N and C domains seems to be weak, as evidenced by its higher than average temperature factor (42.1 Å^2). Two water molecules are located symmetrically around the intramolecular pseudo-2-fold axis that relates the two sub-domains of either the N or C domain, and each water molecule donates two hydrogen bonds to two peptide C=O groups and accepts one from the peptide N–H group, as seen in free BBBI.[14] Similar water molecules have also been observed in structures of other 8 kDa dicotyledonous BBIs.[5,16,20] These water molecules seem to be essential for maintaining the tertiary structures of BBIs. However, these structural water molecules are in markedly different environments
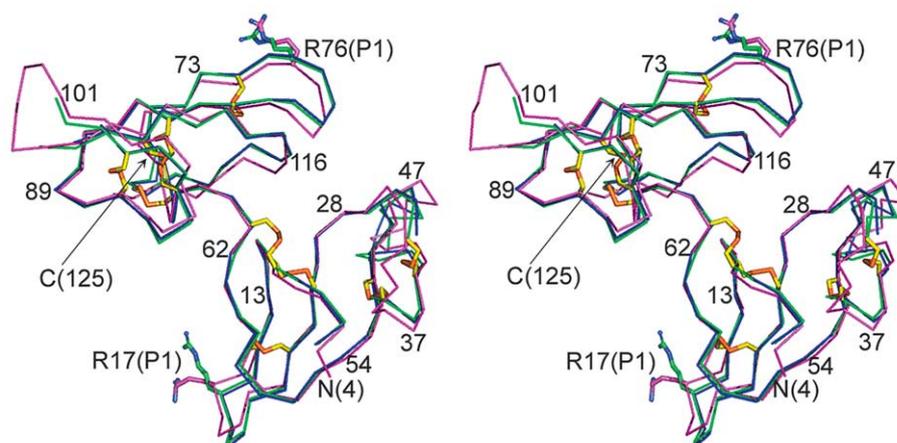
Figure 5. A stereo diagram showing the superposition of C$^{\alpha}$ atoms in the free and trypsin-complexed BBBI models. Magenta, blue, and green lines represent free, I chain, and J chain, respectively. Disulfide bridges are shown in yellow. Side-chain atoms of P1 residues (Arg17 and Arg76) are drawn and labeled. The N and C termini and approximately every tenth residue are also labeled.

in dicotyledonous and monocotyledonous BBIs. In contrast to non-conserved residues in dicot BBIs (Lys6 in SBBI, Glu10 in mung bean protease inhibitor, and Asn9 in peanut protease inhibitor (A-II)), there is a strictly conserved tryptophan residue in BBBI (Trp6 in the N domain and Trp66 in the C domain) near these interesting structural water molecules (Figure 4(b) and (c)). Intriguingly, the side-chain conformation of the tryptophan residue is also different in each domain although the backbone structure is virtually same. The side-chain nitrogen atom of Trp66 in the C domain participates in a hydrogen-bonding network (Figure 4(c)).

Comparison of the free and trypsin-complexed BBBI models

All three BBBI models (two complexed and one free) are superposed in Figure 5. The most striking change in BBBI upon complexation with PPT is a shift in the positions of the inhibitory loops (residues 14–20 and 73–79) and also of the C terminus (Figures 1b and 5), which is invisible in the free BBBI structure.[14] This indicates that the relatively flexible inhibitory loops as well as the C terminus of BBBI become rigid upon binding to trypsin. In general, the inhibitory loop of a standard mechanism protease inhibitor is relatively flexible in its free state and becomes less mobile upon binding to target proteases.[27,28] However, the over-all B-factor of the inhibitory loops of BBBI is relatively low even in the free state (Figure 1(a)). This is partly due to stabilization of the inhibitory loops through intermolecular contacts with symmetry-related molecules in the crystal (Cys15 (P3) O$\cdots$N Arg102#, 3.0 Å; Arg17 (P1) N $\cdots$ O Pro100#, 2.6 Å; Cys74 (P3) O $\cdots$ OE1 Gln52#, 3.2 Å). However, in solution the flexibility of the inhibitory loops of BBBI may be high, as observed

for the inhibitory loops of A-II[16] and other inhibitors.[27,29]

We first analyze the two chains of BBBI in the complex structure. A main-chain superposition of the I and J chains of the complexed BBBI gives a root-mean-square (r.m.s.) difference of 0.44 Å for 441 matching atoms (1.10 Å for all 834 non-hydrogen atoms). A plot of the distance between equivalent main-chain atoms as a function of residue number is given in Figure 1(b). Deviations of greater than 1.0 Å for main-chain atoms are observed for residues 31–32, 43–45, 99, and 125. The large differences are confined to the external loops (some of them might be proteolytically cleaved) and the C-terminal residue Arg125. These regions show high B-factors for the main-chain atoms (Figure 1(a)), indicating that the coordinate errors of these regions are larger than the average. This large discrepancy has contributions from the large coordinate errors in these regions as well as from the different loop conformations due to different crystalline environments. Therefore, the structural difference between the I and J chain models of BBBI in the complexed BBBI can be regarded as minor. The r.m.s. differences between the main-chain atoms of the inhibitory loops (P4–P3'; Val14-Pro20 and Ile73-Pro79) of the I and J chains are only 0.10 Å (0.19 Å for all 52 matching atoms) and 0.11 Å (0.26 Å for all 53 matching atoms) for 28 matching atoms, respectively. This result suggests that the I and J chains have essentially the same conformation in complexed BBBI, except for some differences in the external loops (Figure 5).

