Structural Insights into the Role of the Cyclic Backbone in a Squash Trypsin Inhibitor*

Received for publication, October 17, 2013 Published, JBC Papers in Press, October 29, 2013, DOI 10.1074/jbc.M113.528240

Norelle L. Daly1‡, Louise Thorsholm1†1, Kathryn P. Greenwood1, Gordon J. King1, K. Johan Rosengren1‡3, Begoña Heras2‡4, Jennifer L. Martin1‡, and David J. Craik1‡5

From the 1Institute for Molecular Bioscience and 2School of Biomedical Sciences, The University of Queensland, Brisbane 4072, the 4Centre for Biodiscovery and Molecular Development of Therapeutics and School of Pharmacy and Molecular Sciences, James Cook University, Cairns 4878, and the 1Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Melbourne 3086, Australia

Background: MCoTI-II is a potent serine protease inhibitor with a cyclic protein backbone.

Results: Its three-dimensional structure and dynamics in complex with trypsin were determined using x-ray and NMR methods.

Conclusion: The cyclic backbone of MCoTI-II participates in favorable interactions with trypsin.

Significance: We demonstrate a new role for backbone cyclization in enhancing enzyme inhibitory activity.

MCoTI-I and II are potent trypsin inhibitors isolated from the seeds of Manodrica cochinchinensis and are members of the squash trypsin inhibitor family (1, 2). Peptides in this family contain ~30 amino acids and have a cystine knot (or knottin) arrangement of their three conserved disulfide bonds. In contrast to the majority of squash trypsin inhibitors, MCoTI-I and II both contain an amide-cyclized backbone and consequently are classified as cyclotides (3), ultra-stable plant defense mini-proteins that contain a cystine knot and cyclic backbone. This unique structural motif, present only in the cyclotides, is referred to as a cyclic cystine knot (4).

Cyclotides are gene-encoded, and although the organization of the genes differs in different plants, the common feature is a conserved Asn/Asp residue at the C-terminal of the mature peptide sequence (5–8). This C-terminal residue appears to be critical for cyclization via an asparaginyl endopeptidase. The sequences of selected cyclotides, including MCoTI-I and II, are shown in Fig. 1A. There are six backbone loops between the cysteine residues, with loop 6 forming as a result of the cyclization. MCoTI-I and II were the first cyclic peptides isolated from M. cochinchinensis, but three additional cyclic peptides with homologous sequences to MCoTI-I and II have recently been discovered in this plant (7). A database (Cybase) (9) has been established to collate the sequences of cyclotides, and there are now >250 sequences characterized.

Both MCoTI-I and MCoTI-II are stable in human serum, with MCoTI-I reported to have a half-life of more than 2 days (10, 11). The stability of cyclotides, coupled with their high sequence diversity, has led to the suggestion they might be useful in the design of novel drug leads (12). In support of this suggestion, there are now several examples of the use of cyclotides as templates, including re-engineering kalata B1, the prototypical cyclotide, to incorporate and stabilize a bioactive peptide sequence with anti-cancer activity (13). Similarly, recent studies have shown that MCoTI-II can be used as a template to develop therapeutic angiogenic agents (11), as well as inhibitors of β-tryptase and human leukocyte elastase, therapeutic targets in inflammatory disorders (14, 15). MCoTI-I has been used to design inhibitors of HIV-1 viral replication (16) and a compound that activates the p53 tumor suppressor pathway in vivo (10). The latter example is an intracellular target, consistent with the finding that MCoTI-II is able to penetrate cells (17).

To exploit the potential of cyclotide frameworks such as MCoTI-II in drug delivery, it is important to understand their structure and dynamics. Several three-dimensional structures

* This work was supported by the Australian Research Council (DP0984390) and the National Health and Medical Research Council of Australia (APP1028509).

The atomic coordinates and structure factors (code 4GUX) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 Both authors contributed equally to this work.

2 An Australian Research Council future fellow.

3 A National Health and Medical Research Council Career Development Award fellow.

4 A La Trobe Institute for Molecular Science fellow.

5 An Australian Research Council Australian Laureate Fellow (FL0992138).

6 An National Health and Medical Research Council Professorial Research fellow (APP1026501). To whom correspondence should be addressed: The University of Queensland, St. Lucia, QLD 4072, Australia. Tel.: 61-7-3346-2019; E-mail: d.craik@imb.uq.edu.au.
Structural Analysis of MCoTI-II

