Enzymatic Cyclization of a Potent Bowman-Birk Protease Inhibitor, Sunflower Trypsin Inhibitor-1, and Solution Structure of an Acyclic Precursor Peptide*[S]

Ute C. Marx‡§, Michael L. J. Korsinczky‡, Horst Joachim Schirra‡¶, Alun Jones‡, Barrie Condie‖, Laszlo Otvos, Jr.‖, and David J. Craik‡**

From the ‡Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia, §Lehrstuhl für Biopolymere, Universität Bayreuth, Universitätsstrasse 30, Bayreuth D-95447, Germany, and ‖The Wistar Institute, Philadelphia, Pennsylvania 19104

The most potent known naturally occurring Bowman-Birk inhibitor, sunflower trypsin inhibitor-1 (SFTI-1), is a bicyclic 14-amino acid peptide from sunflower seeds comprising one disulfide bond and a cyclic backbone. At present, little is known about the cyclization mechanism of SFTI-1. We show here that an acyclic permutant of SFTI-1 open at its scissile bond, SFTI-1[6,5], also functions as an inhibitor of trypsin and that it can be enzymatically backbone-cyclized by incubation with bovine β-trypsin. The resulting ratio of cyclic SFTI-1 to SFTI-1[6,5] is ~9:1 regardless of whether trypsin is incubated with SFTI-1[6,5] or SFTI-1. Enzymatic resynthesis of the scissile bond to form cyclic SFTI-1 is a novel mechanism of cyclization of SFTI-1[6,5]. Such a reaction could potentially occur on a trypsin affinity column as used in the original isolation procedure of SFTI-1. We therefore extracted SFTI-1 from sunflower seeds without a trypsin purification step and confirmed that the backbone of SFTI-1 is indeed naturally cyclic. Structural studies on SFTI-1[6,5] revealed high heterogeneity, and multiple species of SFTI-1[6,5] were identified. The main species closely resembles the structure of cyclic SFTI-1 with the broken binding loop able to rotate between a *cis/trans* geometry of the I7–P8 bond with the *cis* conformer being similar to the canonical binding loop conformation. The non-reactive loop adopts a β-hairpin structure as in cyclic wild-type SFTI-1. Another species exhibits an iso-aspartate residue at position 14 and provides implications for possible *in vivo* cyclization mechanisms.

Over recent years there has been much interest in the discovery of circular proteins in higher organisms (1) and in the development in synthetic approaches to cyclize proteins (2). In general, backbone cyclic peptides have several advantages over their non-cyclic counterparts. They are resistant to attack by exopeptidases, making them less vulnerable to degradation

and can have an increased thermal stability (3). Also, unfavorable entropic losses upon binding to target proteins are significantly reduced, resulting in a thermodynamically more efficient binding interaction (1). These biological advantages of backbone cyclized peptides may lead to their use as scaffolds for the design of stable pharmaceuticals and pesticides (4).

The new generation of circular peptides/proteins discovered in the last few years differs from previously known cyclic peptides such as cyclosporins in that the latter are generally not direct gene products but are synthesized in bacteria by multifunctional enzymes and often contain non-conventional amino acids (5). By contrast, recently discovered circular miniproteins such as the plant cyclotides (6) are gene products that are post-translationally processed to cyclize their conventional peptide backbone (7). Although *in vitro* cyclization procedures are now being developed for the synthetic production of circular proteins, little is known about the mechanisms and driving force behind *in vivo* cyclization of naturally occurring cyclic proteins. This is in part because limited knowledge about the locality of the NH_2- and COOH termini of putative linear precursors of cyclic proteins makes predictions about cyclization mechanisms difficult.

More than 50 naturally occurring backbone cyclized proteins have now been reported (1), with three of these being proteinase inhibitors: MCoTI-I and MCoTI-II from *Momordica cochinchinensis* (8, 9), and SFTI-1[1] from sunflower seeds (10). The Bowman-Birk (BB) inhibitor SFTI-1, shown in Fig. 1, is the focus of the current study and is the smallest known cyclic plant peptide. It is the most potent and only known cyclic BB inhibitor, having a K_i value for trypsin in the sub-nanomolar range (10–14).

Interestingly, the cyclic backbone of SFTI-1 does not appear to be essential for the high potency, because we recently showed that SFTI-1[1,14], an acyclic permutant of SFTI-1, with a break in the backbone chain between residues 1 and 14 (see Fig. 1), also has a very high affinity for trypsin (11). Although its K_i value is apparently about 20-fold weaker than the cyclic form, this difference is not highly significant due to the difficulties of accurately determining K_i values for such potent inhibitors (15). Zablotna *et al.* (14) also recently reported that there is essentially no difference between the inhibitory poten-

[1] The abbreviations used are: SFTI-1, sunflower trypsin inhibitor-1; BB, Bowman-Birk; DQF-COSY, double quantum filtered two-dimensional correlation spectroscopy; E.COSY, exclusive COSY; LC-MS, liquid chromatography mass spectrometry; NOE, nuclear Overhauser effect, also used for NOESY cross-peak; NOESY, two-dimensional NOE spectroscopy; TOCSY, two-dimensional total correlation spectroscopy; WATERGATE, water suppression by gradient-tailored excitation; r.m.s.d., root mean square deviation.
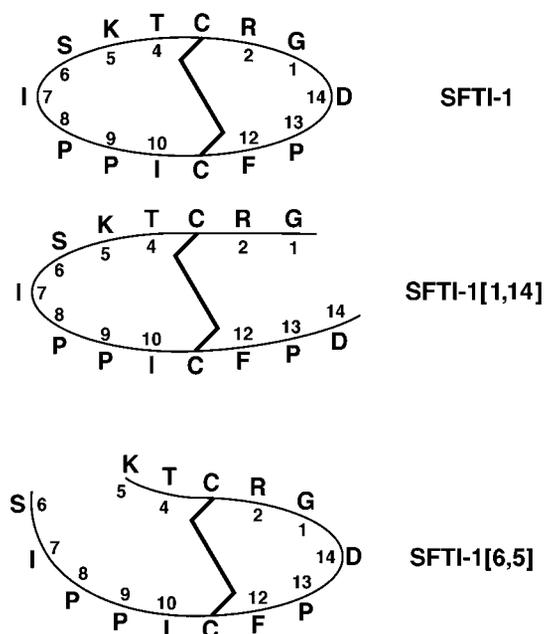
FIG. 1. **Amino acid sequences of SFTI-1, SFTI-1[6,5], and SFTI-1[14,1] shown in one-letter amino acid code.** The peptide backbone is not continuous in SFTI-1[6,5] and SFTI-1[1,14] between S6 and K5 and D14 and G1, respectively. The disulfide bond between C3 and C11 in all molecules is indicated by a *bold line*.

cies of cyclic SFTI-1 and this particular acyclic derivative. Thus, it would appear that structural conservation of the reactive loop of SFTI-1[1,14], rather than the cyclic backbone, is essential for potency and that a slight disordering of the open termini in the non-reactive loop of SFTI-1[1,14] does not significantly reduce binding efficiency to trypsin. Although the role of cyclization is not certain, it has been suggested that the cyclic backbone of native SFTI-1 possibly evolved to protect it against degradation by proteases to allow SFTI-1 to function over an extended *in vivo* lifetime (11). Understanding the mechanism of cyclization for SFTI-1 may therefore have potential applications in the pharmaceutical industry and for the design of improved peptidic drugs.

