Solution Structure of the Squash Trypsin Inhibitor MCoTI-II. A New Family for Cyclic Knottins[†,‡]

Annie Heitz,[§] Jean-François Hernandez,[||] Jean Gagnon,[||] Thai Trinh Hong,[⊥] T. Trân Châu Pham,[⊥] Tuyet Mai Nguyen,[⊥] Dung Le-Nguyen,[#] and Laurent Chiche*[,§]

Centre de Biochimie Structurale, UMR5048 CNRS-Université Montpellier I, UMR554 INSERM-Université Montpellier I, Faculté de pharmacie, 15 avenue Charles Flahault, 34060 Montpellier, France, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS), 41, rue Jules Horowitz, 38027 Grenoble Cedex 1, France, Centre de Biotechnologie, Université Nationale du Vietnam, 90, Nguyen Trai Street, Hanoï-Viet-Nam, and INSERM U376, CHU Arnaud-de-Villeneuve, 371, rue du doyen Gaston Giraud, 34295 Montpellier-France

Received April 2, 2001; Revised Manuscript Received May 17, 2001

ABSTRACT: The "knottin" fold is a stable cysteine-rich scaffold, in which one disulfide crosses the macrocycle made by two other disulfides and the connecting backbone segments. This scaffold is found in several protein families with no evolutionary relationships. In the past few years, several homologous peptides from the Rubiaceae and Violaceae families were shown to define a new structural family based on macrocyclic knottin fold. We recently isolated from *Momordica Cochinchinensis* seeds the first known macrocyclic squash trypsin inhibitors. These compounds are the first members of a new family of cyclic knottins. In this paper, we present NMR structural studies of one of them, MCoTI-II, and of a β-Asp rearranged form, MCoTI-IIb. Both compounds display similar and well-defined conformations. These cyclic squash inhibitors share a similar conformation with noncyclic squash inhibitors such as CPTI-II, and it is postulated that the main effect of the cyclization is a reduced sensitivity to exo-proteases. On the contrary, clear differences were detected with the three-dimensional structures of other known cyclic knottins, i.e., kalata B1 or circulin A. The two-disulfide cystine-stabilized β-sheet motif [Heitz et al. (1999) *Biochemistry 38*, 10615−10625] is conserved in the two families, whereas in the C-to-N linker, one disulfide bridge and one loop are differently located. The molecular surface of MCoTI-II is almost entirely charged in contrast to circulin A that displays a well-marked amphiphilic character. These differences might explain why the isolated macrocyclic squash inhibitors from *M. cochinchinensis* display no significant antibacterial activity, whereas circulins and kalata B1 do.

A number of small, stable disulfide-rich proteins have been found in plants and animals. The corresponding scaffolds have been largely used by nature to achieve a variety of tasks (inhibition, toxicity, defense, regulation, etc.) and thus represent very interesting starting frameworks for building new active molecules, i.e., by grafting active sites or recognition fragments on them (*1*−*4*).

However, only few different structural motifs are found in proteins with very diverse origins and functions and with no apparent evolutionary relationship (*5*). Although this is consistent with the fact that possible protein folds are limited in number and that similar folds can be observed in proteins with essentially no sequence identity (*6*−*8*), it appears necessary to accumulate information on new structural motifs and on as many of their variants as possible in order to rationalize the sequence structure−function relationship.

One such small and stable motif with three disulfide bridges is found in the squash trypsin inhibitors (*9*−*16*). These small disulfide-rich proteins (28−32 amino acids, 6 cysteines) are composed of a small antiparallel triple-stranded β-sheet, one and a half-turn of a 3_{10} helix, two β-turns and the inhibitory loop. These secondary structural elements are organized around the three disulfide bridges that largely participate in stabilizing the protein core. It was observed that one disulfide bridge crosses the macrocycle formed by the two other disulfide bridges and the interconnecting backbone, hence the terms "knottins", "cystine-knot", or "inhibitor cystine-knot" (*17*−*19*). The knottin scaffold is based on the elementary cystine stabilized β-sheet (CSB)[1] motif (*20*), and opens new interesting perspectives for the engineering of small stable proteins with various novel activities (*21*, *22*).

[†] This work was supported by the collaboration program between CNRS (France) and CNST (Vietnam).

[‡] The coordinates for the 30 refined conformers of MCoTI-II have been deposited in the Brookhaven Protein Data Bank (entry 1HA9).

* To whom correspondence should be addressed. Phone: +33 [0]4 67 04 34 32. Fax: +33 [0]4 67 52 96 23. E-mail: chiche@cbs.univ-montp1.fr.

[§] Centre de Biochimie Structurale.

[||] Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS).

[⊥] Centre de Biotechnologie.

[#] INSERM U376.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; CSB, cystine-stabilized β-sheet; COSY, correlated spectroscopy; CPTI-II, *Cucurbita pepo* trypsin inhibitor II; CSI, chemical shift index; EETI II, *Ecballium elaterium* trypsin inhibitor II; H-bond, hydrogen bond; HSQC, heteronuclear single quantum coherence spectroscopy; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; PCI, potato carboxypeptidase inhibitor; RMS, root-mean-square; TI, trypsin inhibitor; TOCSY, total correlated spectroscopy; TSP, 3-(trimethylsilyl)-propionate, sodium salt.

10.1021/bi0106639 CCC: $20.00
Published on Web 06/13/2001

A

```
CPTI-II          -----RVCPKILMECKKDSDCLAECICLEHG-YCG--
CPTI-III         --HEERVCPKILMECKKDSDCLAECICLEHG-YCG--
CMTI-IV          --HEERVCPRILMKCKKDSDCLAECVCLEHG-YCG--
CMTI-III         -----RVCPRILMKCKKDSDCLAECVCLEHG-YCG--
CMTI-I           -----RVCPRILMECKKDSDCLAECVCLEHG-YCG--
LLTI-I           ---<ERRCPRIYMECKHDSDCLADCVCLEHG-ICGG-
LLDTI-II         -----RRCPRIYMECKHDSDCLADCVCLEHG-ICG--
CVTI-I           ----GRRCPRIYMECKRDADCLADCVCLQHG-ICG--
CMCTI-I          ------MCPKILMKCKQDSDCLLDCVCLKEG-FCG--
CMCTI-III        ---<ERMCPKILMKCKQDSDCLLDCVCLKEG-FCG--
CMCTI-II         -----RMCPKILMKCKQDSDCLLDCVCLKEG-FCG--
CSTI-IV          -----MMCPRILMKCKHDSDCLPGCVCLEHIEYCG--
CSTI-IIB         -----MVCPKILMKCKHDSDCLLDCVCLEDIGYCGVS
TTI-I           MMEGVVACPRILMPCKVNDDCLRGCKCLSNG-YCG--
TTI-II          MMEGVVACPRILMPCKVNDDCLRGCKCLSNG-YCG--
HMTI-I           -----VGCPRILMKCKTDDDCLLGCKCLSNG-YCG--
CMeTI-B          -----VGCPRILMKCKTDRDCLTGCTCKRNG-YCG--
TGT-II          SGRHGGICPRILMPCKTDDDCMLDCRCLSNG-YCG--
LATI-1           ------ICPRILMECSHDSDCFGECICLSSG-YCG--
LATI-2           -----IRCPRIYMECKHDSDCLGECICLESG-FCG--
LCTI-II          -----RICPRILMECSSDSDCLAECICLEQDGFCG--
LCTI-I           -----RICPRILMECSSDSDCLAECICLEQG-FCG--
LCTI-III         -----RICPRILMECSSDSDCLAECICLENG-FCG--
TGTI-I           ------ICPRILMPCSSDSDCLAECICLENG-FCG--
MCoTI-II        SGSDGGVCPKILKKCRRDSDCPGACICRGNG-YCG--
MCoTI-I         SGSDGGVCPKILQRCRRDSDCPGACICRGNG-YCG--
MCoTI-III        ---<ERACPRILKKCRRDSDCPGECICKENG-YCG--
MCTI-I           ----ERRCPRILKQCKRDSDCPGECICMAHG-FCG--
MCTI-III         ----ERGCPRILKQCKQDSDCPGECICMAHG-FCG--
EETI-II          ------GCPRILMRCKQDSDCLAGCVCGPNG-FCG--
BDTI-II          -----RGCPRILMRCKRDSDCLAGCVCQKNG-YCG--
ELTI-I           -KEEQRVCPRILMRCKRDSDCLAQCTCQQSG-FCG--
ELTI-II          -----RVCPRILMRCKRDSDCLAQCTCQQSG-FCG--
SATI-I           -----RVCPRILMRCKRDSDCLAECTCQGSG-YCG--
MRTI-I           -----GICPRILMECKRDSDCLAQCVCKRQG-YCG--
MCTI-II(1mct)    -----RICPRIWMECTRDSDCMAKCICVAGH--CG--
TI-A             -----RSCPRIWMECTRDSDCMAKCICVAGH--CG--
MCTI-II(1f2s)    -----RICPRIWMECKRDSDCMAECICVMGH--CG--
MCEI-IV          --EEERICPLIWMECKRDSDCLAQCICVDGH--CG--
MCEI-III         ---EERICPLIWMECKRDSDCLAQCICVDGH--CG--
MCEI-II          ----ERICPLIWMECKRDSDCLAQCICVDGH--CG--
MCEI-I           -----RICPLIWMECKRDSDCLAQCICVDGH--CG--
MCTI-II(sw)      -----RICPRIWMECKRDSDCMAQCICVDGH--CG--
SATI-II          ----GRICPRILMECKRDSDCLAECICQSGY--CG--
SATI-III         -<ERGRICPRILMECKRDSDCLAECICQSGY--CG--
                       ** *   *  : **   * *      **
```

