THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 33, pp. 23668–23676, August 18, 2006

The Absolute Structural Requirement for a Proline in the P3′-position of Bowman-Birk Protease Inhibitors Is Surmounted in the Minimized SFTI-1 Scaffold*

Received for publication, February 14, 2006, and in revised form, May 31, 2006 Published, JBC Papers in Press, June 9, 2006, DOI 10.1074/jbc.M601426200

Norelle L. Daly[‡1], Yi-Kuang Chen[‡], Fiona M. Foley[‡], Paramjit S. Bansal[‡], Rekha Bharathi[‡], Richard J. Clark[‡], Christian P. Sommerhoff[§], and David J. Craik[‡2]

From the [‡]Institute for Molecular Bioscience and Australian Research Council Special Research Centre for Functional and Applied Genomics, the University of Queensland, Brisbane, Queensland 4072, Australia and the [§]Department of Clinical Chemistry and Clinical Biochemistry, Ludwig-Maximilians-University, D-80336 Munich, Germany

SFTI-1 is a small cyclic peptide from sunflower seeds that is one of the most potent trypsin inhibitors of any naturally occurring peptide and is related to the Bowman-Birk family of inhibitors (BBIs). BBIs are involved in the defense mechanisms of plants and also have potential as cancer chemopreventive agents. At only 14 amino acids in size, SFTI-1 is thought to be a highly optimized scaffold of the BBI active site region, and thus it is of interest to examine its important structural and functional features. In this study, a suite of 12 alanine mutants of SFTI-1 has been synthesized, and their structures and activities have been determined. SFTI-1 incorporates a binding loop that is clasped together with a disulfide bond and a secondary peptide loop making up the circular backbone. We show here that the secondary loop stabilizes the binding loop to the consequences of sequence variations. In particular, full-length BBIs have a conserved *cis*-proline that has been shown previously to be required for well defined structure and potent activity, but we show here that the SFTI-1 scaffold can accommodate mutation of this residue and still have a well defined native-like conformation and nanomolar activity in inhibiting trypsin. Among the Ala mutants, the most significant structural perturbation occurred when Asp[14] was mutated, and it appears that this residue is important in stabilizing the *trans* peptide bond preceding Pro[13] and is thus a key residue in maintaining the highly constrained structure of SFTI-1. This aspartic acid residue is thought to be involved in the cyclization mechanism associated with excision of SFTI-1 from its 58-amino acid precursor. Overall, this mutational analysis of SFTI-1 clearly defines the optimized nature of the SFTI-1 scaffold and demonstrates the importance of the secondary loop in maintaining the active conformation of the binding loop.

Protease inhibitors are of significant interest because they play key roles in an extremely wide range of physiological function, including peptide hormone release, blood coagulation, and complement fixation. They have also been implicated in the treatment of various diseases, including some cancers and inflammatory processes (1). Bowman-Birk inhibitors (BBIs),[3] one of at least 18 different families of serine protease inhibitors, are involved in plant defense and have been implicated as cancer chemopreventive agents (2). The focus of this paper is SFTI-1 (sunflower trypsin inhibitor 1), a 14-residue cyclic peptide isolated from sunflower seeds that is related to the Bowman-Birk inhibitors and is one of the most potent inhibitors of trypsin of any naturally occurring peptide (3). SFTI-1 forms a tightly folded scaffold, either when complexed with trypsin or free in solution (3, 4). Its compact structure (shown in Fig. 1) and high potency have led to suggestions that it may serve as a scaffold for the design of novel peptide-based drug leads (5).

The structure of SFTI-1 consists of two β-strands connected at each end by turns and is braced by a single disulfide bond that creates two distinct loops (3, 4). One loop, known as the binding loop, contains the reactive site sequence Lys[5]–Ser[6], although the "secondary" loop contains a β-hairpin turn that essentially cyclizes the binding loop. SFTI-1 inhibits trypsin with a K_i of 0.1 nM (3) and inhibits cathepsin G with a comparable K_i. SFTI-1 is highly selective as it is 74-fold less inhibitory for chymotrypsin, and 3 orders of magnitude less inhibitory for elastase and thrombin. It has no detectable inhibitory activity against factor Xa (3).

Although SFTI-1 is only 14 residues in size, it has remarkable similarity to the active site sequence of the Bowman-Birk trypsin inhibitors that are typically 60–90 amino acids in length. Three key features have been identified that are conserved between BBIs and SFTI-1, namely covalent cyclization of a hairpin loop via a disulfide bridge, a *cis*-Pro at the P3′-site (using the protease active site nomenclature of Schechter and Berger (6)), and an extensive network of hydrogen bonds. The active site sequences of selected BBIs are given in Fig. 2 to highlight some of these features. A range of synthetic peptides based on the BBIs have been synthesized and their activities determined, as reviewed by Korsinczky *et al.* (7). Essentially all of the important features identified from these studies appear to have been

* This work was supported in part by a grant from the National Health and Medical Research Council (to N. L. D.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] National Health and Medical Research Industry Fellow.

[2] Australian Research Council Professorial Fellow. To whom correspondence should be addressed. Tel.: 61-7-3346-2019; Fax: 61-7-3346-2101; E-mail: d.craik@imb.uq.edu.au.

[3] The abbreviations used are: BBI, Bowman-Birk inhibitor; RP-HPLC, reverse phase-high performance liquid chromatography; SFTI, sunflower trypsin inhibitor; TCEP, triscarboxylethylphosphine; TOCSY, total correlation spectroscopy; Boc, *t*-butoxycarbonyl.

