Isolation, Solution Structure, and Insecticidal Activity of Kalata B2, a Circular Protein with a Twist: Do Möbius Strips Exist in Nature?[†,‡]

Cameron V. Jennings,[§] K. Johan Rosengren,[||] Norelle L. Daly,[||] Manuel Plan,[||] Jackie Stevens,[§] Martin J. Scanlon,[⊥] Clement Waine,[||] David G. Norman,[@] Marilyn A. Anderson,[§] and David J. Craik*[,||]

The Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, Australia 4072, Victorian College of Pharmacy, Monash University, Parkville, Australia 3052, School of Life Sciences, The University of Dundee, Dundee, Scotland DD1 5EH, and Department of Biochemistry, La Trobe University, Bundoora, Victoria, Australia 3086

Received October 7, 2004; Revised Manuscript Received November 3, 2004

ABSTRACT: A large number of macrocyclic miniproteins with diverse biological activities have been isolated from the Rubiaceae, Violaceae, and Cucurbitaceae plant families in recent years. Here we report the three-dimensional structure determined using ^{1}H NMR spectroscopy and demonstrate potent insecticidal activity for one of these peptides, kalata B2. This peptide is one of the major components of an extract from the leaves of the plant *Oldenlandia affinis*. The structure consists of a distorted triple-stranded β-sheet and a cystine knot arrangement of the disulfide bonds and is similar to those described for other members of the cyclotide family. The unique cyclic and knotted nature of these molecules makes them a fascinating example of topologically complex proteins. Examination of the sequences reveals that they can be separated into two subfamilies, one of which contains a larger number of positively charged residues and has a bracelet-like circularization of the backbone. The second subfamily contains a backbone twist due to a *cis*-peptidyl−proline bond and may conceptually be regarded as a molecular Möbius strip. Kalata B2 is the second putative member of the Möbius cyclotide family to be structurally characterized and has a *cis*-peptidyl−proline bond, thus validating the suggested name for this subfamily of cyclotides. The observation that kalata B2 inhibits the growth and development of *Helicoverpa armigera* larvae suggests a role for the cyclotides in plant defense. A comparison of the sequences and structures of kalata B1 and B2 provides insight into the biological activity of these peptides.

In recent years, a large number of plant-derived macrocyclic peptides have been reported and defined as the family of plant cyclotides (*1*). They comprise 28−37 amino acids, have a variety of biological activities, and are characterized by a head-to-tail cyclized backbone and six conserved cysteine residues. The cysteine residues form three disulfide bonds that are arranged in a cystine knot motif, in which two disulfides and their interconnecting backbone form an embedded ring through which the third disulfide bond is threaded (*2*, *3*).

Prior to 1993, there were no reported examples of macrocyclic peptides of this size, with all known cyclic peptides being less than half the size of the cyclotides and being derived only from microorganisms or marine invertebrate animals. The early reports on the discovery of macrocyclic peptides in plants arose from screening programs for various bioactivities. For example, circulins A and B were discovered by researchers at the U.S. National Cancer Institute during a screen for anti-HIV activity (*4*), and cyclopsychotride A was discovered by Merck researchers looking for neurotensin antagonist activity (*5*). Violapeptide I was discovered in a screen aimed at uncovering hemolytic activity (*6*), while kalata B1 was discovered on the basis of native medicine applications of an extract from the plant *Oldenlandia affinis* (*7*, *8*). Since then, a large number of related macrocyclic peptides have been reported in plants from the Rubiaceae and Violaceae families (*1*, *9*−*13*).

The sequences of selected cyclotides and their general fold are summarized in Figure 1, from which the absolute conservation of the six cysteine residues is readily apparent. Since none of the cysteine residues are sequentially adjacent to another, and because the peptide backbone is cyclic, there are six intra-cysteine backbone segments in the structure. Examination of the sequences shows that while some loops are highly conserved there is variability of amino acid sequences within other loops. Furthermore, the peptides fall into two subfamilies, termed the bracelet and Möbius cyclotides (*1*). Within each subfamily, there is high level of sequence identity, but there are variations in the size and content of loops between subfamilies.

The first structural characterization of these macrocyclic peptides was reported in 1995 when we determined that the prototypical cyclotide, kalata B1, has a compact global fold

[†] D.J.C. is an Australian Research Council Professorial Fellow. C.V.J. was supported by an Australian Postgraduate Research Award. This work was supported by a grant from the Australia Research Council.

[‡] The coordinates representing the solution structure of kalata B2 have been submitted to the Protein Data Bank (entry 1PT4).

* To whom correspondence should be addressed. Fax: +61-7-33462019. Phone: +61-7-33462029. E-mail: d.craik@imb.uq.edu.au.

[§] La Trobe University.

[||] University of Queensland.

[⊥] Monash University.

[@] The University of Dundee.

10.1021/bi047837h CCC: $30.25
Published on Web 12/24/2004

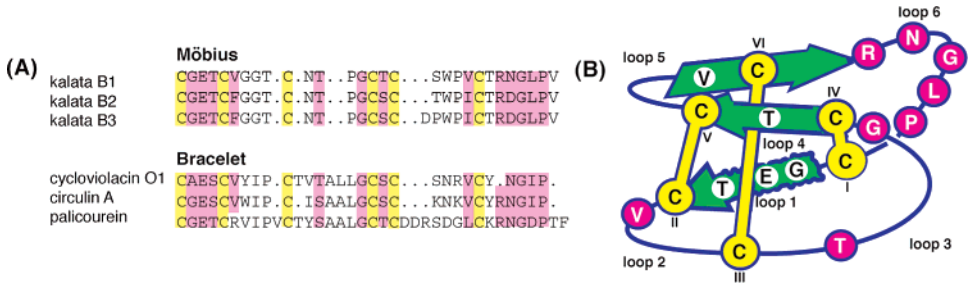

FIGURE 1: Sequences of selected cyclotides of the Möbius and bracelet families (A) and schematic illustration of the cyclotide fold, illustrating the characteristic cystine knot and triple-stranded β-sheet (B). While there is significant conservation within the two subfamilies, only a small set of residues, apart from the six cysteines (shown in yellow), are conserved throughout most of the cyclotide family. These residues are shown in pink in the sequence alignment (A), and their positions in the fold are represented by circles in panel B.

that incorporates the cystine knot motif (*8*). The structure also incorporates a distorted triple-stranded β-sheet and several turns. The nomenclature for the subfamilies is based on the presence or absence of a *cis*-Pro motif in loop 5. As a *cis* peptide bond has been likened to a twist in the backbone, peptides containing the potential for this twist have been named the Möbius cyclotides (*1*). While such a designation is useful in conceptual terms, so far structures have not been determined for any members of the Möbius subfamily other than kalata B1. In contrast, structures of several members of the bracelet cyclotide family, including circulin A (*14*), cycloviolacin 01 (*15*), and MCoTI-II (*16*, *17*), have been determined. Since the *cis*-Pro bond occurs in a β-turn in what has been proposed to be a key region of these molecules, it was of interest to determine the three-dimensional structure of another member of the Möbius family to confirm the generality of the Möbius classification.