B

```
                                  →          →  →
MCoTI-II              SG---SDGGVCPKILKKC--RRDSD---CPGACICRGNGYCG
kalata B1             RNGLP-------VCGETCVGGT--C---NTPGCTCS-WPVCT
circulin A            RNGIP--------CGESCVWIP--C-ISAALGCSCK-NKVCY
cycloviolacin O1      -NGIP--------CAESCVYIP--CTVTALLGCSCS-NRVCY
                                    *           * *     *
```

FIGURE 1: (A) Sequence alignment of members of the squash inhibitors family. Sequences were taken from (*28*) and from the Swiss-Prot, TrEMBL, and PIR databases, except for SATI-I, -II, and -III (*56*). The alignment and sequence order is that given by the CLUSTALW program (*57*). The precursor sequences were truncated at the first corresponding residue in the MCoTI-II sequence. Three sequences are reported for MCTI-II and are followed by the corresponding PDB ID or the indication sw (Swiss-Prot) in parentheses. The bottom line indicates fully conserved residues (*) or physicochemical properties (:). (B) Structural alignment between MCoTI-II and cyclic knottins kalata B1, circulin A, and cycloviolacin O1. The alignment was done manually. The conserved triple-stranded β-sheet is shown as arrows and only structurally conserved residues are aligned.

In the past few years, macrocyclic peptides kalata B1 (*23*), cyclopsychotride A (*24*), circulin A and B (*25*), cycloviolacin O1 (*26*), and cycloviolins A−D (*27*) were shown to share this structural motif, thus defining a new structural family of macrocyclic knottins. All these peptides are homologous and were named plant cyclotides. They were grouped into two subfamilies following sequence comparisons (*26*). However, shortly after, we identified new members of the cyclic knottin structural family from seeds of *Momordica cochinchinensis*, a common cucurbitaceae in Vietnam (*28*). Two major cyclic trypsin inhibitors named MCoTI-I and -II were isolated, along with rearranged forms containing a β-Asp residue (MCoTI-Ib and MCoTI-IIb). All these compounds share large sequence identity with other squash inhibitors but only very low sequence identity, if any, with previously known cyclic knottins (Figure 1). The small size and yet very high stability afforded both by the macrocyclization and by the disulfide bonds, render these molecules very attractive.

In this paper, we report NMR solution structure studies on the most abundant cyclic squash trypsin inhibitor, MCoTI-II, and on the rearranged form containing a β-Asp residue, MCoTI-IIb. The peptide segment that was absent in noncylic squash inhibitors but present in cyclic MCoTI-II and -IIb has been termed the C-to-N linker since it links residues that used to be the C-terminus and the N-terminus in previously known squash inhibitors. The well-defined conformation calculated for MCoTI-II is compared with both the conformation of noncyclic squash inhibitors and the conformation of nonsquash cyclic knottins. Structural similarities and/or differences are detailed and their possible impact on functional aspects are discussed.

MATERIALS AND METHODS

Materials. The proteins were isolated from *M. cochinchinensis* seeds as described previously (*28*). Natural MCoTI-II and MCoTI-IIb were obtained in quantities sufficient for structural studies.

Microbial Strains. *Escherichia coli* D31 was from H. G. Boman (Department of microbiology, University of Stockholm, Stockholm, Sweden). *Micrococcus luteus* A270 was from the Pasteur Institute. *Neurospora crassa* (CBS 327-54) was a gift from W. F. Broekaert (Jansens Laboratory of Genetic, Catholic University of Leuven, Heverlee, Belgium).

Antimicrobial Assays. Antibacterial activities were measured using a liquid-growth inhibition assay as described previously (*29*). Briefly, 10 μL from 2-fold serial dilutions of the peptides (100−0.2 μM, final concentrations) were incubated in 96-wells microtiter plates with a starting OD_{600} of 0.001. After a 24 h incubation at 25 °C, the antibacterial activity was monitored by measuring the culture absorbance at 595 nm using a microplate reader. The antifungal activity against *N. crassa* used liquid-growth inhibition assay (*29*). Briefly, fungal spores (final concentration of 10^{-4} spores/mL) were suspended in a growth medium containing Potato Dextrose Broth [DIFCO, in half-strength, supplemented with tetracycline (10μg/mL) and cefotaxim (100μg/mL)], dispensed by aliquots of 90 μL into wells of a microplate containing 10 μL of the serial dilution of the peptides, and incubated for 48 h at 25 °C in the dark. Growth of fungi was evaluated as above. Positive controls were obtained using insect antimicrobial peptides (thanatin and androctonin).

NMR Spectroscopy. Samples were prepared by dissolving peptides in either 90% H_2O/10% ^2H_2O (v/v) or 100% ^2H_2O to a concentration of approximately 2.5 mM with the pH adjusted to 3.4 by addition of dilute HCl or NaOH. All ^1H NMR spectra were recorded on a Bruker AMX-600 spectrometer. Data were acquired at 12 and 27 °C, and TSP-d4 was used as an internal reference. All 2D experiments, COSY, TOCSY, and NOESY, were performed according to standard procedures (*30*) using quadrature detection in both dimensions with spectral widths of 6849.3 Hz in both dimensions. The carrier frequency was centered on the water signal, and the solvent was suppressed by continuous low power irradiation during the relaxation delay and during the mixing time for NOESY spectra. The 2D spectra were obtained using 2048 or 4096 points for each t_1 value, and 512 t_1 experiments were acquired for COSY, TOCSY, and

NOESY experiments. TOCSY spectra were recorded with spin lock times of 30 and 60 ms. The mixing time was 150 and 300 ms in NOESY spectra. Spectra were processed using XWINNMR (Bruker). The t_1 dimension was zero filled to 1024 points and $\pi/8$- and $\pi/4$-shifted sine bell functions were applied in t_1 and t_2 domains, respectively, prior to Fourier transform. $^3J_{NH-H\alpha}$ coupling constants were measured on 1D spectra. The exchange of amide protons with deuterium was studied at 12 °C on 2.5 mM samples lyophilized from H_2O at pH 3.4 and dissolved in ^2H_2O. A series of 1D, TOCSY, and NOESY spectra were acquired over a 72-h period. ^{1}H-^{13}C HSQC spectra (*31*, *32*) were recorded on the samples in ^2H_2O. Spectral widths were 6849.3 and 25 000 Hz in the ^{1}H and ^{13}C dimensions, respectively. A total of 2048 data points was acquired with 512 t_1 increments.

