Article No. mb982032 *J. Mol. Biol.* (1998) **282**, 447–457

JMB

The Role of Threonine in the P_2 Position of Bowman-Birk Proteinase Inhibitors: Studies on P_2 Variation in Cyclic Peptides Encompassing the Reactive Site Loop

Jeffrey D. McBride, Arnd B. E. Brauer, Marco Nievo and Robin J. Leatherbarrow*

Department of Chemistry Imperial College of Science Technology and Medicine South Kensington, London SW7 2AY, UK

Previously, we have described a template-assisted combinatorial peptide library based on the anti-tryptic reactive site loop of a Bowman-Birk inhibitor (BBI). Sequences that displayed inhibitory activity re-directed towards chymotrypsin were found to have a consensus binding motif, with their most striking feature being that exclusively threonine was found at the P_2 position. The present study investigates the reason for this surprising specificity by maintaining the binding motif but systematically varying the P_2 residue. From analysis of 26 variants, it is found that the requirements for inhibitory activity at P_2 are finely tuned, and in agreement with the library work, threonine at P_2 provides optimal inhibition. In addition, peptides with threonine at P_2 are significantly less susceptible to hydrolysis. Examination of all available BBI sequences shows that threonine is very highly conserved at P_2, which implies that the functional requirement extends to the full-length BBI protein. Our results are consistent with a dual requirement for hydrophobic recognition within the S_2 pocket and maintenance of an inhibitory conformation *via* hydrogen bonding within the reactive-site loop. As the isolated peptide loop reproduces the active region of full-length BBI, these results explain why threonine is well conserved at P_2 in this class of inhibitor. Furthermore, they illustrate that proteinase inhibitor specificity can have characteristics that are not easily predicted from information on the substrate preferences of a proteinase.

Keywords: serine proteinase; proteinase inhibitor; Bowman-Birk inhibitor; canonical loop; combinatorial chemistry

**Corresponding author*

Introduction

Inhibition by serine proteinase inhibitor proteins is often mediated by an exposed reactive site that is fixed in a characteristic ''canonical'' conformation, thought to be similar to that of a productively bound substrate (Laskowski & Kato, 1980; Bode & Huber, 1992). The sequence of the reactive site determines the specificity of the inhibition, which frequently mirrors the specificity of proteinases for their substrates. The conformation is stabilised by many intramolecular interactions between residues flanking the reactive site and the inhibitor core. Unlike a good substrate, however, the energy barrier for hydrolysis is large and unfavourable, resulting in low rates of hydrolysis (Laskowski & Kato, 1980; Read & James, 1986; Longstaff *et al.*, 1990). For most inhibitor proteins, disulphide bonds and/or extensive hydrogen bond networks maintain the structure of the canonical loop. For example, the Bowman-Birk family of serine proteinase inhibitors (BBIs) are small proteins (6 to

Abbreviations used: BBI, Bowman-Birk inhibitor; P_n, the *n*th peptide subsite of a substrate or inhibitor on the carboxyl side of the scissile bond; P'_n, The *n*th peptide subsite of a substrate or inhibitor on the amino side of the scissile bond; DMF, dimethylformamide; DMSO, dimethylsulphoxide; NMP, *N*-methylpyrrolidone; HBTU, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, *N*-hydroxybenzotriazole; r.m.s., root-mean-square.

E-mail address of the corresponding author: R.Leatherbarrow@ic.ac.uk

0022–2836/98/370447–11 $30.00/0

9 kDa) with a symmetrical structure of two tricyclic domains, each containing an independent binding loop (Ikenaka & Norioka, 1986; Chen *et al*., 1992; Werner & Wemmer, 1992a; Lin *et al*., 1993). The stability and inhibitory activity of all members of this family has been attributed to the conserved disulphide bridging pattern (Ikenaka & Norioka, 1986). BBIs are unusual, as they can form a complex with two separate proteinases, typically allowing simultaneous inhibition of chymotrypsin and trypsin. These activities can be separated by fragmentation using a combination of cyanogen bromide and pepsin (Odani & Ikenaka, 1973). Each proteinase binding domain has the same general structure, and is typically contained within a nonapeptide region joined *via* a disulphide between flanking cysteine residues. Synthetic cyclic peptides that incorporate this nonapeptide motif retain inhibitory activity (Nishino *et al*., 1977; Terada *et al*., 1978), and are thought to have the same reactive-site loop structure as full-length BBI protein (Li *et al*., 1994; Pavone *et al*., 1994).

Using molecular graphics and least-squares fitting, Chen *et al*. (1992) have demonstrated that the co-ordinates for the reactive loop of BBI (PI-II) can be superimposed with a member of the Kazal family, PSTI and the Kunitz family, BPTI. The chymotrypsin-binding loop of BBI is also very similar to the binding loop conformation of potato chymotrypsin inhibitor-I, turkey ovomucoid inhibitor third domain and barley chymotrypsin inhibitor-II (Werner & Wemmer, 1992b). In each case, the only common structural feature is the conformation of the reactive site, the cores of the inhibitors being quite different.

The mechanism of action of canonical inhibitors is well established, and the identity of the P_1 residue is considered to be the main determinant of specificity. Redirection of inhibitory activity by site-directed mutagenesis (Longstaff *et al*., 1990; Lu *et al*., 1997) semi-synthesis and peptide synthesis (Nishino *et al*., 1977; Terada *et al*., 1978; Favel *et al*., 1989; Domingo *et al*., 1995) at this locus have been described. This preference is dependent on the type of substrate or inhibitor (Longstaff *et al*., 1990), and therefore re-directing inhibitor specificity might benefit from a more extensive series of changes. One approach to achieve this has been provided by phage display libraries, which have principally been applied to members of the Kunitz family (Roberts *et al*., 1992; Markland *et al*., 1996a,b). We have recently used a synthetic combinatorial protocol to create a library of potential inhibitors (McBride *et al*., 1996). Such an approach has the advantage of allowing incorporation of non-proteinogenic amino acids, but is complicated by the complexity of natural inhibitor proteins. Indeed, one goal of inhibitor research has been the reduction in size and simplification of these proteins to their minimal structural elements (Wenzel & Tschesche, 1995). Our library employed an 11 amino acid residue cyclic peptide template (McBride *et al*., 1996) derived from the anti-tryptic loop domain of the Bowman Birk inhibitor MAI-D4 (Maeder *et al*., 1992). In our studies, this template was used to create a ''one bead, one peptide'' library (Furka *et al*., 1991; Lam *et al*., 1991) in which three locations considered important for proteinase recognition were varied (P_1, P_2 and P_2'). The incorporation of one of 20 possible residues at each locus generated a library with 8,000 different components. When this library was screened for re-directed activity against chymotrypsin, active sequences were found to contain exclusively threonine at the P_2 site. The P_1 site had phenylalanine and tyrosine, with leucine, isoleucine and norleucine found at P_2' (McBride *et al*., 1996). Although the preference for aromatic residues at P_1 is consistent with the substrate preference of chymotrypsin (Schellenberger *et al*., 1991), the sequences obtained at the other locations were less predictable, and it was particularly striking that exclusively threonine was found at the P_2 location.