of cyclotides have been determined using NMR spectroscopy, and they generally highlight the well defined structures present in this family (18). Surprisingly, however, the solution structure of MCoTI-II showed loop 6 to be highly disordered as shown in Fig. 1B (3, 19). Although the structures of a range of cyclotides have been determined, there have been limited studies on the dynamics and interactions of cyclotides with biological targets. The most extensive analysis was done on MCoTI-I (20) where analysis of the dynamics with NMR spectroscopy suggested a lower order parameter for the cyclization loop (loop 6), indicative of greater structural flexibility for this region of the peptide compared with other loops. Interestingly, an analysis of the dynamics of MCoTI-I bound to trypsin was also carried out, and it was reported that several residues, including the active site lysine residue, become more flexible upon binding to trypsin (20). This is a very surprising result given that the MCoTI-I peptide has picomolar Ki values against trypsin (21) and that binding interactions involving localized increases in flexibility, rather than increased flexibility (22, 23). However, a limited number of binding interactions involving localized increases in flexibility reflect enthalpy-entropy compensation are known to occur (23).

In this study, we have analyzed the dynamics of MCoTI-II in solution and determined its crystal structure bound to trypsin. This is the first structure of a peptide containing a cyclic cystine knot motif bound to trypsin and, indeed, the first resolution of the structure of any cyclotide bound to a defined macromolecule. Although MCoTI-II differs from MCoTI-I by just two residues (Fig. 1A), our results differ strikingly with the study by Puttamadappa et al. (20) and definitively reveal a tight rigid complex formed between trypsin and MCoTI-II. We believe that the earlier result of an apparently looser structure on binding probably reflects a previously undetected cleavage reaction in the complex. Furthermore, our study reveals a number of interactions between the cyclization loop and trypsin and provides new insights into the role of the cyclic backbone in enhancing the potency of protease inhibitors.

EXPERIMENTAL PROCEDURES

Extraction of MCoTI-II from M. cochinchinensis—Native MCoTI-II was purified from the seeds of M. cochinchinensis as described previously (19). The mass of the purified peptide was analyzed by electrospray ionization mass spectrometry, and purity was assessed by analytical RP-HPLC.

Synthesis of MCoTI-II—MCoTI-II was synthesized using native chemical ligation with methods established previously for disulfide-rich cyclic peptides (24–26).

NMR Spectroscopy—T1 and heteronuclear NOE experiments on free MCoTI-II (6 mM in 100% D2O) were recorded at 293 K on 500 and 600 MHz Bruker spectrometers using pulse programs based on those described by Farrow et al. (27). A recycle delay of 2 s was used, and for the T1 experiments, variable relaxation delays of 10, 20, 40, 50, 75, 100, 200, 300, 400, 500, 700, 1000, 1250, and 1300 ms were used. The data were analyzed with Tosspin (Bruker), and the experimental 13Ca T1 and NOE data were fitted using the model-free approach (28). The fit of the experimental and calculated T1 and NOE values was assessed using an r.m.s.2 analysis (29).

\[
\text{r.m.s.} = \left( \frac{1}{4} \sum W \left( \frac{\text{exp} - \text{calc}}{\text{exp}} \right)^2 \right)^{0.5}
\]

Here, the experimental values are denoted as “exp,” and the calculated values are denoted as “calc.” The final r.m.s. was weighted toward the T1 fit, with W = 1 for T1 values and W = 0.5 for NOEs, as the T1 is more accurate than the NOE (29). The overall r.m.s. for the model-free fit was 1.97%. Based on the model-free analysis, the average overall correlation time \( \tau_0 \) was calculated to 2.4 ± 0.05 ns, the average internal correlation time \( \tau_s = 20 \pm 0.4 \) ps, and the overall average of the order parameters \( S^2 \) was calculated to be 0.84 ± 0.02. The two parameters used to describe the internal dynamics, \( \tau_s \) and \( S^2 \), were recalculated with the overall correlation time fixed to 2.4 ns for all data points. The resulting average \( \tau_s \) and \( S^2 \) were 22 ± 0.8 ps and 0.84 ± 0.03, respectively, and the overall r.m.s. value was 3.66%. All experimental and calculated T1 and NOE values are shown in Table 1 and Fig. 2.

Analytical HPLC and LC-MS—A solution of 10:1 MCoTI-II/trypsin (0.2 mM MCoTI-II and 0.02 mM L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated bovine pancreatic trypsin (Sigma Aldrich)) in 100 μl of 10 mM KH2PO4, pH 7, was incubated at room temperature for 1 week. The stability of MCoTI-II in the presence of trypsin was monitored by analytical HPLC and LC-MS after 16, 40, 64, 112, and 164 h. At each time point, 20 μl of the reaction mixture was quenched with 4 μl of 4% TFA. The samples were immediately analyzed by analytical HPLC and LC-MS.

