Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 20, Issue Number 5, (2003)

The Structural Basis of a Conserved P2 Threonine in Canonical Serine Proteinase Inhibitors

http://www.jbsdonline.com

Arnd B. E. Brauer[1]
Marco Nievo[1]
Jeffrey D. McBride[2]
Robin J. Leatherbarrow[1]*

[1]Department of Chemistry
Imperial College London
South Kensington Campus
London SW7 2AZ, UK
[2]Department of Immunology and
Molecular Pathology
Windeyer Institute of Medical Sciences
University College London
46 Cleveland Street
London W1T 4JF, UK

Abstract

Bowman-Birk inhibitors (BBIs) are a well-studied family of canonical inhibitor proteins of serine proteinases. In nature, the active region of BBIs possesses a highly conserved Thr at the P2 position. The importance of this residue has been reemphasized by synthetic BBI reactive site loop proteinomimetics. In particular, this residue was exclusively identified for active chymotrypsin inhibitors selected from a BBI template-assisted combinatorial peptide library. A further kinetic analysis of 26 P2 variant peptides revealed that Thr provides both optimal binding affinity and optimal resistance against enzymatic turnover by chymotrypsin.

Herein, we report the ^{1}H-NMR spectroscopic study of a 5-membered sub-set of these reactive site loop peptides representing a stepwise elimination of the Thr side-chain functionalities and inversion of its side-chain chirality. The P2 Thr variant adopts a three-dimensional structure that closely mimics the one of the corresponding region of the complete protein. This validates the use of this template for the investigation of structure-function relationships. While the overall backbone geometry is similar in all studied variants, conformational changes induced by the modification of the P2 side chain have now been identified and provide a rational explanation of the kinetically observed functional differences. Eliminating the γ-methyl group has little structural effect, whereas the elimination of the γ-oxygen atom or the inversion of the side-chain chirality results in characteristic changes to the intramolecular hydrogen bond network. We conclude that the transannular hydrogen bond between the P2 Thr side-chain hydroxyl and the P5´ backbone amide is an important conformational constraint and directs the hydrophobic contact of the P2 Thr side chain with the enzyme surface in a functionally optimal geometry, both in the proteinomimetic and the native protein.

In at least four canonical inhibitor protein families similar structural arrangements for a conserved P2 Thr have been observed, which suggests an analogous functional role. Substitutions at P2 of the proteinomimetic also affect the conformational balance between *cis* and *trans* isomers at a distant Pro-Pro motif (P3´-P4´). Presented with a mixture of *cis*/*trans* isomers chymotrypsin appears to interact preferably with the conformer that retains the *cis*-P3´ Pro-*trans*-P4´ Pro geometry found in the parent BBI protein.

Key words: protein mimetic, β-hairpin peptide, peptide NMR, Bowman-Birk Inhibitor.

Introduction

Bowman-Birk inhibitor (BBI) proteins are a well-studied family of proteins, which are typically comprised of 60 to 90 residues in a pseudo-symmetric arrangement that supports two independent reactive site loops exposed at the extreme ends of the molecule (1). These loops are able to inhibit serine proteinases because within their β-hairpin structure they encompass the 'canonical' backbone conformation, which is complementary to the enzyme surface and widely found in small serine proteinase inhibitor proteins (2). It is a particular feature of the BBI reactive site loop that much of the biological potency and structural integrity is retained following excision from

Phone: +44 207 594 5752
Fax: +44 207 5944 5880
Email: R.Leatherbarrow@imperial.ac.uk

the remainder of the protein (3, 4). Short synthetic peptides have, therefore, employed this loop sequence as a proteinomimetic template to generate inhibitors for a range of serine proteinases (3, 4). This has been studied in some detail in an attempt to understand the unusual sequence-inherent properties of this template. The presence of the canonical backbone geometry is the basis for its versatile inhibitory activity (5, 6). Covalent cyclization and a conserved *cis* peptide bond are key elements for retaining the inhibitory structure (7, 8). In common with natural proteinaceous inhibitors, the identity of the primary specificity or P1 residue (Schechter and Berger nomenclature (9)) is largely responsible for the directed inhibitory activity against a given serine proteinase (3, 4). A particular feature of this residue is that it is hyper-exposed to allow interaction with its target (2). One study has focused on the role of the adjacent P2 position (10). At this position Thr is highly conserved in BBI proteins (≈83% of the protein sequences currently listed in the SWISS-PROT database (11); for a tabulation of sequences see (10)). Furthermore, Thr was exclusively found at the P2 position in active chymotrypsin inhibitors identified from a combinatorial peptide library based on this template (where randomization was performed at P2, P1, and P2´ (12)). The finding was confirmed by the kinetic analysis of a further series of 26 of such proteomimetic peptides, varied exclusively at P2. This revealed that Thr provides both optimal binding affinity and optimal resistance against enzymatic turnover (10). In the investigation reported herein, we have selected a sub-set of these P2-varied peptides for the assessment of their structural properties in solution by ^{1}H-NMR spectroscopy. This has allowed rationalization of the kinetically observed properties in structural terms at atomic level, providing new insights into the design of canonical serine proteinase inhibitors.

Materials and Methods

Peptides were synthesised as previously described (10). Bovine pancreatic chymotrypsin treated with tosyl-lysine chloromethyl ketone was purchased from Sigma-Aldrich. The ^{1}H-NMR analysis was performed in aqueous solution (90% H_2O/10% ^2H_2O and 100% ^2H_2O; 100 mM phosphate buffer for experiments with added enzyme) at a pH* of 3.8. Chemical shifts were referenced to 3-(trimethylsilyl)-1-propane sulfonic acid. DQF-COSY (13), TOCSY (mixing time 80 ms), (14) NOESY (mixing time 300 ms) (15) and ROESY (mixing times of 200 ms and 300 ms) (16) experiments were recorded on a Bruker AMX 600 spectrometer, and processed and analysed by using X-WinNMR and Aurelia software packages on Silicon Graphics work stations. Following sequential assignment (17), amide temperature coefficients (18), diastereotopic proton assignment and χ^1 conformations (19), $^3J_{HNH\alpha}$ coupling constants (20), and $^3J_{H\alpha H\beta}$ coupling constants (21, 22) were analysed as described. All coupling constants were derived from one-dimensional spectra.