In order to evaluate further what features of the P_2 locus are necessary to the properties of the peptide mimetic, we have undertaken a systematic analysis of 26 P_2 variations of the consensus chymotrypsin-binding motif: cyclic SC**X**FSIPPQCY (P_2 position shown in bold, cyclised *via* a disulphide between the Cys residues). Our results allow us to infer why, in BBI loops, a P_2 Thr residue is extremely well conserved. They also allow an evaluation of the results obtained from screening a peptide library.

Results

Effect of P_2 variation on K_i

Using the consensus sequence (cyclic SC**T**FSIPPQCY; P_2 residue in bold, cyclised *via* formation of an intramolecular disulphide) found for re-directed activity of the anti-tryptic loop towards chymotrypsin, a range of P_2 variants was synthesised. Initially, all the amino acids present in the original library were tested and the individual K_i values that were derived are listed in Table 1. It is found that variation at P_2 results in K_i values differing over three orders of magnitude, and ranging from 19 nM to almost 100 μM.

In the light of the K_i values obtained, a further series of analogues was synthesised, which probed the specific role of the threonine side-chain functionalities. The K_i values for these peptides are listed in Table 2.

Effects on hydrolysis rates

We have previously noted that synthetic peptides based on BBI proteins often show marked differences in the rates at which they are hydrolysed by proteinases (Domingo *et al*., 1995; Gariani & Leatherbarrow, 1997). The results of the current study suggested that one factor in the role of threo-

Table 1. Equilibrium dissociation constants (K_i) for P_2 variants of the chymotrypsin re-directed anti-tryptic loop of MAI-D4

P_2 residue	K_i (μM)	Side-chain classification
Thr	0.019	Neutral, hydrophilic
Ser	0.4	
Gln	7.1	
Asn	17	
Ala	0.57	Neutral, hydrophobic, aliphatic
Val	1.3	
Nle	1.3	
Met	2.2	
Leu	2.4	
Gly	4.4	
Ile	13.4	
Phe	8.7	Neutral, hydrophobic, aromatic
Trp	9	
Tyr	17.1	
Glu	84	Acidic, hydrophilic
Asp	98	
His	1.4	Basic, hydrophilic
Arg	1.8	
Lys	2.8	
Pro	5.2	Imino acid

All the amino acids used in the original library screening (McBride *et al*., 1996) were used to generate 20 different P_2 variants of the sequence SC**X**FSIPPQCY (cyclised *via* formation of an intramolecular disulphide), where **X** has been replaced by one of the residues listed in the Table. The K_i values for chymotrypsin inhibition were determined by competitive binding studies using the chromogenic substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. Assays were performed at pH 7.8 in 144 mM Tris-HCl at 298 K. Standard error values for the K_i values, as estimated from the data fitting, were in each case <10%.

Table 2. Equilibrium dissociation constants (K_i) for further P_2 variants of the chymotrypsin re-directed anti-tryptic loop of MAI-D4

P_2 residue	Side-chain structure	K_i (μM)
Ala	$-CH_3$	0.57
Abu	$-CH_2-CH_3$	0.13
Nva	$-CH_2-CH_2-CH_3$	0.28
Nle	$-CH_2-CH_2-CH_2-CH_3$	1.3
Val	$-CH(CH_3)-CH_3$	1.3
t-Butylglycine	$-C(CH_3)_2-CH_3$	560
Thr	$-CH(OH)-CH_3$	0.019
Ser	$-CH_2-OH$	0.4
Hse	$-CH_2-CH_2-OH$	2.6
allo-Thr	$-CH(CH_3)-OH$	130
D-Thr		NI

NI, no detectable inhibition at concentrations up to 0.25 mM.

Inhibition constants are given for P_2 variants of the sequence SC**X**FSIPPQCY (cyclised *via* formation of an intramolecular disulphide), where **X** has been replaced by one of the residues listed in the Table. The K_i values for chymotrypsin inhibition were determined by competitive binding studies using the chromogenic substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. Assays were performed at pH 7.8 in 144 mM Tris-HCl at 298 K. Standard error values for the K_i values, as estimated from the data fitting, were in each case <10%.

nine at P_2 was a facility to form an internal hydrogen bond. As this seemed likely to influence the rate of hydrolysis, a time-course of the hydrolysis was followed for selected analogues using HPLC. Figure 1 records a hydrolysis time-course for P_2 Thr, Ser, Ala and Abu peptides. It is found that the P_2 Thr variant is the most stable to hydrolysis, and that those peptides with a β-hydroxyl function (Thr and Ser) are significantly slower to hydrolyse than those where this group is deleted.

Fitting the data assuming a pseudo first-order process allows the differences in hydrolysis rate constants to be compared. The ratio of the hydrolysis rate constants for the Thr, Ser, Abu and Ala peptides is found to be 1:19:170:500, respectively.

Discussion

Conservation of the P_2 residue in BBI proteins

The activity of full-length BBI protein is dependent upon interactions between the reactive site of the inhibitor and the active site of the target protease. The peptides used in the current study act by reproducing the reactive site, and so it is reasonable to infer that factors important for inhibition by these peptides will be relevant to the full protein. The converse will also apply; information about the protein should be pertinent to the peptide studies.

A large number of sequences are available for the inhibitory loops from BBI proteins, which give some indication of the importance of the P_2 residue. Table 3 gives an alignment for all currently available BBI reactive-site loop sequences that include a nine residue disulphide-linked motif†. From this comparison, it is seen that the P_2 residue is very highly conserved, with 87% of sequences having Thr at this locus. This high level of sequence conservation implies that a P_2 Thr residue is functionally important to the inhibitory properties. This is consistent with our earlier peptide library screening (McBride *et al*., 1996), and

† Members of the peanut A-II family of BBI proteins have one nine-membered and one 11-membered disulphide-linked reactive-site loop, which each have distinct conformations (Suzuki *et al*., 1993). For this reason, we have not included any 11-membered loop in this comparison.

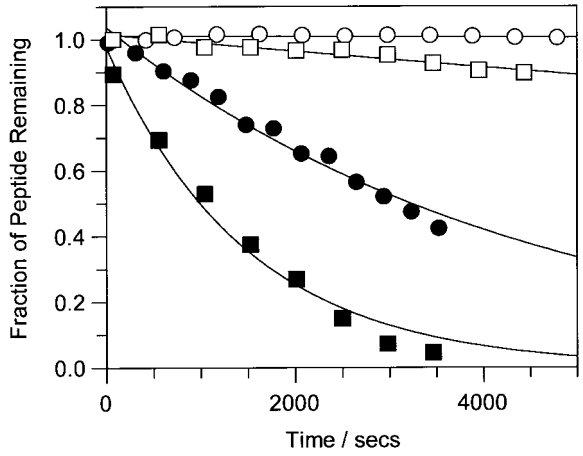

Figure 1. Time-course of chymotrypsin hydrolysis for various inhibitor peptides that differ at the P_2 position. Hydrolysis of peptides with P_2 Thr (○), Ser (□), Abu (●) and Ala (■) was performed in 144 mM Tris, pH 7.8 at 298 K. Degradation of the peak corresponding to the cyclic peptide was monitored by integration of the peak area after separation by HPLC. The data were fitted assuming a pseudo first-order process, and the continuous lines show the theoretical fits.

supports the view that the loop peptides have inhibition properties similar to those of the whole protein.