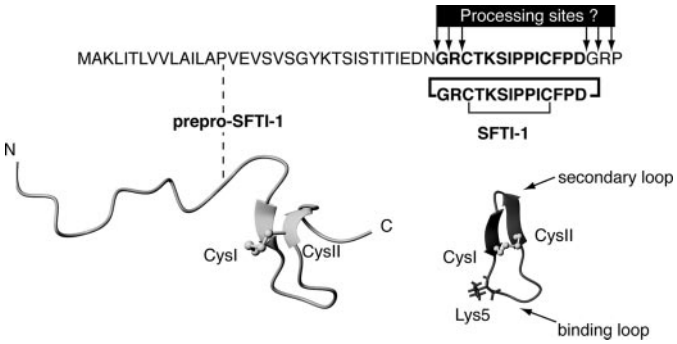

FIGURE 1. **Sequence and structure of native SFTI-1 and its precursor protein.** Ribbon representations of the NMR derived three-dimensional structures of SFTI-1 (4) and prepro-SFTI-1 (8) are shown adjacent to the sequences. The β-strands are shown as *arrows*. The potential cleavage sites involved in excision of the mature peptide from the precursor protein are labeled with *arrows*.

GRCTKSI*P*PICFPD	SFTI-1	sunflower	14 AA
– CCDCL**C**TKSI*P*PQ**C**QCADIR–	BBI	adzuki bean	78 AA
– DECGT**C**TRMI*P*PR**C**TCMDVS–	BBI	wheat	71 AA
– CNFCP**C**TRSI*P*PQ**C**RCTDIG–	BBI	alfalfa	58 AA
– CDEAV**C**TRSI*P*PI**C**TCMDEV–	BBI	barley	124 AA
– CDTCL**C**TKSE*P*PT**C**RCVDVG–	BBI	broad bean	63 AA
– CDSCD**C**TKSI*P*PE**C**HCANIR–	BBI	mung bean	72 AA
– CDKCF**C**TKSN*P*PI**C**QCRDVG–	BBI	coral tree	61 AA

P3 P2 P1 P1′ P2′ P3′ P4′

FIGURE 2. **A comparison of the active site sequences of Bowman-Birk inhibitors with SFTI-1.** The conserved *cis*-proline residue at the P3′-position is highlighted in *italics,* and the two cysteine residues that are linked in a disulfide bond to form the binding loop are highlighted in *boldface.* The standard Schechter and Berger nomenclature (6) for protease active site positions is illustrated below the BBI sequences.

incorporated naturally into the SFTI framework, demonstrating that native SFTI is a highly optimized trypsin inhibitory scaffold.

The precursor protein of SFTI-1 was recently discovered, and the sequence is given in Fig. 1. Interestingly, the sequence of the precursor does not show any similarities with the Bowman-Birk inhibitors, apart from the mature peptide domain, posing questions about the evolution of SFTI-1 (8). The mechanism of cyclization has yet to be elucidated, and the presence of a GR dipeptide motif at both ends of the mature sequence leaves a degree of ambiguity as to the cleavage sites (8), although it is likely that cyclization occurs between an Asp and a Gly residue (see Fig. 1), which opens up the possibility of a role for an asparaginyl endoprotease in the process (8). The mechanism of cyclization of the largest family of backbone cyclic proteins, the plant cyclotides (9), has also been suggested to involve an Asp/Asn-Gly and may also involve an asparaginyl endoprotease, although this has yet to be experimentally verified (10). Circular proteins have also been reported in bacteria and mammals, and the mechanism of cyclization is similarly unknown (11).

As well as having potential applications as a protein engineering scaffold, SFTI-1 also displays potent inhibitory activity against matriptase, an enzyme implicated in prostate cancer, suggesting that it may also have direct therapeutic applications (12–14). Matriptase was first isolated as a novel proteinase expressed by human breast cancer cells and is also highly expressed in prostate, breast, and colorectal cancers *in vitro* and *in vivo* (12, 13). Inhibition of matriptase suppresses both primary tumor growth and metastasis in a rat model of prostate cancer.

Understanding the structure-activity relationships of SFTI-1 will facilitate potential protein engineering and therapeutic applications of this peptide. Mutagenesis studies involving the systematic replacement of individual residues also have the potential to provide insights into their role in the cyclization mechanism. In this study we have synthesized a complete suite of alanine mutants of SFTI-1 and characterized them structurally and functionally. Lys^5 is the only residue that resulted in a significant loss of trypsin inhibitory activity, and most surprisingly, the *cis*-Pro that is highly conserved across the Bowman-Birk inhibitors is not critical for activity.

MATERIALS AND METHODS

Peptide Synthesis—Boc-based solid phase peptide synthesis was carried out using standard protocols. Peptides were assembled using a thioester linker assembled on resin allowing subsequent cyclization by a thiazip mechanism (15). Hydrogen fluoride cleavage was conducted on the deprotected resins using standard protocols (0 °C, 90 min, 90% HF, 8% *p*-cresol, 2% *p*-thiocresol). Crude cleavage products were purified by RP C_{18} HPLC (1%/minute gradient of 90% acetonitrile, 10% water, 0.05% trifluoroacetic acid against 100% water, 0.05% trifluoroacetic acid) to give linear, reduced peptides. Peptides were cyclized and oxidized in 0.1 M ammonium bicarbonate at pH 8 overnight and purified as above. Purity of fractions was assessed using electrospray ionization-mass spectrometry and analytical HPLC using a 2%/minute gradient of the same solvents used for previous steps.

NMR Spectroscopy—Samples for ^{1}H NMR measurements contained ~1 mM peptide in 90% H_2O, 10% D_2O (v/v) at pH ~5. D_2O (99.9 and 99.99%) was obtained from Cambridge Isotope Laboratories, Woburn, MA. Spectra were recorded at 290 K on a Bruker Avance-500 or Avance-600 spectrometer equipped with a shielded gradient unit. Two-dimensional NMR spectra were recorded in phase-sensitive mode using time-proportional phase incrementation for quadrature detection in the t_1 dimension (16). The two-dimensional experiments consisted of a TOCSY (17) using an MLEV-17 spin lock sequence (18) with a mixing time of 80 ms and nuclear Overhauser effect spectroscopy (19) with mixing times of 100–250 ms. Solvent suppression was achieved using a modified WATERGATE sequence (20). Spectra were acquired over 6024 Hz with 4096 complex data points in F_2 and 512 increments in the F_1 dimension. Spectra were processed on a Silicon Graphics Indigo work station using XWINNMR (Bruker) software. The t_1 dimension was zero-filled to 1024 real data points, and 90° phase-shifted sine bell window functions were applied prior to Fourier transformation.