Model building and refinement were performed with the software package TINKER (23) using the AMBER force field. The program DISTGEOM was used to refine model structures against the NMR-derived restraints with the distance geometry and simulated annealing protocol that has been described in detail in the original publication (23). According to the observed cross-peak intensities in NOESY and ROESY spectra, interproton distances restraints were classified into three ranges and implemented with these limits: 1.86-2.50 Å, 1.86-3.30 Å, and 1.86-5.00 Å. For methylene protons that were not diasterotopically assigned the upper bound was increased by 2 Å and the force constant was halved (for methyl protons the force constant was reduced to one-third) (23). For aromatic ring protons, the upper bound was increased by 5 Å and the force constant was halved. The peptide bonds were restrained to a *cis* conformation for Ile6-Pro7 and to *trans* for all other peptide bonds. ϕ-Restraints were implemented in three ranges: -150° to -90°, -75° to -55°, and -180° to -60° for residues with observed $^3J_{HNH\alpha}$ coupling constants ≥ 9.0 Hz, for Pro residues, and for the remaining residues, respectively. Experimentally determined preferred χ^1 rotamers were implemented with bounds of ±30°. Selected hydrogen bonds were restrained in the calculations by implementation of a pair of

interatom distances (O-N 2.5-3.5 Å; O-H 1.86-2.5Å). Structures were visualised and analysed with the programs Swiss PDB Viewer (24), WebLab (Molecular Simulations Inc.), MOLMOL (25), and PROCHECK (26).

Results

Design of Proteinomimetic Peptide Variants

In order to assess the individual contributions of all the features of the P2 Thr side chain, a series of proteinomimetic 11-residue disulfide-cyclized peptides, varied exclusively at P2, were designed and synthesised. The stepwise variation of the P2 side chain is schematically shown in Figure 1, and the biological activities of the corresponding variants are summarised in Table I.

Table I

Sequences and biological activities (10) of the disulfide-cyclized peptides of this study, which are varied only at the P2 position (Schechter and Berger nomenclature (9); Abu, α-aminobutyric acid; n.d., not determined). Different aspects of the biological activity are described by the inhibition constant (K_i), K_i-derived binding free energy relative to the P2 Ala variant ($\Delta\Delta G_{relative\ to\ Ala}$), and the relative rate of hydrolysis by chymotrypsin. These data show that binding affinity and resistance against hydrolysis are not correlated.

P4	P3	P2	P1	P1′	P2′	P3′	P4′	P5′	P6′	P7′	K_i in nM	$\Delta\Delta G_{relativ\ to\ Ala}$ in kcal/mol	rate of hydrolysis relative to Ala
S	C	**T**	F	S	I	P	P	Q	C	Y	19	–2.0	0.002
S	C	**S**	F	S	I	P	P	Q	C	Y	400	–0.2	0.038
S	C	**Abu**	F	S	I	P	P	Q	C	Y	130	–0.9	0.34
S	C	**A**	F	S	I	P	P	Q	C	Y	570		1
S	C	***allo*-T**	F	S	I	P	P	Q	C	Y	130 000	+3.2	n.d.

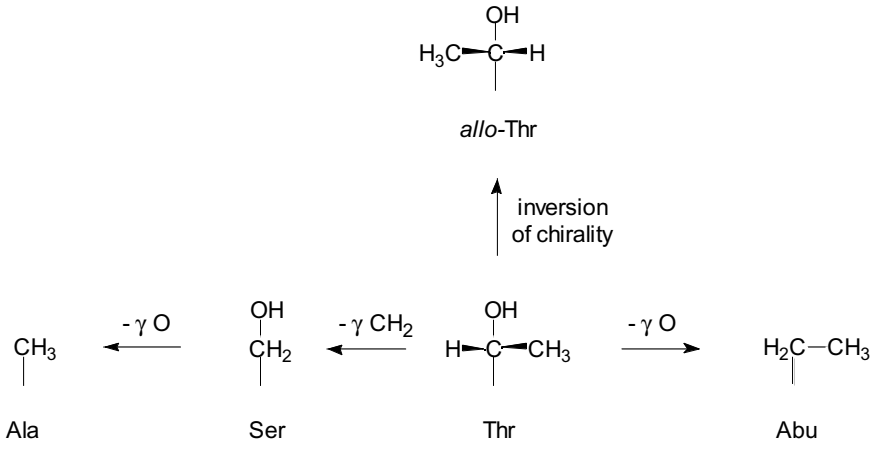


Figure 1: Stepwise changes of the P2 side chain in the variant peptides of this study.

NMR Analysis of the Two Dominant Conformers

The particularly informative amide region of the one-dimensional ^{1}H-NMR spectra of the P2 variant peptides of this study is shown in Figure 2. In all the spectra two sets of amide resonances can be distinguished by their intensities. These originate from two distinct conformers, all of which could be fully assigned with the exception of the second conformer of the P2 Thr variant due to low signal intensity. The relative signal intensities as estimated from integration indicate that the population of the second conformer increases at the expense of the dominant conformer in the order of Thr (≈21 %) ≤ Ser ≤ Abu ≤ Ala ≤ *allo*-Thr (≈34 %).

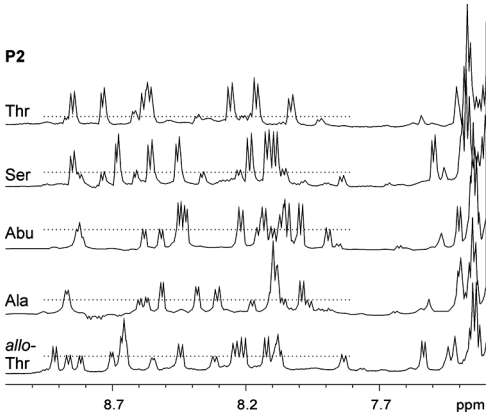


Figure 2: Amide region of one-dimensional 600 MHz ^{1}H-NMR spectra of the P2 variant peptides in aqueous solution at pH* 3.8 and 298 K. Horizontal dotted lines indicate the approximate signal intensity of the conformer with the second highest population.

The NMR parameters that characterise the backbone structures of the assigned conformers are summarised in Figure 3. The position of a characteristic *cis* peptide bond (for which an $H\alpha_i$-$H\alpha_{i+1}$ NOE is diagnostic (17)) allows a systematic classification. In the dominant conformers of all variants the *cis* peptide bond is located at the P3´ Pro (Figure 3a) as observed in all available structures of complete BBI proteins. In the second most populated conformers, the location of the *cis* peptide bond is consistently shifted by one residue towards the C-terminus, which results in a *cis* geometry at P4´ Pro (Figure 3b).