Nothing is known so far about the cyclization mechanism for SFTI-1 *in planta*, and there are no data describing enzymes with the specificity necessary to perform this task. Here we describe an *in vitro* enzymatic mechanism for the cyclization of a linear precursor of SFTI-1, SFTI-1[6,5] (Fig. 1), that is open at the scissile bond, *i.e.* the bond that would normally be cleaved in homologous substrates. Although this mechanism, based on the reversible equilibrium between species containing an intact and a cleaved peptide bond, is generally applicable to many serine proteinase inhibitors, SFTI-1 is the first example of an inhibitor in which the peptide bond resynthesis results in a backbone cyclic peptide. The gene sequence of the putative precursor protein of SFTI-1 is unknown, and the demonstration of enzymatic cyclization of a linear analogue of SFTI-1 potentially provides insight into a theoretically possible *in vivo* precursor.

In the current study we have gained a structural insight into five conformers/isomers of this putative acyclic precursor and have determined full three-dimensional structures for two of the conformers, which we compare with the structure of SFTI-1. In addition, anti-trypsin activity assays were carried out for SFTI-1[6,5], and the products of a 1:1 reaction mixture of trypsin and SFTI-1[6,5] have been analyzed. The results have provided an insight into potential mechanisms of cyclization of SFTI-1 *in vitro* and *in vivo*.

EXPERIMENTAL PROCEDURES

Solid-phase Synthesis of Acyclic SFTI-1[6,5]—The solid phase synthesis of SFTI-1[6,5] was carried out as described for another acyclic SFTI-1 variant (11) using standard Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry. The peptide was air-oxidized in 0.1 M of ammonium bicarbonate solution, pH 8.0, for 12 to 24 h and purified by reversed-phase high performance liquid chromatography using a conventional C_{18} column and a water/acetonitrile/trifluoroacetic acid elution system. Electrospray ionization mass spectrometry verified the mass of SFTI-1[6,5] with an *m/z* value of 766.3 for the doubly charged peak.

NMR Spectroscopy—Samples prepared for NMR spectroscopy contained ~2 mM SFTI-1[6,5] in 20 mM sodium phosphate (H_2O/D_2O, 90/10%, v/v) at pH 4.5. One-dimensional ^{1}H-NMR spectra recorded over the temperature range 280–310 K revealed that the dispersion of peaks was optimal at 286 K. Homonuclear two-dimensional ^{1}H spectra were recorded at 286 and 298 K on a Bruker DMX750 spectrometer. NOESY and TOCSY spectra used a WATERGATE scheme (16) for water suppression. The mixing times for the NOESY spectra were 200 and 300 ms, and the mixing time for the MLEV17 sequence (17) was 80 ms. Water suppression in the DQF-COSY experiment was achieved using selective low power irradiation of the water resonance during a relaxation delay of 1.3 s. 16,384 data points were collected for one-dimensional spectra, and 4,096 data points in the t_2 and 512 increments in the t_1 dimension were collected for two-dimensional experiments. NOESY, TOCSY, and E.COSY spectra were recorded in D_2O at 286 K. Slowly exchanging amide protons were identified by recording a series of one-dimensional spectra and two-dimensional TOCSY spectra at 286 K over a period of 12 h on a sample immediately after dissolution in 99.996% D_2O. A sequence-specific resonance assignment was performed by standard methods (18). *Cis* proline residues were identified by strong $H^\alpha(i)$/Pro-$H^\alpha(i+1)$ NOE signals in the NOESY spectrum, whereas a *trans* conformation was evidenced by unambiguous $H^\alpha(i)$/Pro-$H^\delta(i+1)$ NOE signals and the absence of $H^\alpha(i)$/Pro-$H^\alpha(i+1)$ connectivities (18). $^3J_{NH\alpha}$ coupling constants were obtained from a high resolution one-dimensional spectrum and from line-shape analysis of the anti-phase cross-signal splitting in a high digital resolution DQF-COSY spectrum using a Lorentzian function for peak fitting. The χ_1-angles for side-chain rotamers were obtained from line-shape analysis of H^α–H^β couplings in an E.COSY spectrum together with NOE intensity patterns between resolved H^N–H^β protons. Secondary chemical shifts were obtained using published random coil values (19), and specific corrections for residues preceding a proline (19) were taken into account.

Structure Calculations—Distance constraints were derived from the intensities of cross-peak signals from NOESY spectra at 286 K recorded with a mixing time of 200 ms. The spectrum at 298 K was used to resolve ambiguities due to spectral overlap. Distance restraints were divided into three groups according to the relative NOESY cross-peak intensities: strong, upper distance limit of 2.7 Å; medium, upper distance limit of 3.5 Å; weak, upper distance limit of 5.0 Å. All structures were calculated using X-PLOR 3.851 (20) with a modified *ab initio* simulated annealing protocol (21). The structure calculation strategy was based on a three-stage simulated annealing protocol (22, 23). For details see the Supplementary Material. An iterative approach using several rounds of structure calculations with subsequent distance analysis was used to solve ambiguities in NOE cross-peak assignments. Dihedral angle constraints were introduced in the later rounds of structure calculations. Distance constraints for hydrogen bonds were generated based on deuterium exchange data and by examination of structure families calculated without hydrogen-bond constraints. The 20 structures with lowest overall energies from the final 50 structures were retained for further analysis. Structures were visualized using the program MOLMOL (24) and analyzed with PROCHECK-NMR (25).

Trypsin Assays—Trypsin assays were carried out at 303 K as described by Korsinczky *et al.* (11). The concentration of SFTI-1[6,5] was determined by amino acid analysis. SFTI-1[6,5] (0–150 nM) and trypsin (42 nM) were preincubated together for 3 h at 303 K. The reactions were initiated by addition of substrate and monitored at 405 nm with readings taken every 2 min with shaking between readings using a Molecular Devices SpectraMax 250. All reactions were carried out in duplicate, and the corresponding K_i value was determined using a tight-binding equation (26).

Mass Spectroscopy/Monitoring of the Trypsin-dependent Interconversion of SFTI-1[6,5] to SFTI-1—SFTI-1[6,5] and bovine β-trypsin (Type XIII L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated, EC 3.4.21.4) were incubated together at 303 K in a 1:1 ratio (350 nM each), at pH 8.0, 50 mM Tris HCl, 25 mM $CaCl_2$. 20-μl aliquots were taken in duplicate after 2 min and every 15 min for the first 2 h and every 30 min

for the next 2 h and then every hour. The reaction was stopped by a sudden pH drop to <1 by immediate addition of 5 μl of trifluoroacetic acid to the aliquots with vigorous shaking. The samples were subjected to LC-MS analysis, using an electrospray Qq time of flight mass spectrometer (QSTAR Pulsar, Applied Biosystems, Foster City, CA) operated under positive ion detection conditions: declustering potential of 80 V, focusing potential of 250 V, declustering potential 2 of 10 V, and ion spray sprayer level of 5000 V. Extracted ion chromatograms for the doubly protonated SFTI-1[6,5] (*m/z* 766.3) and SFTI-1 (*m/z* 757.3) were constructed, and the ratio of the two peaks was calculated using both peak area and height. The integrity of the initial measurements was determined by re-measuring the amounts of SFTI-1 and SFTI-1[6,5] from the initial reaction aliquots 1 day after the initial readings. An unchanged result for the two peptides proved that trifluoroacetic acid addition did not degrade either SFTI-1 or SFTI-1[6,5] and did completely stop the trypsin reaction.