Structure Calculations. All calculations were performed on a Silicon Graphics Origin 200 workstation. The structures were displayed and analyzed using either INSIGHT II (MSI, San Diego) on a Silicon Graphics O2 workstation or MOLMOL 2K.1 (*33*) on a Linux box. The NOE intensities were classified as strong, medium, and weak, and converted into distance constraints of 2.5, 3, and 4 Å, respectively. If the connectivity involved side-chain protons, 3.0, 4.0, and 5.0 Å upper bounds were used instead to account for higher mobility. For sequential $d_{\alpha N}$ and d_{NN} connectivities, we used bounds of 2.5, 3.0, and 3.5 Å and 2.8, 3.3, and 4.0 Å, respectively. When necessary, the distance constraints were corrected for pseudoatoms (*34*). ϕ angles of residues with small or large $^3J_{HN-H\alpha}$ coupling constants ($<$4 Hz or $>$8.5 Hz) were constrained in the $-90°$ to $-40°$ or $-160°$ to $-80°$ ranges. χ_1 angles of residues for which stereospecific attribution of the β-protons could be achieved were constrained in the corresponding range. Disulfide bridges were imposed through distance constraints of 2.0−2.1 Å, 3.0−3.1 Å, and 3.75−3.95 Å on Si−Sj, Si−Cβj, Sj−Cβi, and Cβi−Cβj distances, respectively. No H-bond was imposed.

3D structures were obtained from the distance and angle restraints using the torsion angle molecular dynamics method available in the DYANA program (*35*). Preliminary DYANA runs and analyses with the GLOMSA routine (*36*) were used to perform stereospecific assignment whenever possible. Gly^2, Gly^6, and Gly^{34} α-protons could be unambiguously assigned with this method. One thousand structures were then calculated with the standard simulated annealing protocol. The thirty structures with the lowest violation of the target function were selected for further refinement. They were submitted to molecular mechanics energy refinement with the SANDER module of the AMBER 6 program (*37*), using the parm94 force field (*38*) and the GB/SA implicit solvation system (*39*). During the molecular dynamics runs, the covalent bond lengths were kept constant by applying the SHAKE algorithm (*40*) allowing a 1.5 fs time step to be used. The nonbonded pair list was updated every 20 steps, and the temperature was regulated by coupling the system to a heat bath with a coupling constant of 0.2 ps. Pseudoenergy terms taking into account the NMR interproton distance restraints were defined as follows via four threshold distance values: r_1, r_2, r_3, and r_4. In all cases, r_1 and r_2 were set to 1.3 and 1.8 Å, respectively. r_3 was taken as the upper boundary used in the DYANA calculations and r_4 was chosen as $r_3 + 0.5$ Å. For an observed distance lying between r_2 and r_3, no restraint was applied. Between r_1 and r_2 or between r_3 and r_4, parabolic restraints were applied. Outside the r_1 to r_4 range, the restraints were linear with slopes identical at parabolic slopes at points r_1 and r_4. A similar strategy was used for dihedral restraints. When no stereospecific assignment could be achieved for methyl or methylene protons, an $\langle r^{-6}\rangle^{-1/6}$ averaging scheme was used instead of pseudoatoms. Five thousand cycles of restrained energy minimization were first carried out followed by a 30-ps long simulated annealing procedure in which the temperature was raised to 900 K for 20 ps then gradually lowered to 300 K. During this stage, the force constant for the NMR distance and dihedral constraints were gradually increased from 3.2 to 32 kcal mol^{-1} Å^{-2} and from 0.5 to 50 kcal mol^{-1} rad^{-2}, respectively.

Color Figures 8, 9, and 10 were produced with the MOLMOL (*33*) and POV−Ray (http://www.povray.org) programs.

RESULTS AND DISCUSSION

In the late 1980s, we determined the first 3D structure of EETI-II, a squash trypsin inhibitor. This compound, with only 28 amino acids but three disulfide bridges arranged in a pseudo-knotted topology, displayed a particularly high degree of stability along with protease resistance (*9*, *17*). Since then, more than 20 different small disulfide rich protein families were shown to share the same "knottin" topology. More recently, nearly 40 homologous peptides from plants of the Rubiaceae and Violaceae families have been reported to belong to a new cyclic knottins structural family (*23*, *25*−*27*, *41*). We reported recently the first known macrocyclic trypsin inhibitors (TI) from the squash family, MCoTI-I and MCoTI-II, and rearranged β-Asp isoforms that belong to the cyclic knottin family but display no significant sequence similarity with cyclic knottins of the Rubiaceae and Violaceae families. We have determined the solution structure of MCoTI-II and MCoTI-IIb to analyze structural differences with homologous noncyclic squash TIs and with nonhomologous cyclic knottins.

Solution Structure of MCoTI-II. (1) NMR Assignments. The assignment of all the ^{1}H and ^{13}C resonances present in the spectrum was achieved using well-established techniques (*30*) and part of the sequential assignment is presented in Figure 2. The lists of ^{1}H and ^{13}C chemical shifts are available as Supporting Information.

(2) Secondary Structure. Figure 3 summarized the sequential and medium range NOEs, $^3J_{HN-H\alpha}$ coupling constants, slowly exchanging amide protons and the Cα chemical shift index (CSI) (*42*, *43*). The observation of two small $^3J_{HN-H\alpha}$ coupling constants for residues Asp^{18} and Ser^{19}, and the $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+2)$, and $d_{\alpha N}(i,i+3)$ NOEs in the region 16−21 shows the presence of a short 3_{10} helix which is generally detected between the second and the third cysteine of the TIs of the squash family. The NMR parameters measured in the region 22−25 are in agreement with a β-turn [$d_{\alpha N}(i,i+2)$ NOE and slowly exchanging amide proton of residue 25]. It is now well-established that squash TIs share a common structural motif with others cysteine-rich peptides. This motif made of an antiparallel triple-stranded β-sheet is well-defined in MCoTI-II. The three regions of the sequence involved in this motif are 13−15, 26−28 and 32−34. Large $^3J_{HN-H\alpha}$ coupling constants and slowly exchanging amide

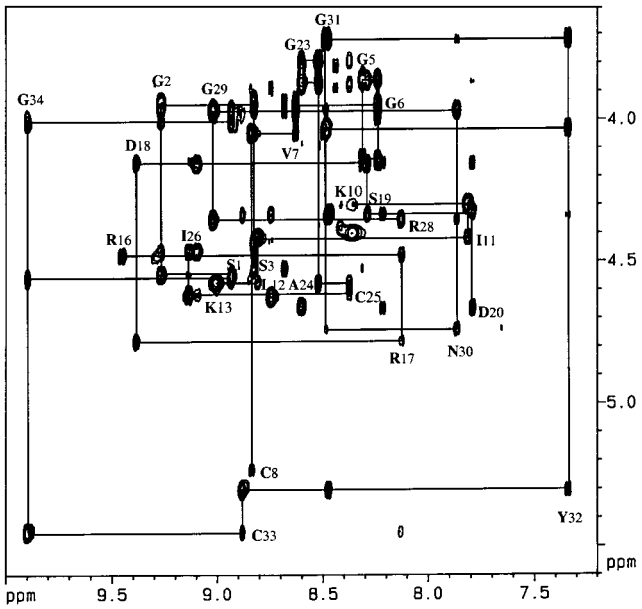

FIGURE 2: Fingerprint region of the 600 MHz NOESY spectrum of MCoTI-II at 12 °C and pH 3.4 in 90% H_2O/10% ^2H_2O. The following sequences are traced: 5−8, 10−13, 16−20, 23−26, and 28−3.