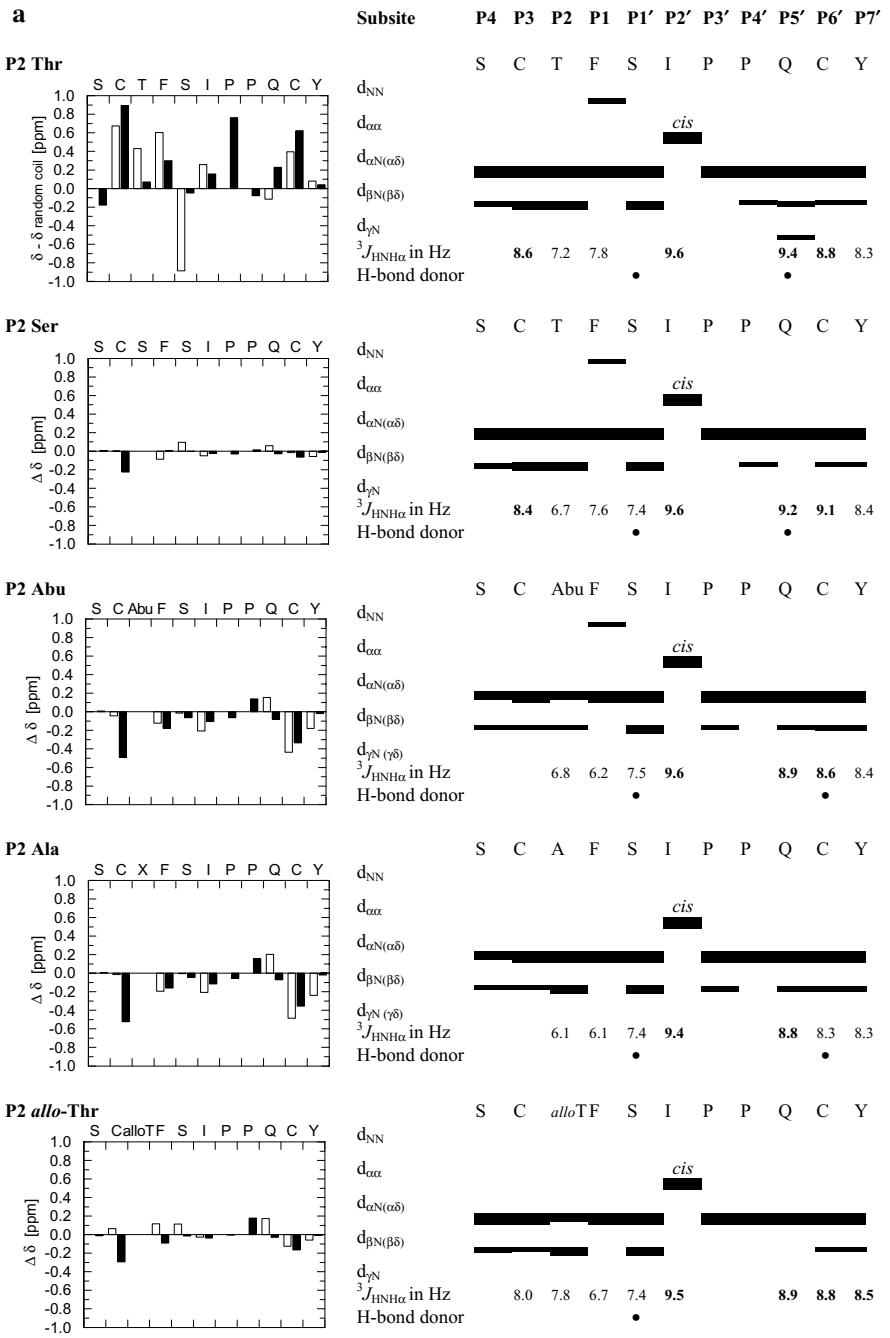
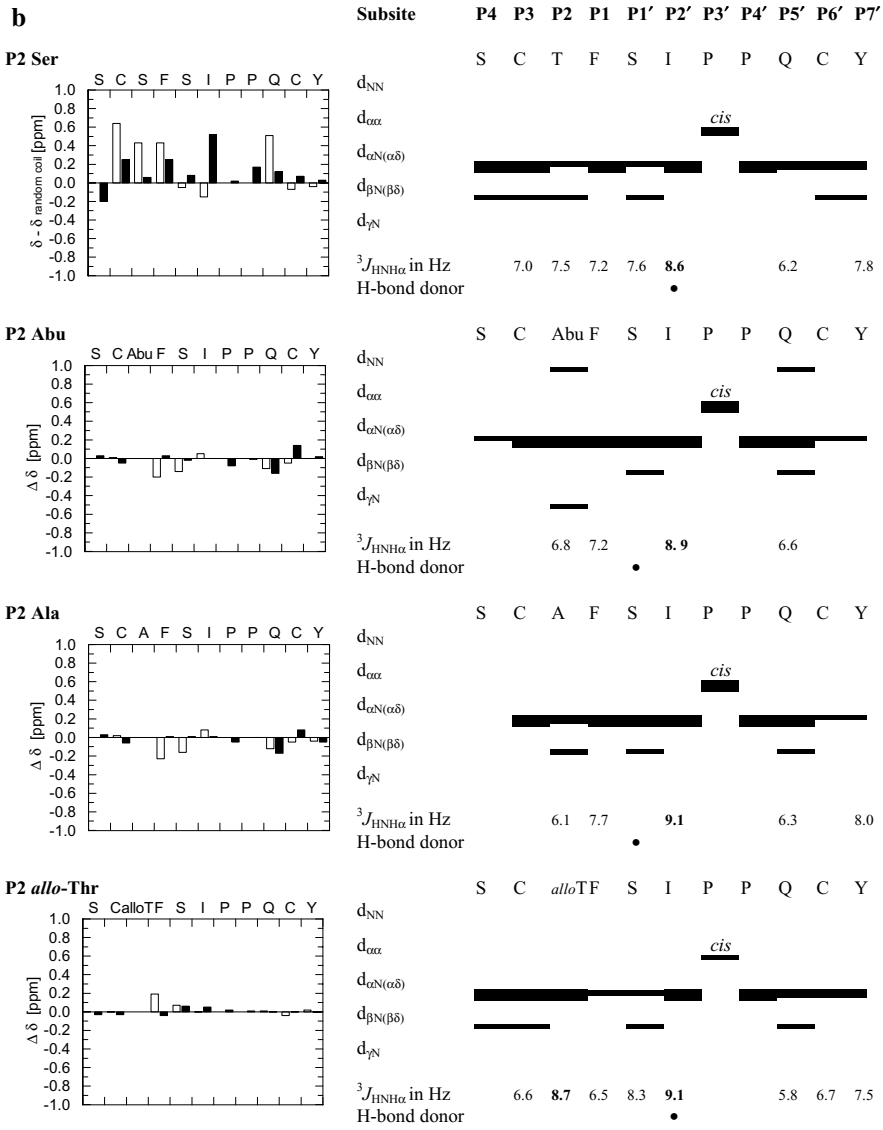


