THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 278, No. 8, Issue of February 21, pp. 6314–6322, 2003

Printed in U.S.A.

Disulfide Folding Pathways of Cystine Knot Proteins

TYING THE KNOT WITHIN THE CIRCULAR BACKBONE OF THE CYCLOTIDES*[S]

Received for publication, October 13, 2002, and in revised form, December 9, 2002
Published, JBC Papers in Press, December 12, 2002, DOI 10.1074/jbc.M210492200

Norelle L. Daly, Richard J. Clark, and David J. Craik‡

From the Institute for Molecular Bioscience, Australian Research Council Centre for Functional and Applied Genomics, University of Queensland, Brisbane, 4072 Queensland, Australia

The plant cyclotides are a fascinating family of circular proteins that contain a cyclic cystine knot motif. The knotted topology and cyclic nature of the cyclotides pose interesting questions about folding mechanisms and how the knotted arrangement of disulfide bonds is formed. In the current study we have examined the oxidative refolding and reductive unfolding of the prototypic cyclotide, kalata B1. A stable two-disulfide intermediate accumulated during oxidative refolding but not in reductive unfolding. Mass spectrometry and NMR spectroscopy were used to show that the intermediate contained a native-like structure with two native disulfide bonds topologically similar to the intermediate isolated for the related cystine knot protein EETI-II (Le-Nguyen, D., Heitz, A., Chiche, L., El Hajji, M., and Castro B. (1993) *Protein Sci.* 2, 165–174). However, the folding intermediate observed for kalata B1 is not the immediate precursor of the three-disulfide native peptide and does not accumulate in the reductive unfolding process, in contrast to the intermediate observed for EETI-II. These alternative pathways of linear and cyclic cystine knot proteins appear to be related to the constraints imposed by the cyclic backbone of kalata B1 and the different ring size of the cystine knot. The three-dimensional structure of a synthetic version of the two-disulfide intermediate of kalata B1 in which Ala residues replace the reduced Cys residues provides a structural insight into why the two-disulfide intermediate is a kinetic trap on the folding pathway.

The cyclotide family comprises macrocyclic peptides that are ~30 residues in length and incorporate a knotted topology of disulfide bonds formed between six conserved cysteine residues (1, 2). They are interesting both structurally and functionally as they represent the largest family of naturally occurring circular proteins (3). Further, a diverse range of biological functions have been reported for these peptides, including uterotonic activity (4), inhibition of neurotensin binding (5), hemolytic (6, 7), anti-HIV (8), anti-microbial (7) and insecticidal activity (9), as well as trypsin inhibitory activity (10).

Cyclotides have been isolated from various species in the Rubiaceae, Violaceae, and Cucurbitaceae plant families. The three-dimensional structures of four members of the family (kalata B1, Ref. 11; circulin A, Ref. 12; cycloviolacin O1, Ref. 1; and MCoTI-II, Refs. 13 and 14), have been elucidated using NMR spectroscopy and have revealed the presence of a motif termed the cyclic cystine knot (CCK)[1] (1). The major structural features of the motif are a cyclic backbone, a distorted triple stranded β-sheet and a cystine knot arrangement of the three disulfide bonds. The cystine knot comprises an embedded ring formed by disulfide bonds I–IV and II–V that is threaded by the III–VI disulfide bond. This knotted topology was originally determined for kalata B1 using NMR data (11) but the corresponding disulfide connectivity of other members of the cyclotide family has since been confirmed by synthetic methods involving directed disulfide bond formation (7, 15). Despite this, a recent study based primarily on analysis of local NOE patterns (16) suggested that an alternative, non-knotted disulfide bonding pattern may be possible for kalata B1. The present study provides additional evidence in favor of the originally proposed knotted topology.