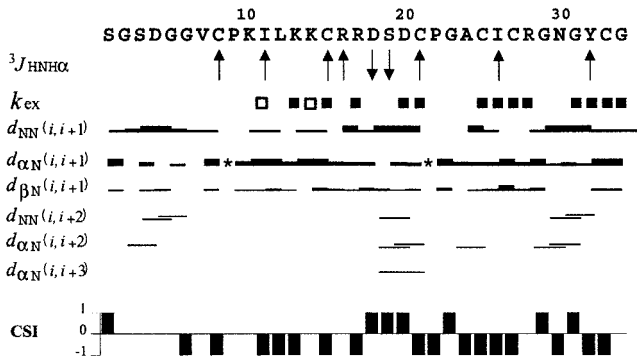

FIGURE 3: NMR data summary of the sequential and medium range NOE connectivities, $^3J_{HN-H\alpha}$ coupling constants and slowly exchanging amide protons observed for MCoTI-II. The asterisk indicate the sequential Hα-H$\delta(i-1)$ connectivities for proline residues. The height of the bar correspond to the strength of the NOE. The values of the ^{3}JHN−Hα coupling constants are indicated by ↓ (<4 Hz) and ↑ (>8.5 Hz). Open and filled squares indicate backbone amide protons that were still observed after 3 and 24 h, respectively, in ^2H_2O. The chemical shift index (CSI) derived from the Cα chemical shifts of MCoTI-II is plotted at the bottom of the figure.

protons measured in these parts of the sequence are in agreement with this structure. In addition, all the characteristic inter-strand NOEs were detected. Region 26−34 is a β-hairpin with a β-turn involving residues 28−31 that was ascertained by the presence of $d_{NN}(i,i+2)$ and $d_{\alpha N}(i,i+2)$ NOEs. Concerning the C-to-N linker, weak $d_{NN}(i,i+2)$ and $d_{\alpha N}(i,i+2)$ NOEs were observed in the region 3−6, but all the amide protons of the residues constituting this loop are rapidly exchanging. This part of the sequence thus appears as rather flexible and poorly structured, in accordance with the CSI.

(3) Structure Calculations. The three-dimensional structure of the cyclic compound MCoTI-II was determined from NMR data using the same strategy previously used for structural studies of native squash inhibitor EETI II and of analogues (*9*, *20*, *44*−*47*). The NMR study led to 86

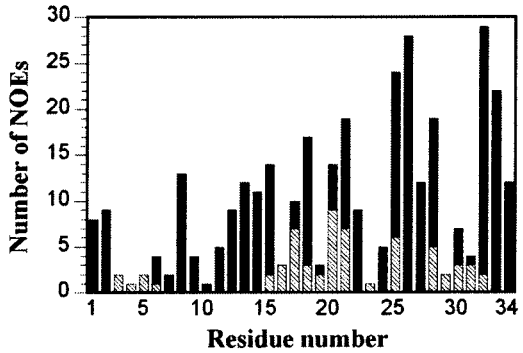

FIGURE 4: Distribution of the number of experimental constraints deduced from medium (hatched bars) and long range (filled bars) NOEs as a function of the sequence of MCoTI-II. Each constraint is counted twice, once for each proton involved.

Table 1. Constraint Violations and Structural Statistics[a]

spatial constraints	
distances[b]	
short	86
medium	31
long-range	140
dihedrals[c]	
ϕ	8,11,15,16,18,19,21,26,32
χ_1	8,15,18,20,21,25,27,30,32,33
constraint violations[d]	
distances	
number > 0.2 Å	0
	2.97 (0.81)
number > 0.1 Å	1.17 (0.83)
	11.46 (1.71)
sum	0.55 (0.17)
	2.49 (0.26)
maximum	0.11 (0.02)
	0.35 (0.07)
dihedral	
number > 5°	0
	0.67 (0.48)
number > 2°	0.77 (0.77)
	1.37 (0.85)
AMBER energies (kcal mol^{-1})	
bond	20.1 (0.47)
angle + dihedral	189.9 (44.2)
van der Waals	−113.7 (3.54)
generalized Born	−523.3 (145.9)
surface based	11.7 (0.28)
total AMBER	−1279.3 (43.96)
constraint	2.51 (0.52)
PROCHECK statistics	
residues in most favored regions (A,B,L)	85%
residues in additional allowed regions (a,b,l,p)	14%
deviations from ideal geometry	
bond	0.012 (10^{-4})
angle	2.15 (0.38)

[a] Values in parentheses indicate standard deviations. [b] Number of constraints. [c] Residue numbers. [d] Values are for refined structures and for DYANA structures in italics.

sequential and 171 medium and long-range NOEs. The distribution of the medium and long-range NOEs along the sequence is displayed in Figure 4. Nine ϕ angles were determined from the $^3J_{HN-H\alpha}$ coupling constants and stereospecific assignment of the Hβ protons was achieved for 10 residues (Figure 3 and Table 1). Although not experimentally determined, the disulfide bridge pattern was assumed to be the same as that derived from the three-dimensional structures of closely related noncyclic homologues (Figure 1) (*28*). The

FIGURE 5: Stereoview of the refined solution structures of MCoTI-II. Thirty structures were superimposed for backbone atoms of residues 13−33. Backbone atoms (N, CA, C, O) and disulfide bridges (atoms CB and SG) are shown, and cysteine residue numbers are displayed.

NMR data were converted into distance and angle constraints as usual. The list of constraints used is available as Supporting Information, and the statistics for constraint violations and for molecular mechanics energies are shown in Table 1. The program DYANA (*35*) was used to compute 1000 3D conformations compatible with the constraints using the torsion angle dynamics method. The large number of calculation was to avoid convergence problems due to the highly constrained knotted topology of the molecule. Using AMBER 6.0 (*37*), the 30 best resulting models were further refined using a molecular dynamics simulated annealing protocol including a 20-ps long heating period (900 K) to better search the conformational space. Molecular dynamics refinements in previous studies used either a cpu-intensive explicit water treatment (*10*, *44*) or a simple distance dependent dielectric function coupled to reduced charges on charged side chains (*20*). In this study, the more accurate GB/SA implicit solvation model (*39*), made available in AMBER 6.0, was used instead.

Structure Analysis and Comparison with Noncyclic Squash Inhibitors. The calculated structures satisfy the NMR data very well with no distance and dihedral violation equal or larger than 0.2 Å or 5°, respectively (Table 1). Statistical analyses using the PROCHECK-NMR software (*48*) show that the overall stereochemistry of the MCoTI-II solution structures is very good with 99% of nonglycine and non-proline residues lying in the most favored and additional allowed regions of the Ramachandran map (Table 1). As expected, the refined models also display large negative molecular mechanics AMBER energies. The stereochemical quality of the calculated structures, is supported by the very good similarity with X-ray structures of homologous proteins (see below). These results constitute a good validation of the refinement protocol using the implicit GB/SA solvation model (*39*). Since it has been shown that the quality of solution structures may be correlated to the refinement protocols or softwares (*49*), then the AMBER-GB/SA combination does appear well suited to the refinement of solution structures.

Table 2. Global RMS Deviations[a] (Å)

residues	1−34	7−34	13−33
MCoTI-II solution structures			
trace	1.53 (0.47)	0.78 (0.30)	0.33 (0.09)
backbone	1.44 (0.45)	0.74 (0.28)	0.29 (0.08)
heavy atoms	2.18 (0.40)	1.72 (0.29)	1.57 (0.33)
MCoTI-II vs CPTI-II			
trace		0.86	0.62
backbone		0.82	0.57

[a] Values in parentheses indicate standard deviations.

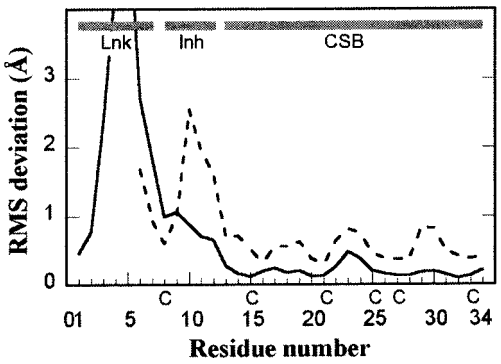

FIGURE 6: Average per residue RMS deviation between calculated structures of MCoTI-II (plain line) and per residue RMS deviation between MCoTI-II and CPTI-II (dashed line). MCoTI-II structures were superimposed pairwise for the backbone atoms of residues 13−33. The solution structure closest to the average conformation of MCoTI-II was superimposed onto the X-ray structure of CPTI-II for backbone atoms of residues 13−33 (MCoTI-II numbering). Grey boxes at the top indicate different segments. Lnk: C-to-N linker; Inh: inhibitory loop; CSB: cystine stabilized β-sheet motif.