The original report about the discovery of kalata B1 dates back to the early 1970s, although at that time only preliminary data on the peptide sequence were reported and neither the fact that the peptide was cyclic nor any structural information was known (*18*). The presence of a second peptide in the plant extract from *O. affinis* was reported, and this component was named kalata B2, although it was not characterized further. We have now isolated and characterized kalata B2 from *O. affinis* leaf tissue, and the amino acid sequence indicates that it is a member of the Möbius subfamily (Figure 1). We find that the expression level of individual peptides varies seasonally but that kalata B2 is consistently expressed at high levels, sometimes at levels higher than that of kalata B1. It was therefore of interest to fully characterize kalata B2 and to determine its three-dimensional structure and biological activity.

We are particularly interested in the biological functions of cyclotides within plants. It has been reported that kalata B1, cyclopsychotride A, and circulins A and B inhibit the growth of microorganisms (*19*). Furthermore, kalata B1 has a potent inhibitory effect on the growth and development of *Helicoverpa punctigera* larvae (*20*). Such activities suggest that the natural role of cyclotides in plants is as defense molecules. To see whether insecticidal activity is unique to kalata B1, we have tested kalata B2 for growth inhibition of *Helicoverpa armigera* larvae, a more important pest against a range of commercially important crop plants than *H. punctigera*. We show that kalata B2 possesses potent insecticidal activity, confirming that the natural role of cyclotides is likely to be in plant defense.

EXPERIMENTAL PROCEDURES

Isolation and Characterization of Kalata B2. Cyclotides were extracted from *O. affinis* leaf material. The initial step was to grind the plant material to a fine powder in liquid nitrogen using a mortar and pestle; 200 mL of a 1:1 MeOH/DCM mixture was added to the ground material and the mixture left to extract for 5 h or overnight. Subsequently, insoluble plant material was removed by filtration through cotton wool, and the MeOH and DCM were separated by the addition of water, resulting in the formation of a DCM and a MeOH/water phase. After several steps of fractionation, the crude extract was passed through a C18 flash coloumn, which was washed with 40% MeOH before elution with 100% MeOH. The fractions were lyophilized and dissolved in 1.5 mL of buffer A and analyzed by RP-HPLC. Kalata B2 was subsequently purified by RP-HPLC on a C18 preparative column using an 80 min 8 mL/min gradient from 0 to 80% buffer B. Buffer A was 0.05% (v/v) trifluoroacetic acid (TFA) in milli-Q H_2O, and buffer B was 0.05% (v/v) TFA and 90% (v/v) acetonitrile in milliQ H_2O. Kalata B2 was reduced with an excess of TCEP and alkylated with maleimide. The reduced and alkylated peptide was cleaved with Endo-Glu C in ammonium bicarbonate buffer at pH 8 overnight and then purified by reverse phase HPLC. The cleaved peptide was N-terminally sequenced using Edman degradation on an Applied Biosystems 477A protein sequencer.

NMR Spectroscopy. Samples prepared for ^{1}H NMR spectroscopy contained typically 1 mM peptide in 70% (v/v) ^2H_2O (Cambridge Isotope laboratories, Woburn, MA) and 30% (v/v) CD_3CN,[1] or 10% (v/v) ^2H_2O, 30% (v/v) CD_3CN, and 60% (v/v) H_2O at pH 3.6. Spectra were recorded at 290 and 298 K on a Bruker ARX-500 spectrometer or on a Bruker DRX-750 spectrometer. Two-dimensional NMR spectra were recorded in the phase sensitive mode using time-proportional phase incrementation (*21*) for quadrature detection in the t_1 dimension. Two-dimensional experiments included DQF-COSY (*22*), TOCSY using an MLEV-17 spin lock sequence (*23*) with a mixing time of 80 ms, and NOESY (*24*) with mixing times of 200 and 250 ms. Solvent suppression for NOESY and TOCSY experiments was achieved using a WATERGATE (water suppression by gradient-tailored excitation) sequence (*25*) and by selective low-power irradiation of the water resonance during a

[1] Abbreviations: CCK, cyclic cystine knot; CD_3CN, deuterated acetonitrile.

relaxation delay of 1.8 s for DQF-COSY spectra. Spectra were acquired with 4096 complex data points in F_2 and 512 increments in the F_1 dimension. Slowly exchanging NH protons were detected by acquiring a series of one-dimensional and TOCSY spectra of the fully protonated peptide immediately after dissolution in ^2H_2O and acetonitrile. Spectra were processed using XWINNMR (Bruker). The F_1 dimension was generally zero-filled to 2048 real data points, and 90° phase-shifted sine bell window functions were applied before Fourier transformation. Chemical shifts were referenced to DSS at 0.00 ppm.

Structure Calculations. Distance restraints were derived from 200 ms NOESY spectra recorded at 290 and 298 K. Cross-peaks were assigned and integrated in XEASY (*26*) and converted to distance restraints using DYANA (*27*). Backbone dihedral restraints were inferred from $^3J_{NH-H\alpha}$ coupling constants with ϕ restrained to $-120 \pm 40°$ for $^3J_{NH-H\alpha}$ values greater than 8 Hz and to $-65 \pm 25°$ for $^3J_{NH-H\alpha}$ values less than 5 Hz (*28*). Side chain χ_1 angles were restrained on the basis of observed NOE and $^3J_{H\alpha-H\beta}$ coupling patterns. For a predicted t^2g^3 side chain conformation, the χ_1 angles were restrained to $-60 \pm 30°$; for a g^2g^3 conformation, the angles were restrained to $60 \pm 30°$, and for a g^2t^3 conformation, the angles were constrained to $180 \pm 30°$. Three-dimensional structures were calculated using a simulated annealing and energy minimization protocol in CNS, version 1.0 (*29*). The starting structures were generated using random (ϕ and ψ) dihedral angles and were energy minimized to produce structures with the correct local geometry. A set of 50 structures was then generated using simulated annealing (*30*). The protocol involved a high-temperature phase comprising 4000 steps of 0.015 ps of torsion angle dynamics, a cooling phase with 4000 steps of 0.015 ps of torsion angle dynamics during which the temperature is lowered to 0 K, and finally an energy minimization phase comprising 5000 steps of Powell minimization. The NOE restraints were checked for violations, and ambiguous cross-peaks were resolved on the basis of interproton distances in the initial family of structures. A further 50 structures were calculated with the inclusion of ϕ and χ_1 dihedral angle restraints derived from spin−spin coupling constants and tightened distance restraints involving stereospecifically assigned protons that had been given pseudoatom correction status in the initial calculations. ω angles were set to *trans* with the exception of the Trp19−Pro20 angle, which was set to *cis* on the basis of local NOE characteristics of this configuration. In a final step, preliminary structures were subjected to energy minimization in a water shell using CNS and parameters developed by Linge *et al*. (*31*). The final set of restraints included 93 long-range, 65 medium-range, and 117 sequential distances, three hydrogen bonds, and 16 dihedral angles.