Purification of MCoTI-II Trypsin Complex—A complex of MCoTI-II bound to L-1-tosylamido-2-phenylethyl chloromethyl ketone-treated bovine pancreatic trypsin (Sigma Aldrich)) in 100 μl of 10 mM KH2PO4, pH 7, was incubated at room temperature for 1 week. The stability of MCoTI-II in the presence of trypsin was monitored by analytical HPLC and LC-MS after 16, 40, 64, 112, and 164 h. At each time point, 20 μl of the reaction mixture was quenched with 4 μl of 4% TFA. The samples were immediately analyzed by analytical HPLC and LC-MS.

7 The abbreviations used are: r.m.s., root mean square; RP-HPLC, reverse phase high performance liquid chromatography; HSQC, heteronuclear single quantum coherence.
ethyl ketone-treated bovine pancreatic trypsin (Sigma Aldrich) was purified by cation exchange chromatography at 4 °C using FPLC (GE Healthcare). MCoTI-II and trypsin were combined in 50 mM NaCl, 50 mM Tris, pH 8, at a molar ratio of 1:2. The trypsin-MCoTI-II complex was loaded onto a 1-ml HiTrap SP HP column (GE Healthcare) and eluted using a 0.5%/min gradient of 0–20% 1M NaCl, 50 mM Tris, pH 8. The absorbance of the eluted complex was monitored at 280 nm, and mass and purity were evaluated by MALDI-TOF mass spectrometry and SDS-PAGE, respectively. FPLC-purified trypsin-MCoTI-II was concentrated in an Amicon Ultra 4 centrifugal filter device with a 5-kDa cut-off. The final protein concentration of 77 mg/ml was measured at 280 nm on a NanoDrop ND-1000 using an extinction coefficient of 37,650 M⁻¹ cm⁻¹.

Crystal Structure Determination of the MCoTI-II Trypsin Complex—Crystals were grown at 4 °C by hanging drop vapor diffusion in drops composed of 1-µl well condition and 1 µl 23 mg/ml trypsin-MCoTI-II complex in 24-well plates containing shallow gradients of PEG 3350 and ammonium acetate in 0.1M BisTris, pH 8.5, buffer. The buffer used for crystallization was 15 mM Tris, pH 8.0, containing 30 mM NaCl, and crystals grown with 28% PEG 3350 and 0.24 mM ammonium acetate in 0.1 M BisTris, pH 6.5, buffer as the well solution were selected for data collection and cryoprotected by the addition of ethylene glycol to 20% (v/v). Data were measured on beamline MX1 (30) of the Australian Synchrotron (wavelength, 0.953645 Å). Data were processed using HKL2000 (31), and the structure was solved by molecular replacement with Phaser (32) from the structure of bovine trypsin (Protein Data Bank code 2UUY (33)). Refinement was performed with Phenix (34), and model building was performed with COOT (35). Statistics are given in Table 2.

RESULTS

Synthesis and Characterization of MCoTI-II—MCoTI-II was synthesized using native chemical ligation (36) with methods established previously for other cyclotides (26). Cyclization and oxidation were achieved in a single step in 0.1 M ammonium bicarbonate, pH 8. Purified samples were analyzed by RP-HPLC and electrospray ionization MS.

Analysis of Backbone Dynamics of MCoTI-II—To determine whether the disorder in the three-dimensional structure of MCoTI-II is a result of flexibility or simply a lack of distance restraints in the structure calculations, the backbone dynamics of MCoTI-II in solution were analyzed using NMR relaxation experiments. Natural abundance 13C relaxation data were recorded at several field strengths on the synthetic version of MCoTI-II. In these spectra, the signals for Cys⁰¹ and Arg²⁴ appear to be overlapped with Lys¹⁰ and Lys⁹, respectively, and hence, these residues were not part of the relaxation analysis, which included 23 non-Gly residues of MCoTI-II. To describe the MCoTI-II backbone flexibility, the 13C relaxation data were analyzed using the Lipari-Szabo model-free approach (37), in which the order parameter, S², reflects the nano- to picosecond dynamics of the α-carbon backbone. S² can be between 0 and 1, with a high value representing small amplitude internal motions and a low value reflecting larger amplitude internal motions.