Effect of P_2 variation on K_i

The initial impetus to the current study was to discover why our earlier library screening resulted in all active sequences having a P_2 threonine residue (McBride *et al.*, 1996). To do this, we retained the consensus sequence found for the remaining residues that were randomised in the library (P_1 Phe; P'_2 Ile), and systematically varied the P_2 residue. The K_i values found for 20 single position variants, which represent the total variation at P_2 present in our original library, are presented in Table 1. The results confirm a significant influence of the P_2 residue on inhibitory activity. Although all sequences were active inhibitors, there was variation of K_i by over three orders of magnitude resulting from the substitutions. Reassuringly, the only residue identified at P_2 by our earlier screening procedure, Thr, was found to give the lowest K_i value (Table 1). This explains why our earlier library studies returned exclusively Thr at this location. The sequence from the original library with the next lowest K_i value, given by Ser at P_2, is a 20-fold poorer inhibitor, which is apparently sufficient discrimination to ensure that no P_2 Ser sequence was returned in our earlier experiments.

From the results, it is found that there is a preference for small residues at P_2. Large aliphatic or aromatic side-chains result in relatively poor inhibitors, suggesting that they are not well accommodated in the S_2 pocket. Particularly noteworthy is the extremely poor inhibition when a negatively charged residue is present at P_2 (Asp or Glu). It is likely that the presence of these side-chains results in unfavourable electrostatic interactions with Asp102 of the catalytic triad, which is in close proximity.

The K_i value for the P_2 Thr variant is only slightly higher than that of the classical soybean BBI (2-IV) for chymotrypsin (6.4 nM, Larionova *et al.*, 1993; 13 nM, Voss *et al.*, 1996). This suggests that the peptide fragment retains most of the interactions that are found with the intact protein, and thus is a good model for this inhibitor.

Why is Thr the best P_2 residue?

From our results, it is found that a P_2 Thr gives the lowest K_i value. To probe the reasons for this further, we synthesised a range of variants that allow more detailed examination of the contribution made by the threonine side-chain. Table 2 lists the K_i values for a number of peptides that have P_2 residues related to threonine. To facilitate comparison, Table 2 also records the values for peptides containing P_2 Ala, Ser, Val and Nle, and shows the structures of the various side-chains.

The results show that the requirements for inhibitory activity at the P_2 locus are finely tuned. Inversion of the chiral β-carbon atom from R to S (allo-Thr) results in a very poor inhibitor, as does the replacement of L-Thr with D-Thr. Loss of either the β-hydroxyl or methyl functions of the threonine residue produces a significant loss of activity. Whilst the loss of the hydroxyl gives a 6.8-fold reduction (Abu), the loss of the methyl group (Ser) gives a 21-fold reduction in activity. In the absence of the hydroxyl function, the homologous series Ala, Abu, Nva and Nle further demonstrate the requirement for a methyl group attached to the β-carbon atom. Relative to Abu, when the aliphatic side-chain is reduced in size to Ala there is an approximate fourfold loss in activity, but further increase of the alkyl length to Nva results in a twofold loss in activity. Increasing the side-chain length by a further methylene group (Nle) results in a K_i value tenfold higher than that for Nva. Overall, these results indicate that the optimal aliphatic side-chain is two carbon units long. Side-chain branching generally results in reduced activity, particularly for branching at the β-carbon atom. Hence, Val has a 4.6-fold higher K_i than Nva, and Ile a 10.3-fold higher K_i than Nle. More substantial branching at the β-carbon atom, as found in the t-butylglycine variant, appears to be very poorly accommodated within the S_2 pocket, resulting in extremely low activity. In a similar fashion, the requirement for a β-hydroxyl group is demonstrated by comparing Ser and Hse, with a γ-hydroxyl group giving a K_i value 6.5-fold higher than the β equivalent.

Energetics of the Thr side-chain interactions

$$\Delta\Delta G = -RT \ln \frac{K_{i(1)}}{K_{i(2)}} \quad (1)$$

The energetics of the interactions made by the methyl and hydroxyl functionalities of the Thr side-chain can be calculated by comparing the K_i values for the Thr, Ser, Ala and Abu peptides. For pairs of inhibitors, the difference in binding free energy, $\Delta\Delta G$, is related to the ratio of the two K_i values as described in equation (1):

The change from Ala to Thr involves insertion of a methylene group and an oxygen atom; Abu and Ser represent stepwise introductions of these units, respectively. Comparison of the $\Delta\Delta G$ values for the γ-methyl and β-hydroxyl interactions shows that their contributions are synergetic, i.e. the combined effect is greater than the sum of the individ-

Table 3. Sequence of the P4-P7′ region of various BBI proteins

	P4	P3	P2	P1	P1′	P2′	P3′	P4′	P5′	P6′	P7′
Soybean I [IBB1_SOYBN]	A	C	T	K	S	N	P	P	Q	C	R
	I	C	A	L	S	Y	P	A	Q	C	F
Soybean CII [IBB2_SOYBN]	M	C	T	A	S	M	P	P	Q	C	H
	A	C	T	R	S	M	P	G	Q	C	R
Soybean DII [IBB3_SOYBN]	M	C	T	R	S	M	P	P	Q	C	S
	M	C	T	R	S	Q	P	G	Q	C	R
Mung bean MBI [IBB_PHAAN]	D	C	T	K	S	I	P	P	Q	C	H
	I	C	T	R	S	M	P	G	K	C	R
Common vetch VAI [IBB_VICAN]	L	C	T	R	S	Q	P	P	T	C	R
	V	C	N	Y	S	N	P	P	Q	C	Q
Cowpea BTCI [IBB_VIGUN]	A	C	T	K	S	I	P	P	Z	C	R
	A	C	T	F	S	I	P	A	Z	C	F
Broad bean FBI [IBB_VICFA]	A	C	T	K	S	I	P	P	Z	C	R
	V	C	R	Y	S	N	P	P	K	C	Q
Apple leaf DE-4 [IBB4_LONCA]	C	C	T	R	S	R	P	P	Q	C	Q
	M	C	T	F	S	I	P	A	Q	C	V
Macrotyloma axillare DE-3	A	C	T	K	S	I	P	P	Q	C	R
[IBB3_DOLAX]	V	C	T	F	S	I	P	A	Q	C	F
Macrotyloma axillare DE-4	T	C	T	K	S	I	P	P	Q	C	H
[IBB4_DOLAX]	I	C	A	L	S	E	P	A	Q	C	F
Barley [IBB_HORVU]	V	C	T	R	S	I	P	P	I	C	T
	I	C	T	R	S	N	P	P	T	C	R
Adzuki bean [IBB1_PHAAN]	S	C	T	K	S	M	P	P	K	C	R
	A	C	T	Y	S	I	P	A	K	C	F
Adzuki Bean IA IB IA′ [IBB2_PHAAN]	L	C	T	K	S	I	P	P	Q	C	Q
	M	C	T	R	S	M	P	G	Q	C	R
Medicago sativa [IWIT_MEDSA]	P	C	T	R	S	I	P	P	Q	C	H
	L	C	T	K	S	I	P	P	Q	C	R
Lima bean [IBB_PHALU]	A	C	T	K	S	I	P	P	Q	C	R
	I	C	T	L	S	I	P	A	Q	C	V
French bean II [IBB_PHAVU]	V	C	T	A	S	I	P	P	Q	C	V
	M	C	T	R	S	M	P	G	K	C	R
Torresea cearensis[a]	A	C	T	K	S	I	P	P	Q	C	H
	A	C	T	H	S	I	P	A	Q	C	R
Vigna unguiculata[b]	E	C	T	K	S	I	P	P	Q	C	R
	A	C	T	F	S	I	P	A	Q	C	G
Erythrina variegata[c]	F	C	T	K	S	N	P	P	I	C	Q
	I	C	A	L	S	Y	P	A	Q	C	H
Peanut AII [IBB1_ARAHY]	V	C	T	R	S	N	P	P	Q	C	R
Peanut BII [IBB2_ARAHY]	V	C	T	R	S	I	P	P	I	C	R
Wheat I-2B [IBB1_WHEAT]	V	C	T	R	S	I	P	P	V	C	R
Wheat II-4 [IBB2_WHEAT]	I	C	T	K	S	F	P	P	M	C	R
Rice RBTI [IBBR_ORYSA]	F	C	N	K	M	N	P	P	T	C	R
Medicago scutellata [IBB_MEDSC]	P	C	T	R	S	I	P	P	Q	C	Q
Foxtail millet FMTI-II [IBB2_SETIT]	T	C	T	K	S	I	P	A	F	C	R
Foxtail millet FMTI-III [IBB3_SETIT]	T	C	T	K	S	I	P	A	F	C	R
Job's tears [IBB1_COILA]	M	C	T	R	S	I	P	P	I	C	R