Figure 3: NMR parameters characterising the backbones of the two dominant conformers of the P2 variant peptides. The first panel of each sub-figure presents the chemical shifts of backbone HN (□) and Hα (■) protons as the deviation from random coil values (δ - δ random coil (39)) for the first variant, and as the deviation from the values of the first peptide (Δδ) for the following variants. The Δδ values in (a) are the deviation from the values of the P2 Thr variant. As the second conformer of the P2 Thr variant could not be fully assigned due to its low population, the second conformer of P2 Ser was used instead to derive the Δδ patterns shown in (b). The second panel of each sub-figure summarises sequential interproton NOE connectivities (relative intensities are represented by the thickness of the lines), $^3J_{HNH\alpha}$ coupling constants (values > 8.5 Hz characteristic of extended or β-strand conformation are highlighted in bold; (40)), and backbone amide temperature coefficient indicative of hydrogen bonding (• (18)). A comparison of the individual patterns indicates that the presence or absence of a γ-oxygen atom in the P2 side chain results in modified conformational preferences. The P2 Thr and the P2 Ser variant exhibit one type of pattern, the P2 Abu and the P2 Ala variants another, and the pattern of the P2 *allo*-Thr variant combines features of both. This classification applies to both sets of conformers, even though the detection of NOEs in the lower populated conformers was hampered by weak signal intensity.

The dominance of Hα chemical shifts with a downfield deviation from random coil values, the patterns of sequential NOE connectivities which are dominated by $H\alpha_i$-HN_{i+1} NOEs with only few HN_i-HN_{i+1} NOEs, and several $^3J_{HNH\alpha}$ coupling constants larger than 8 Hz (in the *cis*-P3´ Pro conformers only) all indicate a prevalence of β-type structures in all the assigned conformers. The deviations from random coil values of both chemical shifts and $^3J_{HNH\alpha}$ coupling constants are more pronounced in the *cis*-P3' conformers and indicate a higher degree of structural integrity.

The patterns of NMR parameters are systematically affected by the presence or absence of a side-chain hydroxyl group and the side-chain chirality, but not by the side-chain methyl group (Figure 3). The dominant conformers of both the P2 Thr and the P2 Ser variants exhibit almost identical patterns, which include the preferential population of the +60° χ^1 rotamer by the P2 side chain. This indicates that the P2 side-chain methyl group is structurally dispensable. For both sets of conformers, the elimination of the side-chain hydroxyl (P2 Thr → Abu and P2 Ser → Ala) induces significant changes of the chemical shift and of the hydrogen bond patterns. The inversion of the P2 Thr side-



chain chirality gives rise to a third pattern of NMR parameters that appears to combine features seen in variants with and without a side-chain hydroxyl group.

Isomer-selective Hydrolytic Pathway

For the anti-tryptic inhibitor peptide from which the anti-chymotryptic variants of this study were originally derived, we have previously demonstrated preferential binding and consecutive hydrolysis of the native-like *cis*-P3′ Pro conformer (8). For the peptides of this study we applied the same type of analysis, which attempts to monitor the action of added proteinase directly by a timed series of one-dimensional NMR spectra. An example for the P2 Abu variant is shown in Figure 4. The observed timed spectra for the Abu, Ala and Ser P2 variants (data not shown for the latter two) as well as for a further P2 Thr variant peptide, which incorporates a strategic ^{19}F NMR label at P1 (unpublished), are all consistent with a preferable interaction of chymotrypsin with the native-like *cis*-P3′ Pro conformer. This implies a hydrolytic pathway in which the *cis*-P4′ Pro conformer isomerises into the *cis*-P3′ Pro conformer prior to the interaction with chymotrypsin.

Three-dimensional Structures

As the elimination of the P2 side-chain hydroxyl group appears to exert the strongest structural effect, this was investigated in more detail by focusing on the P2 Thr and

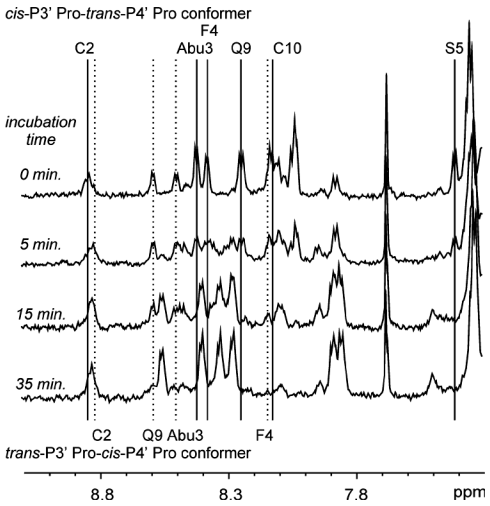


Figure 4: Timed series of one-dimensional ^{1}H-NMR spectra monitoring the hydrolysis of the P2 Abu variant by a catalytic quantity of chymotrypsin in 100 mM phosphate buffer at pH* 3.8 and 298 K. Selected well-resolved resonances of the two most highly populated conformers are labelled. Preferential hydrolysis of the *cis*-P3′ Pro-*trans*-P4′ Pro conformer is most apparent in the spectrum at 15 minutes incubation time: Resonances of the *cis*-P3′ Pro-*trans*-P4′ Pro conformer (solid lines) are at the detection limit, while resonances of the *trans*-P3′ Pro-*cis*-P4′ Pro conformer (dotted lines) remain of substantial intensity. The two sets of resonances of the intact starting material 'collapse' into a single set of resonances of hydrolysed material (these dominate the spectrum at 35 minutes incubation time) as a result of enzymatic turnover.