Trypsin Inhibition—The concentrations of the inhibitory active peptides and the equilibrium dissociation constants K_i were determined with trypsin. Bovine pancreatic trypsin (*N*-*p*-tosyl-L-phenylalanine chloromethyl ketone-treated; Sigma) was standardized by burst titration with *p*-nitrophenyl *p*′-guanidinobenzoate (21). Trypsin (25 nM) was then incubated with serial dilutions of the peptides in 50 mM HEPES, 150 mM NaCl,

0.01% Triton X-100, 0.01% sodium azide, pH 7.4, for 5 min at room temperature. The residual activity was quantified by following the hydrolysis of the substrate carbobenzoxy-L-arginine-7-amino-4-methylcoumarin (125 μM; Sigma) in an HTS 7000 BioAssay Reader (PerkinElmer Life Sciences). The concentrations of the active peptides were calculated by assuming a 1:1 interaction between inhibitor and trypsin; the mutant K5A could not be titrated because of its considerably lower affinity, and HPLC and amino acid analyses were used for quantification. Subsequently similar experiments were performed at a lower enzyme concentration to determine the equilibrium dissociation constant for the complex (22). Thus, the peptides were incubated with trypsin (0.01 nM), and the residual activity was quantified using the substrate *N-p*-tosyl-glycine-proline-arginine-7-amido-4-methylcoumarin (5 μM; Sigma). The K_i values were calculated by fitting the steady state velocities to the equation for tight binding inhibitors (23) using nonlinear regression analysis with the software ProFit (Quantum Soft, Uetikon am See, Switzerland). Mean values ± S.D. of at least four experiments are reported.

RESULTS

Each non-cysteine residue in SFTI-1 was replaced individually with an alanine residue in a suite of peptides synthesized using Boc-based solid phase peptide synthesis. The native peptide was also synthesized for comparison. For convenience the peptides are referred to here by reference to the mutation, *e.g.* G1A for [1-Ala]SFTI-1. For a cyclic sequence of n residues, there are in principle n choices of linear precursors for the synthesis. For the mutants of SFTI-1, the peptide sequences were assembled with the flanking residues being a C-terminal thioester and an N-terminal cysteine residue to facilitate cyclization via an adaptation of native peptide ligation chemistry (24, 25). Fig. 3 outlines the chemical strategy involved. Briefly, nucleophilic attack of the N-terminal Cys on the thioester linker at the C terminus results in a cyclic thiolactam that subsequently rearranges via an *S-* to *N*-acyl transfer to produce a native peptide bond. As there are two cysteine residues in SFTI-1, there are two possible ligation points. The preferred ligation point was chosen to be between Cys^3 and Arg^2, or in the case of R2A between Cys^3 and Ala^2.

Formation of the cyclic backbone and disulfide bond was generally performed in a single step in 0.1 M ammonium bicarbonate at pH 8. For D14A the cyclization reaction did not proceed efficiently and a two-step approach was required. The cyclization reaction was performed in the presence of TCEP, and following purification of the cyclic product, the peptide was oxidized in 0.1 M ammonium bicarbonate. The TCEP was used to keep the two cysteine residues in a reduced state as this appears to facilitate the cyclization reaction. All peptides were purified using preparative reverse phase HPLC (RP-HPLC) and characterized using analytical scale RP-HPLC and mass spectrometry.

Sufficient quantities of the 12 SFTI-1 Ala mutants were isolated and purified for structural analysis with NMR spectroscopy and functional analysis of the trypsin inhibitory activity. The NMR spectra were recorded in aqueous solution at 290 K, and NMR spectral assignments were made using established

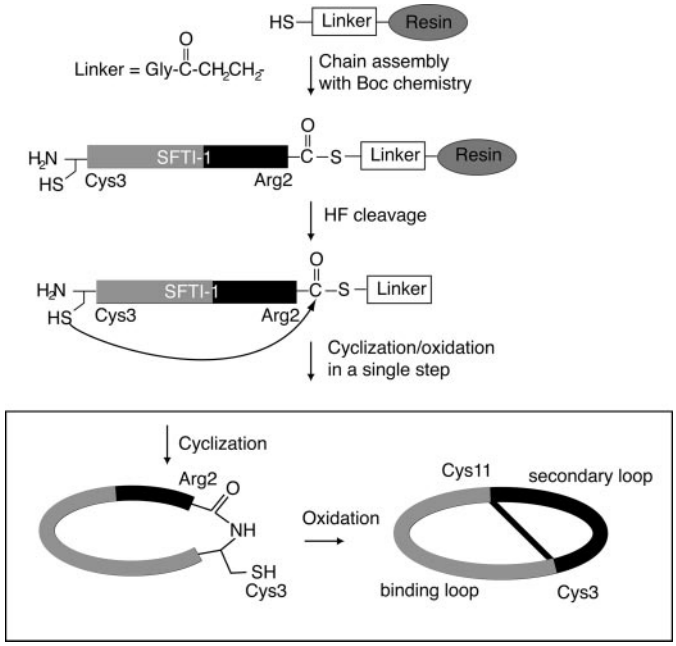

FIGURE 3. **A schematic representation of the strategy for synthesizing the SFTI-1 analogues.** A thioester linker is attached to the solid support (resin) and the peptide chain assembled using Boc chemistry. The peptide with the linker attached is cleaved from the resin using HF and a single step reaction used for cyclizing the backbone and oxidizing the disulfide bond. The binding and secondary loops of the SFTI-1 sequence are shown in *lighter* and *darker shading,* respectively.

techniques (26). Chemical shifts, in particular the αH shifts, are extremely sensitive to structural changes and thus offered a convenient method of assessing structural changes that may have occurred upon alanine substitution.