Unless otherwise referenced, sequence data have been extracted from the University of Geneva ExPasy database (http://expasy.hcuge.ch and references therein) and the sequence identifier code is given in the Table. The inhibitors in the first group (Soybean to *Erythrina*) have two reactive sites that each contain a nine-membered disulphide-bonded loop (P3–P6′), and both sequences are given. The second group of sequences is from BBI proteins that have one disulphide-linked reactive site loop of nine residues, and one 11-membered loop. The 11-membered motifs have a different structure (Suzuki *et al*., 1993), and so only the nine-membered loop sequence is listed for comparison.

[a] Tanaka *et al*. (1997).

[b] Morhy & Ventura (1987).

[c] Kimura *et al*. (1994).

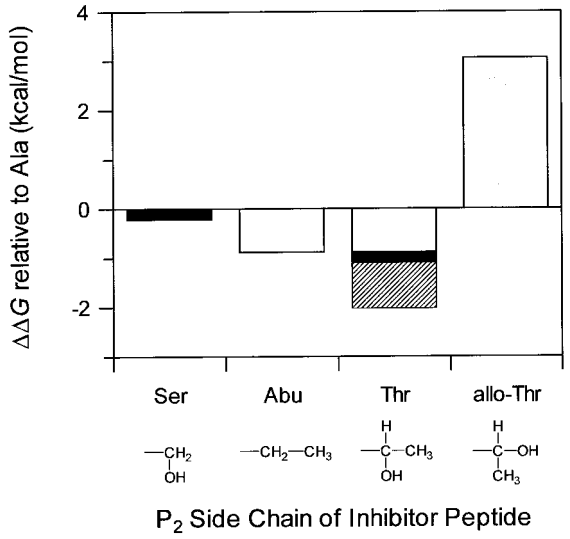

Figure 2. Contributions from the γ-methyl and β-hydroxyl groups at the P_2 position of the inhibitor peptide. The difference in binding free energy, $\Delta\Delta G$, is shown relative to the peptide having a P_2 Ala residue. The binding energy of a P_2 Thr is found to be greater than the sum of the Ser (filled) and Abu (open) combined. The hatched area indicates the additional synergetic contribution. The interactions made by the hydroxyl and methyl portions of the threonine are stereospecific, as illustrated by the extremely unfavourable substitution to allo-Thr, which differs only in the chirality of the side-chain.

ual parts (Figure 2). The γ-methyl interaction is worth −0.88 kcal mol^{-1} comparing Abu with Ala, but −1.81 kcal mol^{-1} when the hydroxyl group is present (comparing Ser with Thr). Similarly, the β-hydroxyl group interaction accounts for only −0.21 kcal mol^{-1} comparing Ser with Ala, but −1.14 kcal mol^{-1} when the methyl group is present (Abu *versus* Thr). The synergetic contribution is therefore −0.93 kcal mol^{-1}. The effect depends upon the stereochemistry of the side-chain, as inversion of the Thr side-chain configuration to allo-Thr exacts a penalty of 5.07 kcal mol^{-1}.

Structural data

The above results, which indicate that important interactions are made involving both the hydroxyl and methyl groups of the P_2 Thr, can be explained in light of the available structural data for BBIs and BBI-proteinase complexes. For uncomplexed BBI, there are structures determined by X-ray crystallography (Chen *et al.*, 1992; Suzuki *et al.*, 1993; Voss *et al.*, 1996) and by NMR (complete protein, Werner & Wemmer, 1992a ; reactive-site peptides, Pavone *et al.*, 1994) In complex with trypsin, there is an X-ray structure for full-length BBI (Lin *et al.*, 1993), for a fragment of BBI (approximately half the molecule; Zhang *et al.*, 1985) and for a tricyclic reactive-site loop peptide (Li *et al.*, 1994). All the complexed structures show essentially the same interactions and the same characteristic conformation for the reactive-site loop, which has *cis*-Pro at P3′. In our structural analysis, we have used the structure reported by Li *et al.* (1994), since this inhibitor most resembles our peptides. There is currently no structure of BBI bound to chymotrypsin, and so we have used structures of this enzyme in complex with two other natural inhibitors: eglin c (Frigerio *et al.*, 1992) and turkey ovomucoid third domain. (Fujinaga *et al.*, 1987). These inhibitors show the typical canonical conformation, and have a P_2 Thr residue. Backbone superimposition of these three structures (Figure 3) shows a high degree of homology in the spatial arrangement of the selected residues. The interactions that are made by the P_2 Thr of the BBI loop peptide with chymotrypsin can therefore be inferred from these structures, which is illustrated in Figure 4.