A main-chain superposition of BBBI models in the free state and from the BBBI–PPT complex gives an r.m.s. difference of 1.22 Å for 433 atom pairs (1.76 Å for all 822 non-hydrogen atoms) in the I chain and 1.09 Å for 448 atom pairs (1.56 Å for all 853 non-hydrogen atoms) in the J chain. Therefore, the J chain of complexed BBBI is more similar to free

Table 2. r.m.s. differences in Å of the main-chain atoms of seven matching residues (P4–P3′) in the inhibitory loop of Bowman–Birk inhibitors

	Free BBBI-C	PPT–BBBI-N(I)	PPT–BBBI-C(I)	PPT–BBBI-N(J)	PPT–BBBI-C(J)
Free BBBI-N	0.58	0.34	0.53	0.34	0.43
Free BBBI-C		0.58	0.31	0.54	0.32
PPT–BBBI-N(I)			0.52	0.10	0.44
PPT–BBBI-C(I)				0.47	0.11
PPT–BBBI-N(J)					0.39

BBBI than is the I chain. A plot of the distance between equivalent main-chain atoms as a function of residue number is given in Figure 1(b). Residues showing deviations greater than 2.0 Å for main-chain atoms are 30–32, 43–47, 75–80, and 123 for the I chain and 30, 43–47, 76–79, 93, and 123 for the J chain. As mentioned earlier, some of the largest differences are observed in the inhibitory loop (Figure 1), although the inhibitory loop is the most conserved part of the sequence.[14] The r.m.s. differences between main-chain atoms of the two inhibitory loops, P4–P3′; Val14-Pro20 and Ile73-Pro79, of free BBBI and those of the I chain from the complex are 0.34 Å (0.85 Å for all 52 matching atoms) and 0.34 Å (0.87 Å for all 53 matching atoms) for 28 matching atoms, respectively. The r.m.s. differences between main-chain atoms of the two inhibitory loops, P4–P3′; Val14-Pro20 and Ile73-Pro79, of free BBBI and those of the J chain from the complex are 0.31 Å (0.69 Å for all 52 matching atoms) and 0.32 Å (0.74 Å for all 53 matching atoms) for 28 matching atoms, respectively. Therefore, the local conformations of the inhibitory loops in BBBI are very similar and are little affected by complex formation, as expected for a canonical inhibitor. Thus, the large displacement in the relative position of the inhibitory loop between free and complexed BBBI is most likely due to crystal packing, as reported.[14] Table 2 presents a summary of the comparison between the inhibitory loops (P4–P3′) from several BBBI models.

Inhibitory loops: conformation and comparison with other inhibitors

BBBI belongs to the family of substrate-like inhibitors, which possess two exposed inhibitory loops with a characteristic canonical conformation. The inhibitory loops of BBBI are constrained by disulfide bridges (Cys15-Cys23 in the N domain and Cys74-Cys82 in the C domain), as in many other proteinase inhibitors,[30] and a characteristic polyproline-II conformation (at the position of P3′ and P4′) that may limit its conformational freedom. The backbone conformation of the inhibitory loops of BBBI does not change significantly upon forming a complex with PPT, as indicated by similar main-chain dihedral angles for the P4–P3′ residues (Table 3). Table 3 lists the conformational angles of several other inhibitors as well as their amino acid sequences around the scissile peptide bond. Between BBBI and BPTI, there is a large deviation

in the ϕ and ψ angles of the P3 and P4 positions. Between BBBI and STI, a large deviation is observed for ϕ and ψ angles of P3 and the ψ angle of P2′. In contrast, BBBI and SBBI, which belong to the same class of inhibitors, show similar conformations at all positions in the inhibitory loops. Thus, not-withstanding some differences in the main-chain conformations at outer positions of the inhibitory loop, the residues that primarily interact with the target proteases are quite similarly oriented in several substrate-like inhibitors.

When the C^α atoms of the P4–P3′ residues of BBBI (J chain) are superposed with the corresponding atoms of other inhibitors, the r.m.s. deviations are 0.24 Å and 0.33 Å for the trypsin and chymotrypsin inhibitory loops of SBBI (PDB ID: 1D6R), respectively. The deviations are 1.74 Å for BPTI (PDB ID: 3TGI) and 2.03 Å for STI (PDB ID: 1AVW). Removing just the P4 residue decreases the r.m.s. deviations dramatically to 0.76 Å and 1.06 Å for BPTI and STI, respectively. Therefore, the confor-mations of the central portions of the inhibitory loops, especially P2–P2′, are very similar. In contrast, the conformations outside the P2–P2′ residues can be very different in different inhibitors.