Under the conditions examined, one predominant conformation was present in solution for the native peptide. Analysis of TOCSY spectra between pH 2.8 and 5.8 indicated that Asp^{14} has a pK_a of ~4, and there were no significant conformational changes over this pH range (as assessed via a lack of αH chemical shift changes). The SFTI-1 mutants also displayed only one predominant conformation, with the exception of D14A, for which two isomers were present. The NMR spectra of the SFTI-1 mutants were recorded at both pH 4 and 5, which gave essentially identical spectra, again showing a lack of pH-induced conformational changes. A comparison of the secondary shifts of native SFTI with the mutants (at pH 5) is shown in Fig. 4. The majority of the mutants display chemical shifts very similar to the native peptide, indicating that no major structural changes occur as a result of the mutations. However, I7A, P8A, and P9A display differences for certain residues that indicate local structural perturbations. I7A differs at residue 7, but this is likely to result from a local effect of the alanine substitution. P8A has changes for residues 6–8, whereas P9A differs at residue 8. T4A also displays minor differences from the native peptide.

A comparison of the TOCSY spectrum of D14A with native SFTI-1 is given in Fig. 5. It is clear that one major conformer is present in native SFTI-1, whereas D14A has two sets of spin systems from two distinct conformations. For some residues significant differences are observed for the amide chemical shifts of the two conformers, and in some cases no differences,

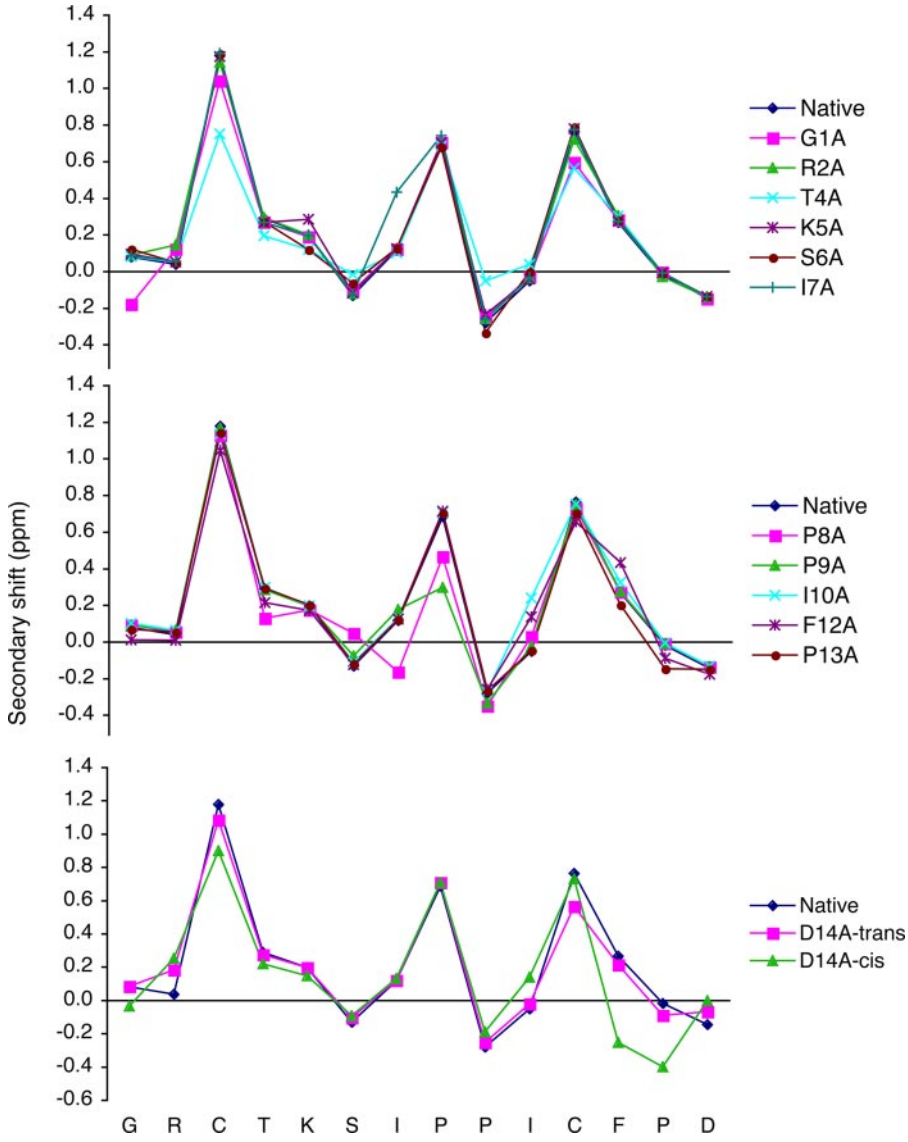

FIGURE 4. **Secondary shift analysis of SFTI-1 alanine mutants compared with native SFTI-1.** The Hα secondary shifts were calculated by subtracting the random coil shift (40) from the Hα shift.

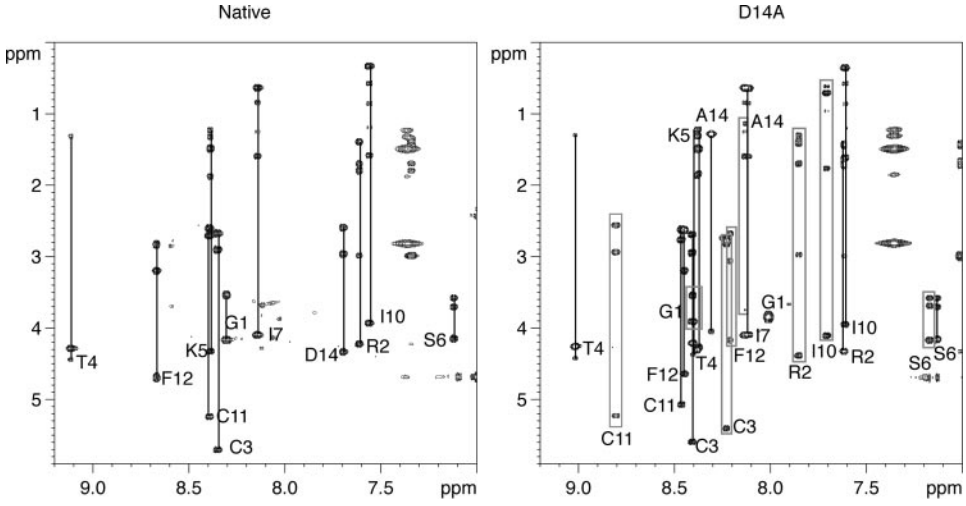

FIGURE 5. **Comparison of the TOCSY spectra of native SFTI-1 and the D14A mutant.** The spin systems are labeled with the single letter amino acid code and residue number. SFTI-1 exists as a single well defined conformation, whereas in D14A the Pro[13] is present in both the *cis* and *trans* forms. The *cis* conformer of D14A is highlighted with *boxes*.