Isolation of SFTI-1 from Sunflower Seeds—Sunflower seeds (about 1 week prior to maturation) were ground with dry ice, and the mixture was defatted with chilled petroleum spirit (10). The mixture was filtered, and a water extraction was made from the dried solid phase. The water extract was divided in two portions, and to eliminate possible protease activity one portion was treated with trifluoroacetic acid to drop the pH to <1. After centrifugation for 10 min at 15,000 rpm the supernatants of both extracts were directly subjected to LC-MS analysis and yielded identical results.

RESULTS

Multiple SFTI-1[6,5] Conformers/Isomers—The conformation of SFTI-1[6,5] in aqueous solution, pH 4.5, was determined using NMR spectroscopy. The two-dimensional spectra contained more signals than expected for a 14-amino acid peptide (Fig. 2). Because the peptide was chemically pure, the additional spin systems must derive from multiple solution conformations (*cis/trans* isomerization of the three proline residues), or co-eluting isomers. Careful analysis showed that five conformers/isomers, called **A1**, **A2**, **B**, **C1**, and **C2** (Fig. 3), were present in sufficient quantities to exhibit signal sets that were detected in both TOCSY and NOESY spectra. These five conformers/isomers can be viewed as three major species, **A**, **B**, and **C**, two of which (**A** and **C**) split further into two conformers each due to *cis/trans* isomerization of P8 or P9 in the NH₂-terminal part of the peptide. Species **A** and **C** are conformers of SFTI-1[6,5], and species **B** is an isomer, with an altered D14 configuration. For consistency and for comparison with cyclic SFTI-1, we retain the original numbering scheme, starting at G1, but note that SFTI-1[6,5] is a linear peptide with a sequence from the NH₂ to COOH termini of S6–I7 . . . T4–K5.

Species **A**, the major species of SFTI-1[6,5], exists as two conformers (**A1** and **A2**) (Fig. 3) evidenced by two sets of spin systems for residues S6 to P9 that differ only in *cis/trans* isomerization of P8, with the *trans* conformation (**A2**) being more abundant. For species **B** all three Xaa-Pro bonds adopt a *trans* conformation as for **A2**, but the backbone chemical shifts are notably different from those of species **A** (Fig. 2). The most striking feature that distinguishes species **B** from all the other species is the presence of an iso-aspartate residue at position 14, as deduced from the clear absence of a $H^{\alpha}(i)H^{N}(i+1)$ sequential NOE, together with strong $H^{\beta}(i)H^{N}(i+1)$ NOEs. Furthermore, the high $^{3}J_{H\alpha HN}$ coupling constant (>9 Hz) of residue 14 of species **B** correlates with an extended ϕ-angle of 120 ± 40°. This angle is easier to fulfill if the non-binding loop is extended by the additional methylene group resulting in a more relaxed ring structure. Species **C** exhibits an F12/P13 *cis* orientation, and the four NH₂-terminal amino acids (S6–P9) adopt two different conformations: one with P8 and P9 both in the *trans* orientation (**C1**), and one with P8 in the *trans* and P9 in the *cis* orientation (**C2**) (Fig. 3). For the NH₂-terminal segment of **C2**, only the two proline residues could be definitively assigned. Out of the three species, **A** exhibits backbone chemical shifts that most closely resemble the corresponding shifts
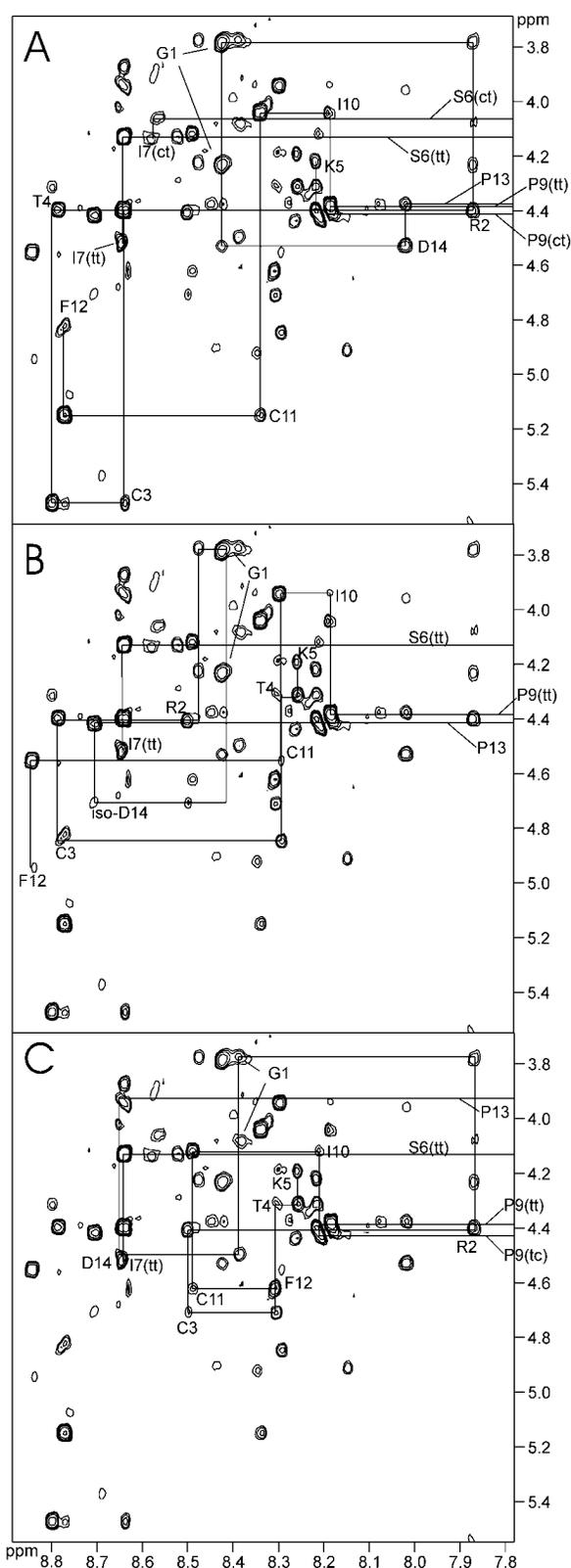
FIG. 2. **Fingerprint region of the NOESY spectrum of SFTI-1[6,5] recorded at 750 MHz and 286 K (pH 4.5).** Species **A**, **B**, and **C** (in *A*, *B*, and *C*, respectively) exhibit three different sequential connectivity patterns for the COOH-terminal residues I10–D14 through G1–K5, respectively. Similarly, two different patterns are indicated for residues S6–I7 of the NH₂-terminal segments *ct* and *tt*. For P8–P9 a third signal set (*ct*) could be assigned in the aliphatic region of the spectra (not shown). The sequential connectivity pattern is only broken at P8, P9, and P13. The sequential connectivities from I10 to P9 are indicated explicitly for all three species, and the designations *ct* and *tt* in *parentheses* describe the conformers of the NH₂-terminal part of the molecule (residues S6–P9).
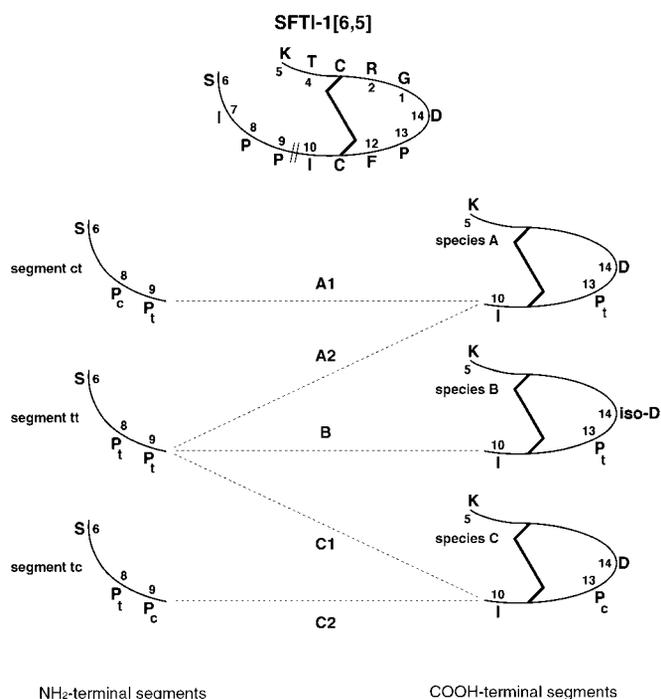
FIG. 3. **Schematic summary of the identified species (A–C) of SFTI-1[6,5].** The *top panel* shows SFTI-1[6,5] in *one-letter amino acid code*, and the disulfide bond between residues C3 and C11 is represented as a *bold line*. The *double line* between residues P9 and I10 indicates the break between the NH$_2$- and COOH-terminal segments of signal sets. The *right lower panel* shows residues K5–I10 of species **A**, **B**, and **C** (COOH-terminal signal sets) highlighting some key residues. The segments *ct*, *tt*, and *tc* (NH$_2$-terminal signal sets), containing residues S6–P9, are shown in the *left lower panel*. *Dotted lines* indicate the sequential connections between I10 and P9 of species **A**, **B**, and **C** with the corresponding segments *ct*, *tt*, and *tc* to form the five different conformers/isomers, including the sub-species of **A** (*A1* and *A2*) and **C** (*C1* and *C2*).