the P2 Abu variants. For the dominant conformers of these two variants three-dimensional structures consistent with the combined NMR information were calculated. The resulting families of final structures are compared in Figure 5 and their statistics are summarised in Table II. For both variants structures were initially calculated without hydrogen-bond restraints. While the resulting structures for the P2 Thr variant converged to a single family (Figure 5a), several significantly different structures were found to be consistent with the NMR data of the P2 Abu variant. This is presumably due to the detection of fewer long-range NOE contacts for the P2 Abu variant. The sub-families of this variant were inspected for geometrically favourable hydrogen-bond acceptors for the two experimentally determined hydrogen bond donors (Figure 3a). The potential hydrogen bonds judged to be most compatible with these donors were implemented as additional restraints in further calculations. The resulting final family of P2 Abu variants exhibited a similar level of convergence as the family of P2 Thr variant structures (Table II; Figure 5b).

Table II

Statistics of the families of the 20 lowest-penalty function distance geometry/simulated annealing structures calculated from NMR data for the dominant conformers of the P2 Thr and P2 Abu variants. Two backbone-to-backbone hydrogen bonds (O P2 Abu–HN P1' Ser and O P4' Pro–HN P7' Cys) were introduced as restraints into the calculations for the P2 Abu variant after these were identified as particularly likely in initial structures that had been calculated without hydrogen bond restraints. The Ramachandran analysis was performed with the program PROCHECK (26).

	P2 Thr variant	P2 Abu variant
Distance restraints		
sequential	18	20
$i + 2$	2	1
$i + 3$	1	1
$i + 6$	5	1
$i + 8$	3	1
Dihedral restraints		
ω	10	10
ϕ	10	10
χ^1	6	3
Hydrogen bond restraints	0	2
Combined penalty function value	0.0078 ± 0.0012	0.0070 ± 0.0010
Ramachandran analysis		
Residues in favoured regions	63.4 %	49.3 %
Residues in additional allowed regions	36.4 %	50.7 %
Residues in generously allowed regions	0 %	0 %
Pairwise root mean square deviations (Å)		
All residues over backbone atoms	0.83 ± 0.21	0.88 ± 0.26
All residues over heavy atoms	1.69 ± 0.27	1.91 ± 0.33
P3 to P6' residues over backbone atoms	0.58 ± 0.15	0.61 ± 0.19
P3 to P6' residues over heavy atoms	1.43 ± 0.29	1.34 ± 0.24

The backbone structures of both variants resemble that of the reactive site loop of intact BBI proteins (Figure 5, Table III). The changes induced by the selective elimination of the P2 side-chain hydroxyl, however, result in significant backbone deviations in the P2 Abu variant. These, in particular, are located in the disulfide-bridged region of the molecule, as consistently indicated by the chemical shifts, the amide temperature coefficients (Figure 3a) and by the calculated structures (Figure 5b). Both variants, however, successfully mimic the P2 to P2´ backbone region that primarily interacts with the enzyme (with root mean square deviations of approximately 0.7 Å or less; Figure 5). This importantly includes a P1 residue that is hyper-exposed for the primary interaction with the enzyme.

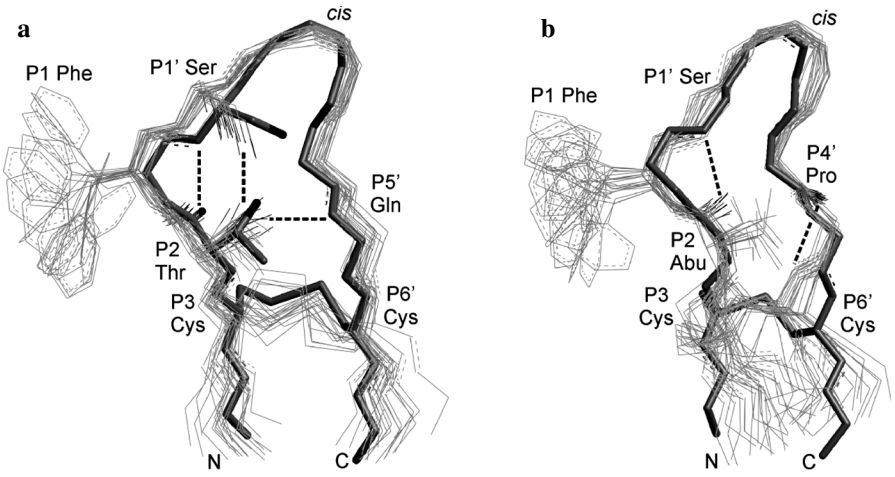


Figure 5: Families of the 20 lowest-penalty function distance geometry/simulated annealing structures (shown as lines) that are consistent with the observed NMR data for the dominant conformers of the P2 Thr (a) and P2 Abu variants (b) are superimposed onto the anti-tryptic reactive site loop of garden pea BBI protein (X-ray crystal structure shown in sticks; this loop was selected for its particular high sequence homology with the peptides; PDB code 1pbi (35)). Only selected side chains are shown for clarity and two slightly different perspectives are chosen to highlight similarities and differences. N- and C-terminus, *cis* peptide bond, selected oxygen atoms (highlighted in black), and selected potential hydrogen bonds (black dashed lines) are indicated. The root mean square deviations over the backbone of the families of peptide structures and crystal structure are 0.68 (± 0.12) Å and 1.10 (± 0.23) Å over the covalently closed loop (P2 to P6´), and 0.47 (± 0.08) Å and 0.69 (± 0.12) Å over the primary contact region (P2 to P2´) for the P2 Thr and P2 Abu variant, respectively.

Discussion

Protein Mimicry

The finding that the 11-residue disulfide-cyclized peptides successfully mimic the reactive site loop structure as present in complete BBI proteins explains the observed biological activity. As the protein reactive site loops directed at chymotrypsin and trypsin are very similar in structure, the peptides mimic the backbone geometry of both types almost equally well (Table III). We have previously reported similar structures for related anti-tryptic and anti-elastase peptides, which have a hydrophilic Lys or a small Ala at the hyper-exposed P1 position (5, 6). It is an important finding of this study that the sequence-inherent stability of this peptide scaffold is sufficient for the presentation of a very hydrophobic residue such as Phe at the hyper-exposed P1 position in aqueous solution (in canonical serine protease inhibitors over 90% of the P1 side chain surface area was found to be solvent accessible (27) with the P1 residue contributing up to half of all contacts with the enzyme (28)). This validates the use of these peptides for the investigation of structure-function relationships.