FIGURE 2. 13C relaxation data of the MCoTI-II backbone. A and B, heteronuclear NOE(13C, 1H) for MCoTI-II at 500 and 600 MHz, respectively. C and D, spin-lattice relaxation values, T₁ (s) for MCoTI-II at 500 and 600 MHz, respectively. Loop residues are shown in gray, and the cysteine residues (separating the loops) are shown in red. Loop numbers are indicated above the columns.
**Structural Analysis of MCoTI-II**

**TABLE 1**

Experimental $^{13}$C T$_1$ and NOE values for MCoTI-II at 500 and 600 MHz, and calculated $T_1$ and NOE values based on the Lipari-Szabo model-free approach

The r.m.s. was calculated according to the equation under “Experimental Procedures” and measures the percent deviation between the experimental value and the calculated/fitted value. The overall average correlation time ($\tau_c$) was calculated to be 2.4 ns in a preliminary free-fitting analysis. The calculation of the internal motional parameters, order parameters ($S^2$), and internal correlation time ($\tau_c$) was then carried out with $\tau_c$ fixed at 2.4 ns.

| Residue | NOE | $\tau_c$ (ns) | 13C-1H NOE | RMSD (Å) |
|---------|-----|---------------|-------------|----------|
| MC | 0.29 | 0.56 | 0.75 | 0.88 |
| 15 | 0.31 | 0.41 | 0.75 | 1.42 |
| 27 | 0.19 | 0.31 | 0.14 | 1.25 |
| 46 | 0.19 | 0.31 | 0.15 | 1.37 |
| 68 | 0.28 | 0.39 | 0.38 | 1.23 |
| 85 | 0.36 | 0.38 | 0.86 | 1.27 |
| 100 | 0.28 | 0.39 | 0.16 | 1.20 |
| 210 | 0.28 | 0.39 | 0.16 | 1.20 |
| 36144 | 0.28 | 0.39 | 0.16 | 1.20 |

The $T_1$ and NOE data for MCoTI-II are given in Table 1 and Fig. 2. The $^{13}$C-1H HSQC NMR spectrum of MCoTI-II and $S^2$ values for the intercysteine loops and individual residues are shown in Fig. 3. The overall correlation time for MCoTI-II is 2.4 ± 0.2 ns, and the average order parameter is 0.84 ± 0.06. Loop 6, with a significantly lower order parameter than the other loops ($S^2 = 0.76$) is more flexible in solution (Fig. 3).

In addition, the order parameter for loop 1 ($S^2 = 0.79$, excluding residues 9 and 10) is smaller than the average order parameter, indicating that this loop is more flexible than other parts of the molecule. Interestingly, the residues on each side of Lys, the active site residue for trypsin inhibitory activity, have order parameters ~0.1 lower than the rest of the loop, with Pro$_5$ and Ile$_7$ at $S^2 = 0.75$. Loops 2, 3 and 4 have order parameters ($S^2 = 0.84$, 0.85, and 0.85, respectively) close to the average order parameter. Loop 5 appears to be the most ordered loop with $S^2 = 0.88$. This high value reflects the conformation of loop 5, which comprises a $\beta$-hairpin, consistent with a lack of flexibility in this loop. The cysteine residues have a considerably higher average order parameter ($S^2 = 0.90$) than the rest of the molecule, indicating that the core of the cystine knot motif is the most rigid part of MCoTI-II.

**Determination of the Crystal Structure of MCoTI-II Bound to Trypsin**—The crystal structure of MCoTI-II bound to trypsin was determined at 1.8 Å resolution (Fig. 4A and Table 2) and indicates that MCoTI-II binds to trypsin in a similar orientation to other squash trypsin inhibitors such as CMTI-I (38), with loop 1, the reactive site loop, forming the primary contact with trypsin, and Lys$^6$ from this loop engaging the active site. Upon binding to trypsin the Lys$^6$ side chain adopts a single conformation, in contrast to the disorder observed in the solution structure. The side chain of the active site Lys residue is shown in Fig. 4B, highlighting its well defined conformation.

Loop 6, the cyclization loop, which is not present in acyclic squash trypsin inhibitors, forms several interactions with trypsin, including a direct hydrogen bond between the side chain of Ser$^{11}$ and the backbone of Gly$^{219}$ of trypsin. Interestingly, despite the interactions of loop 6 with trypsin, the crystallographic B-factors indicate that this region of the structure remains the most flexible (Fig. 4C). The cyclization loop of MCoTI-II, therefore, has a propensity for flexibility whether bound to trypsin or free in solution.

A comparison of the bound and solution structures of MCoTI-II is shown in Fig. 4D. Overall, the structure of bound MCoTI-II is similar to that determined in solution. An overlay of the lowest energy NMR-derived structure with the bound form has an r.m.s. deviation of 1.7 Å over the backbone. The orientation of the active site Lys side chain and the structure of loop 6 differ significantly in the lowest energy solution and bound structures. However, in the structural ensemble representing the solution structure, the Lys side chain and loop 6 have a range of conformations that encompass the orientations found in the bound state.