Knowledge of the geometry of the P1 carbonyl group is important to understanding the interaction between protease and inhibitor during the catalytic reaction. In the earlier and lower resolution structure reported for the STI–PPT complex,[31] a tetrahedral intermediate was proposed, but in subsequent high resolution studies of various complexes, the out-of-plane deformation of the P1 carbonyl group was observed to be small or negligible,[25,26,30,32] while Ser195′ OG and the carbonyl carbon atoms were found to be within the "sub-van der Waals" distance (typically around 2.7 Å).[30] In our BBBI–PPT models, the distance is 2.70 Å (Arg17 (P1) C···OG Ser195′) and 2.94 Å (Arg76) for the I chain and 2.86 Å (Arg17) and 2.96 Å (Arg76) for the J chain, respectively. In order to obtain an unbiased geometry of the P1 carbonyl group, refinements were carried out either without or with the restraints applied to the carbonyl carbon atom. All four P1 carbonyl groups of the two BBBIs in the asymmetric unit display negligible out-of-plane displacement and thus each P1 carbonyl group of BBBI retains a nominal trigonal planar geometry when complexed with PPT. Similar behavior has been reported for many inhibitor–protease complexes, indicating that BBBI belongs to a typical substrate-like inhibitor family in which the

Table 3. Comparison of inhibitory loop dihedral angles (ϕ/ψ) of several protease inhibitors

	P4	P3	P2	P1	P1′	P2′	P3′
BBBI-N	V	C	T	R	S	I	P
BBBI-C	I	C	T	R	S	N	P
SBBI-N	A	C	T	K	S	N	P
SBBI-C	I	C	A	L	S	Y	P
BPTI	G	P	C	K	S	R	I
STI	S	P	Y	R	I	R	F
Free BBBI-N	−100/107	−123/155	−83/153	−87/80	−144/165	−112/109	−81/165
Free BBBI-C	−124/109	−121/154	−81/178	−87/−5	−70/160	−126/106	−76/152
PPT–BBBI-N(I)	−82/113	−125/132	−68/161	−105/40	−99/177	−121/110	−83/166
PPT–BBBI-C(I)	−115/124	−136/160	−79/166	−105/33	−89/171	−134/113	−70/158
PPT–BBBI-N(J)	−80/118	−128/132	−66/164	−106/36	−96/180	−124/110	−79/171
PPT–BBBI-C(J)	−105/118	−132/154	−80/170	−112/36	−92/174	−138/113	−69/165
Free SBBI-N	−137/139	−148/132	−69/147	−78/101	−179/153	−105/118	−81/171
Free SBBI-C	−125/113	−115/149	−60/131	−73/81	−174/150	−97/130	−73/163
BPTI–SBBI-N	−121/135	−139/145	−71/156	−95/45	−96/162	−108/97	−88/171
BPTI–SBBI-C	−118/130	−127/163	−85/159	−95/21	−98/179	−127/106	−91/139
Free BPTI	86/176	−86/−6	−81/165	−104/9	−76/172	−127/76	−105/121
BPT–BPTI	78/174	−77/−29	−70/155	−116/39	−87/164	−112/79	−98/124
Free STI	−112/149	−70/−28	−62/158	−84/21	−60/148	−90/−22	−127/126
PPT-STI	−114/144	−58/−34	−56/139	−89/38	−83/148	−67/−38	−118/155

reactive site loop takes on a canonical conformation in both the free and the complexed state.

Mode of interaction between BBBI and PPT

The interaction between BBBI and PPT is summarized in Table 4 and drawn in Figure 6. There are some minor differences between the interaction patterns of the I and J chain BBBI models. Twenty-seven amino acid residues out of the 125 in BBBI make contact with PPT in the J chain model. They are Glu12 (P6), Ala13 (P5), Val14 (P4), Cys15 (P3), Thr16 (P2), Arg17 (P1), Ser18 (P1′), Ile19 (P2′), Pro21 (P4′), Met43 (P26′), and Gln54 (P37′) in the N domain and Glu67 (P10), Lys71 (P6), Ala72 (P5), Ile73 (P4), Cys74 (P3), Thr75 (P2), Arg76 (P1), Ser77 (P1′), Asn78 (P2′), Pro80 (P4′), Arg83 (P7′), V118 (P32′), Pro120 (P34′), Arg121 (P35′), and Pro124 (P38′) in the C domain. In the I chain model, the three residues Ala72 (P5), Val118 (P32′), and Arg120 (P34′) do not interact with PPT and two additional residues, Gln52 (P35′) and Cys122 (P36′), interact with PPT. However, the pattern for the hydrogen bonding interaction involving the inhibitory loop residues from P4 to P2′ is well conserved among all four trypsin binding sites (Table 4). A detailed description of the interactions is given below.

Primary binding site

Most of the contacts between BBBI and PPT involve the six residues in the inhibitor loop (P4–P2′). In the I chain model, 13 of the 15 hydrogen bonds between the N domain of BBBI and PPT and 16 of the 18 hydrogen bonds between the C domain of BBBI and PPT are accounted for by these six residues, while in the J chain 15 of the 16 hydrogen bonds between the N domain of BBBI and PPT and 16 of the 20 hydrogen bonds between the C domain of BBBI and PPT are accounted for. The P4 residue