The crystal structure of the BBI-trypsin complex shows a hydrogen bond network involving the hydroxyl groups of P_2 Thr and P_1' Ser, and the backbone amide group of P_5' Gln (Figure 5). This interaction is consistent with our K_i and hydrolysis data (see below), both of which suggest a role for the Thr hydroxyl group. In contrast to the crystallographic data (Voss *et al.*, 1996), the NMR structure does not specifically show the P_2 Thr hydroxyl group involved in this hydrogen bonding network

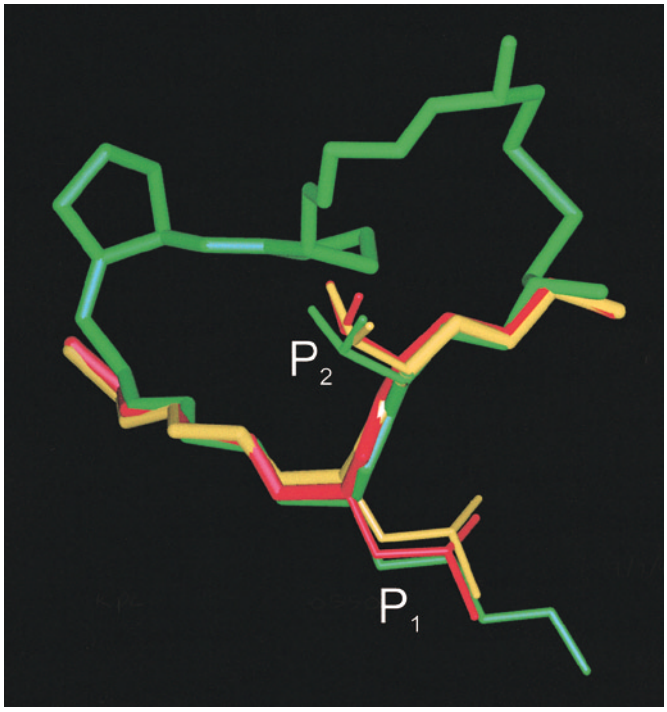

Figure 3. Superimposition of the residues spanning P_3-P_2' from the structures of eglin c (red; Frigerio *et al.*, 1992) and turkey ovomucoid third domain (yellow; Fujinaga *et al.*, 1987) onto the reactive-site loop of the Bowman-Birk inhibitor (green; Li *et al.*, 1994). The side-chains of the P_1 and P_2 residues are shown to demonstrate that the P_2 Thr residues adopt a similar orientation in each structure.

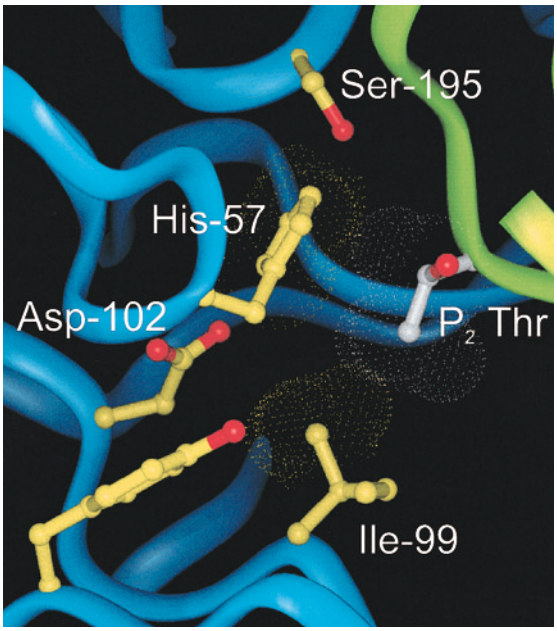

Figure 4. Interactions between the P_2 Thr side-chain and chymotrypsin. The area around the P_2 side-chain in the complex between eglin c and chymotrypsin (Frigerio *et al.*, 1992) is illustrated, showing the contacts between this side-chain and His57 and Ile99 of the enzyme. The backbones of the proteins are displayed as ribbons (chymotrypsin, blue; eglin c, green). The van der Waals surfaces of the interacting residues are shown dotted.

(Werner & Wemmer, 1992a); this would be more difficult to fit with our results.

The crystal structures all show that the β-carbon atom of the P_2 residue rests against the imidazole ring of His57 (Figure 4). Extensive branching at the β-carbon atom would be expected to be poorly accommodated, consistent with our data. The S_2 subsite in chymotrypsin is a relatively shallow hydrophobic pocket (Figure 4), with the P_2 Thr γ-methyl group packing against the side-chain of Ile99. It seems likely that any increase in size would be poorly accommodated within the structure. Overall, the structural data are consistent with the observation that both the hydroxyl and methyl functions of the P_2 threonine residue are important for low K_i, and show why no other residue (of those tested) gives such potent inhibition.

Effects on hydrolysis rates

The presence of an intramolecular hydrogen bond network involving the P_2 Thr hydroxyl group might be expected to stabilise the organisation of the reactive site (Figure 5), and could influence the rate of hydrolysis. It is well known that these peptide inhibitors are also substrates (as indeed are the inhibitor proteins), but that hydrolysis rates are relatively low. During preliminary experiments with the peptides described here, we observed that some sequences appeared to be more prone to hydrolysis than others. To examine the effect of the P_2 Thr hydroxyl group, we performed detailed hydrolysis studies on those peptides with Thr, Ser, Ala or Abu at P_2. Comparison of these sequences should reveal the effect (if any) of the P_2 hydroxyl group on hydrolysis rates. Figure 1 shows the time-course of hydrolysis of these peptides under comparable conditions, as monitored by HPLC. It is found that sequences with a P_2 hydroxyl group are far less prone to hydrolysis than their aliphatic counterparts. This strongly suggests that the hydrogen bonding revealed in the crystal structures leads to lower rates of hydrolysis for the scissile bond. The P_2 Thr sequence is the most stable of all those tested. This observation is presumably a further reason why Thr is highly conserved at this position in BBI proteins (Table 3), and is likely to have been a further factor in selection of sequences from our original library (McBride *et al.*, 1996). It is interesting to

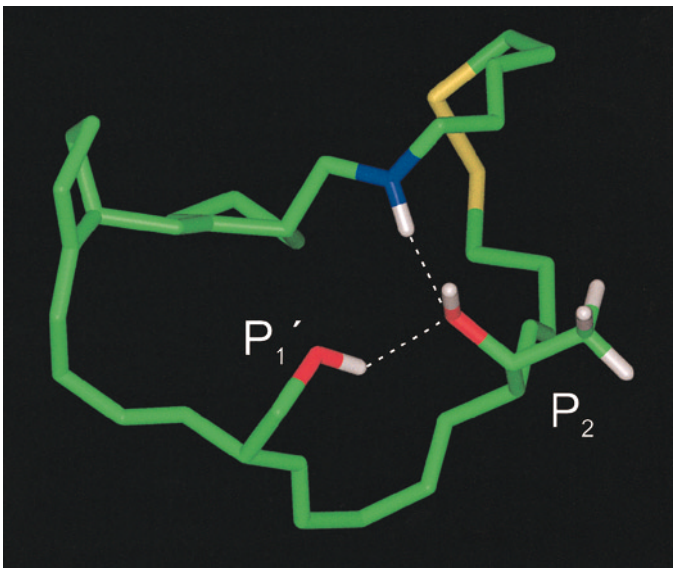

Figure 5. Structure of the reactive-site loop of BBI taken from the X-ray structure of the complex with trypsin (Li *et al.*, 1994). The hydrogen bonding interactions involving the P_2 Thr hydroxyl group are shown as broken lines.

compare the relative contributions of interactions from the γ-methyl and β-hydroxyl groups of the Thr towards K_i values and hydrolysis rate constants. Loss of the β-hydroxyl group, which participates in intramolecular hydrogen bonding, results in much more rapid rates of hydrolysis, yet has only a small effect on the K_i value. In contrast, loss of the γ-methyl group, which makes intermolecular contacts with the proteinase, results in a lesser effect on hydrolysis rates but a more significant increase in K_i.