of native cyclic SFTI-1 (Fig. 4), and **A1** exhibits the same geometry for all three prolines as cyclic SFTI-1 (11) (Protein Data Bank code 1JBL).

Secondary Structure of SFTI-1[6,5] Conformers/Isomers—A summary of the secondary chemical shifts from I10 to T4 for all three SFTI-1[6,5] species relative to those of cyclic SFTI-1 is given in Fig. 4. Consecutive downfield shifted H$^\alpha$ resonances, indicative of extended structure, were found for G1 to C3 of both cyclic SFTI-1 and the SFTI-1[6,5] species **A**, with another distinctive downfield shifted value for C11. Species **C** exhibits consecutive upfield shifted H$^\alpha$ resonances for residues F12 to D14, indicating helical or turn-like structure. For species **B**, no indications of regular secondary structure were found. These findings were supported by an analysis of the $^3J_{NH\alpha}$ coupling constants. For species **A** they were >8 Hz for C3, C11, and I10, representing residues in an extended structure. For species **B** F12 and iso-D14 had coupling constants >8 Hz, and for species **C** only C11 exhibited a coupling constant typical of an extended conformation. No coupling constants <6 Hz, which are typical of an α-helix, were detected for any of the three species. Slow exchanging amide protons were only detected for I10, F12, G1, and R2 of species **A**. Further evidence for a β-sheet like structure for species **A** was deduced by a strong H$^\alpha$–H$^\alpha$ NOE between the two cysteine residues. In summary, evidence for regular secondary structure could only be deduced for species **A**, indicating a small anti-parallel β-sheet.

Structure Calculations—Structure calculations were performed for the four conformers and one isomer of SFTI-1[6,5]. Input data and structure statistics are summarized in Table I. The structure calculations provided well-defined structures for
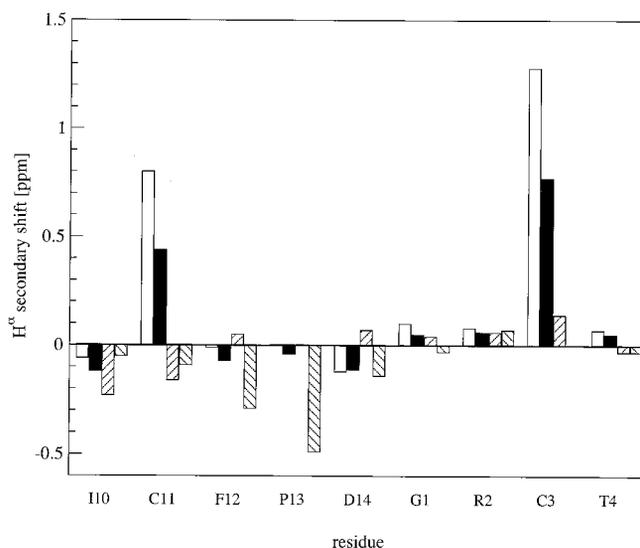
FIG. 4. **H$^\alpha$ secondary chemical shifts of cyclic SFTI-1 in 20% TFE (*white*) (pH 4.5) recorded at 750 MHz and 273 K, species A of SFTI-1[6,5] in water (*black*), species B of SFTI-1[6,5] in water (*hatched lines* sloping down to the left) and species C of SFTI-1[6,5] in water (*hatched lines* sloping down to the right) (pH 4.5) recorded at 750 MHz and 286 K.** K5 is omitted, because it is the open COOH-terminal residue in SFTI-1[6,5], and therefore, a comparison to the closed K5 residue of cyclic SFTI-1 is not meaningful.

the two conformers of species **A** with a backbone r.m.s.d. for the final 20 structures of 0.49 Å for **A1** and 0.45 Å for **A2** (protein data bank codes: **A1**, 1O8Z; **A2**, 1O8Y).

Species **A** exhibits a hairpin structure that is similar for both conformations (**A1/A2**) and closely resembles the corresponding region of cyclic wild-type SFTI-1, *i.e.* a short double-stranded anti-parallel β-sheet connected by the disulfide bridge and stabilized by hydrogen bonds (Fig. 5). The difference between the two conformations is the *cis/trans* isomerization of P8. The I7–P8 bond shows a *cis* orientation in the minor conformer **A1**, essentially maintaining the binding loop orientation, even though the backbone is open between residues K5 and S6. In the major conformer **A2**, P8 is in a *trans* configuration, giving the binding loop an extended conformation (Fig. 5).