Val14 (Ile73) interacts with several hydrophobic residues in PPT (Table 4) and, among these, Trp215′ seems to be important because it is found in all binding interfaces. The main-chain oxygen and nitrogen atoms of the P3 residue Cys15 (Cys74) form hydrogen bonds with those of Gly216′ in PPT. The threonine residue at the P2 position (Thr16 and Thr75 in BBBI) is relatively well conserved although aspartic acid is in the first sub-domain of group IV BBIs and alanine/asparagine residues are also found in the second sub-domain of several BBIs. Extensive derivatization by means of combinatorial chemistry on the P2 of BBI showed that a threonine residue at the P2 position provides optimal inhibition.[33] This residue forms a hydrogen bond network that maintains the characteristic conformation of the inhibitory loops.[5,14] The P2 residue makes extensive van der Waals contact with His57′ in PPT. The P1 residue Arg17 (Arg76) makes the most extensive hydrogen bonds with PPT, forming 10 or 11 hydrogen bonds in total. The side-chain of the P1 arginine residue occupies its expected position in the primary binding pocket of PPT. The guanidinium group makes an ionic interaction with the carboxylate group of Asp189′ in PPT. Hydrogen bonding may occur between the nitrogen atoms of the side-chain of Arg17 (Arg76) and the side-chain oxygen atom (OG) of Ser190′ and the carbonyl oxygen atom of Gly219′. The main-chain oxygen and nitrogen atoms of the P1 residue form hydrogen bonds with those of Gly193′, Asp194′, Ser195′, and Ser214′ in PPT. Although the side-chain oxygen atom (OG) of the P1′ serine residue forms a hydrogen bond with the side-chain oxygen atom of P2 Thr, substitution of this residue with alanine has no effect on the integrity of the inhibitory loop.[34] However, alanine substitution increases the equilibrium dissociation constant for trypsin by a factor of 4. A previous biochemical study using proline variant at the P1′ position suggests that this

Table 4. Total interactions between BBBI and PPT

Site	BBBI	PPT	No. of interactions		Note
			PPT–BBBI(I)	PPT–BBBI(J)	
P6	Glu12	Tyr217′	1, –	2, 1 (OE2–OH: 3.24)	
P5	Ala13	Tyr217′	1, –	1, –	
P4	Val14	Leu99′	1, –		
		Gln175′	1, –		
		Trp215′	2, –	4, –	
		Tyr217′	1, –	1, –	
P3	Cys15	Trp215′	2, –	3, –	
		Gly216′	–, 1 (O–N: 3.34)	–, 2 (N–O: 3.29)	
				(O–N: 3.26)	
P2	Thr16	His57′	7, –	8, –	
		Leu99′	2, –	2, –	
		Gln192′	1, –	3, –	
		Ser214′	1, –	1, –	
		Trp215′		1, –	
P1	Arg17	Asp189′	4, 2 (NH1–OD2: 2.69)	2, 2 (NH1–OD2: 2.97)	Ion
			(NH2–OD1: 2.89)	(NH2–OD1: 3.13)	
		Ser190′	3, 2 (NH2–OG: 2.88)	1, 1 (NH2–OG: 3.03)	
			(NH2–O: 2.82)		
		Cys191′	2, –	2, –	
		Gln192′	6, –	5, –	
		Gly193′	2, 1 (O–N: 2.64)	2, 1 (O–N: 2.45)	
		Asp194′	–, 1 (O–N: 3.32)	–, 1 (O–N: 3.20)	
		Ser195′	8, 2 (N–OG: 2.91)	6, 3 (N–OG: 2.78)	
			(O–N: 2.92)	(O–N: 3.02)	
				(O–OG: 2.94)	
		Ser214′	–, 1 (N–O: 3.07)	–, 1 (N–O: 3.09)	
		Trp215′	1, –		
		Gly216′	1, –	1, –	
		Gly219′	–, 1 (NH1–O: 3.06)	–, 1 (NH1–O: 2.98)	
		Gly226′	3, –	1, 1 (NH1–O: 3.09)	
P1′	Ser18	Cys42′	1, –	1, –	
		His57′	2, –		
		Gln192′	1, –	2, –	
		Gly193′	1, –	1, –	
		Ser195′	2, 1 (N–OG: 3.11)	2, 1 (N–OG: 2.94)	
P2′	Ile19	Phe41′	–, 1 (N–O: 3.09)	1, 1 (N–O: 3.16)	
		Tyr151′	2, –	3, –	
		Gln192′	1, –	1, –	
		Gly193′	3, –	3, –	
P4′	Pro21	Gln192′	1, –	1, –	
P26′	Met43	Thr98′	2, –		
		Pro173′		2, –	
		Gly174′		3, –	
P35′	Gln52	Asn97′	–, 1 (O–ND2: 3.33)		
P37′	Gln54	Gly96′	1, 1 (NE2–O: 2.64)	2, –	
		Asn97′	3, –	1, –	
P10	Glu67	Gln221′	–, 1 (OE2–NE2: 2.82)	4, 1 (OE2–NE2: 2.86)	
		Lys222′	4, –		
P6	Lys71	Tyr217′	5, –	4, –	
P5	Ala72	Gly219′		–, 1 (O–N: 3.40)	
P4	Ile73	Gln175′	1, –		
		Trp215′	2, –	2, –	
P3	Cys74	Trp215′	2, –	2, –	
		Gly216′	–, 2 (N–O: 3.36)	–, 2 (N–O: 3.11)	
			(O–N: 3.22)	(O–N: 3.11)	
P2	Thr75	His57′	3, –	4, –	
		Leu99′	1, –	1, –	
		Gln192′	1, –	1, –	
		Ser214′	1, –	1, –	
		Trp215′	1, –	1, –	
P1	Arg76	Asp189′	2, 2 (NH1–OD2: 3.06)	3, 2 (NH1–OD2: 2.79)	Ion
			(NH2–OD1: 2.97)	(NH2–OD1: 3.15)	
		Ser190′	3, 2 (NH2–OG: 2.97)	2, 2 (NH2–OG: 2.82)	
			(NH2–O: 3.02)	(NH2–O: 3.11)	
		Cys191′	1, –	1, –	
		Gln192′	4, –	3, –	
		Gly193′	2, 1 (O–N: 2.57)	3, 1 (O–N: 2.67)	
		Asp194′	–, 1 (O–N: 3.20)	–, 1 (O–N: 3.19)	
		Ser195′	7, 3 (N–OG: 2.88)	7, 3 (N–OG: 3.00)	