The structure of MCoTI-II is particularly well resolved with a very low global backbone RMS deviation of 0.29 ± 0.08 Å for superimposition of backbone atoms of core residues 13−33 (Table 2 and Figures 5 and 6). Even the inhibitory loop displays rather low (<1 Å) RMS deviations. Only the C-to-N linker displays high RMS deviations well

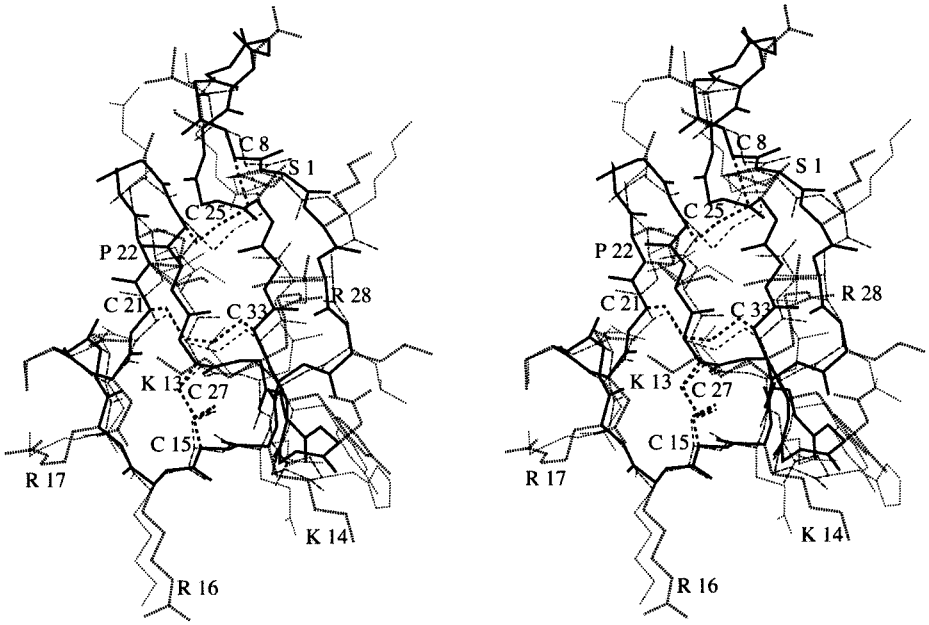

FIGURE 7: Comparison of MCoTI-II with noncyclic squash inhibitors. Structural superimposition of the solution structure closest to the average conformation of MCoTI-II (thick lines) onto the X-ray structure of CPTI-II (thin lines). The two structures were superimposed for backbone atoms of residues 13−33 (MCoTI-II numbering). Side chains are shown as dashed lines with longer dashes for disulfide bridges. Ser^{1}, cysteines and basic residues are labeled.

above 1 Å and therefore clearly constitutes the most mobile part of the molecule.

The conformation of MCoTI-II closest to the average has been superimposed onto X-ray structure of the noncyclic squash TI CPTI-II (*15*) (Figure 7). CPTI-II has been selected for comparison because this structure has been determined with good accuracy, and because its sequence is closer to MCoTI-II as compared with other squash inhibitors with known 3D structure. The sequence identity for the 28 C-terminal residues is 64% between MCoTI-II and CPTI-II whereas it is 57 and 54% between MCoTI-II and CMTI-I or EETI-II, respectively (Figure 1). Global RMS deviations between structures are reported in Table 2 and local RMS deviation along the sequence is displayed in Figure 6. The two structures are strikingly similar with low RMS deviations of 0.57 Å for backbone superimposition of residues 13−33 that correspond to the CSB motif. Superimposition of residues 7−34 leads to a slightly larger value of 0.82 Å. This increase is essentially the result of a rigid group motion of the inhibitory loop (residues 8−12) versus the CSB motif in the MCoTI-II solution structure when compared to the CPTI-II X-ray structure. It is worth noting, however, that the local conformation of the inhibitory loop in MCoTI-II is still strikingly similar to the conformation of the loop of CPTI-II in the complex with trypsin. The RMS deviation for superimposition of the backbone atoms of only residues 8−12 is 0.69 Å, and the only conformationally significant difference between structures in this region is an approximately 180° flip of the peptide bond that links Pro^{9} and Lys^{10} (residue P1, MCoTI-II numbering) due to change in the ψ dihedral angle of Pro^{9}. Pro^{9} ψ value is 149° in CPTI-II, whereas 19 of 30 MCoTI-II structures display values between −18° and +13°. The 11 remaining structures display values between 69° and 120°. Whether this difference is a true consequence of trypsin binding would be difficult to ascertain. The only related NMR data is sequential NOEs for the HN proton of Lys^{10} (Figure 3), which is compatible with the two sets of ψ values. It must be remembered that structure calculations were performed with an implicit solvation model [the GB/SA model (*39*)]. However, a water molecule has been determined in X-ray structures of squash inhibitors which is located inside the inhibitory loop and is held in place via H-bonds from carbonyls of strictly conserved residues Pro^{9} and Ile^{11} of the inhibitory loop (*11*, *15*). In MCoTI-II, water molecules were not treated explicitly, thus precluding observation of such effect and the carbonyl of Pro^{9} preferentially turns outward to the bulk solvent. Nevertheless, the inhibitory loop of MCoTI-II in solution exhibits a high conformational similarity with the loop conformation of homologous inhibitors in complex with trypsin, strongly supporting previous observations that these binding loops do not change their conformation upon binding to trypsin (*44*, *50*).

Detailed comparison of MCoTI-II with noncyclic CPTI-II (Figure 7) reveals a very good structure conservation, showing that the knottin structural motif is barely affected by peptide cyclization between the N- and C-termini. All three disulfide bridges display similar conformations, even the Cys^{8}−Cys^{25} bridge that is sequentially and spatially close to the C-to-N linker. All prolines in MCoTI-II are in the trans conformation. Pro^{22} that occupies position 1 of a β-turn has dihedral angle values rather close to values for the corresponding Leu 16 in CPTI-II ($\phi/\psi = -80°/+166°$ and $-83°/+149°$, respectively).

All H-bonds that define the elements of secondary structure in noncyclic squash inhibitors have been observed in MCoTI-II

β-sheet: $O^{13}\cdots HN^{33}$, $O^{31}\cdots HN^{15}$, $O^{26}\cdots HN^{34}$, $O^{34}\cdots HN^{26}$, $O^{32}\cdots N^{28}$