Protein structures were analyzed using PROMOTIF (*32*) and PROCHECK-NMR (*33*) and displayed using MOLMOL (*34*). The coordinates representing the solution structure of kalata B2 have been submitted to the Protein Data Bank as entry 1PT4.

Insect Feeding Trials. H. armigera larvae were raised on an artificial soybean flour diet consisting of the following ingredients: soy flour (7.42%), wheat germ (6.55%), torula yeast (5.46%), water (52.4%), 4% agar solution (27.29%), antibacterial agents, including methyl paraben (0.33%), sorbic acid (0.11%), and ascorbic acid (0.33%), and antifungal solution (0.11%), which consisted of 42% propionic acid, 4% phosphoric acid, and 54% water. The test diets were supplemented with either kalata B2, kalata B1, or casein (negative control) at a concentration of 0.825 μmol/g. Forty neonates were raised on each of the diets, and each larva was placed in an individual plastic container with a lid. Weight gain was recorded at the fourth day and every second day thereafter, and mortality was recorded every day. Larvae were initially fed small amounts of diet that were replaced as required to provide a continuous supply (*20*). The larvae were kept at 25 ± 1 °C with a cycle of light for 16 h and dark for 8 h.

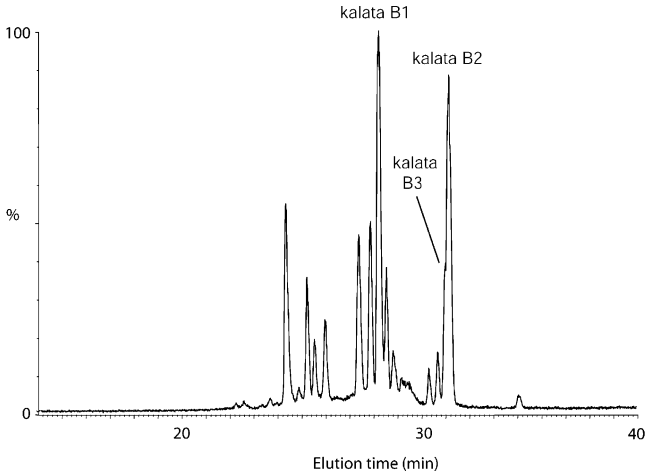

FIGURE 2: HPLC profile of the crude extract from *O. affinis* indicating the positions of the main components kalata B2 and kalata B1. As is evident from the trace, kalata B2 elutes significantly later than kalata B1, indicating a more hydrophobic surface nature. Also evident is the similar elution time of kalata B2 and kalata B3, which results in difficulties separating the two components into pure fractions.

RESULTS

Isolation and Purification of Kalata B2. In the initial reports on kalata B1, the peptide was isolated by boiling an aqueous mixture of *O. affinis* leaf material followed by butanol fractionation (*7*). For our studies, we isolate cyclotides from the above ground parts of Violaceae and Rubiaceae plants by methanol/dichloromethane extractions. The dichloromethane dissolves unwanted lipid components, and the cyclotides are partitioned preferentially into the methanol layer. Typically, the procedure involves homogenization of plant material and overnight extraction in a MeOH/DCM mixture followed by addition of water to separate the DCM from the MeOH/water layer and several steps of fractionation and purification using C18 flash columns and finally purification by RP-HPLC (*1*). Broadly similar procedures have been used in a range of laboratories (*4*, *5*, *10*, *35*). This strategy is similar to procedures for extraction of nonprotein bioactive components from plants but works well for the proteinaceous cyclotides because they are very stable. Furthermore, they are readily soluble in methanol. Figure 2 shows a HPLC trace of an extract from *O. affinis* leaves and stem tissue and identifies various peptides, including the major components kalata B1 and B2. Kalata B2 had a mass of 2954.8 Da, consistent with the sequence reported previously from both peptide sequencing and cDNA (*20*).

NMR Spectroscopy and Resonance Assignment. A sample of kalata B2 was prepared for NMR analysis by dissolution of the peptide to a concentration of ~0.5 mM. However, from inspection of the solution as well as from preliminary spectra, it was clear that, unlike kalata B1, which is freely soluble up to a concentration of ~6 mM in water, kalata B2 aggregates in aqueous solution at pH ~4. This pH is typically obtained on dissolution of HPLC-purified peptides due to the presence of TFA in the HPLC buffer and is convenient for NMR analysis. In subsequent experiments, the solubility was increased by adjusting the pH from ~4 to 7 or by addition of CD_3CN. While the increase in pH led to broadening of some amide signals due to enhanced amide exchange, the addition of cosolvent did not significantly affect the chemical shifts, or by implication the structure. The spectra used for structural analysis of kalata B2 were thus recorded on a sample containing ~1 mM peptide in a 70% H_2O/30% CD_3CN mixture (v/v). Under these conditions, kalata B2 gave rise to high-quality spectral data with sharp lines and good dispersion, indicating that the peptide was essentially monomeric.

For the determination of the structure of kalata B2, two-dimensional NMR spectra (500 and 750 MHz) were recorded and assigned using the sequential assignment procedure (*36*). The assignment was relatively straightforward as there was minimal overlap of amide protons. Of particular interest was a determination of the conformation of the peptide bond preceding Pro20 for validation of the naming scheme for the Möbius cyclotides. Similar to that of kalata B1, this bond is in the *cis* conformation as is evident from a strong sequential $H_\alpha-H_{\alpha+1}$ NOE between Trp19 and Pro20. The two other prolines in the sequence, Pro13 and Pro28, are in the *trans* conformation, as confirmed by strong sequential $H_\alpha-H_{\delta+1}$ NOEs. Interestingly, some unusual chemical shifts were observed for certain amino acid residues; e.g., the β-protons of Pro20 occur at the high-field values of 0.67 and −0.91 ppm. Similar trends were observed previously in kalata B1 and were thought to arise from the ring current effect of the adjacent Trp residue. This effect appears to be even more pronounced in kalata B2 than in kalata B1.

Disulfide Connectivity. The knotted disulfide connectivity of kalata B1 was confirmed recently by Rosengren *et al.* using NMR methods (*15*), and by Göransson and Craik (*37*) using a chemical approach following a suggestion that another disulfide arrangement might be possible (*38*). As the cystines are tightly packed in the molecular core, inter-cysteine NOE connections cannot be used to predict the disulfide connectivity. However, analysis of the structural data in combination with cystine side chain dihedral angles provided sufficient proof for the presence of the cyclic cystine knot motif (*15*). The high degree of sequence homology between kalata B1 and kalata B2 strongly suggested an identical disulfide connectivity, and this was confirmed by our NMR data showing that the cystine side chain orientations are identical to those seen in kalata B1.

Structure Determination. Figure 3 summarizes the sequential, medium-range, and long-range NOE information that was used to define the secondary structure of kalata B2. Several medium-range NOEs, indicating the presence of turns, and a large number of long-range contacts consistent with interstrand interactions across β-sheets were identified. On the basis of these data, it was concluded that kalata B2 adopts a fold similar to that of the previously reported cyclotide structures, which is characterized by a compact structure comprising a cystine knot combined with elements of β-sheet structure.