Table 4 (*continued*)

Site	BBBI	PPT	No. of interactions		Note
			PPT–BBBI(I)	PPT–BBBI(J)	
			(O–N: 2.98)	(O–N: 2.86)	
			(O–OG: 3.15)	(O–OG: 2.85)	
		Val213′			
		Ser214′	–, 1 (N–O: 3.14)	1, 1 (N–O: 3.07)	
		Trp215′	2, –	2, –	
		Gly216′	3, –	1, –	
		Gly219′	–, 1 (NH1–O: 3.05)	–, 1 (NH1–O: 2.89)	
		Gly226′	1, –	1, –	
P1′	Ser77	Phe41′	1, –	1, –	
		Cys42′	1, –	1, –	
		Gln192′	1, –	1, –	
		Gly193′	1, –	2, –	
		Ser195′	2, 1 (N–OG: 3.23)	2, 1 (N–OG: 3.04)	
P2′	Asn78	Phe41′	2, 1 (N–O: 3.02)	2, 1 (N–O: 3.02)	
		Tyr151′	2, 1 (OD1–OH: 3.45)	2, 1 (OD1–OH: 3.25)	
		Gly193′	2, –	2, –	
P4′	Pro80	Gln192′	2, –	2, –	
P7′	Arg83	Asn97′	1, 1 (NH1–O: 2.86)	–, 1 (NH1–O: 3.41)	
P32′	Val118	Gln192′		2, –	
P34′	Pro120	Lys224′		1, –	
P35′	Arg121	Asn223′	1, –	2, 1 (O–ND2: 3.07)	
		Gly219′		1, –	
P36′	Cys122	Asn223′	1, –		
P38′	Pro124	Leu185′	1, –	1, –	
		Asn223′	1, –	3, –	

mutation is deleterious for protein folding, but not for protease inhibitory activity[35] and the proline residue at the P1′ position is also found in rice BBI. The main-chain nitrogen atom makes a hydrogen bond with the side-chain oxygen (OG) of Ser195′. A detailed analysis of the binding interface shows that, after the P1 residues, the P2′ residues (Ile19 and Asn78) make the second most significant interactions with trypsin, which is consistent with mutagenesis results showing that the P2′ residue is also important in determining the specificity of the inhibitor towards its cognate enzyme.[36] The side-chain oxygen atom (OD1) of Asn78 forms a hydrogen bond with the hydroxyl group of tyrosine 151′ in PPT. However, the side-chain of the P2′ residue of SBBI (Asn18), which points a different direction, does not form any specific interactions with the trypsin molecule.[24] Beyond the inhibitory loop, the P5 and P6 residues are involved in the interaction with trypsin but their contribution seems to be relatively weak, and the P3′ and P4′ proline residues maintain a characteristic polyproline II conformation solely for the structural role of inhibitor, and do not participate directly in binding to trypsin.

