Article No. jmbi.1999.2891 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 290, 525–533

JMB

High-Resolution Structure of a Potent, Cyclic Proteinase Inhibitor from Sunflower Seeds

S. Luckett[1,4], R. Santiago Garcia[2], J. J. Barker[1], Al. V. Konarev[3], P. R. Shewry[4], A. R. Clarke[1] and R. L. Brady[1]*

[1]*Department of Biochemistry University of Bristol, Bristol BS8 1TD, UK*

[2]*Department of Analytical Chemistry, University of Oviedo, Julian Claveria 8 Oviedo, Asturias 33006, Spain*

[3]*All Russian Institute for Plant Protection, St. Petersburg Russia*

[4]*IACR-Long Ashton Research Station, Department of Agricultural Sciences University of Bristol, Long Ashton, Bristol BS41 9AF, UK*

**Corresponding author*

Proteinaceous serine proteinase inhibitors are widespread throughout the plant kingdom where they play an important role in protection against pests and pathogens. Here, we describe the isolation and characterisation of a novel 14 amino acid residue cyclic peptide from sunflower seeds, which is a potent inhibitor of trypsin ($K_i = 100$ pM). The crystal structure of this peptide in complex with bovine β-trypsin shows both sequence and conformational similarity with the trypsin-reactive loop of the Bowman-Birk family of serine proteinase inhibitors. This inhibitor, however, is unique in being monofunctional, cyclic and far shorter (14 amino acid residues) than inhibitors belonging to this family (typically 60-70 amino acid residues). The high potency of this peptide is likely to arise from the considerable structural rigidity achieved through its cyclic nature which is further stabilised by a single internal disulphide bond. This study helps delineate the minimal unit required for effective peptide inhibitors of serine proteinases, and will assist in the further design of inhibitors to this widespread class of enzymes.

Keywords: Bowman-Birk inhibitor; trypsin inhibitor; cyclic peptide; sunflower seeds; X-ray crystallography

Introduction

Protein inhibitors of proteinases are widely distributed in plants, being particularly abundant in storage tissues such as seeds and tubers (Richardson, 1991; Shewry, 1998). In addition, their synthesis may be induced by infection or wounding, particularly in vegetative tissues (Ryan, 1990). These characteristics, together with the fact that many are active against exogenous rather than endogenous enzymes, suggest that they play a role in plant defence, conferring a broad spectrum of resistance to pests and pathogens. In some cases this role has been confirmed by *in vitro* tests or expression in transgenic plants, as discussed by Shewry & Lucas (1997).

Plant proteinase inhibitors can be classified into at least 12 families (Shewry, 1999), based on their amino acid sequences and the mechanistic class of their target proteinases. In general, they inhibit only proteinases of one mechanistic class and, not surprisingly, the majority of inhibitors are specific for the most widespread class of enzymes, the serine proteinases. At least ten families of plant inhibitors for this class of proteinases have been described and, although structurally unrelated, all appear to react with the target enzymes according to a standard substrate-like mechanism (Laskowski & Kato, 1980) and possess one or more inhibitory sites on exposed loops.

The Bowman-Birk inhibitor (BBI) family (Laskowski & Kato, 1980) are small serine proteinase inhibitors found in seeds of legumes and in many other plants. Characteristically, their molecular masses range between 6000 and 9000 Da, and they contain seven disulphide bonds that help stabilise their active configurations. All members of the Bowman-Birk inhibitor family have two tandem homology regions on the same polypeptide chain, each homology region comprising a consensus motif of three β-strands with a kinetically independent reactive site on the outermost

Abbreviations used: BBI, Bowman-Birk inhibitor; SFTI-1, sunflower trypsin inhibitor; CNBr, cyanogen bromide; TFA, trifluoroacetic acid; L-BAPNA, Nα-benzoyl-L-arginine *p*-nitroanilide; DMF, *N,N*-dimethylformamide; ESMS, electrospray mass spectroscopy; BPTI, bovine pancreatic trypsin inhibitor.

E-mail address of the corresponding author: L.Brady@bris.ac.uk

0022-2836/99/270525–9 $30.00/0

loop. It has been demonstrated that splitting the homology regions by partial peptic digestion yields two active fragments (Odani & Ikenaka, 1973). The ability of the Bowman-Birk type inhibitors to inhibit two proteinases simultaneously and independently has led to them being called ''double-headed'' inhibitors. Most Bowman-Birk-type inhibitors inhibit trypsin at the first reactive site (N-terminal) and chymotrypsin at the second reactive site (C-terminal). Using the convention described by Schechter & Berger (1967), it is the P1 residue of the reactive site that confers inhibitory specificity: alanine for elastase; arginine and lysine for trypsin; and leucine, phenylalanine and tyrosine for chymotrypsin. The sequences of both reactive sites are well conserved, being located within a nine-residue loop with a disulphide bond between residues 1 and 9. Synthetic nonapeptides based on the antitryptic reactive site from soybean BBI have been shown to demonstrate antitryptic activity, but inhibition was significantly weaker than with the native BBI (Nishino *et al*., 1977; Terada *et al*., 1980).