The first cyclotides discovered fell into two distinct subfamilies based on sequence similarities within, but not between, the subfamilies. These were termed the Möbius and bracelet subfamilies based on topological considerations (1), and comprised peptides isolated from the Rubiaceae and Violaceae plant families. A third subfamily has now been proposed based on a determination of the three-dimensional structure of MCoTI-II (13). This macrocyclic trypsin inhibitor was discovered in the seeds of a plant from the Cucurbitaceae, or squash, family (10), and its structure contains the CCK motif (13, 14) seen in the previously reported cyclotides despite having no sequence homology except for the six cysteine residues.

Acyclic forms of squash trypsin inhibitors have been known for many years and include EETI-II (17) and CMTI-I (18). The structures (19, 20) of these molecules are very similar to that determined for MCoTI-II, apart from the cyclic backbone. Interestingly, no acyclic forms of the cyclotides from the Rubiaceae and Violaceae plant families have been found. A sequence comparison of kalata B1, MCoTI-II, and EETI-II is given in Fig. 1 and illustrates the common cystine knot motif in the three molecules. Fig. 1 also defines and highlights the

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

[S] The on-line version of this article (available at *http://www.jbc.org*) contains supplementary materials.

The atomic coordinates and structure factors (code 1N1U) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ An Australian Research Council Senior Fellow. To whom correspondence should be addressed. Tel.: 61-7-3365-4945; Fax: 61-7-3365-2487; E-mail: d.craik@imb.uq.edu.au.

[1] The abbreviations used are: CCK, cyclic cystine knot; DTT, dithiothreitol; TCEP, Tris(2-carboxyethyl)phosphine; RP-HPLC, reversed-phase high performance liquid chromatography; LCMS, liquid chromatography mass spectrometry; TOCSY, two-dimensional total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; DQF-COSY, double quantum filtered correlation spectroscopy; ECOSY, exclusive correlation spectroscopy; WATERGATE, water suppression by gradient-tailored excitation; RMSD, root mean square deviation.

diversity of the backbone loops, defined as the regions between successive Cys residues in these peptides. Owing to their cyclic nature kalata B1 and MCoTI-II have six loops while the acyclic homologue EETI-II has five loops. In the current study it was of interest to compare the oxidative folding of kalata B1 with those of linear cystine knot peptides such as EETI-II because they have some interesting topological similarities and differences.

The cyclotides have recently been shown to be translated gene products (9) that are processed from linear precursor proteins; however, the mechanism of cyclization, and the folding pathways required to achieve the correct three-dimensional shape and disulfide bond connectivity are yet to be elucidated. Examination of the gene sequence suggests that loop 6 is where cyclization occurs and this is consistent with expectations based on a sequence comparison of cyclotides with acyclic cystine knot proteins such as EETI-II (Fig. 1). Effectively, loop 6 may be regarded as the linker region that is clasped together to form the cyclic peptide backbone.

The cyclic nature and knotted topology of the cyclotides gives them remarkable stability and there is much interest in determining how the unique knotted fold is formed. We have previously shown that prior cyclization of kalata B1 appears to facilitate oxidative folding into the native conformation (6). An acyclic form of kalata B1 required a hydrophobic solution environment to produce significant amounts of the correctly folded form, whereas the cyclic form folded efficiently into the native isomer in aqueous solution. We have also explored the importance of the cyclic backbone in the folding of cyclotides by synthesizing all of the topologically distinct acyclic permutants (21). This is equivalent to opening the backbone of kalata B1 in each of the six loops between successive Cys residues. We found that four of six permutants folded into native-like conformations, but two did not; these were precisely the two that involved breaking the embedded ring of the cystine knot. That study showed that a native-like fold can be produced without the cyclic backbone, but highlighted the importance of the cystine knot in forming the native conformation.