Additional binding sites

In addition to the inhibitory loops and residues in their vicinity, several other regions are involved in the interaction between BBBI and PPT (Figure 6). In the N domain, the hydrophobic P26′ residue (Met43), which is a counter residue for chymotrypsin inhibition in BBIs from dicots, makes van der Waals contact with trypsin. Two glutamine residues, P35′ and P37′, in the N domain form hydrogen bonds with Asn97′ and Gly96′ in PPT, respectively (Figure 6(d)). However, the residues in the N domain that interact with PPT are not consistent with the I and J chain and, therefore, it is not clear whether they are critical for specific interactions. In the C domain, more specific and extensive interaction between BBBI and trypsin is observed (Figure 6(e)). The side-chain oxygen atom of the P10 residue (Glu67) forms a hydrogen bond with the side-chain nitrogen atom of Gln221′ in PPT. The side-chain nitrogen atom of the P7′ residue (Arg83) forms a hydrogen bond with the main-chain oxygen atom of Asn97′ in PPT. Interestingly, the segment of C-terminal residues (residues Pro124 and Arg125) that is invisible in the free structure could be built with confidence, most probably due to stabilization of its conformation by interaction with the trypsin molecule. P32′ (Val118), P34′ (Pro120), P35′ (Arg121), P36′ (Cys122′), and P38′ (Pro124) make contact with PPT, although the details are somewhat different for each chain (Table 4). The intriguing feature of this tail is its simultaneous interaction with two trypsin molecules (Figure 6(e)). The C-terminal tails of several BBIs seem to play an important role in their function. The C-terminal tail of the pea seed trypsin inhibitor isoform IVb (PsTI-IVb) also makes contact with the trypsin molecule in the model of the PsTI-Ivb–trypsin–chymotrypsin ternary complex.[18] In the crystal structures of the azuki bean trypsin
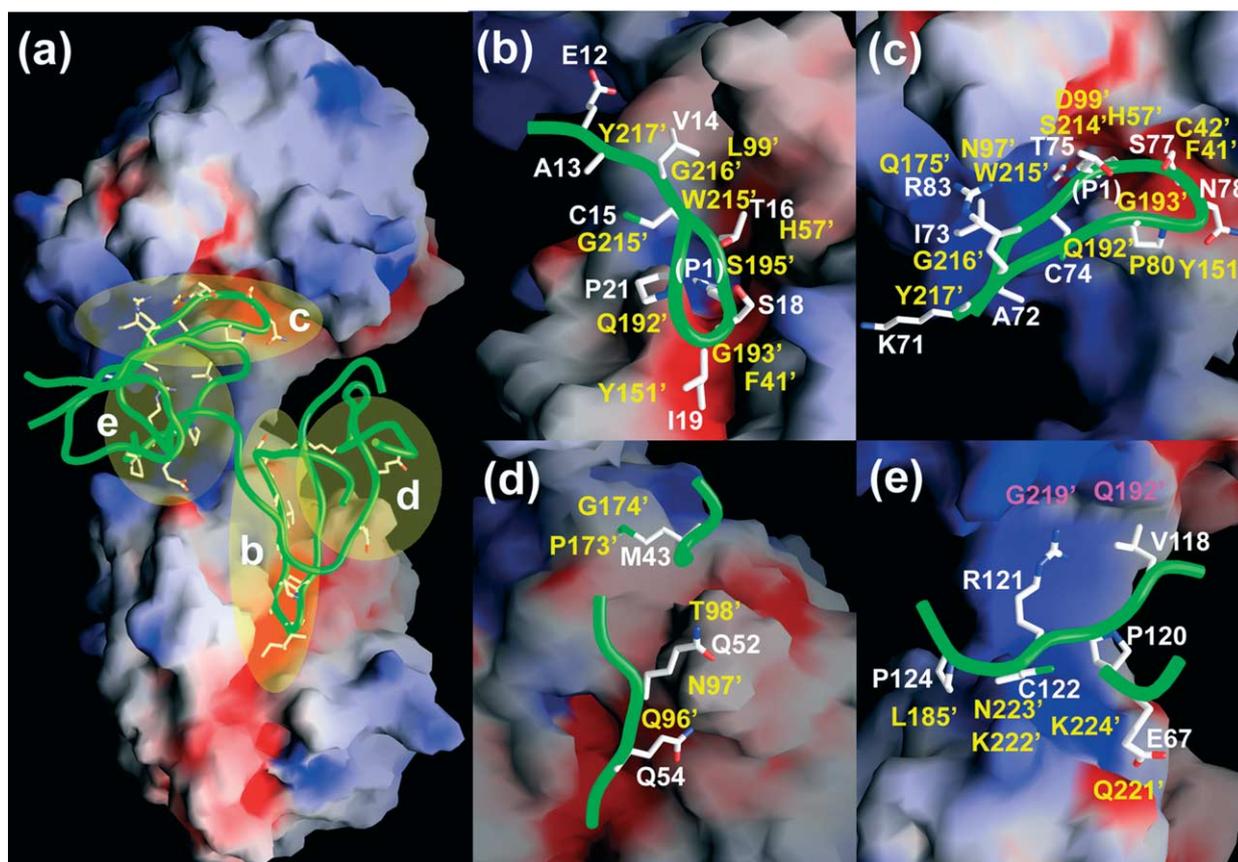
Figure 6. (a) The backbone model of BBBI (green tubes) and the electrostatic potential surface of trypsin molecules are drawn and four interacting regions are indicated ((b), (c), (d), and (e)). The view of this Figure is the same as that of Figure 3(a). (b) and (c) The contact region between (b) the inhibitory loop in the N domain of BBBI and PPT, and (c) the inhibitory loop in the C domain of BBBI and PPT. (d) Additional binding regions in the N domain of BBBI. (e) The region around C-terminal tail of BBBI. In (b), (c), (d) and (e), side-chain atoms of interacting residues in BBBI are drawn and labeled in white. The residues in trypsin are labeled in yellow. In (e) the residues in the adjacent trypsin molecule are shown in magenta. Each view is obtained for minimal overlap of the side-chain atoms. Positively charged regions are blue and negatively charged regions are red. This Figure was generated using GRASP.[45]

inhibitor (AB-I)–trypsin and SBBI–trypsin complexes, the C-terminal tails were not visible in the electron density maps and thus do not seem to be involved in any protein–protein interaction.[19,24] However, the AB-I–trypsin complex has been determined as a 1 : 1 complex structure, indicating the possibility that the invisible C-terminal tail of AB-I is stabilized by additional interactions in the 1 : 2 ternary complex. Interestingly, there is a very recent report that the C-terminal tail in the BBI from horsegram (*Dolichos biflorus*) seeds plays a pivotal role in dimerization mediated by a zinc ion.[37] Indeed, the C-terminal tail is the most variable region among BBIs, suggesting that it might have evolved for unique functions in each inhibitor.

Materials and Methods

Phasing

The procedures for purifying, crystallizing, and collecting X-ray data from the 1 : 2 complex between BBBI and

PPT have been reported.[38] The structures were determined by the molecular replacement method using the program MOLREP.[22] The starting models for PPT and BBBI were the previously reported structures of inhibitor-complexed PPT (PDB ID: 1AVW)[27] and free BBBI (PDB ID: 1C2A),[14] respectively. For the calculations of the rotation function, only non-hydrogen protein atoms were included. Orienting and positioning of the PPT model alone according to the molecular replacement solutions gave an *R*-factor of 52.4% and a correlation coefficient of 0.448 for 30.0–3.5 Å data. The complex model was generated using the free BBBI model with careful consideration of electron density and the interaction mode of trypsin with the inhibitor.