Here, we describe the isolation of a peptide, sunflower trysin inhibitor (SFTI-1), with a molecular mass of 1513 Da from sunflower seeds which specifically inhibits trypsin. This peptide has been characterised by determining its three-dimensional structure in complex with bovine β-trypsin. The sequence and conformation identified from this structure shows similarity to the reactive site loop of the Bowman-Birk inhibitors. This peptide, however, is significantly shorter than members of the BBI family, exhibits a novel cyclic structure, and has considerably (at least two orders of magnitude) enhanced potency relative to other peptides of similar length.

Results

Purification and characterisation of SFTI-1

Separation of an aqueous extract of defatted sunflower seeds by trypsin affinity chromatography (Figure 1(a)) followed by reversed-phase HPLC (Figure 1(b)) gave a single homogeneous component with an M_r, determined by electrospray mass spectroscopy (ESMS), of 1513. Attempts to determine the sequence by Edman degradation gave low yields (less than 10%) of a sequence starting at position Ser6-I. This corresponds to the active site bond, indicating that partial cleavage may have occurred during trypsin affinity chromatography. The sequence of the peptide was subsequently determined directly from the electron density map (see below). This showed a cyclic structure of 14 amino acid residues (shown below), hence explaining the difficulty encountered in N-terminal sequencing:

G-R-C-T-K-S-I-P-P-I-C-F-P-D

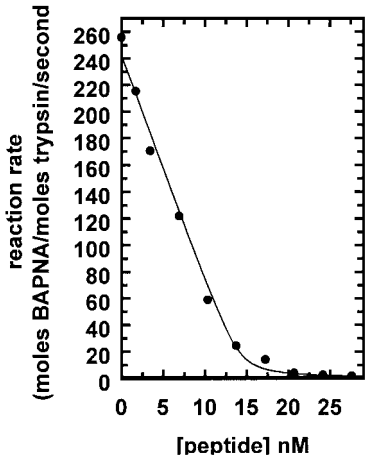

Figure 1. Isolation and purification of SFTI-1. (a) The elution profile of SFTI-1 from a trypsin affinity column, which was then further purified (b) by reversed-phase HPLC. The SFTI-1 peak is indicated by an arrow.

The calculated M_r, 1513, agrees with that from ESMS, and the sequence was further confirmed by amino acid composition analysis (Table 1). Reduction of the peptide and pyridylethylation followed by digestion with Lys-C produced a single fragment confirming the cyclic nature of the peptide.

Sequence comparisons demonstrated that SFTI-1 is related to a fragment present in members of the Bowman-Birk family of proteinase inhibitors, showing a high degree of sequence identity to the disulphide-stabilised loop region which contains the first active site (which is usually specific for trypsin). However, SFTI-1 is the smallest naturally occurring plant protein inhibitor reported to date, corresponding to the reactive loop and parts of the two adjacent strands. It contains a single disulphide bond and has a novel cyclic structure.

Inhibitory properties

SFTI-1 inhibited bovine β-trypsin with a stoichiometric ratio of 1:1 and a K_i of 0.1 nM (100 pM) being derived from the competition assays (Figure 2). SFTI-1 was also observed to inhibit cathepsin G ($K_i \sim < 0.15$ nM), elastase ($K_i \sim 105$ ($\pm$12) μM), chymotrypsin ($K_i \sim 7.4(\pm 1.5)$ μM) and thrombin ($K_i \sim 136(\pm 21)$ μM) (data not shown). No inhibition ($K_i > 1$ mM) was observed with Factor Xa.

Structure of SFTI-1

The 14 amino acid residues of SFTI-1 form two antiparallel β-strands connected at the reactive site end by an extended loop region, and by a hairpin turn at the opposite end (Figure 3). These strands are constrained by the single disulphide bond (between Cys3-I and Cys11-I), dividing the peptide into a nine-residue loop region (the ''reactive

Table 1. Comparison of the directly determined amino acid composition of SFTI-1 with that determined by X-ray crystallography

Amino acid	Directly determined Mol (%)	Directly determined Residues/mol	From X-ray structure Residues/mol
D	8.4	1	1
Z	1.5	0	0
S	7.5	1	1
G	13.0	1	1
H	0.3	0	0
R	7.6	1	1
T	7.8	1	1
A	0.7	0	0
P	20.6	3	3
Y	0.3	0	0
V	0.4	0	0
M	0.2	0	0
C[a]	-	2	2
I	13.8	2	2
L	0.9	0	0
F	7.5	1	1
W	0.0	0	0
K	9.6	1	1

[a] Cysteine residues were determined by mass spectrometry after reduction and alkylation using vinylpyridine, to form *S*-pyridylethylcysteine (PEcys).

loop''), and a five-residue turn (the ''cyclic loop''). There is a sharp turn in the peptide chain at Pro8-I, with the Ile7-I-Pro8-I peptide bond adopting a *cis* conformation. The electron denisty is contiguous for the main-chain throughout the peptide (Figure 3), and all of the residues adopt reasonable conformations as determined by Procheck (Laskowski *et al*., 1993) including Ramachandran plot analysis. The mean temperature factor for inhibitor main-chain atoms is 25.4 Å^2, compared with 16.7 Å^2 for the main-chain atoms of trypsin.