or only minor differences, are evident. Assignment of both conformers showed that Pro[13] is present in a *trans* conformation in one of the conformers and a *cis* conformation in the other. The discrimination of the *trans* and *cis* forms was based on detection of d$\alpha_{i-1}\delta$ and d$\alpha\alpha$ nuclear Overhauser effects, respectively, for the X-Pro peptide bonds. The secondary shifts (*i.e.* the differences between observed and random coil αH chemical shifts) of the non-native *cis* conformation of D14A differ from the native peptide at residues 12–14, in contrast to the *trans* conformation, which is very similar to the native peptide (Fig. 4). Prolines 8 and 9 are present as *cis* and *trans* forms, respectively, in both conformers. Clearly, mutating Asp[14] to an alanine residue has a significant effect on the stability of the turn region involving residues 12–14 and 1. (Note that residues 14 and 1 are sequentially adjacent in the cyclic peptide.)

Although the geometry of Pro[13] is affected in D14A, the native *trans* conformation is maintained in all other Ala mutants containing this Asp residue. When present, Pro[8] and Pro[9] also maintain their native geometries (*i.e. cis* for Pro[8] and *trans* for Pro[9]) in the Ala mutants.

The ability of SFTI-1 and the analogues to inhibit trypsin was compared by determining the equilibrium dissociation constants K_i of the complexes (Table 1). The measured K_i value of SFTI-1 (0.03 nM) is in good agreement with that reported for SFTI-1 isolated from sunflower seeds (0.1 nM (3)). Not surprisingly, the P1 reactive site mutant K5A is a considerably weaker inhibitor (K_i 190 nM). All of the mutants inhibit trypsin at low nanomolar concentration, the K_i values differing less than 50-fold from that of SFTI-1.

DISCUSSION

SFTI-1 is one of the most potent BBIs known and has exciting potential therapeutic applications based on its inhibitory activity against matriptase, an enzyme implicated in

TABLE 1
Equilibrium dissociation constant K_i for the inhibition of bovine trypsin by SFTI-1 and its alanine mutants

Peptide	Sequence	K_i[a]
		nM
SFTI-1	-GRCTKSIPPICFPD-	0.03 ± 0.01
G1A	-**A**RCTKSIPPICFPD-	0.08 ± 0.02
R2A	-G**A**CTKSIPPICFPD-	0.47 ± 0.05
T4A	-GRC**A**KSIPPICFPD-	0.26 ± 0.05
K5A	-GRCT**A**SIPPICFPD-	190 ± 39
S6A	-GRCTK**A**IPPICFPD-	0.35 ± 0.04
I7A	-GRCTKS**A**PPICFPD-	0.84 ± 0.25
P8A	-GRCTKSI**A**PICFPD-	1.49 ± 0.19
P9A	-GRCTKSIP**A**ICFPD-	0.08 ± 0.04
I10A	-GRCTKSIPP**A**CFPD-	0.11 ± 0.02
F12A	-GRCTKSIPPIC**A**PD-	0.53 ± 0.06
P13A	-GRCTKSIPPICF**A**D-	0.05 ± 0.01
D14A	-GRCTKSIPPICFP**A**-	0.24 ± 0.05

[a] Values are mean ± S.D., $n \geq 4$.

prostate cancer (14). Furthermore, because of its tightly constrained structure, it has been suggested that it may serve as a scaffold for peptide-based drug development (5, 27–29). In this study, a suite of alanine mutants of SFTI-1 has been synthesized to facilitate a determination of structure-activity relationships, as such an analysis is critical if the therapeutic potential of the SFTI-1 framework is to be realized.

We employed Boc chemistry and a thioester-based native chemical ligation approach to cyclize the SFTI-1 mutants. Such an approach has been successfully applied to another family of macrocyclic plant proteins, namely the cyclotides, which contain three disulfide bonds and hence are intrinsically more complex (9). The method worked efficiently for the suite of SFTI-1 mutants, and in all but one case only a single step reaction was required. The D14A mutant required a two-step procedure to facilitate cyclization and oxidation. Our approach represents a novel and efficient method of synthesizing SFTI-1 analogues. In the past, several methods have been used to synthesize SFTI-1 and analogues, including an on-resin cyclization approach (4) and Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry followed by cyclization in solution (30). All have been successful, thus emphasizing the synthetic accessibility of the framework, but the thioester approach has the advantage of a single step cyclization and oxidation reaction.

The similarity of the αH chemical shifts between the mutants and native SFTI-1 suggests that no major structural changes occur in the majority of the mutants. The most significant differences from the native peptide are for mutants I7A, P8A, and P9A, indicating that the Ile^7–Pro^9 region of the sequence is important for structural integrity. These residues are in a turn region of the binding loop adjacent to the active site Lys^5–Ser^6 peptide bond, and include the *cis*-Pro residue conserved in the P3′-position of all BBIs. The other significant mutation from a structural perspective occurs in the secondary loop, where mutation of Asp^{14} to an Ala destabilizes a turn region and leads to *cis-trans* isomerization of Pro^{13}.