Refinement

The complex of the BBBI and PPT molecules refined using the program CNS.[39] A rigid-body refinement was initially carried out with 30.0–5.0 Å data to further improve the positional and orientational parameters. The high-resolution limit of the diffraction data was increased stepwise from 5.0 Å to 3.0 Å. The *R*-factor at this stage was 39.0% for 30.0–3.0 Å data. Atomic positions were refined by conventional conjugate gradient

minimization, with higher resolution data to 2.2 Å being added in steps. The electron density for the BBBI part was improved during the successive refinement steps. After a round of simulated annealing refinement employing the torsion angle dynamics protocol, the electron density for the missing C-terminal residues (Pro124, Arg125) in the free BBBI model was clear. Individual isotropic *B*-factors, initially set to 20 Å^2, were refined in the last stages of the refinement with restraints. Solvent molecules were placed by searching the model-phased $(F_o - F_c)$ maps and a bulk solvent correction was applied. A summary of the refinement statistics is given in Table 1.

Model building and structure analysis

The $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps were used as guides for manual rebuilding of the model. Models were displayed with O,[40] CHAIN,[41] and PyMOL†. At each step of the model rebuilding and refinement, the stereochemistry of the model was assessed by the program PROCHECK.[42] Structural comparisons were made using the program LSQKAB in the CCP4 program package.[43] To analyze the protein–protein interface of the BBBI–PPT complex structure, we used the protein–protein interaction server‡[44] and the program GRASP.[45]

Protein Data Bank accession number

The atomic coordinates have been deposited with the RCSB Protein Data Bank (entry code: 1TX6).

Acknowledgements

References

1. Birk, Y. (1985). The Bowman–Birk inhibitor. Trypsin- and chymotrypsin-inhibitor from soybeans. *Int. J. Pept. Protein Res.* **25**, 113–131.
2. Bowman, D. E. (1946). Differentiation of soybean antitrypsin factors. *Proc. Soc. Expt. Biol. Med.* **63**, 547–550.
3. Birk, Y., Gertler, A. & Khalef, S. (1946). A pure trypsin inhibitor from soya beans. *Biochem. J.* 1946;, 87.
4. Birk, Y. (1993). Protease inhibitors of plant origin and role of protease inhibitors in human nutrition. In *Protease Inhibitors as Cancer Chemopreventive Agents* (Troll, W. & Kennedy, A. R., eds), pp. 97–106, Plenum Press, New York.
5. Voss, R. H., Ermler, U., Essen, L. O., Wenzl, G., Kim, Y. M. & Flecker, P. (1996). Crystal structure of

† http://pymol.sourceforge.net
‡ http://www.biochem.ucl.ac.uk/bsm/PP/server

the bifunctional soybean Bowman–Birk inhibitor at 0.28-nm resolution. Structural peculiarities in a folded protein conformation. *Eur. J. Biochem.* **242**, 122–131.
6. Philipp, S., Kim, Y. M., Durr, I., Wenzl, G., Vogt, M. & Flecker, P. (1998). Mutational analysis of disulfide bonds in the trypsin-reactive subdomain of a Bowman–Birk-type inhibitor of trypsin and chymotrypsin–cooperative *versus* autonomous refolding of subdomains. *Eur. J. Biochem.* **251**, 854–862.
7. Kennedy, A. R. (1993). Anticarcinogenic activity of protease inhibitors. In *Protease Inhibitors as Cancer Chemopreventive Agents* (Troll, W. & Kennedy, A. R., eds), pp. 9–64, Plenum Press, New York.
8. Prakash, B., Selvaraj, S., Murthy, M. R., Sreerama, Y. N., Rao, D. R. & Gowda, L. R. (1996). Analysis of the amino acid sequences of plant Bowman–Birk inhibitors. *J. Mol. Evol.* **42**, 560–569.
9. Mikola, J. & Kirsi, M. (1972). Differences between endospermal and embryonal trypsin inhibitors in barley, wheat, and rye. *Acta Chem. Scand.* **26**, 787–795.
10. Ogiso, T., Noda, T., Sako, Y., Kato, Y. & Aoyama, M. (1975). Studies on trypsin inhibitor in barley. I. Purification and some properties. *J. Biochem. (Tokyo)*, **78**, 9–17.
11. Boisen, S. & Djurtoft, R. (1982). Protease inhibitor from barley embryo inhibiting trypsin and trypsin-like microbial proteases. Purification and characterization of two isoforms. *J. Sci. Food Agric.* **33**, 431–440.
12. Odani, S., Koide, T. & Ono, T. (1983). The complete amino acid sequence of barley trypsin inhibitor. *J. Biol. Chem.* **258**, 7998–8003.
13. Nagasu, A., Fukamachi, H., Ikenaga, H. & Funatsu, G. (1988). The amino acid sequence of barley rootlet trypsin inhibitor. *Agric. Biol. Chem.* **52**, 1505–1514.
14. Song, H. K., Kim, Y. S., Yang, J. K., Moon, J., Lee, J. Y. & Suh, S. W. (1999). Crystal structure of a 16 kDa double-headed Bowman–Birk trypsin inhibitor from barley seeds at 1.9 Å resolution. *J. Mol. Biol.* **293**, 1133–1144.
15. Chen, P., Rose, J., Love, R., Wei, C. H. & Wang, B. C. (1992). Reactive sites of an anticarcinogenic Bowman–Birk proteinase inhibitor are similar to other trypsin inhibitors. *J. Biol. Chem.* **267**, 1990–1994.
16. Suzuki, A., Yamane, T., Ashida, T., Norioka, S., Hara, S. & Ikenaza, T. (1993). Crystallographic refinement of Bowman–Birk type protease inhibitor A-II from peanut (*Arachis hypogaea*) at 2.3 Å resolution. *J. Mol. Biol.* **234**, 722–734.
17. Werner, M. H. & Wemmer, D. E. (1992). Three-dimensional structure of soybean trypsin/chymotrypsin Bowman–Birk inhibitor in solution. *Biochemistry*, **31**, 999–1010.
18. Li de la Sierra, I., Quillien, L., Flecker, P., Gueguen, J. & Brunie, S. (1999). Dimeric crystal structure of a Bowman–Birk protease inhibitor from pea seeds. *J. Mol. Biol.* **285**, 1195–1207.
19. Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C. *et al.* (1986). Structure of the trypsin-binding domain of Bowman–Birk type protease inhibitor and its interaction with trypsin. *J. Biochem. (Tokyo)*, **100**, 1637–1646.
20. Lin, G., Bode, W., Huber, R., Chi, C. & Engh, R. A. (1993). The 0.25-nm X-ray structure of the Bowman–Birk-type inhibitor from mung bean in ternary complex with porcine trypsin. *Eur. J. Biochem.* **212**, 549–555.
21. Raj, S. S., Kibushi, E., Kurasawa, T., Suzuki, A., Yamane, T., Odani, S. *et al.* (2002). Crystal structure of