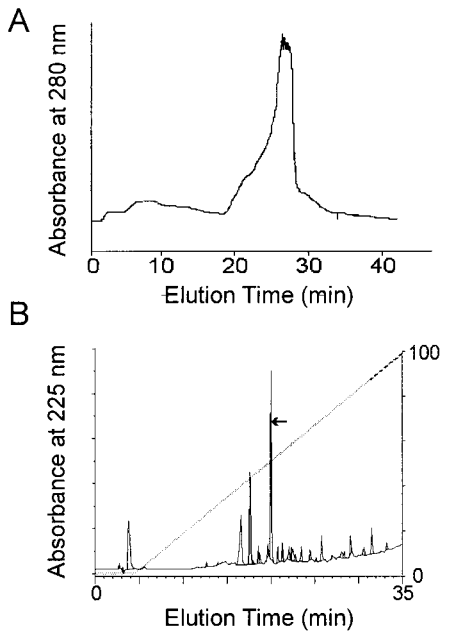

Figure 2. Inhibition curve showing the inhibition of bovine trypsin by SFTI-1. Trypsin activity was assayed using BAPNA as substrate, as described in Materials and Methods, using a trypsin concentration of 14 nM.

This is consistent with the surface location of the peptide within the crystal complex.

The overall conformation of SFTI-1 is extremely similar to the reported structures for the equivalent trypsin-inhibitory loop fragments from within the structures of reported Bowman-Birk inhibitors. Figure 4 shows superimpositions of SFTI-1 with the mung bean inhibitor (Li *et al*., 1994), adzuki bean inhibitor (Tsunogae *et al*., 1986), and soybean inhibitor (Werner & Wemmer, 1992). The root-mean-square deviations (rmsd) of 13 equivalent main-chain atoms are 0.14, 0.39 and 0.66 Å^2, respectively. There is also similarity with the bitter gourd inhibitor (Huang *et al*., 1993), although this is restricted to the first seven residues of SFTI-1 as this inhibitor diverges considerably from the consensus fold after this point. Note that in all cases, SFTI-1 represents only a subset of the fold observed in these other inhibitors. The exception is the mung bean inhibitor (1SMF), but this is a synthetic fragment of a larger inhibitor, and the construct co-crystallised is actually 22 amino acid residues, with only the core fragment observed in the electron density maps.

A network of internal hydrogen bonds, in concert with the disulphide bond, appear to stabilise the fold of the inhibitor (Figure 3). There are three intramolecular main-chain hydrogen bonds stabilising the inhibitor backbone: these connect residues Gly1-I to Phe12-I and Arg2-I to Phe12-I within the cyclic loop, and Thr4-I to Ile10-I within the reactive loop. Additionally, within the reactive loop region the side-chain hydroxyl group of Thr4-I forms a bifurcated hydrogen bond with the main-chain nitrogen atom of Ile10-I and side-chain hydroxyl group of Ser6-I. The latter is also bonded to the main-chain carbonyl group of Pro8-I. This pattern of intramolecular bonding within the reac-

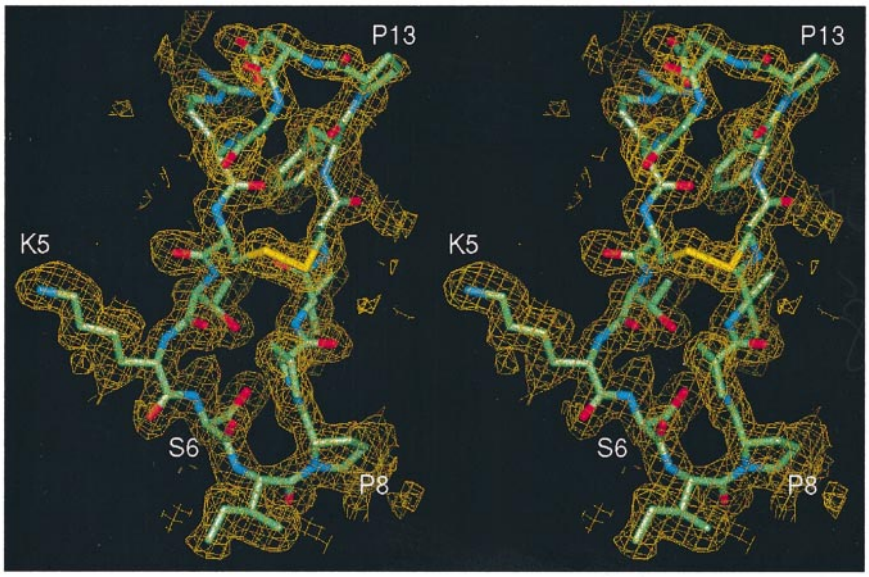

Figure 3. Stereoview showing the bound conformation of SFTI-1 in the trypsin-SFTI1 complex and the omit map electron density corresponding to SFTI-1. The map was calculated from $F_o - F_c$ coefficients, calculated using the final refined coordinates of the complex from which the inhibitor was removed, and the remaining trypsin coordinates were subjected to several rounds of further refinement. The map is contoured at 1.5 sigma. The small additional density at Lys5-Ser6 was not observed in later maps calculated from the refined structure.

tive loop appears to replicate the stabilisation previously described for other similar inhibitors (Papamokos *et al.*, 1982). In SFTI-1, the cyclic loop conformation is further stabilised by the side-chain of Asp14-I, which acts as proton acceptor for the amide nitrogen atoms of both Gly1-I and Arg2-I. Furthermore, extensive interactions between the inhibitor and trypsin also stabilise the observed inhibitor conformation.