A full structure determination of kalata B2 was achieved through the collection of structural restraints in the form of inter-proton distances based on NOE connectivities and dihedral angle restraints based on spin−spin coupling constants. These restraints were subjected to simulated annealing and energy minimization protocols within CNS (*29*). This procedure involved generation of preliminary structures by torsion angle dynamics and further refinement by Cartesian dynamics and energy minimization in a water shell (*31*). Figure 4 shows a superposition of the 20 lowest-energy structures in stereo representation. The structures are of excellent precision, with an rmsd of 0.44 Å over the whole backbone region. Furthermore, they have good covalent geometry as is evident from low deviations from ideal bond lengths and angles and the Ramachandran analysis, which indicates that 75% of the residues are in the most favored regions and the remaining residues are in the additionally allowed regions (Table 1). Among the 25% that lie outside the most favored areas are Cys5, Cys17, and Trp19. These residues are found in similarly displaced positions in the NMR structure of kalata B1, which may indicate that the associated regions of the protein are slightly strained due to the compact cystine knot fold. In particular, Cys5 and Cys17 are connected in a disulfide bond that forms part of the embedded ring of the cystine knot, and Trp19 is on a hairpin turn directly adjacent to this disulfide bond.

Description of the Three-Dimensional Structure. The three-dimensional structure of kalata B2 is highly compact and characterized mainly by the three disulfide bonds forming the cystine knot motif, which to a large extent fills the molecular core. Elements of secondary structure include a triple-stranded β-sheet and several turns. Strand 1 (residues 16−18) and strand 2 (residues 21−23) are connected by a type VIa β-turn comprising the Trp19-*cis*-Pro20 motif, which thus forms a well-defined hairpin. Associated with the hairpin is a distorted β-bulge involving residues 5−29. Finally, loop 2 (residues 5−8) forms a type I β-turn; loop 3 (residues 12−15) forms a type II β-turn, and residues 24−27 in loop 6 adopt a type I′ β-turn. These structural elements are stabilized by a number of hydrophobic interactions and hydrogen bonds, which were identified by structural analysis in combination with amide exchange rates. In total, 10 amide backbone protons were slow exchanging, and for all these, main chain−main chain hydrogen bonds were identified in the structure.

An interesting structural feature that has been identified previously in both kalata B1 and cycloviolacin O1 is the interaction between the Glu3 side chain and several backbone amide protons in loop 3. This interaction appears, as predicted (*15*), to be conserved throughout the family as the amide protons of both Asn11 and Thr12 are close to the carboxyl group of Glu3. This interaction is particularly evident from the high-pH sensitivity of these protons. With an increase in pH, the Glu3 side chain becomes deprotonated, which results in a downfield shift of the Asn NH proton of ~3 ppm. A short stretch of helix in loop 3 has been reported for cycloviolacin O1 and circulin A, but kalata B2, like kalata

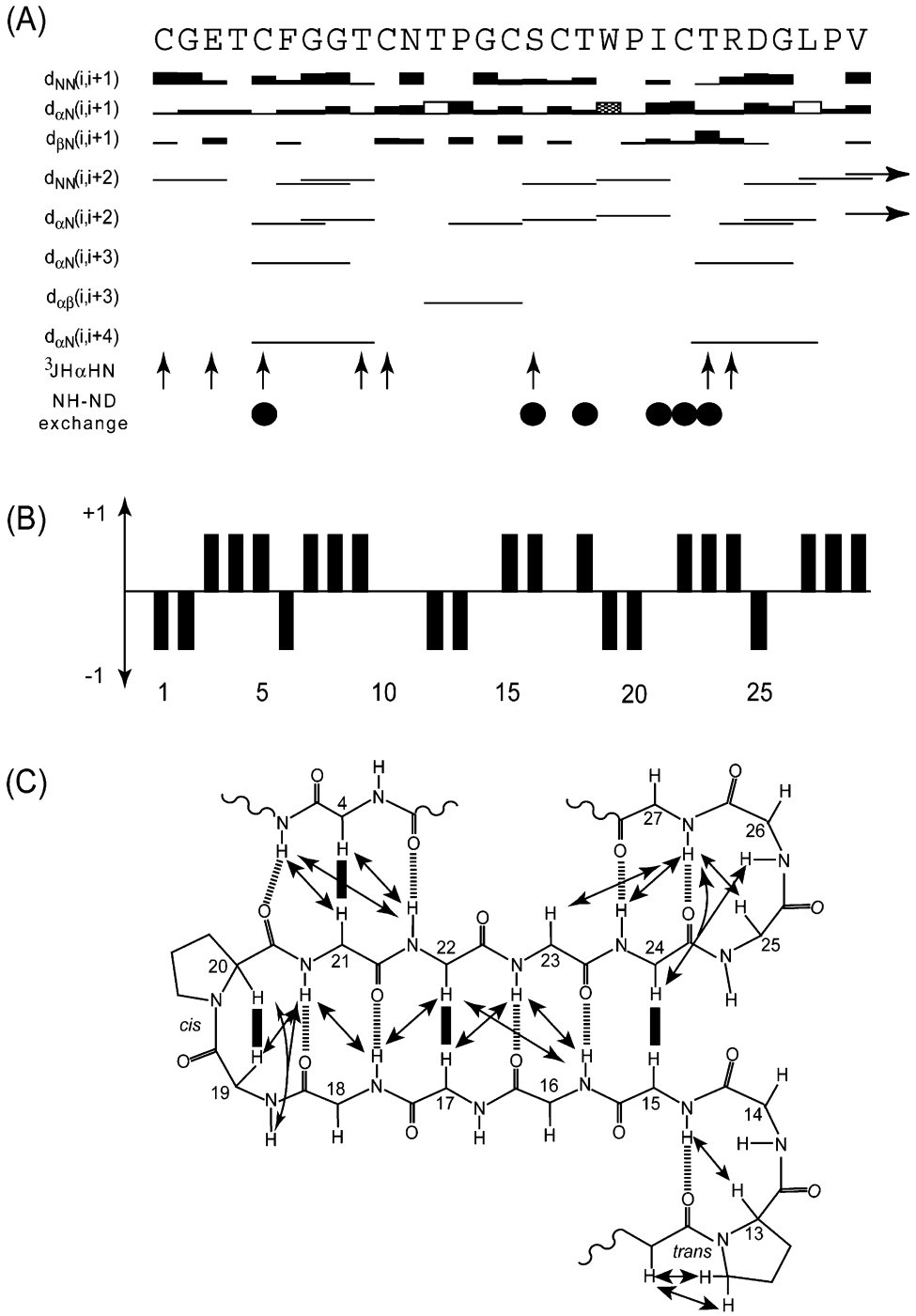

FIGURE 3: (A) Summary of coupling constants, amide exchange, and sequential and medium-range NOE connectivities for kalata B2. Black bars represent sequential NOE connectivities observed in a 200 ms NOESY spectrum from a 500 MHz spectrometer at 290 K. White bars represent $H_\alpha-H_\delta$ connectivities for the proline residues in the *trans* configuration, and the checkered box represents the $H_\alpha-H_\alpha$ connectivity for the proline in the *cis* conformation. The height of the bar represents the strength of the NOE cross-peak, classified as either strong, medium, or weak. $^3J_{NH-H\alpha}$ coupling constants of ≥ 8 Hz ($-120°$) are represented as upward arrows. Filled circles indicate slowly exchanging amide protons still present 16 h after dissolution in ^2H_2O at 298 K. (B) Chemical shift indices derived from H_α chemical shifts of kalata B2. An index of -1 or $+1$ indicates a shift deviation from the random coil value of greater than 0.1 ppm upfield or downfield, respectively. (C) Schematic representation of the secondary structure within kalata B2. Long-range NOE data from NOESY experiments (in ^2H_2O or H_2O) are summarized with intrastrand NOEs presented at arrows and hydrogen bonds presented as dashed lines.