The importance of stabilising intramolecular hydrogen bonds in binding loop stability and inhibitory mechanism has been implicated by point mutations of eglin c. A single point mutation of P_1 from Leu to Arg can re-direct specificity from elastase to trypsin (Heinz *et al*., 1991). However, in an attempt to improve this redirection, a further mutation to the P_1' from Asp to Ser resulted in a shift to substrate-like behaviour (Heinz *et al*., 1992). Although NMR studies showed virtually the same conformation as the wild-type inhibitor, there were indications that the internal rigidity of the binding loop was significantly weakened due to the loss of a hydrogen bond. Increased vulnerability to hydrolysis is found also for Arg50 and Arg52 mutations of CMTI-V (Cai *et al*., 1996), where removal of hydrogen bonding of the loop to the protein core was demonstrated to result in increased conformational flexibility.

Model studies have suggested that the release of processed substrate from a serine proteinase acyl intermediate requires a substantial change in the χ_1 of His57 (Asbóth & Polgár, 1983; Deslongchamps, 1983; Gronenstein & Taira, 1984). It is interesting to speculate that the close hydrophobic contact seen in the crystal structure between the P_2 Thr and the His57 imidazole ring would act to restrict such movement, and so reduce the hydrolysis rate. If this were indeed the case, it might further explain the conservation of this residue and would be consistent with our hydrolysis data.

Comparison of P_2 inhibitor and P_2 substrate specificity

It is usually considered that natural proteinase inhibitors bind to serine proteinases in a substrate-like manner (Laskowski & Kato, 1980; Bode & Huber, 1992; Voss *et al*., 1996), and the P_1 residues of inhibitors typically follow the P_1 substrate specificity of the protease (Lu *et al*., 1997). However, the marked P_2 specificity found for these BBI inhibitor sequences is not reflected in the available substrate specificity data. In an examination of chymotrypsin specificity towards denatured polypeptide substrates, it was found that Val, Ile, Lys and Pro in P_2 favour hydrolysis (Neil *et al*., 1966). More recently, Schellenberger *et al*. (1991) surveyed all available quantitative data to calculate fragment contribution values (log k_{cat}/K_M) of substrates for chymotrypsin subsites. The study, however, relied upon a rather limited range of available synthetic substrate data, and so the S_2 specificity was poorly defined. In contrast to the P_1 and P_3 residues, which form hydrogen bonds with chymotrypsin (from the main-chain P_1 amide group, and the P_3 carbonyl group to Ser214 and Gly216 of the enzyme, respectively (Segal *et al*., 1971; Blow *et al*., 1972)), proline is considered to be well tolerated in the P_2 position, since no such bond is made (Schellenberger *et al*., 1991). A generalised canonical inhibitor has Pro at P_2 (Bode & Huber, 1992); however, there are several examples of inhibitors, including the BBI proteins, that have Thr at P_2. There is some evidence for Thr being the preferred residue in these proteins (Laskowski, 1980), and substitution of the P_2 Thr in eglin c by Pro results in an elastase inhibitor with a tenfold higher K_i (Heinz *et al*., 1992).

Our results, therefore, suggest that the BBI inhibitor shows P_2 specificity that is not a simple reflection of the enzyme-substrate specificity. There are two factors that are involved in overall inhibitor potency: tight binding and (for peptides) the rate of hydrolysis. A good inhibitor will maximise binding (i.e. have a low K_i); it can also be advantageous to minimise hydrolysis. However, tight binding (or a low K_M) can be deleterious to enzyme catalysis, and an optimal substrate would actually maximise K_M for a given k_{cat}/K_M (Fersht, 1974). It should therefore not be too surprising if the sequences that are best for a substrate are not necessarily those that lead to the best inhibitor. This has obvious implications for inhibitor design, which often takes the optimal substrate as the starting point. Such a route would probably not have generated the best P_2 sequence for chymotrypsin inhibition in this particular study.

Conclusions

We have previously shown that a Thr residue at P_2 is exclusively selected from a randomised library derived from the reactive-site loop of BBI protein (McBride *et al*., 1996). The current study confirms that, of 26 P_2 variants tested, Thr gives both the lowest K_i value and the slowest hydrolysis rate. These results account for our earlier findings, and explain why a Thr residue is found at P_2 in 87% of known BBI reactive loop sequences. Examination of the available structural information shows that interactions are made with both the hydroxyl and methyl groups of the Thr, and this is why no other residue is so effective at this location.

Materials and Methods

Materials

N-α-9-Fluorenylmethoxycarbonyl (Fmoc) protected amino acids and derivatives (all L-configuration unless stated otherwise) were purchased from Novabiochem (UK) or Advanced ChemTech Europe with the following side-chain protecting groups: Abu, Ala, Arg(2,2,5,7,8-pentamethylchroman-6-sulfonyl), Asn(trityl), Asp(tBu),

Cys(trityl), Gln(trityl), Glu(tBu), Gly, t-butylglycine, His (trityl), Ile, Leu, Lys(tBoc), Met, Nle, Nva, Phe, Pro, Ser(t-Bu), Thr(tBu), allo-Thr(tBu), D-Thr(tBu) Trp(tBoc), Tyr(t-Bu), Val. Dimethylformamide (DMF) and *N*-methylpyrrolidone (NMP) were peptide synthesis grade from Rathburn Chemicals (Walkerburn, UK) and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and *N*-hydroxybenzotriazole (HOBt) were from SNPE (Croydon, UK). Bovine pancreatic α-chymotrypsin (treated with *N*α-*p*-tosyl-L-lysine chloromethyl ketone), phenol, ethanedithiol, thioanisole, dimethylsulphoxide (DMSO), and trifluoroacetic acid were purchased from Sigma.

Peptide synthesis and purification

Peptides for kinetic analysis were synthesised by standard solid-phase techniques using base-labile Fmoc for the α-amino group protection (Atherton & Sheppard, 1989) with HBTU/HOBt activation (Knorr *et al*., 1989) as described (Domingo *et al*., 1995) using a Shimadzu PSSM-8 peptide synthesiser at 0.02 mmol scale with five-fold excess of amino acids.

Cleavage of the peptides from the resin and deprotection of the side-chain protecting groups was performed using a mixture of 2.5% (v/v) 1,2-dithiolethane 2.5% (v/v) water in trifluoracetic acid for 2.5 to 3 hours. Where the peptide contained Arg and or Met, the deprotection cocktail was 7.5% (w/v) phenol, 2.5% (v/v) 1,2-dithiolethane, 2.5% (v/v) thioanisole, 5% (v/v) water in trifluoracetic acid. Peptide was separated from resin, precipitated in 14 ml of ice-cold tert-methyl butyl ether (Aldrich) and collected by centrifugation. The precipitation was repeated five times with fresh solvent, and the peptide was dried under vacuum over silica gel.