Trypsin-inhibitor association

As expected, SFTI-1 binds in the active site of trypsin in a very similar manner to the equivalent fragment from the BBI inhibitors. Central to this interaction is the extended conformation of the lysine side-chain at position 5 (P1). This residue, which is invariably Lys or Arg in either trypsin peptide inhibitors or substrates, projects into the S1 pocket of the enzyme. Both direct and water-mediated contacts are made with the specificity-determining Asp189 at the base of this pocket, to the hydroxyl group of Ser190 and to the main-chain carbonyl group of Gly219. Again in common with other BBIs, the carbonyl and amine groups of the ''scissile'' bond between Lys5-I and Ser6-I are within hydrogen-bonding distance of the Ser195 and His57 of the trypsin catalytic triad. This orientation is the result of the intramolecular interactions of the Ser6-I side-chain in the inhibitor (see above). There is an extensive network of hydrogen bonds and ion pairs formed between the inhibitor and enzyme, predominantly with residues within

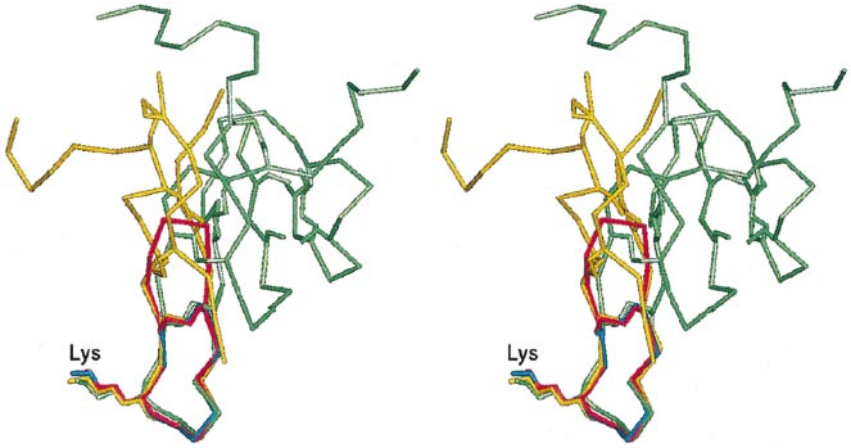

Figure 4. C^{α} traces showing superimpositions of SFTI-1 (red) with the Mung bean inhibitor (Li *et al.*, 1994) shown in blue, Adzuki bean inhibitor (Tsunogae *et al.*, 1986) shown in yellow, and soybean inhibitor (Werner & Wemmer, 1992) shown in green. The P1 lysine residue is indicated.

the reactive loop region of the inhibitor, and these are summarised in Table 2.

As can be seen in Table 2, many of the interactions between inhibitor and enzyme previously observed for the Mung bean inhibitor (1SMF; Li *et al*., 1994) and Adzuki bean inhibitor (1TAB; Tsunogae *et al*., 1986) are highly conserved in the SFTI-1/trypsin complex. These contacts are made predominantly with residues within the reactive loop (residues 3-11), where there is only one amino acid difference between SFTI-1 and the most closely related structure, that of the synthetic peptide based on the Mung bean inhibitor (position 10 is Ile in SFTI-1, and Glu in 1SMF). In the structure of the SFTI-1/trypsin complex, the isoleucine sidechain in this position is close to the surface and makes no contacts with the enzyme. Hence this change does not appear to affect the interaction of the inhibitor and the enzyme. We note that the majority of the hydrogen bonds between the inhibitor and trypsin are formed within the primary binding segment of the inhibitor, Cys3-I-Thr4-I-Lys5-I-Ser6-I-Ile7-I, which occupies the P3-P2′ positions in the enzyme active site. This binding loop is extremely well defined in the electron density map, reflecting its firm immobilisation through these interactions with the highly complementary enzyme surface. The spatial arrangement of the distant side of the reactive loop of the inhibitor (Pro9-I-Cys11-I) is very similar to the relationship between the ''secondary contact region'' (Papamokos *et al*., 1982) of the other inhibitors and their reactive sites. The antiparallel β-sheet region Ile18-Val34 in BPTI, is replaced by a *cis*-proline at the P4′ position in the Bowman-Birk inhibitors. This proline residue (equivalent to Pro8-I in SFTI-1) produces a sharp turn in the inhibitor chain. This distant side of the inhibitor, which through its cyclic nature is especially rigid in SFTI-1 compared to similar regions in the BBI inhibitors, is likely to play an important role in the inhibitory activity. By stabilising the overall inhibitor conformation, this region prevents the necessary structural change that would normally accompany proteolysis. Despite the relatively small size of the inhibitor, extensive solvent-accessible surface area (1422 Å^2, calculated with QUANTA (MSI)) is occluded at the inhibitor/enzyme interface.