bovine trypsin and wheat germ trypsin inhibitor (I-2b) complex (2 : 1) at 2.3 Å resolution. *J. Biochem. (Tokyo)*, **132**, 927–933.

22. Vagin, A. & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallog. sect. D*, **56**, 1622–1624.

23. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein–protein recognition sites. *J. Mol. Biol.* **285**, 2177–2198.

24. Koepke, J., Ermler, U., Warkentin, E., Wenzl, G. & Flecker, P. (2000). Crystal structure of cancer chemo-preventive Bowman–Birk inhibitor in ternary complex with bovine trypsin at 2.3 Å resolution. Structural basis of Janus-faced serine protease inhibitor specificity. *J. Mol. Biol.* **298**, 477–491.

25. 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.

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

27. Song, H. K. & Suh, S. W. (1998). 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**, 347–363.

28. Flecker, P. (1989). A new and general procedure for refolding mutant Bowman–Birk-type proteinase inhibitors on trypsin-Sepharose as a matrix with complementary structure. *FEBS Letters*, **252**, 153–157.

29. 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.

30. Bode, W. & Huber, R. (1992). Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451.

31. 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**, 4212–4228.

32. Read, R. J., Fujinaga, M., Sielecki, A. R. & James, M. N. (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.

33. McBride, J. D., Brauer, A. B., Nievo, M. & Leatherbarrow, R. J. (1998). The role of threonine in the P2 position of Bowman–Birk proteinase inhibitors: studies on P2 variation in cyclic peptides encompassing the reactive site loop. *J. Mol. Biol.* **282**, 447–458.

34. Brauer, A. B. & Leatherbarrow, R. J. (2003). The conserved P1$'$ Ser of Bowman–Birk-type proteinase inhibitors is not essential for the integrity of the reactive site loop. *Biochem. Biophys. Res. Commun.* **308**, 300–305.

35. Flecker, P. (1995). Template-directed protein folding into a metastable state of increased activity. *Eur. J. Biochem.* **232**, 528–535.

36. Gariani, T., McBride, J. D. & Leatherbarrow, R. J. (1999). The role of the P2$'$ position of Bowman–Birk proteinase inhibitor in the inhibition of trypsin. Studies on P2$'$ variation in cyclic peptides encompassing the reactive site loop. *Biochim. Biophys. Acta*, **1431**, 232–237.

37. Kumar, P., Rao, A. G. A., Hariharaputran, S., Chandra, N. & Gowda, L. R. (2004). Molecular mechanism of dimerization of Bowman–Birk inhibitors: pivotal role of Asp76 in the dimerization. *J. Biol. Chem.* **279**, 30425–30432.

38. Kim, Y. S., Song, H. K. & Suh, S. W. (1999). Crystallization and preliminary X-ray analysis of a complex between the Bowman–Birk trypsin inhibitor from barley and porcine pancreatic trypsin. *Acta Crystallog. sect. D*, **55**, 1244–1246.

39. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.

40. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.

41. Sack, J. S. (1988). CHAIN: a crystallographic modeling program. *J. Mol. Graph.* **6**, 244–245.

42. Laskowski, R., MacArthur, M., Hutchinson, E. & Thorton, J. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

43. Collaborative Computational Project No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.

44. Jones, S. & Thornton, J. M. (1996). Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA*, **93**, 13–20.

45. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP—graphical representation and analysis of surface properties. *Biophys. J.* **64**, A166.

46. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

47. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. In *Methods in Enzymology* (Carter, J. C. W. & Sweet, R. M., eds), vol. 277, pp. 505–524, Academic Press, New York.

Edited by R. Huber