The most significant loss in activity was observed in K5A. The importance of the Lys is of course expected as it is the P1 residue and primarily responsible for recognition of the protease. However, the lack of importance of other residues is surprising given that the conserved *cis*-proline at P3′ has been found to be also critical for function in studies of related peptides. For instance, when the active sequence of SFTI-1 was grafted onto a D-Pro-L-Pro template and each residue in the binding loop was subsequently replaced with alanine, both Lys^5 and Pro^8 were found to be critical for activity, with substitution of Pro^8 having a greater effect on activity than substitution of Lys^5 (31). The binding loop template used for those studies was essentially a truncated SFTI-1 molecule in which the secondary loop is omitted, as illustrated in Fig. 6*A*. Pro^8 of SFTI-1 is equivalent to the *cis*-proline conserved throughout the known Bowman-Birk inhibitors, and so it is interesting that we find in this study that it is not as crucial in maintaining activity in the full SFTI-1 scaffold.

The discrepancy between the role of the *cis*-proline in the truncated *versus* full SFTI-1 scaffold appears to be related to the influence of the secondary loop on the structure of the binding loop. Structure-activity studies of the proline residues in disulfide-cyclized peptides corresponding to BBI reactive site loops (Fig. 6*B*) showed that replacement of the P3′ Pro with Ala resulted in poorly defined structure and poor inhibitory activity (27). Furthermore, mutation of the Pro at the P4′-position with Ala resulted in *cis-trans* isomerization of the Pro at P3′ (27). In this study we have shown that mutation of the prolines at the P3′- and P4′-positions of SFTI-1 (*i.e.* Pro^8 and Pro^9) does not result in multiple conformations, and indeed Pro^8 still adopts a single well defined conformation with a *cis*-peptide bond in the P9A mutant. Clearly, the full SFTI-1 scaffold stabilizes the active conformation and can accommodate sequence variations to a greater extent than the disulfide-linked or D-Pro-L-Pro capped binding loop sequences alone. In other words, the secondary loop in SFTI-1 clearly plays an important role in stabilizing the binding loop.

We explored the structural features of the secondary loop that makes it a useful stabilizing template. Most residues are relatively tolerant to substitution, but mutating Asp^{14} to Ala in the secondary loop results in two distinct conformations that differ by the presence of a *cis*- or *trans*-proline at position 13. The conformations of the other two prolines are the same as in the native peptide (*i.e.* Pro^8 and Pro^9 are in *cis* and *trans* conformations, respectively). As two sets of peaks are observed for the majority of residues in D14A, the alternative conformations are in slow exchange on the NMR time scale, consistent with the energy barrier normally seen between *cis*- and *trans*-Pro isomers in peptides. The stabilizing role of Asp^{14} appears to be related to its ability to hydrogen-bond to the backbone of the secondary loop. Specifically, in native SFTI-1 a hydrogen bond is present between the side chain of Asp^{14} and the backbone amide of Arg^2 (4), as shown in Fig. 6*C*. In the D14A mutant, this hydrogen bond is clearly not possible, and its absence most likely accounts for the two distinct conformations observed. The importance of the Asp^{14}–Arg^2 side chain to NH hydrogen bond is reinforced by the fact that the Arg^2 carbonyl group in turn hydrogen-bonds to the backbone NH of Phe^{12}. Thus, the secondary loop has a well defined hydrogen bonding network that helps establish it as a template for defining the structure of the binding loop. This hydrogen bond network is schematically illustrated in Fig. 6*D*.

As an aside, the presence of a Phe residue preceding Pro^{13} is likely to be an important factor influencing the presence of