Table 2. SFTI-1 trypsin contacts

Inhibitor residue/atoms	Trypsin residue/atoms	Distance (Å)
Arg2-I NH2	Asn97 O	3.1
Cys3-I N	Gly216 O	3.1 *(3.0, 3.4)*
Cys3-I O	Gly216 N	3.2 *(3.2, 3.5)*
Thr4-I O	Gln192 NE2	3.0 *(2.7, -)*
Lys5-I N	Ser214 O	3.3 *(3.3, 3.4)*
Lys5-I NZ	Ser190 OG	3.0 *(3.0,2.9)*
Lys5-I NZ	Ser190 O	3.1 *(2.6, 3.1)*
Lys5-I NZ	Asp189 OD1	3.2 (additional bonding through bridging water molecules) *(3.3, 3.0)*
Lys5-I O	Gly193 N	2.6 *(2.6, 2.8)*
Lys5-I O[a]	Asp194 N	3.4 *(3.4, 3.1)*
Lys5-I O	Ser195 N	3.1 *(2.9, 2.7)*
Lys5-I C[a]	Ser195 OG	3.0 *(2.8, 2.4)*
Lys5-I N	Ser195 OG	2.9 *(2.9, 3.0)*
Ile7-I N	Phe41 O	3.0 *(3.2, 3.1)*
Asp14-I OD2-	Gln175 NE2	2.5

Distances in parentheses and italics refer to the equivalent distances reported in the crystal structures of trypsin in complex with the Mung bean inhibitor mimic (Li *et al*., 1994) and the Adzuki bean inhibitor (Tsunogae *et al*., 1986) respectively.

[a] Indicates non-hydrogen bonded contacts.

Conformation of the enzyme active site

Although the electron density confirms the scissile bond is uncleaved within the ''average'' structure observed in the crystal, the lability of a small fraction of the sample to N-terminal cleavage at this site suggests up to 10% of the sample may be cleaved. This implies slow hydrolysis of the peptide by trypsin or related enzymes, although the equilibrium suggests binding of the substrate is preferred to its conversion to product. The crystal structure therefore represents an enzyme-substrate inhibitory complex. Although the omit map electron density could be interpreted to suggest a small modification at the nitrogen atom of this scissile bond (Figure 3), this unusual density was not observed in later maps from the refined structure, and the partial hydrolysis supports our conclusion that this in an unmodified peptide bond. In common with many other peptide inhibitors, the carbonyl group from the P1 residue (Lys5-I) is in close contact with the active site Ser195 OG (3.0 Å; Figure 5). This leaves the His57 NE2 and Ser195 OG of the trypsin catalytic triad in close contact (2.7 Å), implying that the conformation of the inhibited enzyme is close to the active conformation. This differs from the arrangements observed in crystal structures of bovine trypsin in the absence of an inhibitor (Bartunik *et al*., 1989) and in the presence of benzamidine or related inhibitors (Marquart *et al*., 1983) where the active site His and Ser residues are separated by 3 Å or greater.

Similarly, the geometry of the Ser195CB-OG-His57NE2 angle is again related to the proximity of the inhibitor to the reactive serine residue. In SFTI-1 this angle is very similar to the equivalent angle in the Mung bean inhibitor (both approxi-

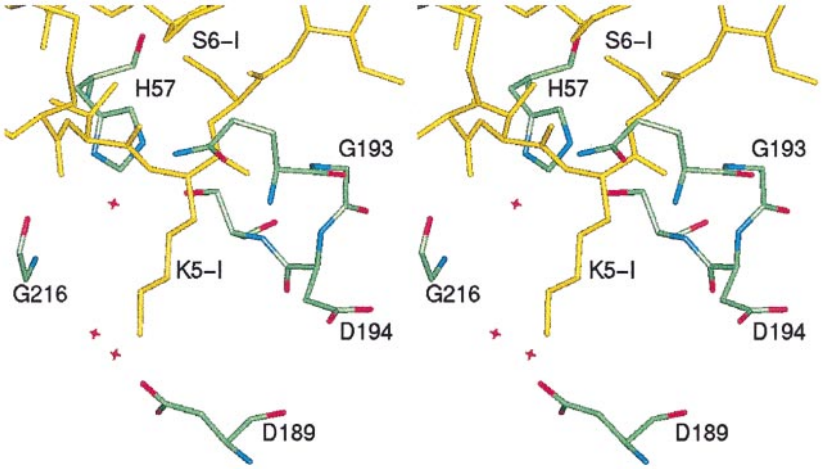

Figure 5. Stereoview showing the interactions of the active site loop of the inhibitor with bovine trypsin. Lys5-I projects into the S1 pocket of the enzyme, delineating specificity for trypsin-like serine proteinases. SFTI-1 residues are shown in yellow and ordered water molecules in red.

mately 96°). In the equivalent non-inhibited and benzamidine inhibited structures, this angle appears significantly different (88° in both cases), although some of these structures are of limited resolution. These observations further confirm SFTI-1 inhibits catalysis in trypsin by an identical mechanism to other peptidic inhibitors, in particular those of the BBI family.

Discussion

Amino acid sequences of about 100 Bowman-Birk type inhibitors are available on protein sequence databases, with three-dimensional structures determined for a small subset of these proteins. The conformation of the trypsin reactive loop region is highly conserved, and changes very little on binding to the enzyme (Bode & Huber, 1992). In the present study we have isolated and characterised a new trypsin inhibitor from sunflower seeds. This inhibitor, SFTI-1, shows clear parallels with the trypsin-reactive loop region of the BBI family of inhibitors in amino acid sequence, conformation and mechanism of inhibition, but differs from this family in size and its cyclic nature. Cyclization has previously been observed in the natural product serine proteinase inhibitor cyclotheonamide A (Ganesh *et al*., 1996) isolated from sponges. SFTI-1, however, is wholly peptidic and displays a greater degree of homology with the extensive BBI family of inhibitors.