Although an element of conformational heterogeneity is present in all variants (which are >95% homogenous by chemical criteria; data not shown), the preferential interaction of chymotrypsin with the major, native-like *cis*-P3´ Pro conformer strongly suggests that this conformer dominates the biological activity. We therefore correlate the structure of this conformer with the experimentally observed biological activity.

Table III

Comparison of the sequences (the sequential number of the P1 residue is indicated) and structures of the covalently closed reactive site loops of complete BBI proteins and two P2 proteinomimetic peptide variants (root mean square deviation over the P3 to P6' backbone in Å). Values above the diagonal compare two anti-chymotyptic loops. Diagonal values compare the anti-tryptic and anti-chymotyptic loops of the same BBI protein. Values below the diagonal compare the anti-tryptic loop of a BBI protein with the anti-tryptic loop of another protein structure or with an anti-chymotryptic proteinomimetic loop.

Sequences										BBI proteins					Proteinomimetics		Reference
										PDB	1pbi free	1bbi free	1k9b free	1d6r in trypsin complex	P2 Thr free	P2 Abu free	
Subsite	P3	P2	P1	P1'	P2'	P3'	P4'	P5'	P6'								
Garden pea BBI protein (1pbi)											X-ray	NMR	X-ray	X-ray	NMR	NMR	
anti-tryptic loop	C	T	K_{16}	S	N	P	P	T	C								
anti-chymotryptic loop	C	A	Y_{43}	S	N	P	P	K	C	1pbi	0.49	0.65	0.58	0.49	0.55	1.15	(35)
Soybean BBI protein (1bbi, 1k9b, 1d6r)										1bbi	0.84	0.89	0.92	0.91	0.85	0.90	(36)
anti-tryptic loop	C	T	K_{16}	S	N	P	P	Q	C	1k9b	0.62	0.74	0.31	0.54	0.77	1.17	(37)
anti-chymotryptic loop	C	A	L_{43}	S	Y	P	A	Q	C	1d6r	0.41	0.85	0.46	0.39	0.55	1.25	(38)
Proteinomimetic peptides																	
P2 Thr variant	C	T	F_4	S	I	P	P	Q	C	This	0.60	0.96	0.72	0.65	-	1.20	
P2 Abu variant	C	Abu	F_4	S	I	P	P	Q	C	study	1.29	1.00	1.17	1.12	-	-	

The presence of a P2 side-chain hydroxyl in the P2 Thr and the P2 Ser variants results in the retention of the protein-like transannular hydrogen bond with the backbone amide of the P5' residue. This correlates with a high level of resistance against chymotryptic hydrolysis. The absence of such a side-chain hydroxyl in P2 Abu and the P2 Ala variants results in a changed hydrogen bond pattern, notably in the loss of the hydrogen bond donor characteristic of the P5' amide (Figure 3). As this is associated with accelerated chymotryptic hydrolysis (by over an order of magnitude; Table I), it provides experimental evidence for our earlier hypothesis that the P2-side-chain to P5'-backbone hydrogen bond enhances the level of resistance against enzymatic hydrolysis (10). This particular hydrogen bond appears to direct and stabilise a close contact of the aliphatic part of the P2 side-chain with the imidazole of the enzyme's catalytic His57 as also observed in the X-ray crystal structures of BBI - proteinase complexes ((10); Figure 6b). While this contact does not appear to be designed to displace the His57 imidazole from the position observed in uncomplexed enzymes, it seems plausible that this hydrogen-bond-stabilized contact is particularly suited to interfere with the action of the catalytic machinery, specifically with a possible movement of His57. In models of serine proteinase catalysis (29-31) a His57 movement was invoked to accommodate its role as a proton shuttle between the hydroxyl of Ser195 and the leaving group nitrogen during the acylation phase, and between the hydrolytic water and Ser195 during the deacylation phase. The very recent analyses of an acylated serine proteinase and its tetrahedral intermediate by high-resolution crystallography provide experimental evidence for a rotational movement of the His57 side chain about χ^2 in the course of catalysis (32-34).