Analysis of the crystal contacts in each of the three distinct complexes observed in the unit cell indicate that in complex 3, loop 6 is more ordered than in the other two complexes because it interacts with a neighboring cyclotide. In all three complexes, loop 2 has limited contacts with neighboring trypsin molecules. Loop 1, including the active site residue, is buried in the complex and does not have crystal contacts with neighboring molecules. This excludes the possibility that any changes in the mobility of loop 1 on binding are the result of crystal contacts.

**HPLC Analysis of the MCoTI-II Complex with Trypsin**—To determine the stability of the MCoTI-II complex in solution, MCoTI-II was incubated with trypsin and subsequently analyzed with RP-HPLC and mass spectrometry. Although highly resistant to trypsin, MCoTI-II is slowly cleaved after extended incubation with trypsin as shown in Fig. 5. After a 40-h incubation, a peak eluting at 20 min became evident, and the intensity of this peak increased over time. The LC-MS analysis of this peak indicates an increase in mass of 18 Da, compared with cyclic MCoTI-II, which is consistent with conversion to hydrolyzed MCoTI-II, presumably with the backbone linearized C-terminally to the active site Lys. After 112 h of incubation, another peak with a mass of 3245.4 Da appears at 18.5 min. This peak corresponds to MCoTI-II with the binding site Pro-Lys fragment excised. Trypsin has been shown to cleave at the N terminus of Pro in other proteins (39) and appears to do so in the hydrolyzed form of MCoTI-II subsequent to initial cleavage at Lys.

**DISCUSSION**

In this study, we analyzed the dynamics of MCoTI-II, a promising scaffold in drug design, and determined its structure bound to trypsin. Despite flexibility in loops 1 and 6 in solution, a well defined complex is formed upon binding to trypsin.
TABLE 2
Data collection and refinement statistics

| Data collection          | C2 |
|--------------------------|----|
| Space group              | 3  |
| No. of trypsin complexes/a.u. | 3 |
| Cell dimensions          | a = 136.1, b = 71.9, and c = 108.5 Å; |
|                          | α = 90.0, β = 119.8, and γ = 90.0° |
| Resolution (Å)           | 31.69–1.80 (1.83–1.80) |
| Rsym                     | 0.103 (0.449) |
| I/I₀                     | 10.7 (2.8) |
| Completeness (%)         | 98.7 (97.7) |
| Redundancy               | 3.7 (3.7) |

| Refinement               | 1.8 |
|--------------------------|-----|
| No. of reflections (R_work/R_free) | 78,423/4,122 |
| R_work/R_free            | 0.157 (0.220)/0.193 (0.246) |
| No. of atoms             | Protein 5,677 |
|                          | Water 1,233 |
|                          | Other 11 |
| Ramachandran Favored     | 97.4% |
| Outliers                 | 0.0% |
| r.m.s. deviations        | Bond lengths (Å) |
|                          | 0.006 |
|                          | Bond angles 1.02° |

a.a.u., asymmetric unit.

A. 13C NMR relaxation analysis of MCoTI-II. A, 13C-HSQC spectrum of MCoTI-II with chemical shift assignments of the CoHα backbone; glycine residues are not shown. (The unlabeled peak is a Cβ from Ser15). B, average order parameters for loop 1–6 (gray) and the cysteine residues in the cystine knot (black) of MCoTI-II. C, order parameters for each assigned residue. Residues in the loops are shown in gray, and the cysteine residues are shown in black. The dashed line indicates the overlapped Cys31 residue that was not included in the analysis.

FIGURE 4. Three-dimensional structure of MCoTI-II bound to bovine trypsin. A, MCoTI-II is shown in magenta, and trypsin is shown in gray. B, the active site lysine residue is highlighted in blue. C, the crystallographic B-factors are mapped onto the MCoTI-II structure. The thicker lines represent higher B-factors. D, a superposition of the bound (magenta) and free (green) structures of MCoTI-II. The positions of the intercysteine loops are labeled with Arabic numbers in B–D. The coordinates for the complex of MCoTI-II bound to trypsin have been deposited in the Protein Data Bank code 4GUX.