B1, comprises a shorter loop 3, which adopts a well-defined turn rather than a helical conformation.

Insecticidal Activity. On the basis of previous experiments showing that the cyclotide kalata B1 possesses potent insecticidal activity and thus potentially serves as a plant defense molecule, we tested kalata B2 for such activity. *H. armigera* larvae fed an artificial diet containing either kalata B1 or kalata B2 in amounts comparable to endogenous levels in plants exhibited significantly reduced growth and higher rates of mortality compared to insects fed a control diet containing no cyclotides, as shown in Table 2 and Figure 5. The insecticidal potency of the two peptides appears to be very similar, despite five amino acid substitutions between them.

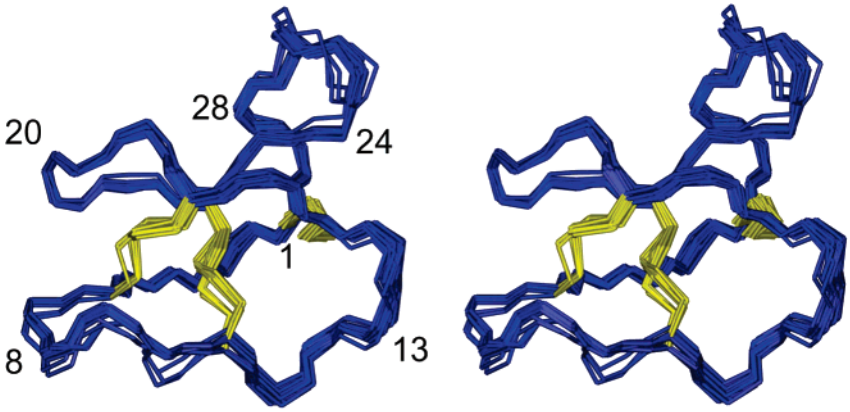

FIGURE 4: Overlay of 20 low-energy structures representing the solution structure of kalata B2 superimposed over all the backbone N, Cα, and C atoms. Selected residues are identified by residue number.

Table 1: Geometric and Energetic Statistics for the Family of 20 Kalata B2 Structures[a]

mean pairwise rmsd (Å)[b]	
backbone	0.43 ± 0.15
heavy atom	1.01 ± 0.21
mean rmsd from experimental restraints	
NOE (Å)	0.02 ± 0.002
dihedral angles (deg)	0.33 ± 0.15
mean rmsd from idealized covalent geometry	
bonds (Å)	0.004 ± 0.0002
angles (deg)	0.55 ± 0.034
impropers (deg)	0.54 ± 0.063
mean energies (kcal/mol)	
E_{NOE}[c]	6.49 ± 1.5
E_{dih}[c]	0.13 ± 0.1
E_{L-J}	−79.37 ± 6.34
E_{bond}	5.81 ± 0.56
$E_{improper}$	9.22 ± 2.26
E_{angle}	33.1 ± 4.16
E_{total}	−906.1 ± 20.54
Ramachandran plot statistics	
residues in most favored regions	75%
residues in additional allowed regions	25%
residues in generously allowed and disallowed regions	0%

[a] The values are given as the mean ± the standard deviation. [b] The rms deviation measured for the whole molecule. [c] Force constants for the calculation of square-well potentials for the NOE and dihedral angle restraints were 50 kcal mol^{-1} Å^{-1} and 200 kcal mol^{-1} rad^{-2}, respectively.

Table 2: Weight[a] and Mortality Rates of *H. armigera* Fed with a Diet Containing Kalata B1 or Kalata B2

	weight (mg) ± standard deviation	mortality (%)
kalata B1	112 ± 60	20
kalata B2	135 ± 70	28
control	462 ± 60	0

[a] Weight recorded at day 14 of the feeding trial.

DISCUSSION

Until recently, the only naturally occurring cyclic peptides known were of microbial or marine invertebrate origin. Because of their favorable properties, cyclic peptides (or mimics of them) have been used widely in pharmaceutical applications, perhaps best exemplified by the immunosuppressant cyclosporin. However, these classic cyclic peptides, which are not directly gene-encoded, invariably comprise fewer than 15 amino acids, usually lack disulfide bonds, and generally do not have well-defined three-dimensional structures. By contrast, the cyclotides are on the order of 30 amino acids in size, exhibit folded structures that are more characteristic of true proteins, and are direct gene products.

FIGURE 5: Effects of kalata B1 and kalata B2 on the growth and development of *H. armigera* larvae. Insects in the middle were fed a control diet, while those on the left and right were fed diets containing trace amounts of kalata B1 and kalata B2, respectively.

In this study, we have isolated and characterized kalata B2, a member of the Möbius family of cyclotides. This molecule is probably the peptide first named by Gran in 1970 (*18*) but not described in detail at the time. Here we show that along with kalata B1 it is one of the major peptide-based constituents of *O. affinis*. It differs from kalata B1 in five amino acid residues but adopts a similar three-dimensional structure.

To date, a large number of cyclotides have been isolated from plants, and the presence of dozens of different cyclotides in a single species (*39*) raises the question of their natural role in plants. Numerous bioactivities have been reported, but perhaps the most exciting finding is that kalata B1 possesses potent insecticidal activity. In this study, we have shown that a second member of the cyclotide family, kalata B2, also possess insecticidal activity against *Helicoverpa* larvae, confirming that the activity is not specific to kalata B1. The concentration of kalata B1 and kalata B2 used in the feeding trial is similar to that present naturally in *O. affinis*, supporting the idea that the cyclotides provide efficient protection from leaf-eating predators.