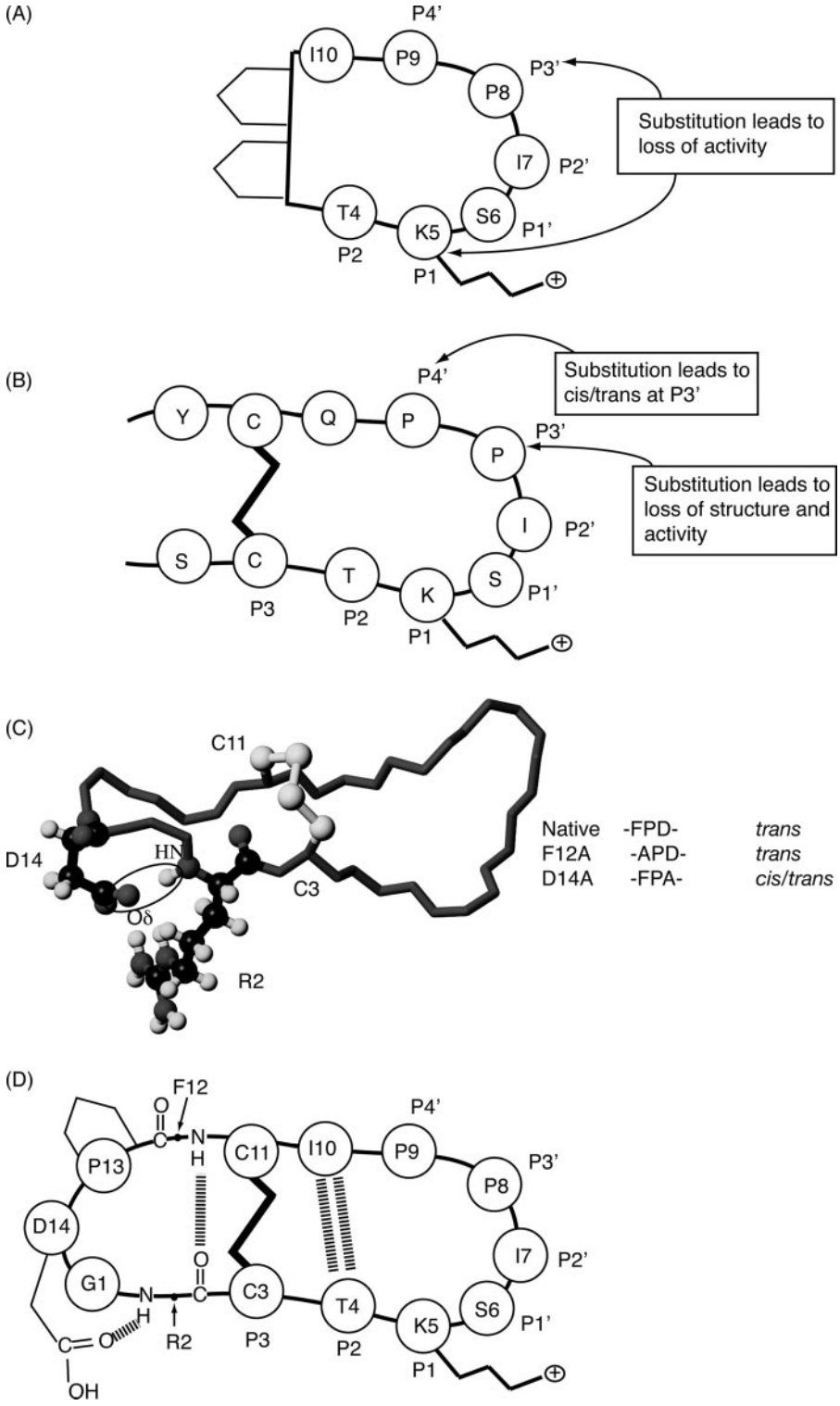

FIGURE 6. **SFTI-1 variants that have been used in structure-activity studies.** *A,* truncated derivative in which a D-Pro-L-Pro template was used to display the binding loop (31). *B,* disulfide-cyclized peptide was used to determine structure-activity relationships of the Pro residues of the reactive site loop of the BBIs. *C,* hydrogen bonding of SFTI-1 involving Arg^2 and Asp^{14}. The hydrogen bond between the side chain of Asp^{14} and the backbone of Arg^2 is highlighted in *ball* and *stick format*. Pro^{13} is in a *trans* conformation in the native peptide and retains this conformation in F12A but has two conformations in D14A. *D,* schematic illustration of the hydrogen bonding network in SFTI-1.

cis and *trans* isomers in D14A, but the importance of the Asp^{14}–Arg^2 hydrogen bond is re-emphasized by considering the role of the Phe residue. Previous studies in a diverse range of peptides have shown that aromatic residues have a high propensity of favoring *cis* peptide bond conformations when preceding proline residues (32, 33). It is interesting that the Phe^{12}–Pro^{13} peptide bond in native SFTI-1 is not in the *cis* conformation despite the high propensity for such a conformation in model peptides, but this may be explained by the influence of the Asp-Arg hydrogen bond. Removal of the hydrogen bond between Asp^{14} and Arg^2 destabilizes this turn region and allows measurable populations of the more favorable *cis* conformation to be detected. Thus, the tightly folded nature of SFTI-1 may have a greater influence on the geometry of the proline residues than is normally observed in model peptides. Overall, the stabilizing hydrogen bond network may be summarized as follows: the Asp^{14} side chain hydrogen bonds to the backbone NH of Arg^2, facilitating orientation of its CO to hydrogen bond to Phe^{12}, which in turn assists in maintaining a *trans* conformation for Pro^{13}, which in turn orients Asp^{14}, and the cycle continues.

The role of Asp^{14} in maintaining the structure of the cyclic scaffold is intriguing given that this residue is likely to be involved in the processing reaction to form the native cyclic peptide from its precursor protein (8). How is it that a residue implicated in cyclization, and hence likely to change its conformation in the transition from precursor to mature cyclic peptide, can play such an important structural role? One could speculate that SFTI-1 is evolutionarily related to an ancient protein encoding a BBI-like active site domain that at some point acquired a mutation leading to Asp at position 14. This predisposed the molecule to the potential for cyclization, in part driven by the thermodynamic stability that results when Asp^{14} is able to hydrogen-bond to the residue that in the cyclic molecule is only two residues sequentially adjacent (Arg^2) but is 12 residues away in sequence in the linear version.

With respect to speculation on processing mechanisms, it is of interest to note that although it appears likely that Asp^{14} is indeed implicated in cyclization, this has not yet been established with certainty, and consideration of the precursor sequence of SFTI-1 alone still allows the possibility of three alternative processing sites, involving cleavage before or after Gly or Arg as illustrated in Fig. 1. A similar ambiguity occurred originally for the cyclotides, where the presence of GLP triplets flanking the mature peptide sequence led to four potential cleavage sites (10, 34). The latter ambiguity was resolved when consideration of numerous cyclotide sequences showed that the first of the two GLP triplets is incorporated into the mature peptide and that cleavage after an Asn or Asp residue excises the second GLP from the mature peptide during processing from the precursor (10, 35, 36). By analogy, in SFTI-1 it seems likely that the first GR doublet is incorporated, and the second is excised because of cleavage after Asp. However, SFTI-1 differs from the cyclotides in that only a single variant has been discovered so far, whereas there are many cyclotide variants (37).

The proposal that Asp or Asn residues are likely involved as C-terminal processing points in cyclization reactions is supported by the common occurrence of asparaginyl endopeptidases in plants. Presumably, such enzymes could be recruited by linear precursor proteins to catalyze cyclization as a supplemental activity to their usual proteolytic function. In the case of SFTI-1, we recently demonstrated the principle of this "reverse" use of an enzyme by showing that an acyclic permutant of SFTI-1 with the backbone broken at the Lys^5–Ser^6 peptide bond is able to be efficiently cyclized by trypsin (38).

The lack of natural variants of SFTI-1 in sunflower seeds or indeed other plants supports the notion that it has an extremely optimized framework for trypsin inhibitory activity. This is in stark contrast to the plant cyclotides where there are more than 80 published sequences and thousands of predicted sequences (9, 39). This fundamental difference between the two types of cyclic plant peptides is likely to be related to their mechanisms of action. Cyclotides are insecticidal agents and are thought to interact with membranes as part of their mode of action (10). A nonspecific mode of action perhaps lends itself to more sequence variation than the highly specific nature of a protease inhibitor such as SFTI-1.

In summary, despite the highly optimized framework of SFTI-1, we have shown that it can accommodate sequence variations and still allow the native-like fold to be maintained with significant levels of trypsin inhibitory activity. This result is in contrast to previous studies on truncated derivatives that only contain the binding loop of SFTI-1 (27, 31) and demonstrates that the secondary loop is a crucial stabilizing factor in SFTI-1. In particular, Asp^{14} from the secondary loop is a key residue in maintaining the well defined native structure. As this residue is likely to be involved in the cyclization process, its role in stabilizing the cyclic scaffold may have implications for the evolution of SFTI-1. This study has significantly broadened our understanding of the structure-activity relationships of SFTI-1, and the finding that a flanking loop can play a crucial stabilizing role provides important information for the design of novel trypsin inhibitors in general.