Although SFTI-1 is the smallest BBI-related inhibitor to be purified from plants, smaller inhibitory synthetic peptides based on BBI structures have been reported. Nishino *et al*. (1977) and Tereda *et al*. (1980) reported that a range of heterodetic nonapeptides based on the inhibitory loops of BBI were active against trypsin or chymotrypsin. Table 3 shows the amino acid composition and reported K_i values for a series of peptides similar to SFTI-1. Note that the naturally occurring SFTI-1 exhibits the greatest potency as an inhibitor of trypsin, with a K_i value in the sub-nanomolar range. Despite the small size of this peptide, its inhibitory activity exceeds that of the 82 amino acid residue soybean BBI, to date the most potent inhibitor reported from the BBI family. There is only a single amino acid residue change between these inhibitors within the reactive loop region (the P5′ amino acid residue, which makes no contacts with the enzyme, is isoleucine in SFTI-1 and glutamine in soybean BBI), although a comparison of

Table 3. Comparison of the amino acid sequences and reported K_i values against trypsin for a series of similar peptides to SFTI-1

Inhibitor	Sequence	K_i (μM)	Ref.
SFTI-1	GRCTKSIPPICFPD	0.0001	This paper
Synthetic	Ac-CTKSNPPQC-NH_2	0.75	Terada *et al*. (1980)
Synthetic	H-CTKSNPPQC-OH	3.6	Terada *et al*. (1980)
Synthetic	H-SCTKSIPPQCT-OH	0.0095	Gariani *et al*. (1997)
Synthetic	SCTKSIPPQCY	0.01	Domingo *et al*. (1995)
SMF	EPCCDSCRCTKSIPPQCHCANI	12	Li *et al*. (1994)
BBI	---DQCACTKSNPPQCRCS---	0.00061	Voss *et al*. (1996)
CtA	cyclic V-Tyr-Dpr-Pro-K-Arg-Phe	0.0002	Ganesh *et al*. (1996)

SMF is an artificial trypsin inhibitor based on Mung bean trypsin inhibitor. BBI, soybean Bowman-Birk inhibitor. CtA (cyclotheonamide A) is a cyclic pentapeptide containing unusual amino acid residues: V-Tyr, vinologous tyrosine; Dpr, β-linked-diaminopropanoic acid; K-Arg, α-ketohomoarginine.

their structures shows significant shifts (up to 1 Å movements of equivalent main-chain atoms) within this region. These changes, however, are likely to result from the determination of the soybean BBI structure in the absence of trypsin, rather than as an enzyme/inhibitor complex as for SFTI-1. Assuming this region adopts an identical binding mode in both cases, it is intriguing that the additional amino acid residues adjoining this region in the 82 amino acid residue soybean BBI, which are potentially able to make a plethora of additional contacts with the enzyme at its surface, do not lead to superior inhibitory activity. There are limited contacts made between SFTI-1 residues outside of the reactive loop and the enzyme surface, implying that the predominant role of this region is to complete the cyclization of the peptide, hence contributing to its overall stability.

The cyclic nature and internal disulphide bond of SFTI-1 are likely to reduce its conformational flexibility significantly. It is possible the solution structure of the peptide is very similar to the well-defined conformation observed in the crystal structure. This expected rigidity correlates with the considerable inhibitory activity (and hence tight binding) of the peptide, implying (as enthalpic contributions are similar to those in other BBIs) that there is no significant entropic cost associated with binding. This is also consistent with previous conclusions (Werner & Wemmer, 1992) that the thermodynamic stability of soybean BBI derives not from the hydrophobic effect but from intramolecular disulphide and hydrogen bonds. This route to improvement of peptide activity by restriction of flexibility *via* cyclisation might have general applicability in inhibitor design. It is likely that the equivalent very high potency reported for cyclotheonamide A also derives from its structural rigidity achieved by cyclisation. This inhibitor, however, is smaller than SFTI-1 (five rather than 14 residues) and additionally incorporates a number of non-peptide bonds and features.

The structure of the SFTI-1/trypsin complex adds to a growing database of inhibitor/serine proteinase complexes. This widespread family of enzymes shares a conserved active site and mechanism, and with its detailed characterisation and frequent occurrence in biologically important enzymes is often targeted by drug designers. Peptide inhibitors can take advantage of a series of well-developed mechanisms for peptide transport in organisms to ensure their bioavailability. Structures of peptide inhibitor/enzyme complexes also provide excellent examples of inhibitor/receptor complementarity, highlighting interactions that might be mimicked in the development of non-peptidic inhibitors. The relatively small size of SFTI-1 helps define the minimal requirement for proteinaceous proteinase inhibitors, and approaches the crossover between these two methods of proteinase inhibitor design.

Materials and Methods

Isolation and purification of SFTI-1

Freeze-dried sunflower seeds (*Helianthus annus* L. Sunbred 246) were ground in a Waring blender and defatted with petroleum spirit (60-80°) at 4 °C using 1.5 l of solvent per 100 g of meal. After vacuum filtration, washing with cold petroleum spirit and air drying, 50 g of flour was extracted with 250 ml of water for 30 minutes at room temperature. The supernatant from the centrifugation (15,000 *g* for 15 minutes) was dissolved in 0.2 M ammonium acetate (pH 6.8), and incubated with a trypsin-Sepharose affinity column (prepared by incubating 30 mg of trypsin with 3.5 ml CNBr-Sepharose 4B gel according to the manufacturer's instructions). Non-specifically bound proteins were removed by washing with 200 ml of 0.2 M ammonium acetate, and 200 ml of 50 mM Tris-HCl buffer (pH 8.0) with 0.5 M NaCl. The bound fraction was eluted using 15 mM HCl, and further fractionated by reversed-phase HPLC at 45 °C on a Vydac 218TP54 column using a 0-100 % acetonitrile gradient. Solvent A was 0.07 % (v/v) trifluoroacetic acid (TFA) and solvent B was 0.05 % (v/v) TFA in acetonitrile.