Figure 6 shows a comparison of the structures of kalata B2 and kalata B1 and illustrates that both the molecular core and many side chain groups have well-defined structural

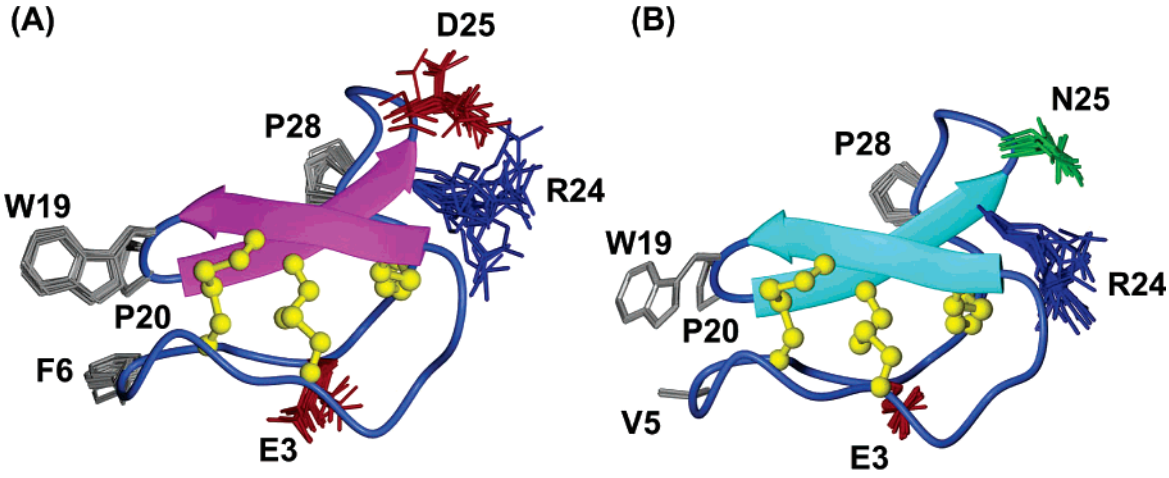

FIGURE 6: Ribbon diagram of kalata B2 (A) and kalata B1 (B) illustrating the cystine knot, the central β-sheet, and the positions of selected amino acid side chains. The structures are both well-defined with the exception of parts of loop 6, in particular, the Arg24 and Asp/Asn25 side chains. This region of the molecule is involved in the post-translational processing that generates the cyclic protein backbone.

roles. The backbones of the structures can be superimposed with an rmsd of 0.96 Å. The main structural features are the compact core comprising the cystine knot and the central β-sheet region. Recently, we described the interactions of each amino acid in kalata B1 in detail (*15*), and many features deduced in that study are consistent with the structure of kalata B2 presented here. In summary, it would appear that apart from the cystines, Glu3 and a number of Gly residues are the most important for retaining the structure of the loop regions of kalata B2. Glu3 interacts closely with loop 3 in a network of hydrogen bonds. Gly2, Gly8, Gly14, and Gly26 all adopt positive ϕ angles, a conformational requirement strongly favoring Gly residues, and are presumably necessary for achieving the tightly folded structure that is braced by three disulfide bonds. It is interesting to note that in both structures loop 6 appears to be the most flexible, highlighted by the lack of a preferred orientation of the side chains of Arg24 and Asn/Asp25, despite most residues in this region being highly conserved. On the basis of the known gene sequences, this region is involved in biosynthetic cyclization, which involves the formation of a peptide bond between the natural N-terminal Gly26 and the C-terminal Asn/Asp25 (*20*). The likely reason for the high level of conservation in this region is that certain residues are required for the processing of the precursor rather than the stability of the final fold. As there are no data available on the mechanism of biological action of the cyclotides, it is difficult to speculate about the residues that may be important for the insecticidal activity, although we are currently synthesizing a series of Ala mutants to gain an insight into structure−activity relationships.

One of the key features of kalata B2 is a Trp−*cis*-Pro peptide bond that occurs in a β-turn region of the molecule, as shown in Figure 7. The existence of this *cis* bond is confirmed by a characteristic Hα−Hα NOE between the Trp and the following Pro. *Cis* amide bonds are relatively rare in proteins because of the significant energy penalty resulting from such an arrangement. However, the steric hindrance associated with *cis*-peptide bonds is greatly reduced when the second residue is a proline, resulting in approximately 6−7% of all X−Pro bonds in known protein structures adopting a *cis* conformation (*40*). Furthermore, NMR studies of short proline-containing peptide fragments and analyses

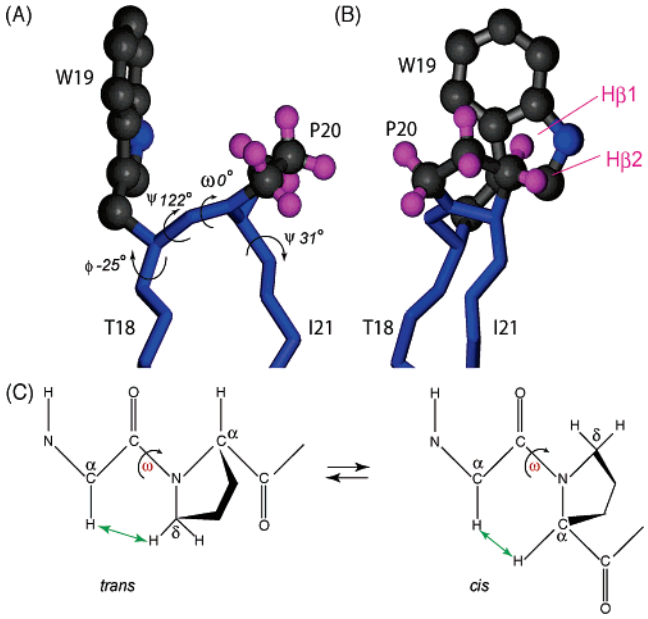

FIGURE 7: Geometry of loop 5 in kalata B2. Panels A and B show the lowest-energy structure of kalata B2 and illustrate the *cis* geometry of the Trp19−Pro20 imide bond and the close interaction between the Trp and Pro side chains. Panel C illustrates the differences in the *trans* and *cis* geometries and the effects on the sequential interactions observed across the peptide bond.

of the structural information available in the Protein Data Bank have indicated that a *cis*-Pro conformation is stabilized by the presence of an aromatic residue in the position immediately preceding the proline. In solution, the GFPG, GYPG, and GWPG tetrapeptides exhibit 17, 21, and 25% of the *cis*-Pro isomers, respectively *(40)*. Among protein structures deposited in the Protein Data Bank, a similar proportion (20−25%) of Tyr−Pro dipeptide motifs adopt a *cis* conformation, but only 6−9% of all Phe−Pro bonds do so. For Trp, it is difficult to estimate a statistically relevant figure as Trp−Pro motifs are rare. One explanation for the lower *cis* ratio for the Phe−Pro motif is that most *cis* amide bonds occur in turn regions and turns in general are found at the protein surface, but it is unusual to find hydrophobic residues such as Phe and Trp exposed on the surface of proteins. The Tyr side chain has a more hydrophilic nature