The structure calculations for species **B** and the two conformations of species **C** did not result in convergent structures due to limited NOE information. However, it could be deduced, from chemical shift data and the absence of certain characteristic NOEs, that the small β-sheet found for both conformations of species **A** is not present in species **B** and **C**.

Trypsin Catalyzed Cyclization of SFTI-1[6,5]—Although trypsin is normally used to facilitate proteolysis, *i.e.* peptide bond cleavage, this equilibrium reaction is in principle reversible, and we were interested to see if presentation of an acyclic precursor species would result in cyclization. Therefore, a 1:1 mixture of SFTI-1[6,5] with bovine β-trypsin was incubated at 303 K (pH 8). Aliquots of the reaction mixture were taken over a 24-h period, and LC-MS analysis confirmed that cyclic SFTI-1 was indeed produced. The ratio of the intact and cleaved form of SFTI-1 was determined by quantitative analysis of the peaks with masses corresponding to SFTI-1[6,5] and cyclic SFTI-1 (Fig. 6). After 2 min of reaction time, cyclic SFTI-1 could be detected. A steady state was reached after ~130 min and resulted in ~90% SFTI-1 to 10% SFTI-1[6,5] (Fig. 6). Coelution experiments with added synthetic SFTI-1 and SFTI-1[6,5] confirmed that the resulting molecules are indeed the cleaved and intact inhibitor.

The above procedure was repeated starting with cyclic SFTI-1 under the same conditions. After a reaction time of ~2

TABLE I
Structural and energetical statistics for the conformers/isomers of SFTI-1[6,5]

Conformers/isomers	Species				
	A		B	C	
	A1	A2	B	C1	C2
NOEs	63	73	36	47	35
θ-Angle restraints	3	3	2	1	1
χ_1-Angle restraints	3	3	2	2	2
Hydrogen-bonds	4	4			
Energy statistics over 20 structures					
			kcal/mol		
Overall	30.05 ± 0.17	29.91 ± 0.03	29.93 ± 0.05	32.86 ± 0.92	34.09 ± 1.12
Bond length	1.07 ± 0.01	1.05 ± 0.005	1.05 ± 0.008	1.25 ± 0.11	1.29 ± 0.08
Bond angles	26.22 ± 0.10	26.12 ± 0.03	26.16 ± 0.07	27.74 ± 0.50	28.86 ± 0.86
Improper	1.88 ± 0.01	1.89 ± 0.002	1.89 ± 0.004	1.99 ± 0.13	2.08 ± 0.16
van der Waals	0.84 ± 0.04	0.84 ± 0.02	0.83 ± 0.03	1.54 ± 0.35	1.64 ± 0.31
NOEs	0.04 ± 0.06	0.005 ± 0.01	0.003 ± 0.01	0.33 ± 0.32	0.23 ± 0.23
Angle restraints	0.0 ± 0.0	0.00001 ± 0.00006	0.000003 ± 0.00001	0.0005 ± 0.0021	0.00004 ± 0.0001
Atomic RMSDs values from the average structures					
			Å		
Overall					
Backbone	0.492	0.450	2.324	2.104	2.288
Heavy atoms	1.206	1.067	3.462	3.152	3.089
C11–3					
Backbone	0.134	0.093	1.157	1.002	0.916
Heavy atoms	0.940	0.832	2.361	2.077	1.906
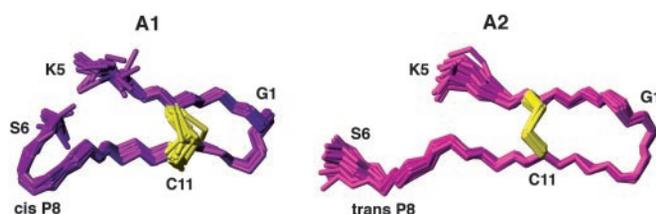
FIG. 5. **Ensemble of 20 structures of the most abundant species (species A) of SFTI-1[6,5].** Conformer **A1** (*cis* I7–P8 conformation, pdb code 1O8Z) and conformer **A2** (*trans* I7–P8 conformation, pdb code 1O8Y) superimposed over the backbone heavy atoms of all residues. The backbone of the conformers **A1** and **A2** are colored *purple* and *pink*, respectively, and the disulfide bond between C3 and C11 is colored *yellow*. The P1–P1′ site occurs at the break in the backbone between residues K5 and S6.
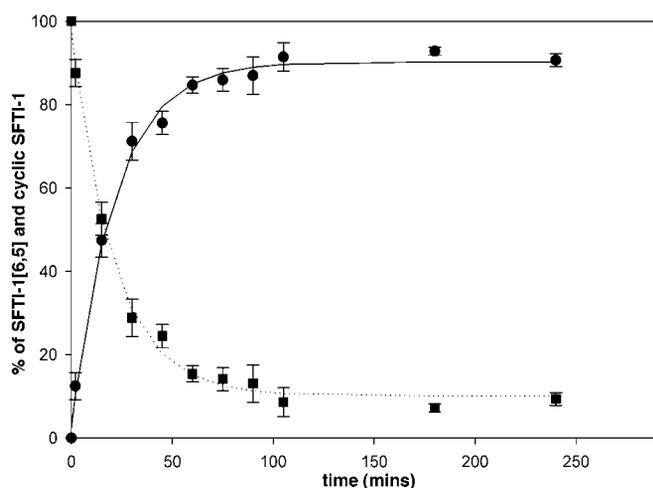
FIG. 6. **Trypsin catalyzed cyclization of SFTI-1[6,5].** Formation of cyclic SFTI-1 over time is represented by a *solid line* and *filled circles* with *error bars*. Conversion of SFTI-1[6,5] over time is represented by a *dotted line* and *filled squares* with *error bars*.

h, cleaved SFTI-1[6,5] in a yield of about 7% was recovered. Therefore, a ratio of cyclic SFTI-1 to SFTI-1[6,5] of ~9:1 was reached regardless of whether the reaction was initiated with SFTI-1 or SFTI-1[6,5].

Isolation and Verification of the Cyclic Nature of SFTI-1 in Sunflower Seeds—The previously reported isolation procedure of SFTI-1 from sunflower seeds (10, 27) involved the incubation of SFTI-1 on a trypsin affinity column for up to 2 h. Given that SFTI-1[6,5] can be 90% converted into the cyclic form when incubated with trypsin in a similar time frame and that Luckett *et al.* (10) reported a ratio of ~9:1 of intact to cleaved SFTI-1 after trypsin affinity column purification, the question was raised as to whether the peptide is actually acyclic in its natural state and whether the exposure to trypsin on the affinity column may have made it cyclic. We therefore repeated the isolation of SFTI-1 from fresh sunflower seeds with a method similar to that described by Luckett *et al.* (10) but omitted the trypsin affinity column step. In the water extract the mass of the doubly charged ion of cyclic SFTI-1 ($M^{++} = 757.32$) was detected, with neither the single nor doubly charged ion of acyclic SFTI-1[6,5] being present. In a separate experiment, we added synthetic cyclic SFTI-1 to the water extract, and subsequent LC-MS analysis of this mixture resulted in one sharp peak with the correct mass for SFTI-1. This confirms unambiguously that SFTI-1 is present naturally in cyclic form in the seed.