Determination of antitrypsin activity

Inhibitory activity was assayed against bovine β-trypsin (Type III, Sigma) using the colourimetric trypsin substrate outlined by Erlanger *et al* (1961). A 25 µl sample of a solution of trypsin (0.01 mg/ml) and increasing amounts of SFTI-1 (0-200 µl of 0.1 µM inhibitor stock) was incubated for five minutes at 20 °C in a total volume of 720 µl of 0.05 M Tris-HCl (pH 8.2) containing 0.025 M $CaCl_2$. A 5 µl sample of L-BAPNA (7.4 mM stock) was added, mixed well and the residual trypsin activity determined by following the change of absorbance at 410 nm over 60 minutes. The dissociation constant, K_i, was calculated using the tight ligand binding equation described by Ranson *et al*. (1997).

Crystallisation

Low molecular density trypsin crystals were obtained by the vapour diffusion technique as described (Bartunik *et al*., 1989). Bovine trypsin (Type III, Sigma) was dissolved to approximately 60 mg/ml in 0.3 M ammonium sulphate, 6 mM calcium chloride, 0.1 M Tris (pH 8.15) and 60 mM benzamidine. The final protein concentration was adjusted to 30 mg/ml. Prior to sealing the wells, 0.5 µl of DMF was added to 10 µl of the protein solution. The hanging drops were suspended over 1.6 M to 2.1 M ammonium sulphate with 0.05 M Tris-HCl buffered to a pH of 8.15. Crystals of suitable size grew within two weeks. Using sitting-drop plastic crystallization vessels, crystals containing benzamidine where soaked in 20 µl of harvesting solution containing 1.5 mg/ml SFTI-1, 2.2 M ammonium sulphate, 1 mM calcium chloride and 0.1 M Tris-HCl (pH 8.0). The reservoir solution was the same as the harvesting solution with the exception of SFTI-1. Crystals containing SFTI-1 were obtained after incubation at 18 °C for 18 hours. Crystals are orthorhombic and belong to the space group $P2_12_12_1$ with unit cell dimensions $a = 60.61$ Å, $b = 64.43$ Å, $c = 70.62$ Å, contain one complex molecule per asymmetric unit and diffract to at least 1.65 Å.

Data collection

Two sets of diffraction data were collected, both at 100 K and from the same, single crystal. This crystal was frozen to 100 K in the presence of 25 % (v/v) glycerol to ensure vitriolic freezing. Data to 1.9 Å resolution were collected using the home source (Nonius DIP2020 image plate with Cu Kα radiation ($\lambda = 1.5418$ Å) and a rotating anode generator). A second, higher resolution data set (1.65 Å resolution) was collected at the Daresbury SRS synchroton (beamline PX7.2, $\lambda = 1.488$ Å). All data were processed using DENZO and SCALEPACK (Otwinowski & Minor, 1997) and manipulated with the CCP4 suite of programs (Collaborative Computing Project Number 4, 1994). Statistics for the synchrotron data used to refine the final model are summarised in Table 4.

Structure solution and refinement

The structure was solved by molecular replacement using the program AMoRe (Navaza, 1994) with coordinates from the refined 1.5 Å resolution structure of bovine trypsin PDB entry 1TLD (Bartunik *et al*., 1989) as the search model. This structure is from isomorphous crystals, although the unit cell dimensions vary by about 4 %. A single solution was readily obtained with a correlation coefficient of 67.6 % and $R_{cryst} = 34.8$ % using data between 10 and 3 Å resolution. The structure was refined initially with the 1.9 Å data using iterative cycles of simulated annealing with X-PLOR (Brünger, 1992) and manual rebuilding using O (Jones *et al*., 1991). Solvent molecules were identified and inserted using waterpick (X-PLOR package). As there was no sequence data available for the inhibitor peptide, a polyalanine chain was initially built into the density, and the side-chains altered as residues were identified from the electron density maps. This was further facilitated by the greater clarity of maps calculated using the 1.65 Å data when this became available. Refinement with this data proceeded with the use of REFMAC (Murshudov *et al*., 1997), starting with the refined 1.9 Å structure with all solvent molecules and the inhibitor removed. Final statistics for the model refined at 1.65 Å are summarised in Table 4.

Protein Data Bank accession number

Coordinates and structure factors have been deposited with the Brookhaven Protein Data Base (accession number 1SFI).

Table 4. Summary of X-ray diffraction data and final model statistics

Resolution range (Å)	20-1.65
Completeness (%)	96.7
Total number of reflections	82,087
Total number of unique reflections	32,839
Completeness (highest resolution shell 1.78 - 1.65 Å) (%)	91
R_{merge} (%)	6.7
Final model	
Number of protein atoms	1629
Number of solvent atoms	376
Number of inhibitor atoms	105
RMS standard deviation: bond lengths (Å)	0.014
RMS standard deviation: bond angles (deg.)	2.46
Final *R*-value	0.175
Final free *R*-value	0.203

Acknowledgements

We thank the staff at the Daresbury SRS for access to synchrotron data collection facilities. R.S.G. is supported by a grant from the Ministerio de Education y Ciencia, Spain, and S.L. by a BBSRC/CASE studentship. Al.V.K. and P.R.S. acknowledge the Royal Society for a Joint Project grant (683072.P810). The authors thank Mervyn Lewis (IACR-Long Ashton) and Dr Phillip Bonner (Nottingham Trent University) for additional mass spectroscopy. IACR (S.L. and P.R.S.) receives grant-aided support from BBSRC.