DECEMBER 13, 2013 • VOLUME 288 • NUMBER 50 • JOURNAL OF BIOLOGICAL CHEMISTRY

Structural Analysis of MCoTI-II
Unexpectedly, the cyclization loop is involved in the interaction with trypsin, providing an explanation for the highly potent subnanomolar inhibitory activity observed for MCoTI-II (21). The similarity between the crystal structure and solution structures of MCoTI-II suggests that no major conformational changes occur on binding, but there are minor local changes, including the selection of a specific conformation of the active site Lys residue in the complex. Analysis of the dynamics of MCoTI-II free in solution suggests that the flexibility observed for the residues adjacent to the active site residue and for residues in loop 6 might contribute to the peptide being able to adopt a conformation allowing optimal binding. The flexibility observed for the cyclization loop in MCoTI-II is consistent with the presence of several small amino acids such as glycine and serine and is also consistent with the dynamics analysis of MCoTI-I in solution (20). Despite this localized flexibility in solution, MCoTI-II is very stable in human serum, as is the case for other cyclotides such as kalata B1 (11). Presumably, the well defined structural core, conferred primarily by the cystine knot motif, is sufficient to prevent degradation by proteases. The NMR relaxation data reported here unequivocally demonstrate the rigidity of the cystine knot core.

Although cyclization is thought to have evolved primarily because of advantages conferred by enhanced structural stability, this study demonstrates that another benefit of cyclization is that it creates an additional structural loop that can be involved in target interactions as shown in Fig. 6. These additional interactions probably explain the enhanced trypsin inhibitory potency of the cyclic peptide compared with linear forms; MCoTI-II has a $K_i$ of 0.03 nM compared with 0.3 nM for a version devoid of loop 6 (21). A role in target interaction has previously been postulated for the cyclic backbone of MCoTI-II (40), but here, we provide the first experimental evidence to support this suggestion. The detailed structural information regarding the binding of MCoTI-II to trypsin should be valuable for subsequent modeling and the design of novel peptides with potential therapeutic activities, particularly for the design of novel enzyme inhibitors. Such modeling and design approaches have already been proven successful for modifying the smaller cyclic trypsin inhibitor sunflower trypsin inhibitor-1 to target kallikrein-4, a protease overexpressed in prostate cancers (41).

The crystal structure of MCoTI-II bound to trypsin appears to be inconsistent in several ways, with the surprising recent report that MCoTI-I apparently becomes more flexible upon binding to trypsin than it is in solution (20). That study originally reported a generalized increase in flexibility of the active site loop (loop 1) of MCoTI-I on binding to trypsin but was modified in a corrigendum (where an error in NMR data analysis was corrected) to conclude that only five residues have increased flexibility upon binding to trypsin, including the active site Lys residue. This interpretation seems to be questionable, given that the side chain amino group of the active site residue in MCoTI-I is clearly detectable in the HSQC spectrum of the complex. Conformationally exchanging or solvent-exposed primary amino groups are not normally visible in NMR spectra due to exchange, so the clear appearance of the Lys amino group in the spectrum is consistent with protection from solvent (19) and a well defined (i.e. not flexible) bound conformation. Furthermore, the other putatively flexible residues are very well defined in the complex of MCoTI-II with trypsin and are involved in either intra- or intermolecular hydrogen bonds. Finally, given that MCoTI-I and II differ by only two residues and as we show that there is a well defined complex between
MCoTI-II and trypsin, it seems highly unusual that the active site residue in MCoTI-I would have enhanced flexibility upon binding to trypsin. Indeed, we propose here that this enhanced flexibility is not intrinsic to the binding process but reflects flexibility in a cleaved product. We found that MCoTI-II bound to trypsin is partially cleaved over a period of days; it is likely that MCoTI-I is also cleaved over a similar time scale, and the presence of an undetected open form of MCoTI-I in the Puttamadappa et al. (20) study could provide an explanation for the apparent flexibility of some residues, as the termini of the open form are likely to be highly flexible. Further study is required to confirm if this is the reason for the discrepancy with the previously published work on MCoTI-I.

Analysis of the structure of MCoTI-II bound to trypsin reported here also provides significant new insights and a different interpretation of the chemical shift changes observed for MCoTI-I upon binding to trypsin to that reported earlier (20). Several chemical shift changes in MCoTI-I were observed upon binding to trypsin. For instance, the chemical shift of the amide proton of Ala20 in loop 3 changes by 0.45 ppm upon binding with trypsin (20) and that of Cys17 changes by 0.46 ppm. Our crystal structure reveals that several MCoTI-II residues are involved in binding with trypsin (Fig. 6, A and B) but that Ala20 forms an intramolecular hydrogen bond with Val3, as shown in Fig. 6C, and Asp16 forms a hydrogen bond with Lys9 (Fig. 6D). Therefore, although these chemical shift changes were interpreted (20) to be associated with direct interactions with trypsin, our study indicates that they are actually the result of intra-molecular structural changes at sites either involving or adjacent to these residues and are not directly a consequence of interaction with trypsin.