Trypsin Inhibition Assays—Trypsin assays were carried out as previously described (11). SFTI-1[6,5] inhibited bovine β-trypsin with a K_i of 0.54 ± 0.1 nM (Fig. 7), a result identical to that for cyclic SFTI-1. Interestingly though, cyclic SFTI-1[6,5] shows inhibitory activity with minimal incubation, whereas SFTI-1[6,5] requires a 2- to 3-h preincubation with trypsin to show substantial inhibitory activity, thus making SFTI-1[6,5] a less effective inhibitor. This delay in complex formation is further examined in the discussion section.

DISCUSSION

Three-dimensional Structure of SFTI-1[6,5]—In this report we have synthesized an acyclic permutant of SFTI-1, SFTI-1[6,5], determined the three-dimensional structures of two of its solution conformers and obtained valuable structural information on three other conformers/isomers. Although SFTI-1[6,5] consists of only 14 amino acids, it exhibits substantial structural heterogeneity caused by *cis/trans* isomerization of the three proline residues and by the presence of an iso-aspartate residue in one isomer.
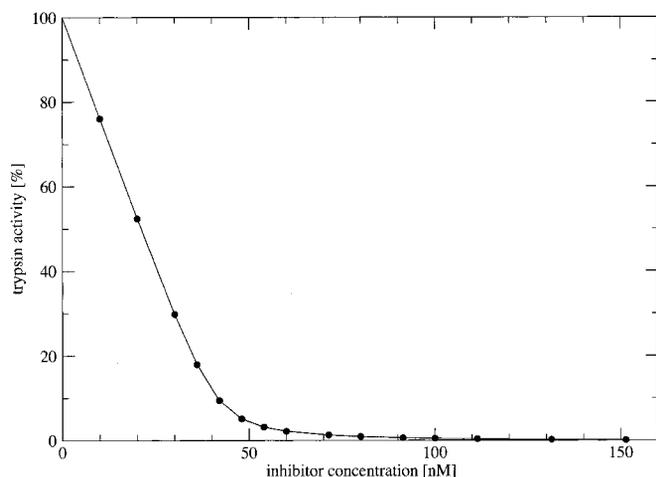
FIG. 7. **Inhibition of bovine β-trypsin by SFTI-1[6,5].** Trypsin activation was assayed using the substrate N^α-benzoyl-L-arginine-*p*-nitroanilide, as described under "Experimental Procedures."

The major species present in solution (**A**) exhibits a β-hairpin structure from C11 to C3 similar to wild-type cyclic SFTI-1 and forms two conformers (**A1** and **A2**) that differ only by a *cis/trans* isomerization about the I7/P8 peptide bond. The three-dimensional structure of the *cis* conformer **A1** of SFTI-1[6,5] closely resembles the structure of cyclic SFTI-1. The Cα atom r.m.s.d. value between the solution structures of **A1** and SFTI-1 (1JBL) is 0.98 Å and that between **A1** and the crystal structure of SFTI-1 in complex with trypsin (1SFI) is 1.21 Å, indicating that, although the backbone is open between residues K5 and S6, the binding loop orientation is essentially maintained (Fig. 5).

The second most abundant species (**B**) features backbone chemical shifts that differ notably from those of species **A** and also from those of cyclic SFTI-1. In this species all three Xaa-Pro peptide bonds are in a *trans* orientation and there is an iso-aspartate residue at position 14. The number of structural restraints for species **B**, however, was not sufficient to result in a family of converging structures, same as for species **C**.

Trypsin-catalyzed Cyclization of SFTI-1[6,5]—In addition to defining the conformational properties of SFTI-1[6,5] we examined its reactivity. We have shown that the proteolytic enzyme trypsin is able to resynthesize the peptide bond between the P1/P1′ residues (nomenclature according to Ref. 28) and consequently cyclizes the linear precursor SFTI-1[6,5] to form native cyclic SFTI-1. The cyclization reaction occurs very efficiently: a stable ratio of more than 9:1 of cyclic SFTI-1 to acyclic SFTI-1[6,5] is reached within 2 h. Resynthesis of a peptide bond between P1/P1′ residues upon interaction with a suitable proteinase has previously been demonstrated for various proteinase inhibitors (29, 30), but the resynthesis of this bond in SFTI-1[6,5] is unique because it leads to a cyclic peptide chain rather than ligation of two chains, as is the case for other classic inhibitors. No matter which form of the inhibitor was incubated with enzyme, after a sufficient reaction time over 90% of cyclic inhibitor was present. This result proves that cyclic SFTI-1 and the cleavage product SFTI-1[6,5] are interconvertible and that the enzyme-inhibitor complex **C** is the same, no matter whether it was made from the proteinase **E** and the virgin inhibitor **I** or **E**, and the modified inhibitor **I*** (29), in accordance to the equilibrium nature of the canonical mechanism of inhibition given by Finkenstadt and Laskowski (8),

$$\mathbf{E} + \mathbf{I} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \mathbf{C} \underset{k_{\text{on}}^*}{\overset{k_{\text{off}}^*}{\rightleftharpoons}} \mathbf{E} + \mathbf{I}^*$$
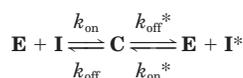
REACTION 1

Kinetically controlled dissociations of other proteinase-inhibitor complexes generally yield ratios, [**I**]/[**I***], of virgin to modified inhibitor of at least 9, which is in accordance with the ratio we have found for the SFTI-1/SFTI-1[6,5] system. This high ratio, together with the fast reaction, provides an efficient mechanism for *in vitro* cyclization of SFTI-1.

From the crystal structure of SFTI-1 in complex with trypsin it is known that the peptide bond between K5 and S6 is intact in the inhibitory complex (10). Our experimental results show that this inhibitory complex can also be formed using SFTI-1[6,5] instead of SFTI-1, in accordance with Reaction 1. Consequently, both peptides have identical inhibition constants, and SFTI-1[6,5] itself is a strong trypsin inhibitor. However, complex formation using SFTI-1[6,5] is significantly slower than complex formation using cyclic SFTI-1, mostly because peptide bond formation has to occur as a preceding step when complex formation starts with SFTI-1[6,5], in accordance with an extended version of Reaction 1,

$$\mathbf{E} + \mathbf{I} \rightleftharpoons \mathbf{L} \rightleftharpoons \mathbf{C} \rightleftharpoons \mathbf{X}^* \rightleftharpoons \mathbf{L}^* \rightleftharpoons \mathbf{E} + \mathbf{I}^*$$

REACTION 2

where **L** and **L*** are loose complexes, and **X*** is an additional intermediate (31). This is evidenced by the similar time frame for the trypsin-catalyzed cyclization of SFTI-1[6,5] and the appearance of inhibitory activity of SFTI-1[6,5].