3_{10}-helix: $O^{17}\cdots HN^{20}$, $O^{18}\cdots HN^{21}$

β-turns: $O^{22}\cdots HN^{25}$, $O^{28}\cdots HN^{31}$,HN^{32}

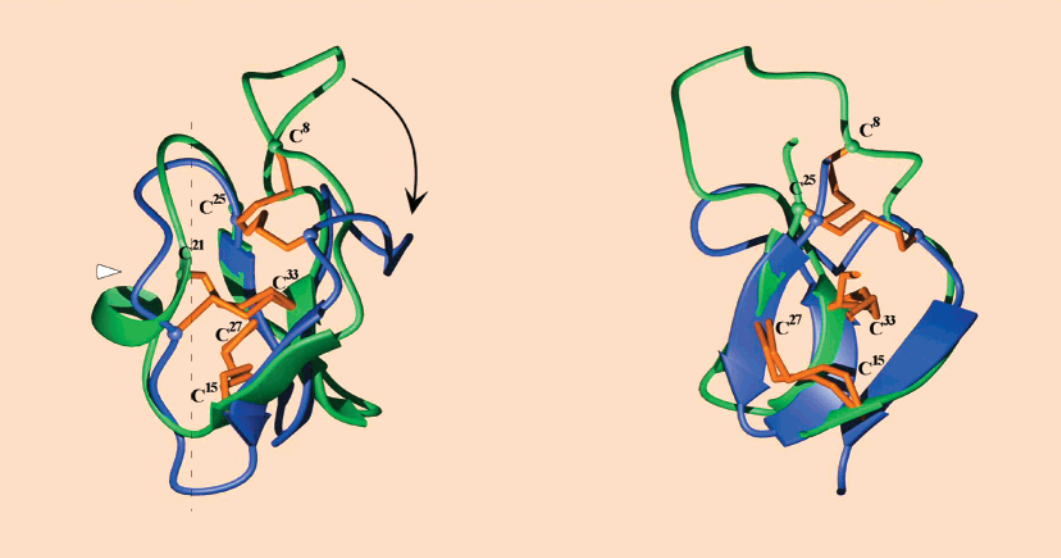

FIGURE 8: Comparison of backbone conformations of cyclic squash inhibitors with other cyclic knottins. Schematic view of MCoTI-II (green) superimposed onto kalata B1 (blue). β-Strands are displayed as flat rounded arrows, whereas the 3_{10} helix in MCoTI-II is displayed as a flat ribbon. Disulfide bridges are shown as orange sticks and cysteines of MCoTI-II are labeled. (Left) The displacement of the C-to-N linker is indicated by an arrow. (Right) Perpendicular view after a 90° rotation around a vertical axis (i.e., viewed from the white triangle viewpoint in the left image). For sake of clarity everything on the left of the dashed line in the left-handed view has been omitted in the right-handed view.

Also, two H-bonds between carboxylic acid side chains and backbone atoms present in CPTI-II are conserved in MCoTI-II calculated structures, Asp^{18}···HN^{27} and Asp^{20}···HN^{16},HN^{17}. It should be noted that the latter bifurcated H-bond was not observed in the NMR structures of CMTI-I [PDB ID 3cti (*51*)], but this might well be a result of the refinement protocol rather than a true specific difference between MCoTI-II and CMTI-I. Indeed, the NMR structures of CMTI-I did not reproduce either the salt bridge between Arg^{1} side chain and the carboxy-terminus that is present in the X-ray structures of CMTI-I and of CPTI-II (note that there is no corresponding arginine and associated salt bridge in MCoTI-II).

Superimposition of MCoTI-II onto CPTI-II in complex with trypsin provides a reasonable model of the interactions between MCoTI-II and trypsin. This model is close to the model presented in our previous report on MCoTI-II (*28*) that was homology modeled from the crystal structure of CMTI-I complexed with trypsin [PDB ID 1ppe (*11*)]. Analysis of this model complex shows that Asp^{4} of the C-to-N linker in MCoTI-II is reasonably close to Lys^{224} and Lys^{222} of trypsin with Asp−Lys Cα−Cα distances of 6.9 and 8.9 Å, respectively. On the other hand, it is worth noting that none of the six positively charged residues in MCoTI-II is in direct contact with trypsin. Indeed, the interaction site of trypsin with the inhibitor is essentially composed of polar, positively charged, and hydrophobic residues.

Structural Impact of the β-Asp^{4} Modification. During the isolation of MCoTI-II, two derived compounds were also identified. One included a succinimide cyclization at the Asp^{4}−Gly^{5} bond, the other was shown to display a β-Asp^{4} residue, probably due to reopening of the succinimide (*28*). The latter form called MCoTI-IIb was obtained in sufficient quantities and submitted to NMR analyses. All the parameters measured on the spectra of MCoTI-IIb appeared to be very similar compared to those measured for MCoTI-II. The most important proton chemical shift differences were detected for residues flanking the modified aspartyl residue, but the carbon chemical shifts remained quite insensitive to this modification. The NOEs detected for the two peptides were only different in the region 3−6 of the sequence. The two NOEs $d_{NN}(i,i+2)$ between residues 3 and 5 and residues 4 and 6 and one $d_{\alpha N}(i,i+2)$ NOE between residues 3 and 5 detected in MCoTI II were no longer observed in MCoTI-IIb. This can be simply explained by the introduction of a CH_2 group in the main-chain, leading to a still larger flexibility of the already flexible C-to-N linker.

Structural Comparison with Other Cyclic Knottins. Schematic drawing of MCoTI-II and comparison with kalata B1 is shown in Figure 8. Although the global folding is similar, clear structural differences are immediately apparent.

The loops between cysteines have different lengths: 5 residues vs 4 residues between the second and third cysteine and 3 residues vs 4 residues between the third and the fourth cysteine, in MCoTI-II and kalata B1, respectively (Figure 1). In the former loop, there is no 3_{10} helix in kalata B1. In the latter loop, a proline occupies position 1 of the β-turn in MCoTI-II but position 2 of the turn in kalata B1. Although prolines are usually considered as preferred residues in position 2 of β-turns, statistical analysis of segments in the Protein Data Bank with conformations similar to the 21−25 loop of MCoTI-II indicates that prolines are frequent in position 1 of this turn as well (data not shown).

Only one disulfide bridge is structurally superimposable with very close χ_1 dihedral angles for the cysteine side chains. This disulfide bridge is between Cys^{15} and Cys^{27} and connects the two external strands of the triple-stranded β-sheet present in both structures. The Cys^{21}−Cys^{33} disulfide bridge is slightly modified in kalata B1, essentially because the loop 15−21 is shorter and has a clearly different conformation (no 3_{10} helix). This results in cysteine 21 being differently located in kalata B1 and close to Asp^{20} of MCoTI-II (Figures 1B and 8). Cys^{21} must therefore adopt a different χ_1 dihedral angle in order to link to Cys^{33} that is well

conserved between the two structures (the χ_1 angle of Cys^{21} is approximately $-60°$ and 180° in MCoTI-II and kalata B1, respectively). The Cys^8-Cys^{25} disulfide bridge of MCoTI-II is the most largely displaced in kalata B1, although Cys^{25} is quite well conserved between the two structures. However, the Cα−Cα distance of Cys^8 from the corresponding cysteine in kalata B1 is about 8.2 Å. This is clearly in relation with the very large displacement of the C-to-N linker in kalata B1 as shown by an arrow in Figure 8. The C-to-N linker between the last and the first cysteine is one residue shorter with one more hydrophobic residue in kalata B1 (GSGSDGGV vs TRNGLPV). The linker in MCoTI-II has been shown to be the most flexible part of the protein. This is not surprising given the presence of four glycines in this eight-residue long segment. Examination of the kalata B1 NMR structures [PDB ID 1kal (*23*)] shows that the C-to-N linker does not display particular flexibility, but rather that the flexibility is equally distributed among all loops. This observation also holds for circulin A [PDB ID 1bh4 (*25*)] and for cycloviolacin O1 [PDB ID 1df6 (*26*)]. More generally, the C-to-N linker of homologous peptides from the Rubiaceae and the Violaceae families is shorter than the linker of MCoTI-II, and contains one proline but only one glycine. These differences in sequence are significant enough to explain, at least in part, the lower flexibility of the C-to-N linker in plant cyclotides. Thus, the cyclic squash inhibitors display a well-defined conformation for most residues except those of the highly flexible C-to-N linker. On the contrary, the C-to-N linker of plant cyclotides displays a flexibility that is similar to the rest of the molecule.

Despite the large modifications of the linker and of the Cys^8-Cys^{25} disulfide bridge discussed above, the antiparallel triple-stranded β-sheet and associated typical H-bonds are well-conserved in MCoTI-II and kalata B1 (Figure 8). The 22−25 β-turn and the H-bond O22···N25 are present in both molecules, although it is a type I turn in MCoTI-II but a type II turn in kalata B1. In the other cyclic knottins with known 3D structure, circulin A and cycloviolacin O1, this loop is longer and includes one turn of helix.

From this analysis, it is clear that the structurally conserved regions between MCoTI-II and other cyclic knottins correspond to the elementary cystine stabilized β-sheet (CSB) motif (*20*). It is worth noting that the cysteines that belong to the triple stranded β-sheet (cysteines 15, 27, and 33) or close to it (cysteine 25) are remarkably well conserved between the two structures with similar side-chain conformations. Outside the CSB motif, large atomic deviations are observed, although the overall topology and the disulfide connectivities are conserved. These structural differences are likely to be necessary to accommodate the very different biological activities.