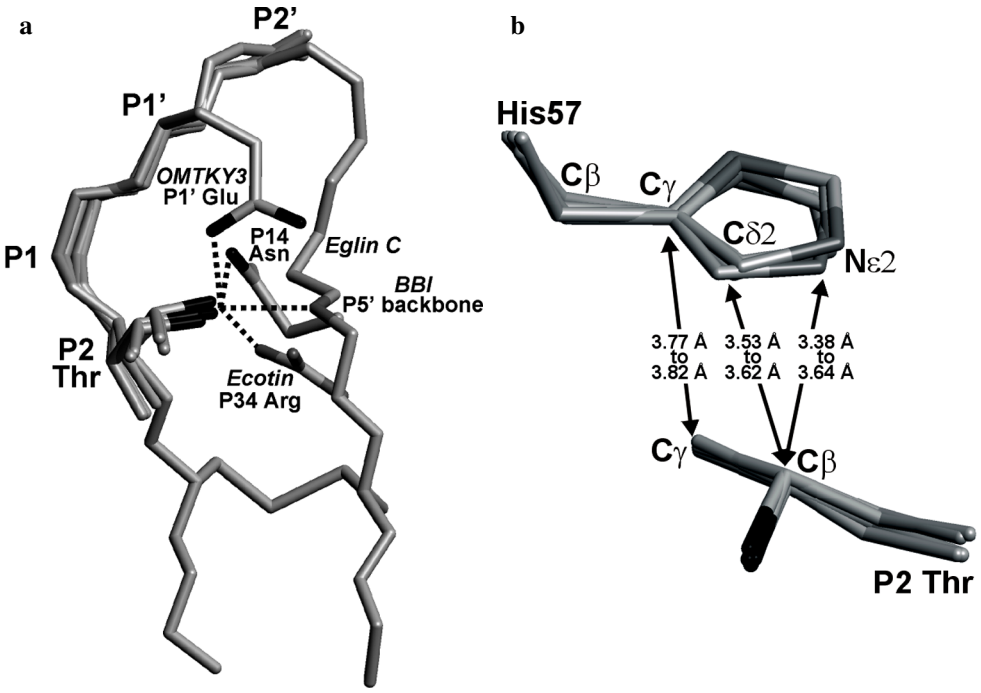


Figure 6: (a) The superimposition of the P2´ to P2 stretch of the reactive site loop of proteins that represent four different families of canonical serine proteinase inhibitors shows that the P2 Thr adopts almost identical arrangements (selected oxygen and nitrogen atoms are highlighted in black and dark grey, respectively). The hydrogen bond partners, however, are different, and for clarity only one is shown for each example. The figure is based on the X-ray crystal structures of mung bean BBI in complex with trypsin ((41); the BBI loop that corresponds to the studied peptides is shown; a second hydrogen bond branch that connects to the side chain of P1´ Ser has been omitted), ecotin (ecotin family) in complex with thrombin (1id5, (42); a second hydrogen bond branch that connects to the backbone of P33 Leu has been omitted), eglin c (potato 1 family) in complex with chymotrypsin (1acb, (43); a second hydrogen bond branch that connects to the side chain of P1´ Asp has been omitted), and turkey ovomucoid third domain (Kazal family) in complex with chymotrypsin (OMTKY3, 1cho (44)). (b) The hydrophobic part of the P2 Thr side chain is in close contact with the imidazole ring of the catalytic His57 of the cognate proteinase. The same structures as in (a) were used to prepare the figure and the observed ranges of distances are indicated for selected pairs of heavy atoms.

Proteinase Affinity

It has been noted from the binding studies of these P2 variants that the side-chain methyl group contributes significantly more to the binding free energy than the side-chain hydroxyl, and that the combination of both groups in a Thr results in a gain that is almost twice the sum of the individual contributions (Table I (10)). This synergistic effect can be rationalised by the directing force exerted by the intramolecular transannular hydrogen bond. Contributions to the synergistic effect, however, may not exclusively arise from the direct interaction of the P2 side-chain with the enzyme, but also from more distant interactions because the inhibitor's backbone geometry is modulated by the elimination of the P2 side-chain hydroxyl (Figure 5).

This contrasts with the situation in the P2 *allo*-Thr variant. All NMR parameters clearly indicate that the backbone of this variant does not differ significantly more from that of the P2 Thr peptide than any of the other studied variants. We, therefore, ascribe the energetic penalty of approximately 5 kcal/mol exerted by the inversion of the P2 side-chain chirality mainly to unfavourable interactions between this side chain and the enzyme. This particular inversion of chirality appears to be incompatible with both aspects of the dual role of the natural P2 Thr, namely to be complementary to the enzyme surface and to form an intramolecular hydrogen bond with the P5' backbone.

Implications for Canonical Proteinase Inhibitor Proteins

This study, in combination with the previous kinetic analysis of 26 anti-chymotryptic P2 variants (10), clearly demonstrates that Thr is an optimal residue with respect to both affinity and stability. It is therefore surprising that the rare examples of non-Thr residues (Ala, Ser, Asn, or Arg) at the P2 position in BBI protein sequences occur in loops that are referred to as anti-chymotryptic by virtue of their large hydrophobic P1 residues such as Leu, Phe and Tyr (two sequence examples are listed in Table III). While this may indicate that the peptide system does not completely mirror the situation in the complete protein, the alternative explanation seems equally reasonable. Bovine pancreatic chymotrypsin, which is commonly used in studies of this type, may not fully reflect the spectrum of chymotypsin-like proteinases that are the natural target of these reactive site loops. The deliberate introduction of an element of flexibility (e.g. by dispensing with the P2 to P5´ hydrogen bond or by mutating the P4´ Pro, the conformational restraining role of which has been demonstrated previously (8)) and a P1 residue compatible with a variety of active sites (such as Leu) may achieve a broad spectrum of specificity and compatibility rather than maximum inhibition of a selected proteinase.

A review of the sequences of canonical serine protease inhibitor proteins found that at least six families have a conserved or common Thr residue at the P2 position (27). For the representatives of four of these families X-ray crystal structures are available. These consistently exhibit the same geometrical pattern for the P2 Thr side chain: a χ^1 conformation close to +50° and intramolecular hydrogen bonding (Figure 6a). This shows that the dual functionality of the P2 Thr side chain is widely utilised in canonical inhibitors of serine proteinases and is not isolated to BBI proteins. On this basis it seems likely that in future crystal structures of P2 Thr bearing canonical serine proteinase inhibitors this residue will conform to the geometric pattern outlined here.

Conclusions

This study has demonstrated that this 11-residue disulfide-cyclized peptide is a useful proteinomimetic for the dissection of structure-function relationships of anti-chymotryptic BBI reactive site loops. Some of the smallest possible changes at molecular level, such as the elimination of a single side-chain atom or its inversion of chirality, resulted in significant effects on the three-dimensional structure, the conformational balance, and the biological activity. Clear relationships between these properties have been established and generated new insights into the parent protein and more generally into a large sub-set of canonical serine proteinase inhibitors. These relationships should also be useful for the design of proteinase inhibitors.

Supplementary Material

Tables with the assignments of the ^{1}H chemical shifts, backbone amide temperature coefficients, and coupling constants of the studied peptides are available from the authors.

Acknowledgements

The authors wish to thank Dr. H. Toms and P. Haycock of the University of London 600 MHz NMR facility at Queen Mary and Westfield College for their expert advice and assistance, Prof. W. Bode for sending the coordinates of trypsin-complexed mung bean BBI, and Prof. C. Schofield, Dr. S. Matthews, and Dr. R. Cooke for helpful discussions. We gratefully acknowledge the financial support of GlaxoSmithKline.