Disulphide bridge formation was performed using DMSO oxidation (Tam *et al*., 1991) as described (McBride *et al*., 1996). Peptide purification was performed by reverse-phase HPLC using a Gilson system with a Waters C_{18} Radial Pak column (25 mm × 10 cm, 6μm particle size) using a gradient between mobile phases of water and acetonitrile, each containing 0.1% trifluoroacetic acid. All peptides were characterised by mass spectroscopy to confirm the correct mass and were determined to be >85% pure by analytical reverse-phase HPLC (Vydac C_{18}, 4.6 mm × 15 cm, 5 μm particle size).

Inhibition kinetics

The inhibition kinetics were determined by competitive binding studies using the chromogenic substrate *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Sigma). Assays were performed at pH 7.8 in 144 mM Tris-HCl at 298 K. The active chymotrypsin concentration was determined by active-site titration with *p*-nitrophenyl acetate (Bender *et al*., 1966) and peptide concentrations by measuring absorbance at 280 nm (Edelhoch, 1967). Substrate hydrolysis was monitored at 410 nm and substrate concentration determined by the final absorbance ($\varepsilon = 8800\ M^{-1}\ cm^{-1}$; Erlanger *et al*., 1961). Initial velocity data were fitted using the GraFit software package (Leatherbarrow 1990, 1992). K_i values were corrected for substrate competition using the formula $K_i = K_{i(app)}(1 + [S]/K_M)$, where $K_{i(app)}$ is the apparent K_i for a given substrate concentration.

Hydrolysis rates

The hydrolysis of the P_2 variants Thr, Ser, Ala and Abu was measured using 20-fold excess to chymotrypsin in 144 mM Tris (pH 7.8 at 298 K). Degradation of the peak corresponding to the cyclic peptide was monitored by integration of the peak area, monitored at 214 nm, after separation by HPLC. This was performed using a Hewlett Packard HP1100 binary pump system equipped with a Micra NPS RP18 reverse-phase column (4.6 × 33 mm, 1.5 μm particle size) with 0.3% trifluoroacetic acid and 90% acetonitrile as solvents.

Structural comparison

Residues spanning the P_3-P_2' segment and the catalytic triads of the complexed structures trypsin-BBI reactive-site loop (Li *et al*., 1994), chymotrypsin-eglin c (Frigerio *et al*., 1992) and chymotrypsin-turkey ovomucoid third domain (Fujinaga *et al*., 1987), PDB entries 1smf, 1cho and 1acb, respectively, were selected for a backbone superimposition using the Insight II software (Biosym) for molecular graphics. The r.m.s. deviations for the 1smf-1acb, 1smf-1cho, 1acb-1cho superimpositions are 0.31, 0.44 and 0.26 Å, respectively, with the resulting structures illustrated in Figure 3.

Acknowledgements

This work was supported by the BBSRC and Glaxo-Wellcome Research. We thank Norman Gray of Glaxo-Wellcome for his assistance with hydrolysis assays, Malcolm Weir and Rob Cooke of Glaxo-Wellcome for their encouragement, Dr Wolfram Bode for providing co-ordinates of the BBI-trypsin complex, and Steve Matthews for helpful discussions.

References

Asbóth, B. & Polgár, L. (1983). Transition-state stabilization at the oxyanion binding sites of serine and thiol proteinases: hydrolysis of thiono and oxygen esters. *Biochemistry*, **22**, 117–122.

Atherton, E. & Sheppard, R. (1989). *Solid Phase Peptide Synthesis, A Practical Approach*, IRL Press, Oxford.

Bender, M. L., Begué-Canton, M. L., Blakley, R. L., Brobacher, L. J., Feder, J., Gunter, C. R., Kézdy, F. J., Killheffer, J. V., Jr., Marshall, T. H., Miller, C. G., Roeske, R. W. & Stoops, J. K. (1966). The determination of the concentrations of hydrolytic enzyme solutions: α-chymotrypsin, trypsin, papain, elastase, subtilisin and acetylcholinesterase. *J. Am. Chem. Soc.* **88**, 5890–5913.

Blow, D. M., Wright, C. S., Kukla, D., Ruehlmann, A., Steigemann, W. & Huber, R. (1972). Model for the association of bovine pancreatic trypsin inhibitor with chymotrypsin and trypsin. *J. Mol. Biol.* **69**, 137–144.

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

Cai, M., Huang, Y., Prakash, O., Wen, L., Dunkelarger, S. P., Huang, J.-K., Liu, J. & Krishnamoorthi, R. (1996). Differential modulation of binding loop flexibility and stability by Arg^{50} and Arg^{52} in *Cucurbitia maxima* trypsin inhibitor-V deduced by trypsin-

catalyzed hydrolysis and NMR spectroscopy. *Biochemistry,* **35**, 4784–4794.

Chen, P., Rose, J., Loves, R., Wei, C. H. & Wang, B. (1992). Reactive sites of an anticarcinogenic Bowman-Birk proteinase inhibitor are similar to other trypsin inhibitors. *J. Biol. Chem.* **267**, 1990–1994.

Deslongchamps, P. (1983). *Stereoelectronic Effects in Organic Chemistry*, Pergamon Press Ltd., Oxford.

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. Pept. Protein Res.* **46**, 79–87.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry,* **6**, 1948–1954.

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

Favel, A., Le-Nguyen, D., Coletti-Previero, M. A. & Castro, B. (1989). Active site chemical mutagenesis of *Ecballium elaterium* trypsin inhibitor II: new microproteins inhibiting elastase and chymotrypsin. *Biochem. Biophys. Res. Commun.* **162**, 79–82.

Fersht, A. R. (1974). Catalysis, binding and enzyme-substrate complementarity. *Proc. Roy. Soc. ser. B,* **187**, 397–407.

Frigerio, F., Coad, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P. & Bolognesi, M. (1992). Crystal and molecular structure of the bovine α-chymotrypsin-eglin *c* complex at 2.0 Å resolution. *J. Mol. Biol.* **225**, 107–123.

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr. & James, M. N. G. (1987). Crystal and molecular structures of the complex of α-chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution. *J. Mol. Biol.* **195**, 397–418.

Furka, A., Sebestyen, F., Asgedom, M. & Dibo, G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493.

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

Gronenstein, D. G. & Taira, K. (1984). Stereoelectronic control in peptide bond formation. *Ab initio* calculations and speculations on the mechanism of action of serine proteases. *Biophys. J.* **46**, 749–761.

Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S. & Grütter, M. G. (1991). Refined crystal structures of subtilisin Novo in complex with wild-type and two mutant eglins. *J. Mol. Biol.* **217**, 353–371.

Heinz, D. W., Hyberts, S. G., Peng, J. W., Priestle, J. P., Wagner, G. & Grütter, M. G. (1992). Changing the inhibitory specificity and function of the proteinase inhibitor eglin *c* by site-directed mutagenesis: functional and structural investigation. *Biochemistry,* **31**, 8755–8766.

Ikenaka, T. & Norioka, S. (1986). Bowman-Birk family serine proteinase inhibitors. In *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G., eds), pp. 361–374, Elsevier Science Publishers, Amsterdam.