Biological functions of cyclic knottins from the Rubiaceae and the Violaceae families in the plants are not known. However, they are supposed to participate in a defense mechanism, and interestingly, antimicrobial activities were reported for several of these cyclic knottins (*52*). Kalata B1 and circulin A were shown to be effective specifically against Gram-positive bacteria, whereas circulin B and cyclopsychotride A displayed activity against both Gram-positive and Gram-negative bacteria. These cyclic peptides also displayed moderate activity against two strains of fungi (*52*). Initial interaction with the microbial surfaces are usually supposed to be electrostatic via exposed cationic residues (at least two excess positive charges) on the peptide surface (*53*). Then amphiphilicity due to hydrophobic cluster (about 50% hydrophobic residues) appears as an essential feature for antimicrobial activity (*53*). The biological role of trypsin inhibitors from plants is not fully understood either but could also participate in defense mechanisms. Indeed, it has recently been shown that an antifungal protein from *Helianthus annuus* flowers displays an associated activity against trypsin (*54*). All this prompted us to check macrocyclic trypsin inhibitors from *M. cochinchinensis* for antimicrobial or antifungal activity. However, despite the structural homology of MCoTI-II with other cyclic knottins, antimicrobial and antifungal activity measurements on this compound were unsuccessful. The peptides present no significant antimicrobial activity against microbial strains chosen for their very high sensitivity to antibiotics.

Therefore, the proposition that the cyclic knottin topology may represent a molecular structure of antimicrobials and may provide a useful template for the design of novel peptide antibiotics, as suggested by Tam and co-workers (*52*), should be tempered. The existence of noncyclic knottins with antimicrobial or antifungal properties is also of interest (*55*). It was then interesting to search for the structural differences that are responsible for the absence of antimicrobial activity of macrocyclic squash inhibitors from *M. cochinchinensis*. Electrostatic potentials at the surface of MCoTI-II and circulin A were compared (Figure 9). MCoTI-II sequence contains six positively charged residues and three negatively charged residues resulting in a net charge of +3. This is more positively charged than circulin A and cyclopsychotride A (net charge +2). Thus initial interaction with microbial surface should not be restricted by the specific sequence of MCoTI-II. However, with nine charged side chains and only four clearly hydrophobic residues (Val^7, Ile^{11}, Leu^{12}, and Ile^{26}), MCoTI-II displays a molecular surface which is almost entirely charged with no significant hydrophobic cluster (Figure 9). This is in clear contrast with circulin A and cyclopsychotride A that display distinct hydrophobic and positively charged clusters on their surface with seven and eight clearly hydrophobic residues, respectively (Figure 9). Therefore, the lack of large hydrophobic patch on the surface of *M. cochinchinensis* cyclic squash inhibitors might be part of the explanation of the absence of antimicrobial activity of these compounds. Kalata B1 appears as a peculiar antimicrobial peptide with only one positively charged residue and no net charge, and with only five clearly hydrophobic residues (Table 1B), and suggests that this peptide might use a different mechanism of action.

CONCLUSION

A large number of small disulfide-rich proteins have been isolated from plants and animals in the last two decades. NMR solution structures were reported for many of them, and the number of structures for small disulfide-rich proteins has grown exponentially. Despite this fact, the number of folds available in this class of proteins remains limited (*5*), and it becomes apparent from the very diverse functions satisfied by similar folds that these small architectures constitute extremely interesting models for drug design. We have shown recently that knottin fold consists of an essential structural submotif containing only two disulfide bridges that

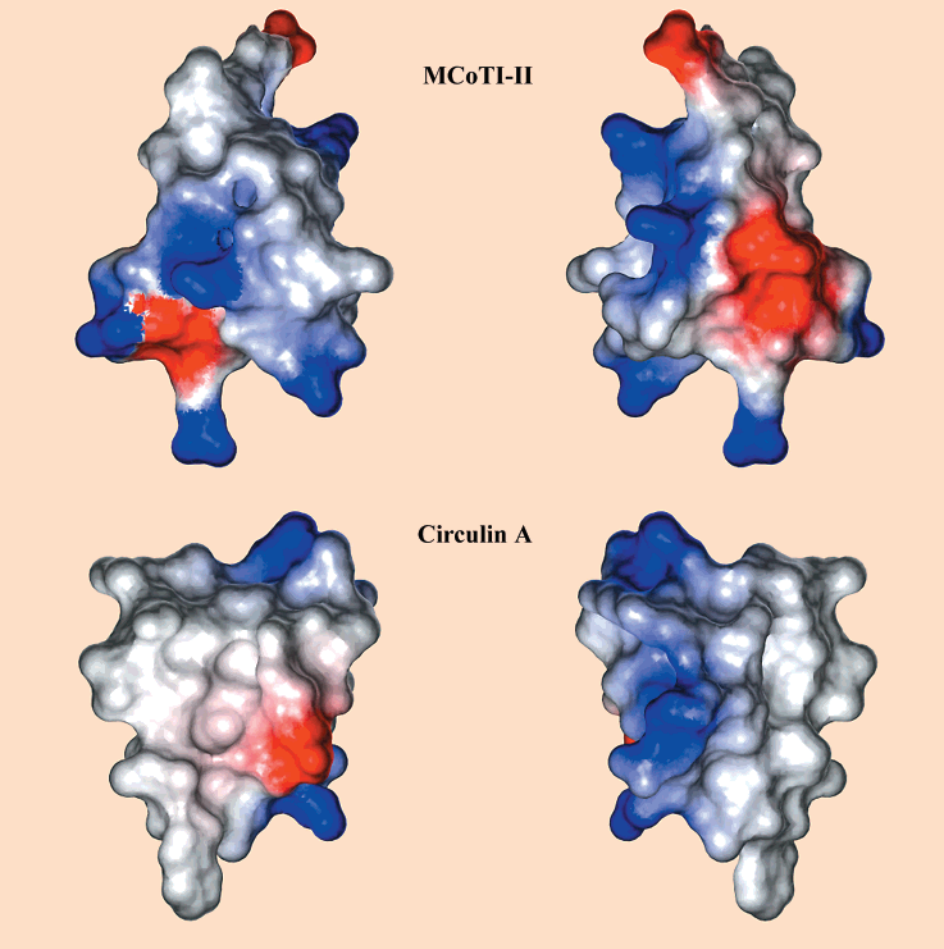

FIGURE 9: Comparison of the electrostatic potential at the molecular surface of MCoTI-II (top) and circulin A (bottom). The right-handed views are after a 180° rotation around the vertical axis from the left-handed views.

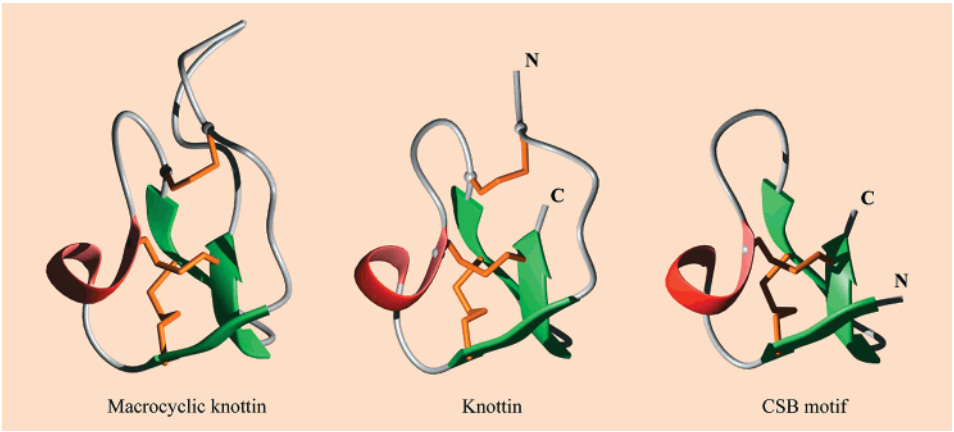

FIGURE 10: Schematic view of knottin structural folds of decreasing complexity.

we called the CSB motif and that this elementary motif is an autonomous folding unit (*20*). It is a possibility that the numerous small disulfide rich protein families containing the CSB motif have evolved from an ancestral smaller protein, although no such natural protein has yet been observed.