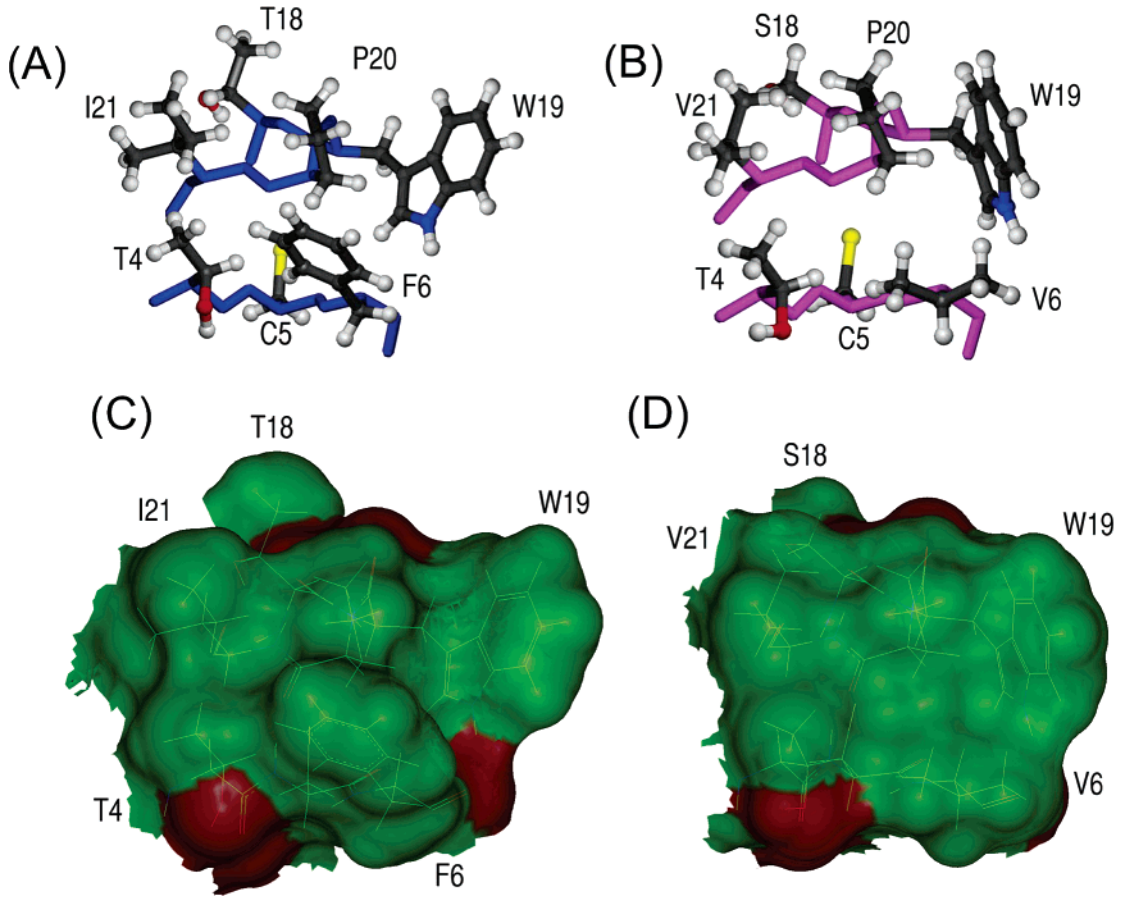

FIGURE 8: Comparison of the major surface-exposed hydrophobic patch in kalata B2 (A and C) and kalata B1 (B and D). While the hydrophobic nature of this region is conserved throughout the cyclotide family, there are differences in the size and hydrophobicity of the individual amino acids. The Val to Phe, Ser to Thr, and Val to Ile substitutions at positions 6, 18, and 21, respectively, shown in panels A and B result in an extended solvent-exposed surface as shown in panels C and D. This is likely the reason for the overall more hydrophobic nature and the tendency for aggregation of kalata B2 in comparison to kalata B1.

and is more likely to be surface-exposed (*40*). A special feature of the cyclotides is that the molecular core is filled by the cystines, leaving most hydrophobic groups, including Trp19, exposed on the surface. Its location in a surface-exposed turn makes it not surprising that the Trp−Pro bond in kalata B1 and kalata B2 adopts a *cis* configuration.

In an earlier report (*1*), we introduced the names "Möbius" and "Bracelet" to describe two major subfamilies of cyclotides based on the suggestion that a *cis* X−Pro bond may be considered a conceptual twist in the backbone of a cyclic peptide. An all-*trans* backbone leads to the bracelet topology, while molecules with a single *cis*-Pro can be regarded as Möbius strips. At the time, only one structure had been determined for a Möbius cyclotide (kalata B1), and it was assumed by homology that the SWPV segment in loop 5 would contain such a *cis* bond in other members. The structure reported here for kalata B2 supports this hypothesis.

Although there are many similarities between kalata B2 and kalata B1, there are also some interesting differences. Kalata B2 elutes significantly later than kalata B1 on RP-HPLC, and as mentioned earlier, kalata B2 is less soluble in aqueous solutions. Both findings suggest a more hydrophobic surface of kalata B2 despite it having an additional charge in the form of Asp25. The presence of a surface-exposed hydrophobic patch is a conserved feature of the cyclotide family, as illustrated by Figure 8, which shows the clustering of the hydrophobic side chains mainly in loops 2 and 5 in both kalata B2 and kalata B1. From here, it is clear that the hydrophobic patch is extended in kalata B2 by the presence of more hydrophobic residues Phe6, Thr18, and Ile21 in place of Val6, Ser18, and Val21, respectively, in kalata B1. It is likely these differences lead to the increased propensity for aggregation of kalata B2. We recently demonstrated by ultracentrifugation studies that while kalata B1 is essentially monomeric in aqueous solution at millimolar concentrations, kalata B2 forms ~30% monomers, ~42% tetramers, and ~25% octamers (*41*).

To date, nothing is known about the mechanisms through which kalata B2 and kalata B1 exert their insecticidal activities. A likely mode of action would be membrane disruption, but it is unclear whether it will be a result of the formation of pores or simply a generalized disruption of the membrane structure. Many peptidic toxins, in particular, those with extensive surface-exposed hydrophobic patches, have been shown to form pores in membranes, including mellitin from bee venom (*42*). Interestingly, mellitin forms a mixture of monomers and tetramers in solution (*43*) similar to what has been observed for kalata B2 (*41*). However, it is unclear whether the formation of tetramers is required for the biological activity in the case of the cyclotides. On one hand, it could be argued that it is not because although there are differences in the oligomerization behavior of kalata B2 and kalata B1, they appear to have similar insecticidal activity. On the other hand, it may be that while kalata B1

does not form tetramers in aqueous solution, it may do so in the presence of membranes.

In conclusion, the cyclotides are fascinating molecules that make up the largest family of naturally occurring cyclic proteins (*44*). Although containing only 30 amino acids, they adopt all of the features of larger proteins in that they have elements of secondary structure, extensive protection of amide protons from solvent via hydrogen bonding, and a compact global fold associated with a range of biological activities. They fall into two highly conserved subfamilies that were named the bracelet and Möbius cyclotides (*1*), based on the presence of a *cis*-peptidyl−proline bond in kalata B1. Here we described the solution structure of kalata B2, which is only the second member of the Möbius subfamily to be isolated and characterized. We have shown that this protein contains a *cis*-peptidyl−proline bond, thus validating the naming scheme for the subfamilies of cyclotides. Furthermore, we have shown that kalata B2 displays insecticidal activity against *H. armigera*, with a potency similar to that of kalata B1. This demonstration suggests the cyclotides may be useful as "natural" insecticides if applied to crop plants (*45*). Given the cyclotides are encoded by genes, the production of transgenic plants may be possible in much the same way that the *Bacillus thuringiensis* (Bt) toxin has been used to protect corn, cotton, and soybeans.