In summary, we report the structure of a cyclotide bound to a protein target and provide mechanistic insights into the biological role of the cyclic backbone. We show that MCoTI-II forms a tight complex with trypsin and that the flexibility in the solution state around the active site residue and throughout cyclization, loop 6, relative to the rest of the molecule, allows the peptide to modify its conformation in these regions upon binding to trypsin to maximize favorable interactions. These interactions provide an explanation for the high potency of MCoTI-II. Interestingly, backbone cyclization appears not only assist in the stabilization of proteins but also facilitates extra binding interactions that can enhance potency. Naturally occurring cyclic proteins such as cyclotides have significant potential as templates in drug design and, in particular, MCoTI-II has been used as a scaffold in the design of a range of novel protease inhibitors. The data obtained from the current study that define the interactions of MCoTI-II with a serine protease might therefore be useful in the design of more potent inhibitors and in understanding how to tailor the specificity of this inhibitor to other related proteases of pharmaceutical interest.

Acknowledgments—We acknowledge the use of the Queensland NMR network 900 MHz spectrometer, the Australian Synchrotron, and the UQ ROCX diffraction Facility.

REFERENCES

1. 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) Squash trypsin inhibitors from Momordica cochininchinensis exhibit an atypical macrocyclic structure. Biochemistry 39, 5722–5730

2. Chiche, L., Heitz, A., Gelly, J. C., Gracy, J., Chau, P. T., Ha, P. T., Hernandez, J. F., and Le-Nguyen, D. (2004) Squash inhibitors: from structural motifs to macrocyclic knotting. Curr. Protein Pept. Sci. 5, 341–349

3. Felizmenio-Quimio, M. E., Daly, N. L., and Craik, D. J. (2001) Circular proteins in plants: Solution structure of a novel macrocyclic trypsin inhibitor from Momordica cochininchinensis. J. Biol. Chem. 276, 22875–22882

4. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. J. Mol. Biol. 294, 1327–1336

5. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis. Proc. Natl. Acad. Sci. U.S.A. 98, 10614–10619

6. Poth, A. G., Colgrave, M. L., Lyons, R. E., Daly, N. L., and Craik, D. J. (2011) From the Cover: Discovery of an unusual biosynthetic origin for circular proteins in legumes. Proc. Natl. Acad. Sci. U.S.A. 108, 10127–10132

7. Mynle, J. S., Chan, L. Y., Chanson, A. H., Daly, N. L., Schaefer, H., Bailey, T. L., Nguyencong, P., Cascales, L., and Craik, D. J. (2012) Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase-mediated biosynthesis. Plant Cell 24, 2765–2778

8. Nguyen, G. K., Zhang, S., Nguyen, N. T., Nguyen, P. Q., Chiu, M. S., Hardjojo, A., and Tam, J. P. (2011) Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-I chain a and cyclotide domains in the fabaceae family. J. Biol. Chem. 286, 24275–24287

9. Kaas, Q., and Craik, D. J. (2010) Analysis and classification of circular proteins in CyBase. Biopolymers 94, 584–591

10. Ji, Y., Majumder, S., Millard, M., Borra, R., Bi, T., Elngar, A. Y., Neamati, N., Shekhtman, A., and Camarero, J. A. (2013) In vivo activation of the p53 tumor suppressor pathway by an engineered cyclotide. J. Am. Chem. Soc. 135, 11623–11633

11. Chan, L. Y., Gunasekera, S., Henriques, S. T., Worth, N. F., Le, S. J., Clark, R. I., Campbell, J. H., Craik, D. J., and Daly, N. L. (2011) Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds. Blood 118, 6709–6717

12. Craik, D. J., Cemazar, M., and Daly, N. L. (2006) The cyclotides and related macrocyclic peptides as scaffolds in drug design. Curr. Opin. Drug Discov. Devel. 9, 251–260

13. Gunasekera, S., Foley, F. M., Clark, R. J., Sando, L., Fabri, L. J., Craik, D. J., and Daly, N. L. (2008) Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides. J. Med. Chem. 51, 7697–7704

14. Thongyoo, P., Bonomelli, C., Leatherbarrow, R. J., and Tate, E. W. (2009) Potent inhibitors of β-tryptase and human leukocyte elastase based on the MCoTI-II scaffold. J. Med. Chem. 52, 6197–6200

15. Sommerhoff, C. P., Avrutina, O., Schmoldt, H. U., Gabrijelcic-Geiger, D., Diederichsen, U., and Kolmar, H. (2010) Engineered cystine knot mini-proteins as potent inhibitors of human mast cell tryptase B. J. Mol. Biol. 395, 167–175

16. Aboye, T. L., Ha, H., Majumder, S., Christ, F., Debyser, Z., Shekhtman, A., Neamati, N., and Camarero, J. A. (2012) Design of a novel cyclotide-based CXCR4 antagonist with anti-human immunodeficiency virus (HIV)-1 activity. J. Med. Chem. 55, 10729–10734