In this paper, we have described the 3D conformation of the first known member of a new family of cyclic knottins. Disulfide bridges and cyclization may be responsible for the protease resistance and well-defined structure of this compound. These knottin fold containing compounds are small and easily accessible to synthesis, but thanks to disulfide bridges, they still display remarkable stability and protease resistance. Also, their well-defined conformation allows accurate geometrical analyses and predictions in the course of design strategies.

Three different topological frameworks of CSB motif containing molecules with increasing complexity are compared in Figure 10. The simplest CSB motif has the advantage of a smaller size and of including only two disulfide bridges. This limits the number of potential disulfide isomers to three, and may facilitate selective synthesis of the desired correct disulfide bridges for modified sequences that would otherwise give rise to disulfide isomers. The knottins possess one more disulfide bridges increasing the

number of potential disulfide isomers to 15. This may bring synthesis and/or purification problems. However, the additional disulfide also affords a higher stability and a supplementary loop. This arrangement is also the most frequent in nature. Finally, the cyclic knottins display a still higher stability and the advantage of good resistance to exoproteases, at the expense of more complex chemical synthesis. A significant number of natural proteins with this arrangement have been identified in the last few year, but the process by which the in vivo cyclization occurs remains to be determined. Understanding this process would be of interest if one wants to use this very interesting scaffold in combinatorial approaches such as the phage display technology.

ACKNOWLEDGMENT

Authors warmly thank Dr. Charles Hetru for kindly performing antimicrobial assays.

SUPPORTING INFORMATION AVAILABLE

Tables of ^{1}H and ^{13}H chemical shifts of MCoTI-II and -IIb and constraints used for structure calculation of MCoTI-II. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cunningham, B. C., and Wells, J. A. (1997) *Curr. Opin. Struct. Biol. 7*, 457−462.
2. Nygren, P. A., and Uhlen, M. (1997) *Curr. Opin. Struct. Biol. 7*, 463−469.
3. Vita, C., Roumestand, C., Toma, F., and Menez, A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6404−6408.
4. Vita, C. (1997) *Curr. Opin. Biotechnol. 8*, 429−434.
5. Tamaoki, H., Miura, R., Kusunoki, M., Kyogoku, Y., Kobayashi, Y., and Moroder, L. (1998) *Protein Eng. 11*, 649−659.
6. Orengo, C. A., Flores, T. P., Taylor, W. R., and Thornton, J. M. (1993) *Protein Eng. 6*, 485−500.
7. Lesk, A. M. (1995) *J. Mol. Graphics 13*, 159−164.
8. Wang, Z.-X. (1998) *Protein Eng. 11*, 621−626.
9. Heitz, A., Chiche, L., Le-Nguyen, D., and Castro, B. (1989) *Biochemistry 28*, 2392−2398.
10. Chiche, L., Gaboriaud, C., Heitz, A., Mornon, J. P., Castro, B., and Kollman, P. A. (1989) *Proteins 6*, 405−417.
11. Bode, W., Greyling, H. J., Huber, R., Otlewski, J., and Wilusz, T. (1989) *FEBS Lett. 242*, 285−292.
12. Likos, J. J. (1989) *Int. J. Pept. Protein Res. 34*, 381−386.
13. Krishnamoorthi, R., Lin, C. L., Gong, Y. X., VanderVelde, D., and Hahn, K. (1992) *Biochemistry 31*, 905−910.
14. Huang, Q., Liu, S., and Tang, Y. (1993) *J. Mol. Biol. 229*, 1022−1036.
15. Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P., and Smalas, A. O. (1999) *Acta Crystallogr., Sect. D 55*, 139−148.
16. Zhu, Y., Huang, Q., Qian, M., Jia, Y., and Tang, Y. (1999) *J. Protein Chem. 18*, 505−509.
17. Le Nguyen, D., Heitz, A., Chiche, L., Castro, B., Boigegrain, R. A., Favel, A., and Coletti-Previero, M. A. (1990) *Biochimie 72*, 431−435.
18. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) *Protein Sci. 3*, 1833−1839.
19. Narasimhan, L., Singh, J., Humblet, C., Guruprasad, K., and Blundell, T. (1994) *Nat. Struct. Biol. 1*, 850−852.
20. Heitz, A., Le-Nguyen, D., and Chiche, L. (1999) *Biochemistry 38*, 10615−10625.
21. Pereira, P. J., Lozanov, V., Patthy, A., Huber, R., Bode, W., Pongor, S., and Strobl, S. (1999) *Struct. Folding Des. 7*, 1079−1088.
22. Christmann, A., Walter, K., Wentzel, A., Kratzner, R., and Kolmar, H. (1999) *Protein Eng. 12*, 797−806.
23. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) *Biochemistry 34*, 4147−4158.
24. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) *J. Nat. Prod. 57*, 1619−1625.
25. Daly, N. L., Koltay, A., Gustafson, K. R., Boyd, M. R., Casas-Finet, J. R., and Craik, D. J. (1999) *J. Mol. Biol. 285*, 333−345.
26. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol. 294*, 1327−1336.
27. Hallock, Y. F., Sowder, I. R., Pannell, L. K., Hughes, C. B., Johnson, D. G., Gulakowski, R., Cardellina, I. J., and Boyd, M. R. (2000) *J. Org. Chem. 65*, 124−128.
28. 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.
29. Hetru, C., and Bulet, P. (1997) *Methods Mol. Biol. 78*, 35−49.
30. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons Inc, NY.
31. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
32. Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., and Tschudin, R. (1990) *J. Magn. Reson. 86*, 304−318.
33. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−5, 29−32.
34. Wüthrich, K., Billeter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
35. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
36. Guntert, P., Braun, W., and Wuthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
37. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1999) University of California San Francisco.
38. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
39. Tsui, V., and Case, D. A. (2000) *J. Am. Chem. Soc. 122*, 2489−2498.
40. van Gunsteren, W. F., and Berendsen, H. J. C. (1977) *Mol. Phys. 34*, 1311−1327.
41. Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon 39*, 43−60.
42. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
43. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
44. Chiche, L., Heitz, A., Padilla, A., Le-Nguyen, D., and Castro, B. (1993) *Protein Eng. 6*, 675−682.
45. Le-Nguyen, D., Heitz, A., Chiche, L., El Hajji, M., and Castro, B. (1993) *Protein Sci. 2*, 165−174.
46. Heitz, A., Chiche, L., Le-Nguyen, D., and Castro, B. (1995) *Eur. J. Biochem. 233*, 837−846.
47. Heitz, A., Le-Nguyen, D., Castro, B., and Chiche, L. (1997) *Lett. Pept. Sci. 4*, 245−249.
48. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
49. Doreleijers, J. F., Rullmann, J. A., and Kaptein, R. (1998) *J. Mol. Biol. 281*, 149−164.

50. Holak, T. A., Bode, W., Huber, R., Otlewski, J., and Wilusz, T. (1989) *J. Mol. Biol. 210*, 649−654.
51. Nilges, M., Habazettl, J., Brunger, A. T., and Holak, T. A. (1991) *J. Mol. Biol. 219*, 499−510.
52. Tam, J. P., Lu, Y. A., Yang, J. L., and Chiu, K. W. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 8913−8918.
53. Hancock, R. E., and Scott, M. G. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 8856−8861.
54. Giudici, A. M., Regente, M. C., and delaCanal, L. (2000) *Plant Physiol. Biochem. 38*, 881−888.
55. Shao, F., Hu, Z., Xiong, Y. M., Huang, Q. Z., WangCg, Zhu, R. H., and Wang, D. C. (1999) *Biochim. Biophys. Acta 1430*, 262−268.
56. Mar, R. I., Carver, J. A., Sheil, M. M., Boschenok, J., Fu, S., and Shaw, D. C. (1996) *Phytochemistry 41*, 1265−1274.
57. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.

BI0106639