REFERENCES

1. 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.
2. Craik, D. J., Daly, N. L., and Waine, C. (2001) The cystine knot motif in toxins and implications for drug design, *Toxicon 39*, 43−60.
3. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides, *Protein Sci. 3*, 1833−1839.
4. Gustafson, K. R., Sowder, R. C., II, Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, J. H., II, McMahon, J. B., Buckheit, R. W., Jr., Pannell, L. K., and Boyd, M. R. (1994) Circulins A and B: Novel HIV-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*, *J. Am. Chem. Soc. 116*, 9337−9338.
5. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) Cyclopsychotride A, a biologically active, 31-residue cyclic peptide isolated from *Psychotria longipes*, *J. Nat. Prod. 57*, 1619−1625.
6. Schöpke, T., Hasan Agha, M. I., Kraft, R., Otto, A., and Hiller, K. (1993) Hämolytisch aktive komponenten aus *Viola tricolor* L. und *Viola arvensis* Murray, *Sci. Pharm. 61*, 145−153.
7. Gran, L. (1973) Oxytocic principles of *Oldenlandia affinis*, *Lloydia 36*, 174−178.
8. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1, *Biochemistry 34*, 4147−4158.
9. Derua, R., Gustafson, K. R., and Pannell, L. K. (1996) Analysis of the disulfide linkage pattern in circulin A and B, HIV-inhibitory macrocyclic peptides, *Biochem. Biophys. Res. Commun. 228*, 632−638.
10. Göransson, U., Luijendijk, T., Johansson, S., Bohlin, L., and Claeson, P. (1999) Seven novel macrocyclic polypeptides from *Viola arvensis*, *J. Nat. Prod. 62*, 283−286.
11. 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 cochinchinensis* exhibit an atypical macrocyclic structure, *Biochemistry 39*, 5722−5730.
12. Bokesch, H. R., Pannell, L. K., Cochran, P. K., Sowder, R. C., II, McKee, T. C., and Boyd, M. R. (2001) A novel anti-HIV macrocyclic peptide from *Palicourea condensata*, *J. Nat. Prod. 64*, 249−250.
13. Broussalis, A. M., Göransson, U., Coussio, J. D., Ferraro, G., Martino, V., and Claeson, P. (2001) First cyclotide from Hybanthus (Violaceae), *Phytochemistry 58*, 47−51.
14. Daly, N. L., Koltay, A., Gustafson, K. R., Boyd, M. R., Casas-Finet, J. R., and Craik, D. J. (1999) Solution structure by NMR of circulin A: A macrocyclic knotted peptide having anti-HIV activity, *J. Mol. Biol. 285*, 333−345.
15. Rosengren, K. J., Daly, N. L., Plan, M. R., Waine, C., and Craik, D. J. (2003) Twists, knots, and rings in proteins. Structural definition of the cyclotide framework, *J. Biol. Chem. 278*, 8606−8616.
16. 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 cochinchinensis*, *J. Biol. Chem. 276*, 22875−22882.
17. 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 knottins, *Biochemistry 40*, 7973−7983.
18. Gran, L. (1970) An oxytocic principle found in *Oldenlandia affinis* DC, *Medd. Nor. Farm. Selsk. 12*, 173−180.
19. 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.
20. 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.
21. Marion, D., and Wüthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of $^1H-^1H$ spin−spin coupling constants in proteins, *Biochem. Biophys. Res. Commun. 113*, 967−974.
22. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering, *Biochem. Biophys. Res. Commun. 117*, 479−485.
23. Bax, A., and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy, *J. Magn. Reson. 65*, 355−360.
24. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy, *J. Chem. Phys. 71*, 4546−4553.
25. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
26. Eccles, C., Guntert, P., Billeter, M., and Wüthrich, K. (1991) Efficient analysis of protein 2D NMR spectra using the software package EASY, *J. Biomol. NMR 1*, 111−130.
27. Guntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
28. Pardi, A., Billeter, M., and Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-Cα proton coupling constants, $^3J_{HNH\alpha}$, in a globular protein. Use of $^3J_{HNH\alpha}$ for identification of helical secondary structure, *J. Mol. Biol. 180*, 741−751.
29. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54* (Part 5), 905−921.
30. Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation, *J. Magn. Reson. 124*, 154−164.
31. Linge, J. P., and Nilges, M. (1999) Influence of non-bonded parameters on the quality of NMR structures: A new force field for NMR structure calculation, *J. Biomol. NMR 13*, 51−59.
32. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF: A program to identify and analyze structural motifs in proteins, *Protein Sci. 5*, 212−220.
33. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR:

Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
34. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
35. Claeson, P., Göransson, U., Johansson, S., Luijendijk, T., and Bohlin, L. (1998) Fractionation protocol for the isolation of polypeptides from plant biomass, *J. Nat. Prod. 61*, 77−81.
36. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
37. Göransson, U., and Craik, D. J. (2003) Disulfide mapping of the cyclotide kalata B1. Chemical proof of the cystic cystine knot motif, *J. Biol. Chem. 278*, 48188−48196.
38. Skjeldal, L., Gran, L., Sletten, K., and Volkman, B. F. (2002) Refined structure and metal binding site of the kalata B1 peptide, *Arch. Biochem. Biophys. 399*, 142−148.
39. Trabi, M., and Craik, D. J. (2004) Tissue-specific expression of head-to-tail cyclized miniproteins in Violaceae and structure determination of the root cyclotide *Viola hederacea* root cyclotide1, *Plant Cell 16*, 2204−2216.
40. Wu, W. J., and Raleigh, D. P. (1998) Local control of peptide conformation: Stabilization of cis proline peptide bonds by aromatic proline interactions, *Biopolymers 45*, 381−394.
41. Nourse, A., Trabi, M., Daly, N. L., and Craik, D. J. (2004) A comparison of the self-association behavior of the plant cyclotides kalata B1 and kalata B2 via analytical ultracentrifugation, *J. Biol. Chem. 279*, 562−570.
42. Dempsey, C. E. (1990) The actions of melittin on membranes, *Biochim. Biophys. Acta 1031*, 143−161.
43. Iwadate, M., Asakura, T., and Williamson, M. P. (1998) The structure of the melittin tetramer at different temperatures: An NOE-based calculation with chemical shift refinement, *Eur. J. Biochem. 257*, 479−487.
44. Craik, D. J., Daly, N. L., Mulvenna, J., Plan, M. R., and Trabi, M. (2004) Discovery, structure and biological activities of the cyclotides, *Curr. Protein Pept. Sci. 5*, 297−315.
45. Craik, D. J., Anderson, M., and Daly, N. L. (2003) Circular proteins tied in knots, *Today's Life Sci. 15*, 28−32.

BI047837H