References

Bartunik, H. D., Summers, L. J. & Bartsch, H. H. (1989). Crystal structure of bovine β-trypsin at 1.5 Å resolution in a crystal form with low molecular packing density. *J. Mol. Biol.* **210**, 813-828.

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

Brünger, A. T. (1992). *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Collaborative Computing Project Number 4 (1994). Programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.

Domingo, G. J., Leatherbarrow, R. J., Freeman, N., Patel, S. & Weir, M. (1995). Synthesis of a mixture of cyclic-peptides based on the Bowman-Birk reactive-site loop to screen for serine-protease inhibitors. *Int. J. Peptide Protein Res.* **46**, 79-87.

Erlanger, B. F., Kokowsky, N. & Cohen, W. (1961). The preparation and properties of two new chromogenic substrates of trypsin. *Arch. Biochim. Biophys.* **95**, 271-278.

Ganesh, V., Lee, A. Y., Clardy, J. & Tulinsky, A. (1996). Comparison of the structures of the cyclotheonamide A complexes of human α-thrombin and bovine β-trypsin. *Protein Sci.* **5**, 825-835.

Gariani, T. & Leatherbarrow, R. J. (1997). Stability of protease inhibitors based on the Bowman-Birk reactive site loop to hydrolysis by proteases. *J. Peptide Res.* **49**, 467-475.

Huang, Q., Liu, S. & Tang, Y. (1993). Refined 1.6 Å resolution crystal 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.

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

Laskowski, M. & Kato, I. (1980). Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593-626.

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.

Li, Y., Huang, Q., Zhang, S., Liu, S., Chi, C. & Tang, Y. (1994). Studies on an artificial trypsin inhibitor peptide derived from the Mung bean trypsin inhibitor: chemical synthesis, refolding, and crystallographic analysis of its complex with trypsin. *J. Biochem.* **116**, 18-25.

Marquart, M., Walter, J., Drisenhofer, J., Bode, W. & Huber, R. (1983). The geometry of the reactive site and of the peptide groups in trypsin, tryspsinogen and its complex with inhibitors. *Acta Crystallog. sect. B*, **39**, 480-490.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240-255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.

Nishino, N., Aoyagi, H., Kato, T. & Izumiya, N. (1977). Studies on the synthesis of proteinase inhibitors. *J. Biochem.* **82**, 901-909.

Odani, S. & Ikenaka, T. (1973). Studies on soybean trypsin inhibitors. VIII. Disulphide bridges in soybean Bowman-Birk proteinase inhibitors. *J. Biochem.* **74**, 697-715.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I. & Lakowski, M., Jr. (1982). Crystallographic refinement of Japanese quail ovomucoid, a Kazal-type inhibitor, and model-building studies of comlexes serine. *J. Mol. Biol.* **158**, 515-537.

Ranson, N. A., Burston, S. G. & Clarke, A. R. (1997). Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. *J. Mol. Biol.* **266**, 656-664.

Richardson, M. (1991). Seed storage proteins: the enzymes inhibitors. Methods in Plant. *Biochemistry*, **5**, 259-305.

Ryan, C. A. (1990). Protease inhibitors in plants: genes for improving defences against insects and pathogens. *Annu. Rev. of Phytopath.* **28**, 425-449.

Schechter, I. & Berger, A. (1967). On the size of the active site proteases. I. Papain. *Biochim. Biophys. Res. Commun.* **27**, 157-162.

Shewry, P. R. (1999). Enzyme inhibitors of seeds: types and properties. In *Seed Proteins* (Shewry, P. R. & Casey, R., eds), Kluwer Academic Publishers, Dordrecht, Holland.

Shewry, P. R. & Lucas, J. A. (1997). Plant proteins that confer resistance to pests and pathogens. *Advan. Botan. Res.* **26**, 135-192.

Terada, S., Sato, K., Kato, T. & Izumiya, N. (1980). Studies on the synthesis of proteinase inhibitors. *Int. J. Peptide Protein Res.* **15**, 441-454.

Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C., Watanabe, K., Nakamura, S. & Takahashi, K. (1986). Structure of the trypsin-binding domain of Bowman-Birk type protease inhibitor and its interaction with trypsin. *J. Biochem.* **100**, 1637-1646.

Voss, R., Ermler, U., Essen, L., Wenzl, G., Kim, Y. & Flecker, P. (1996). Crystal structure of the bifunctional soybean Bowman-Birk inhibitor at 0.28-nm resolution. *Eur. J. Biochem.* **242**, 122-131.

Werner, M. H. & Wemmer, D. E. (1992). Three dimensional structure of soybean trypsin/chymotrypsin Bowman-Birk inhibitor in solution. *Biochemistry*, **31**, 999-1010.

Edited by I. A. Wilson

(Received 20 January 1999; received in revised form 12 May 1999; accepted 20 May 1999)