In the case of SFTI-1[6,5] as **I*** an even more complex equilibrium is present, contributing further to its slower association rate, because it is not a structurally homogeneous compound but a mixture of conformers/isomers that are in mutual equilibrium. A consideration of geometric factors suggests that the orientation of the S6 to I10 binding loop arm, in particular the *cis* conformation of P8 in the P3′ position, is a prerequisite for an efficient interaction of SFTI-1[6,5] with trypsin. Because our studies show that none of the multiple SFTI-1[6,5] conformers/isomers are present at more than 40% in solution, there must be equilibration between the isomers/conformers to produce >90% of cyclic SFTI-1 after reaction with trypsin, namely *cis/trans* isomerization of the proline residues. There has to be equilibration between the iso-D14 and D14 isomers as well, because about 30% of all molecules contain the iso-aspartate residue. It appears that, at the beginning of the incubation with trypsin, only a fraction of all SFTI-1[6,5] molecules present in solution, those adopting the **A1** conformation, fulfills the requirements to fit into the active site pocket of the proteinase.

Mimicry of Larger Bowman-Birk Inhibitors—Like all canonical serine proteinase inhibitors, proteins of the BB inhibitor family interact with the enzymes they inhibit via an exposed surface loop that adopts a canonical proteinase inhibitory conformation. A particular feature of full-length BB inhibitor proteins is that the interacting loop is stabilized by a well-defined disulfide-linked β-sheet region. For small nonapeptides comprising the 9-amino acid active site loop and the disulfide bridge, this β-sheet is shortened and the reduced stabilization is accompanied by poorer inhibitory activities (32). These small synthetic peptide mimics of BB inhibitor proteins are able to bind to and inhibit target proteinases, but when cleaved at the active site peptide bond, they are incapable of being resynthesized by the proteinase (32, 33). NMR studies on such nonapeptides did not result in stable structures (34). In contrast to these small peptides, the β-sheet structure is intact in SFTI-1 and in both species **A** conformers of SFTI-1[6,5], even if the active site peptide bond is hydrolyzed. This stabilization is sufficient to make SFTI-1[6,5] amenable to resynthesis of the active site peptide bond. Thus, the intact β-sheet structure seems to be another prerequisite for the pre-orientation of the two "arms" of the cleaved inhibitory loop for an efficient inter-

Enzymatic Cyclization of SFTI-1

action with the proteinase. The importance of flanking residues for the short β-sheet is also supported by the fact that a synthetic undecapeptide derived from a combinatorial peptide library screening also exhibits a very stable β-hairpin structure across the active site loop (35). One goal of studies on BB inhibitors has been a determination of the minimal requirements necessary to retain the proteinase inhibitory activity of full-length inhibitors (32, 33). All investigations so far have kept the canonical loop intact, and variations have been made in specific side chains. Here we report that it is not necessary to keep the P1/P1′ site intact to maintain activity, as long as the scaffold of the loop, stabilized by the β-hairpin loop located opposite to the binding loop is maintained.

In Vivo Synthesis of Cyclic SFTI-1—Little is known about the *in vivo* cyclization mechanisms of naturally occurring cyclic peptides/proteins, and it is likely that different cyclization mechanisms might apply to different peptides. Here we have shown that the cyclic nature of SFTI-1 is not an artifact of the previously used purification scheme (10) and that SFTI-1 occurs naturally in the cyclic form in the seed. The possibility that a small linear precursor (*e.g.* SFTI-1[6,5]) is cyclized via a reaction with native proteinases present in the seed constitutes one hypothesis for the *in vivo* synthesis of cyclic SFTI-1. Various serine proteinases that might be suitable for this task are indeed present in plant seeds (36–40). However, the gene encoding SFTI-1 is not known and, hence, nor is the location of the native termini of a linear precursor peptide. Analysis of the genes encoding the cyclic peptide RTD-1 from *Macaca mulatta* has shown that more complicated cyclization mechanisms are conceivable and that the precursor is not necessarily one continuous peptide chain (41).

A second hypothesis for the synthesis of cyclic SFTI-1 can be derived from the fact that all known cyclic mini-proteins from plants appear to be derived from longer precursor proteins (1), and, therefore, both cleavage and cyclization steps seem to be involved in their synthesis. The majority of BB inhibitors from dicot plants comprise 60–90 amino acid residues and have a symmetrical double-headed structure with independent inhibitory activity, either against trypsin or chymotrypsin (32). In contrast, SFTI-1 consists of only 14 amino acid residues and has only one inhibitory loop. Its small size leads to the suggestion that SFTI-1 is possibly derived from a "normal size" BB inhibitor as a result of peptide bond cleavage followed by a transpeptidation reaction or a peptide bond double cleavage followed by a ligation reaction. A likely location for the joined ends of cyclic SFTI-1 is the D14–G1 position within the nonbinding loop (Fig. 8). The detection of an iso-aspartic acid residue in position 14 of one of the structural species of SFTI-1[6,5] is an indication of the high reactivity of this functional residue and supports the possibility of participation of this residue in a transpeptidation reaction. However, up to now, no putative precursor protein of the BB inhibitor family is known with D and G or alternatively N and G in the relevant positions.

From the existence of the iso-Asp residue a third alternative suggestion for *in vivo* cyclization of SFTI-1 is conceivable, starting from the putative precursor SFTI-1[1,14] (Fig. 8). In general, peptide bonds are subject to both acid and base hydrolysis. Most peptide bonds are stable except those in Asp-Xaa sequences, with cleavage being particularly rapid at D–G and D–P bonds due to a reduced energy barrier (42). Because peptide bond hydrolysis/synthesis is an equilibrium reaction, it can theoretically occur in both directions. For linear peptides the reversal reaction is highly unlikely due to separation of the reaction products. However, formation of the D-G bond in the case of a linear SFTI-1 precursor open between G1 and D14
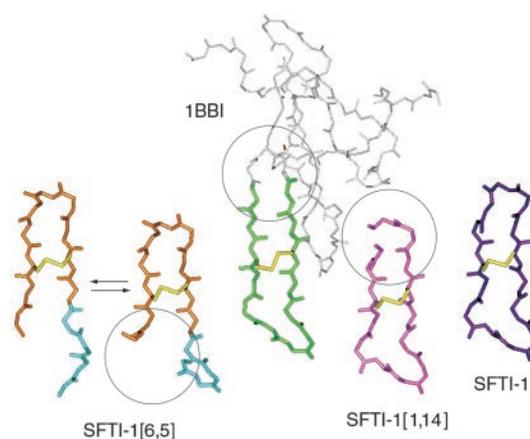
FIG. 8. **Possible mechanisms for the synthesis of cyclic SFTI-1 *in vivo*.** *Stick representation* of the solution structures of SFTI-1[6,5] conformer **A1** with the *cis* I7-P8 bond (1O8Z), SFTI-1[6,5] sub-species **A2** with the *trans* I7-P8 bond (1O8Y), the mean structure of Bowman-Birk inhibitor from Soybean (1BBI) (43), SFTI-1[1,14] (1JBN (11)), and cyclic SFTI-1 (1JBL (11)) are depicted. The *circle* around the trypsin cleavage product SFTI-1[6,5] represents the region where cyclization does occur *in vitro* to form SFTI-1 (first hypothesis). The *circle* on 1BBI represents the region where putative cyclization may occur to form the smaller protein SFTI-1 (second hypothesis), and the *circle* on the putative gene product SFTI-1[1,14] (or alternatively a peptide with a COOH-terminal N instead of D) represents the region where cyclization may occur to form SFTI-1 (third hypothesis).