Acknowledgment—We thank Sabine Streicher for technical assistance.

REFERENCES

1. Grasberger, B. L., Clore, G. M., and Gronenborn, A. M. (1994) *Structure (Lond.)* **2,** 669–678
2. Kennedy, A. R. (1997) *Am. J. Clin. Nutr.* **68,** S1406–S1412
3. Luckett, S., Garcia, R. S., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) *J. Mol. Biol.* **290,** 525–533
4. Korsinczky, M. L., Schirra, H. J., Rosengren, K. J., West, J., Condie, B. A., Otvos, L., Anderson, M. A., and Craik, D. J. (2001) *J. Mol. Biol.* **311,** 579–591
5. Korsinczky, M. L., Schirra, H. J., and Craik, D. J. (2004) *Curr. Protein Pept. Sci.* **5,** 351–364
6. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162
7. Korsinczky, M. L., Clark, R. J., and Craik, D. J. (2005) *Biochemistry* **44,** 1145–1153
8. Mulvenna, J. P., Foley, F. M., and Craik, D. J. (2005) *J. Biol. Chem.* **280,** 32245–32253
9. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294,** 1327–1336
10. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10614–10619
11. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27,** 132–138
12. Lin, C. Y., Anders, J., Johnson, M., Sang, Q. A., and Dickson, R. B. (1999) *J. Biol. Chem.* **274,** 18231–18236
13. Friedrich, R., Fuentes-Prior, P., Ong, E., Coombs, G., Hunter, M., Oehler, R., Pierson, D., Gonzalez, R., Huber, R., Bode, W., and Madison, E. L. (2002) *J. Biol. Chem.* **277,** 2160–2168
14. Long, Y. Q., Lee, S. L., Lin, C. Y., Enyedy, I. J., Wang, S., Li, P., Dickson, R. B., and Roller, P. P. (2001) *Bioorg. Med. Chem. Lett.* **11,** 2515–2519
15. Tam, J. P., and Lu, Y.-A. (1997) *Tetrahedron Lett.* **38,** 5599–5602
16. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113,** 967–974
17. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson.* **53,** 521–528
18. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65,** 355–360
19. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys.* **71,** 4546–4553
20. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2,** 661–665
21. Chase, T., Jr., and Shaw, E. (1970) in *Methods in Enzymology* (Perlmann, G. E. and Lorand, L., eds) pp. 20–27, Academic Press, New York
22. Bieth, J. G. (1980) *Bull. Eur. Physiopathol. Respir.* **16,** (Suppl.) 183–197
23. Morrison, J. F. (1969) *Biochim. Biophys. Acta* **185,** 269–286
24. Tam, J. P., Lu, Y.-A., and Yu, Q. (1999) *J. Am. Chem. Soc.* **121,** 4316–4324
25. Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) *Biochemistry* **38,** 10606–10614
26. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* pp. 130–161, Wiley Interscience, New York
27. Brauer, A. B., Domingo, G. J., Cooke, R. M., Matthews, S. J., and Leatherbarrow, R. J. (2002) *Biochemistry* **41,** 10608–10615
28. Craik, D. J., Simonsen, S., and Daly, N. L. (2002) *Curr. Opin. Drug Discovery Dev.* **5,** 251–260
29. Craik, D. J., Cemazar, M., and Daly, N. L. (2006) *Curr. Opin. Drug Discovery Dev.* **9,** 251–260
30. Zablotna, E., Kazmierczak, K., Jaskiewicz, A., Stawikowski, M., Kupryszewski, G., and Rolka, K. (2002) *Biochem. Biophys. Res. Commun.* **292,** 855–859
31. Descours, A., Moehle, K., Renard, A., and Robinson, J. A. (2002) *Chembiochem.* **3,** 318–323

32. Wu, W. J., and Raleigh, D. P. (1998) *Biopolymers* **45,** 381–394
33. Yao, J., Feher, V. A., Espejo, B. F., Reymond, M. T., Wright, P. E., and Dyson, H. J. (1994) *J. Mol. Biol.* **243,** 736–753
34. Dutton, J. L., Renda, R. F., Waine, C., Clark, R. J., Daly, N. L., Jennings, C. V., Anderson, M. A., and Craik, D. J. (2004) *J. Biol. Chem.* **279,** 46858–46867
35. Jennings, C. V., Rosengren, K. J., Daly, N. L., Plan, M., Stevens, J., Scanlon, M. J., Waine, C., Norman, D. G., Anderson, M. A., and Craik, D. J. (2005) *Biochemistry* **44,** 851–860
36. Daly, N. L., Clark, R. J., Plan, M. R., and Craik, D. J. (2006) *Biochem. J.* **393,** 619–626
37. Simonsen, S. M., Sando, L., Ireland, D. C., Colgrave, M. L., Bharathi, R., Göransson, U., and Craik, D. J. (2005) *Plant Cell* **17,** 3176–3189
38. Marx, U. C., Korsinczky, M. L., Schirra, H. J., Jones, A., Condie, B., Otvos, L., Jr., and Craik, D. J. (2003) *J. Biol. Chem.* **278,** 21782–21789
39. Trabi, M., and Craik, D. J. (2004) *Plant Cell* **16,** 2204–2216
40. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81

The Absolute Structural Requirement for a Proline in the P3′-position of Bowman-Birk Protease Inhibitors Is Surmounted in the Minimized SFTI-1 Scaffold

Norelle L. Daly, Yi-Kuang Chen, Fiona M. Foley, Paramjit S. Bansal, Rekha Bharathi, Richard J. Clark, Christian P. Sommerhoff and David J. Craik

J. Biol. Chem. 2006, 281:23668-23675.
doi: 10.1074/jbc.M601426200 originally published online June 9, 2006