Kimura, M., Kouzuma, Y. & Yamasaki, N. (1994). On a Bowman-Birk family proteinase inhibitor from *Erythrina variegata* seeds. *J. Biochem.* **115**, 369–372.

Knorr, R., Trzeciak, A., Bornwarth, W. & Gillessen, D. (1989). New coupling reagents in peptide chemistry. *Tetrahedr. Letters,* **30**, 1927–1930.

Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, J. J. (1991). A new type of synthetic library for identifying ligand-binding activity. *Nature,* **354**, 82–84.

Larionova, N. I., Gladysheva, I. P., Tikhonova, T. V. & Kazanskaya, N. F. (1993). Inhibition of cathepsin G and human granulocyte elastase by multiple forms of Bowman-Birk type soybean inhibitor. *Biochemistry (Moscow),* **58**, 1046–1052.

Laskowski, M., Jr (1980). An algorithmic approach to sequence → reactivity of proteins. Specificity of protein inhibitors of serine proteases. *Biochem. Pharmacol.* **29**, 2089–2094.

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

Leatherbarrow, R. J. (1990). Using linear and non-linear regression to fit biochemical data. *Trends Biochem. Sci.* **15**, 455–458.

Leatherbarrow, R. J. (1992). *GraFit (Version 3.0)*, Erithacus Software Ltd., UK, Staines.

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. (Tokyo),* **116**, 18–25.

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.

Longstaff, C., Campbell, A. F. & Fersht, A. R. (1990). Recombinant chymotrypsin inhibitor 2: expression, kinetic analysis of inhibition with α-chymotrypsin and wild-type and mutant subtilisin BPN′, and protein engineering to investigate inhibitory specificity and mechanism. *Biochemistry,* **29**, 7339–7347.

Lu, W., Apostol, I., Qasim, M. A., Warne, N., Wynn, R., Zhang, W. L., Anderson, S., Chiang, Y. W., Ogin, E., Rothberg, I., Ryan, K. & Laskowski, M., Jr (1997). Binding of amino acid side-chains to S_1 cavities of serine proteinases. *J. Mol. Biol.* **266**, 441–461.

Maeder, D. L., Sunde, M. & Botes, D. P. (1992). Design and inhibitory properties of synthetic Bowman-Birk loops. *Int. J. Pept. Protein Res.* **40**, 97–102.

Markland, W., Ley, A. C., Lee, S. W. & Ladner, R. C. (1996a). Iterative optimization of high-affinity protease inhibitors using phage display. 1. Plasmin. *Biochemistry,* **35**, 8045–8057.

Markland, W., Ley, A. C. & Ladner, R. C. (1996b). Iterative optimization of high-affinity protease inhibitors using phage display. 2. Plasma kallikrein and thrombin. *Biochemistry,* **35**, 8058–8067.

McBride, J. D., Freeman, N., Domingo, G. J. & Leatherbarrow, R. J. (1996). Selection of chymotrypsin inhibitors from a conformationally constrained combinatorial peptide library. *J. Mol. Biol.* **259**, 819–827.

Morhy, L. & Ventura, M. M. (1987). The complete amino acid sequence of the *Vigna unguiculata* (L) Walp. seed, trypsin and chymotrypsin inhibitor. *An. Acad. Brasil Ci.* **59**, 71–81.

Neil, G. L., Niemann, C. & Hein, G. E. (1966). Structural specificity of α-chymotrypsin: polypeptide substrates. *Nature,* **210**, 903–907.

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

Odani, S. & Ikenaka, T. (1973). Scission of soybean Bowman-Birk proteinase inhibitor into two small fragments having either trypsin or chymotrypsin inhibitory activity. *J. Biochem.* **74**, 857–860.

Pavone, V., Isernia, C., Saviano, M., Falcigno, L., Lombardi, A., Paolillo, L., Pedone, C., Buøen, S., Naess, H. M., Revheim, H. & Eriksen, J. A. (1994). Conformational studies on peptides as enzyme inhibitors: chymotrypsin inhibitors using Bowman-Birk type as models. *J. Chem. Soc. Perkin Trans. 2,* 1047–1053.

Read, R. & James, M. N. G. (1986). Introduction to the proteinase inhibitors: X-ray crystallography. *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G. S., eds), pp. 301–335, Elsevier Science Publishers, Amsterdam.

Roberts, B. L., Markland, W., Ley, A. C., Kent, R. B., White, D. W., Gutterman, S. K. & Ladner, R. C. (1992). Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. *Proc. Natl Acad. Sci. USA,* **89**, 2429–2433.

Schellenberger, V., Braune, K., Hofmann, H.-J. & Jakubke, H.-D. (1991). The specificity of chymotrypsin. *Eur. J. Biochem.* **199**, 623–636.

Segal, D. M., Powers, J. C., Cohen, G. H., Davies, D. R. & Wilcox, P. E. (1971). Substrate binding site in bovine chymotrypsin A_γ. A crystallographic study using peptide cholormethyl ketones as site-specific inhibitors. *Biochemistry,* **10**, 3728–3738.

Suzuki, A., Yamane, T., Ashida, T., Norioka, S., Hara, S. & Ikenazka, 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.

Tam, J. P., Wu, C. R., Liu, W. & Zhang, J. W. (1991). Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. *J. Am. Chem. Soc.* **113**, 6657–6662.

Tanaka, A. S., Sampaio, M. U., Marangoni, S., de Oliveira, B., Novello, J. C., Oliva, M. L. V., Fink, E. & Sampaio, C. A. M. (1997). Purification and primary structure determination of a Bowman-Birk trypsin inhibitor from *Torresea cearensis* seeds. *Biol. Chem.* **378**, 273–281.

Terada, S., Sato, K., Kato, T. & Izumiza, N. (1978). Inhibitory properties of nonapeptide loop structures related to reactive sites of soybean Bowman-Birk inhibitor. *FEBS Letters,* **90**, 89–92.

Voss, R.-H., Ermler, U., Essen, L.-O., Wenzl, G., Kim, M.-Y. & Flecker, P. (1996). Crystal structure of the bifunctional soybean Bowman-Birk inhibitor at 0.28 nm resolution. Structural peculiarities in a folded conformation. *Eur. J. Biochem.* **242**, 122–131.

Wenzel, H. R. & Tschesche, H. (1995). Reversible inhibitors of serine proteinases. In *Peptides, Synthesis, Structures, and Applications* (Gutte, B., ed.), pp. 321–362, Academic Press, San Diego, CA.

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

Werner, M. H. & Wemmer, D. E. (1992b). Identification of a protein-binding surface by differential amide hydrogen-exchange measurements. *J. Mol. Biol.* **225**, 873–889.

Zhang, R. G., Lin, G. D., Yan, Y. W., Tang, W. Z., Tan, F. L., Chi, C. W. & Tsao, T. C. (1985). The crystallographic study of the complex of the Lys active fragment of mung bean inhibitor with trypsin. *Sci. Sin.* **28**, 1163–1166.

Edited by A. R. Fersht

(Received 9 February 1998; received in revised form 16 June 1998; accepted 30 June 1998)