leading to cyclic SFTI-1 is not as difficult, because both termini are in close proximity due to the disulfide bond and the continuous sequence of the putative precursor. Because the side-chain carbonyl group of D14 is involved in a hydrogen bond with the amide proton of R2, it is available for a nucleophilic attack of the amide nitrogen of G1 for succinimide formation. Hydrolysis on either side of the succinimide nitrogen atom results in an D or iso-D residue within the peptide chain. A feasible mechanism for cyclization of SFTI-1 from SFTI-1[1,14] is the formation of a cyclic anhydride within D14, which would lead after nucleophilic attack of G1 to a peptide bond between D14 and G1. However, this reaction does not seem to occur spontaneously as no cyclic product was detected in our investigation of this putative precursor of SFTI-1 (11) indicating that a catalytic enzyme might be necessary. Alternatively, a COOH-terminal N instead of D residue within a putative precursor peptide could form an intra-residual succinimide that might more readily be attacked by the NH_2-terminal G1 than the cyclic anhydride of aspartate. Subsequent deamidation of the asparagine residue would lead to the formation of cyclic SFTI-1 with its D–G peptide bond. Interestingly, the cyclic inhibitor MCoTI-II from *Momordica cochinchinensis* exhibits also an D–G peptide bond that gives rise to two isomers of MCoTI-II (9).

CONCLUSIONS

We have shown in the current study that SFTI-1[6,5], an acyclic permutant of the cyclic peptide SFTI-1, adopts several equilibrating conformers/isomers in solution, one of which is very similar to the cyclic native peptide. Furthermore, treatment of this acyclic permutant with trypsin results in cyclization to produce SFTI-1. This process represents a potential biosynthetic pathway for SFTI-1, in addition to being an efficient *in vitro* reaction for backbone cyclization. Homology with larger BB inhibitors suggests, however, that a more likely possibility would involve processing near residues D14–G1, opposite to the reactive site (Fig. 8). Further studies are under way to determine the biosynthetic origin of this potent cyclic inhibitor.

REFERENCES

1. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27,** 132–138
2. Li, P., and Roller, P. P. (2002) *Curr. Top. Med. Chem.* **2,** 325–341
3. Iwai, H., and Pluckthun, A. (1999) *FEBS Lett.* **459,** 166–172
4. Craik, D. J., Simonsen, S., and Daly, N. L. (2002) *Curr. Opin. Drug Discov. Devel.* **5,** 251–260
5. Gewolb, J. (2002) *Science* **295,** 2205–2207
6. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294,** 1327–1336
7. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10614–10619
8. Finkenstadt, W. R., and Laskowski, M., Jr. (1967) *J. Biol. Chem.* **242,** 771–773
9. Hernandez, J. F., Gagnon, J., Chiche, L., Nguyen, T. M., Andrieu, J. P., Heitz, A., Trinh Hong, T., Pham, T. T., and Le Nguyen, D. (2000) *Biochemistry* **39,** 5722–5730
10. 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
11. 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
12. 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
13. Descours, A., Moehle, K., Renard, A., and Robinson, J. A. (2002) *Chembiochem.* **3,** 318–323
14. Zablotna, E., Kazmierczak, K., Jaskiewicz, A., Stawikowski, M., Kupryszewski, G., and Rolka, K. (2002) *Biochem. Biophys. Res. Commun.* **292,** 855–859
15. Waltham, M. C., Holland, J. W., Robinson, S. C., Winzor, D. J., and Nixon, P. F. (1988) *Biochem. Pharmacol.* **37,** 535–539
16. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. A* **102,** 241–245
17. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65,** 355–360
18. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York
19. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81
20. Brünger, A. T. (1996) *X-PLOR Manual Version 3.851*, Yale University, New Haven, CT
21. Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M., and Pastore, A. (1995) *EMBO J.* **14,** 3572–3584
22. Nilges, M., and O'Donoghue, S. I. (1998) *Prog. NMR Spectrosc.* **32,** 107–139
23. Neudecker, P., Schweimer, K., Nerkamp, J., Scheurer, S., Vieths, S., Sticht, H., and Rösch, P. (2001) *J. Biol. Chem.* **25,** 22756–22763
24. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.* **14,** 51–55
25. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
26. Ranson, N. A., Burston, S. G., and Clarke, A. R. (1997) *J. Mol. Biol.* **266,** 656–664
27. Konarev, A. V., Anisimova, I. N., Gavrilova, V. A., Rozhkova, V. T., Fido, R., Tatham, A. S., and Shewry, P. R. (2000) *Theor. Appl. Genet.* **100,** 82–88
28. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162
29. Ardelt, W., and Laskowski, M. J. (1985) *Biochemistry* **24,** 5313–5320
30. Otlewski, J., Zbyryt, T., Dryjanski, M., Bulaj, G., and Wilusz, T. (1994) *Biochemistry* **33,** 208–213
31. Laskowski, M., and Qasim, M. A. (2000) *Biochim. Biophys. Acta* **1477,** 324–337
32. McBride, J. D., and Leatherbarrow, R. J. (2001) *Curr. Med. Chem.* **8,** 909–917
33. Nishino, N., Aoyagi, H., Kato, T., and Izumiya, N. (1977) *J. Biochem. (Tokyo)* **82,** 901–909
34. Pavone, P., Isernia, C., Saviano, M., Falcigno, L., Lombardi, A., Paolillo, L., Pedone, C., Buoen, S., Naess, H. M., Revheim, H., and Eriksen, J. A. (1994) *J. Chem. Soc. Perkin Trans.* **2,** 1047–1053
35. Brauer, A. B., Kelly, G., McBride, J. D., Cooke, R. M., Matthews, S. J., and Leatherbarrow, R. J. (2001) *J. Mol. Biol.* **306,** 799–807
36. Morita, S., Fukase, M., Hoshino, K., Fukuda, Y., Yamaguchi, M., and Morita, Y. (1994) *Plant Cell Physiol.* **35,** 1049–1056
37. Morita, S., Fukase, M., Hoshino, K., Fukuda, Y., Yamaguchi, M., and Morita, Y. (1996) *J. Biochem. (Tokyo)* **119,** 711–718
38. Morita, S., Fukase, M., Yamaguchi, M., and Morita, Y. (1996) *J. Biochem. (Tokyo)* **119,** 1094–1099
39. Wolf, M. J., and Storey, R. D. (1990) *Phytochemistry* **29,** 2419–2423
40. Dryjanski, M., Otlewski, J., Polanowski, A., and Wilusz, T. (1990) *Biol. Chem. Hoppe-Seyler* **371,** 889–895
41. Tang, Y. Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) *Science* **286,** 498–502
42. Wang, W. (1999) *Int. J. Pharm.* **185,** 129–188
43. Werner, M. H., and Wemmer, D. E. (1992) *Biochemistry* **31,** 999–1010