17. Woodward, K. P., Daly, N. L., Brown, D. L., Stow, J. L., and Craik, D. J. (2007) The cyclic cystine knot mini-protein MCoTI-II is internalized into cells by macropinocytosis. Int. J. Biochem. Cell Biol. 39, 2252–2264

18. Craik, D. J., and Daly, N. L. (2007) NMR as a tool for elucidating the structures of circular and knotted proteins. Mol. Bioyst. 3, 257–265

19. Heitz, A., Hernandez, J. F., Gagnon, J., Hong, T. T., Pham, T. T., Nguyen, T. M., Le-Nguyen, D., and Chiche, L. (2001) Solution structure of the squash trypsin inhibitor MCoTI-II. A new family for cyclic knotting. Biochemistry 40, 7973–7983
Structural Analysis of MCoTI-II

20. Puttamadappa, S. S., Jagadish, K., Shekhtman, A., and Camarero, J. A. (2010) Backbone dynamics of cyclotide MCoTI-I free and complexed with trypsin. *Angew. Chem. Int. Ed. Engl.* **49**, 7030–7034

21. Avrutina, O., Schmoldt, H. U., Gabrijelcic-Geiger, D., Le Nguyen, D., Sommerhoff, C. P., Diederichsen, U., and Kolmar, H. (2005) Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II. *Biol. Chem.* **386**, 1301–1306

22. Creighton, T. E. (1992) *Proteins: Structures and Molecular Properties*, WH Freeman and Company

23. Arumugam, S., Gao, G., Patton, B. L., Semenchenko, V., Brew, K., and Van Doren, S. R. (2003) Increased backbone mobility in β-barrel proteins with bound ligands. *J. Mol. Biol.* **327**, 719–734

24. Tam, J. P., Lu, Y.-A., and Yu, Q. (1999) Thia zip reaction for synthesis of large cyclic peptides: Mechanisms and applications. *J. Am. Chem. Soc.* **121**, 4316–4324

25. Tam, J. P., Lu, Y. A., Yang, J. L., and Chiu, K. W. (1999) An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8913–8918

26. Daly, N. L., Clark, R. J., and Craik, D. J. (2003) Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides. *J. Biol. Chem.* **278**, 6314–6322

27. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984–6003

28. Lipari, G., and Szabo, A. (1981) Nuclear magnetic resonance relaxation in nucleic acid fragments: models for internal motion. *Biochemistry* **20**, 6250–6256

29. Jarvis, J. A., Craik, D. J., and Wilce, M. C. (1993) X-ray diffraction studies of fibrils formed from peptide fragments of transthyretin. *Biochem. Biophys. Res. Commun.* **192**, 991–998

30. McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Solits, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macro-molecular crystallography beamlines. *J. Synchrotron. Radiat.* **9**, 401–406

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

32. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674

33. Paesen, G. C., Siebold, C., Harlos, K., Peacey, M. F., Nuttall, P. A., and Stuart, D. I. (2007) A tick protein with a modified Kunitz fold inhibits human trypsin. *J. Mol. Biol.* **368**, 1172–1186

34. Debatin, J.-M., Jirikowski, G. F., and Baur, J. (2007) The guinea pig Tryptase complexed with its specific inhibitor. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 6250–6256

35. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132

36. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779

37. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559

38. Bode, W., Greylung, H. J., Huber, R., Otleskwi, J., and Wilusz, T. (1989) The refined 2.0 A X-ray crystal structure of the complex formed between bovine β-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (Cucurbita maxima). Topological similarity of the squash seed inhibitors with the carboxypeptidase A inhibitor from potatoes. *FEBS Lett.* **242**, 285–292

39. Rodriguez, J., Gupta, N., Smith, R. D., and Pevzner, P. A. (2008) Does trypsin cut before proline? *J. Proteome Res.* **7**, 300–305

40. Heitz, A., Avrutina, O., Le-Nguyen, D., Diederichsen, U., Hernandez, J. F., Gracy, J., Kolmar, H., and Chiche, L. (2008) Knottin cyclization: impact on structure and dynamics. *BMC Struct. Biol.* **8**, 54

41. Swedberg, J. E., Nigon, L. V., Reid, J. C., de Veer, S. J., Walpole, C. M., Stephens, C. R., Walsh, T. P., Takayama, T. K., Hooper, J. D., Clements, J. A., Buckle, A. M., and Harris, J. M. (2009) Substrate-guided design of a potent and selective kalikrein-related peptidase inhibitor for kalikrein-4. *Chem. Biol.* **16**, 633–643

42. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55