References and Footnotes

1. B. Prakash, S. Selvaraj, M. R. Murthy, Y. N. Sreerama, D. R. Rao, and L. R. Gowda, *J. Mol. Evol. 42*, 560-569 (1996).
2. W. Bode and R. Huber, *Eur. J. Biochem. 204*, 433-451 (1992).
3. J. D. McBride and R. J. Leatherbarrow, *Curr. Med. Chem. 8*, 909-917. (2001).
4. J. D. McBride, E. M. Watson, A. B. E. Brauer, A. M. Jaulent, and R. J. Leatherbarrow, *Biopolymers 66*, 79-92 (2002).
5. A. B. E. Brauer, G. Kelly, J. D. McBride, R. M. Cooke, S. J. Matthews, and R. J. Leatherbarrow, *J. Mol. Biol. 306*, 799-807 (2001).
6. A. B. E. Brauer, G. Kelly, S. J. Matthews, and R. J. Leatherbarrow, *J. Biomol. Struct. Dyn. 20*, 59-70. (2002).
7. G. J. Domingo, R. J. Leatherbarrow, N. Freeman, S. Patel, and M. Weir, *Int. J. Pept. Protein Res. 46*, 79-87 (1995).
8. A. B. E. Brauer, G. J. Domingo, R. M. Cooke, S. J. Matthews, and R. J. Leatherbarrow, *Biochemistry 41*, 10608-10615 (2002).
9. I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun. 27*, 157-162 (1967).
10. J. D. McBride, A. B. E. Brauer, M. Nievo, and R. J. Leatherbarrow, *J. Mol. Biol. 282*, 447-457 (1998).
11. C. O'Donovan, M. J. Martin, A. Gattiker, E. Gasteiger, A. Bairoch, and R. Apweiler, *Brief. Bioinform. 3*, 275-284 (2002).
12. J. D. McBride, N. Freeman, G. J. Domingo, and R. J. Leatherbarrow, *J. Mol. Biol. 259*, 819-827 (1996).
13. A. E. Derome and M. P. Williamson, *J. Magn. Reson. 88*, 177-185 (1990).
14. A. Bax and D. G. Davis, *J. Magn. Reson. 65*, 355-360 (1985).
15. J. Jeener, B. H. Meier, P. Bachman, and R. R. Ernst, *J. Chem. Phys. 71*, 4546-4553 (1979).
16. C. Griesinger and R. R. Ernst, *J. Magn. Reson. 75*, 261-271 (1987).
17. K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York (1986).
18. N. H. Andersen, J. W. Neidigh, S. M. Harris, G. M. Lee, Z. H. Liu, and H. Tong, *J. Am. Chem. Soc. 119*, 8547-8561 (1997).
19. G. Wagner, W. Braun, T. F. Havel, T. Schaumann, N. Go, and K. Wüthrich, *J. Mol. Biol. 196*, 611-639 (1987).
20. A. Pardi, M. Billeter, and K. Wüthrich, *J. Mol. Biol. 180*, 741-751 (1984).
21. A. DeMarco, M. Llinás, and K. Wüthrich, *Biopolymers 17*, 617-636 (1978).
22. M. G. Cai, Y. Huang, J. H. Liu, and R. Krishnamoorthi, *J. Biomol. NMR 6*, 123-128 (1995).
23. M. E. Hodsdon, J. W. Ponder, and D. P. Cistola, *J. Mol. Biol. 264*, 585-602. (1996).
24. N. Guex and M. C. Peitsch, *Electrophoresis 18*, 2714-2723 (1997).
25. R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graph. 14*, 51-55, 29-32. (1996).
26. R. A. Laskowski, M. W. Macarthur, D. S. Moss, and J. M. Thornton, *J. Appl. Cryst. 26*, 283-291 (1993).
27. W. Apostoluk and J. Otlewski, *Proteins Str. Func. Gen. 32*, 459-474 (1998).
28. D. Krowarsch, M. Dadlez, O. Buczek, I. Krokoszynska, A. O. Smalas, and J. Otlewski, *J. Mol. Biol. 289*, 175-186 (1999).
29. D. G. Gronenstein and K. Taira, *Biophys. J. 46*, 749-761 (1984).
30. W. W. Bachovchin, *Biochemistry 25*, 7751-7759 (1986).
31. E. L. Ash, J. L. Sudmeier, R. M. Day, M. Vincent, E. V. Torchilin, K. C. Haddad, E. M. Bradshaw, D. G. Sanford, and W. W. Bachovchin, *Proc. Natl. Acad. Sci. USA 97*, 10371-10376 (2000).
32. R. C. Wilmouth, K. Edman, R. Neutze, P. A. Wright, I. J. Clifton, T. R. Schneider, C. J. Schofield, and J. Hajdu, *Nat. Struct. Biol. 8*, 689-694. (2001).
33. G. Katona, R. C. Wilmouth, P. A. Wright, G. I. Berglund, J. Hajdu, R. Neutze, and C. J. Schofield, *J. Biol. Chem. 277*, 21962-21970. (2002).
34. M. Topf, P. Varnai, C. J. Schofield, and W. G. Richards, *Proteins Str. Funct. Gen. 47*, 357-369. (2002).
35. I. L. de la Sierra, L. Quillien, P. Flecker, J. Gueguen, and S. Brunie, *J. Mol. Biol. 285*, 1195-1207 (1999).
36. M. H. Werner and D. E. Wemmer, *Biochemistry 31*, 999-1010 (1992).
37. R. H. Voss, U. Ermler, L. O. Essen, G. Wenzl, Y. M. Kim, and P. Flecker, *Eur. J. Biochem. 242*, 122-131 (1996).
38. J. Koepke, U. Ermler, E. Warkentin, G. Wenzl, and P. Flecker, *J. Mol. Biol. 298*, 477-491 (2000).
39. D. S. Wishart, B. D. Sykes, and F. M. Richards, *J. Mol. Biol. 222*, 311-333 (1991).

40. L. J. Smith, K. A. Bolin, H. Schwalbe, M. W. MacArthur, J. M. Thornton, and C. M. Dobson, *J. Mol. Biol. 225*, 494-506 (1996).
41. G. D. Lin, W. Bode, R. Huber, C. W. Chi, and R. A. Engh, *Eur. J. Biochem. 212*, 549-555 (1993).
42. S. X. Wang , C. T. Esmon and R. J. Fletterick, *Biochemistry 40*, 10038-10046 (2001).
43. F. Frigerio, A. Coda, L. Pugliese, C. Lionetti, E. Menegatti, G. Amiconi, H. P. Schnebli, P. Ascenzi, and M. Bolognesi, *J. Mol. Biol. 225*, 107-123 (1992).
44. M. Fujinaga, A. R. Sielecki, R. J. Read, W. Ardelt, M. Laskowski, Jr., and M. N. James, *J. Mol. Biol. 195*, 397-418 (1987).

Date Received: January 9, 2003

Communicated by the Editor Ramaswamy H Sarma