In the current study we have examined the oxidative refolding and reductive unfolding of kalata B1 and characterized a major folding intermediate. The properties of this intermediate, when compared with a topologically similar intermediate characterized for EETI-II (22), reveal fundamental differences in the folding pathways of cystine knot proteins with cyclic backbones relative to those with conventional, acyclic, backbones. Part of the study required the synthesis of a mutant in which a single disulfide bond is removed and we show that this adopts a native-like fold, consistent with expectations from the identified folding intermediate.

EXPERIMENTAL PROCEDURES

Isolation of Kalata B1—Kalata B1 was extracted from the above ground parts of *Oldenlandia affinis* with dichloromethane/methanol (1:1; v/v) overnight at room temperature. The extract was partitioned with chloroform and water and the water/methanol layer concentrated on a rotary evaporator prior to lyophilization. The dried product was re-dissolved in H_2O and purified using preparative RP-HPLC on a C_{18} column (Vydac) at 8 ml/min. Masses were analyzed using an ES-TOF Micromass LCT mass spectrometer.

Reductive Unfolding—Kalata B1 (0.5 mg/ml) was dissolved in 0.1 M ammonium bicarbonate, pH 8.5, containing varying concentrations of dithiothreitol (DTT) or TCEP (0.5–300 mM). To monitor the unfolding, aliquots were removed at various time intervals and quenched with 4% aqueous trifluoroacetic acid. Samples were analyzed using RP-HPLC and LCMS. A large-scale reduction was performed for the subsequent oxidative refolding experiments. Kalata B1 was reduced with an excess of DTT (1 M) in ammonium bicarbonate (pH 8.5) at room temperature for 2 h. The reduced material was purified on semi-preparative RP-HPLC at 3 ml/min.

Oxidative Refolding—Large scale refolding reactions were performed in 50% isopropyl alcohol (v/v) in 0.1 M ammonium bicarbonate, pH 8.5 and 1 mM reduced glutathione. To isolate the folding intermediate the refolding reaction was allowed to proceed for 15 min and then immediately purified on a semi-preparative reverse phase column. Following purification the fractions were lyophilized and stored at −20 °C. Small scale reactions were performed under various solution conditions and analyzed by RP-HPLC and mass spectrometry.

Synthesis of [Ala^{1,15}]Kalata B1—[Ala^{1,15}]Kalata B1 was synthesized using a native chemical ligation strategy. This methodology requires the synthesis of a linear precursor peptide containing an N-terminal Cys residue. This forms a thioester with a functionalized C terminus and subsequently undergoes an *S,N* acyl migration to form a native peptide bond in the resultant circular protein. Of the four possible Cys residues in the di-Ala mutant peptide, Cys^5 was chosen as the N-terminal residue. The linear precursor was synthesized using solid phase peptide synthesis with BOC/HBTU chemistry and cleaved from the resin with HF. The cyclization reaction was performed in 0.1 M sodium phosphate (pH 7.4) with an excess of TCEP over a period of ~30 min. The cyclic reduced precursor was then oxidized in 0.1 M ammonium bicarbonate, pH 8.5, 50% v/v isopropyl alcohol and 1 mM reduced glutathione at room temperature overnight. The peptide was purified as described above.

NMR Spectroscopy—Samples for ^{1}H NMR measurements contained ~1 mM peptide in 90% H_2O/10% ^2H_2O (v/v) at ~pH 3. D_2O (99.9% and 99.99%) was obtained from Cambridge Isotope Laboratories, Woburn, MA. Spectra were recorded at 290 and 298 K on a Bruker ARX-500 or Bruker DMX 750 spectrometer equipped with a shielded gradient unit. two-dimensional NMR spectra were recorded in phase-sensitive mode using time-proportional phase incrementation for quadrature detection in the t_1 dimension (23). The two-dimensional experiments consisted of a TOCSY (24) using a MLEV-17 spin lock sequence (25) with a mixing time of 80 ms, DQF-COSY (26), ECOSY (27), and NOESY (28) with mixing times of 100–250 ms. Solvent suppression was achieved using a modified WATERGATE sequence (29). Spectra were acquired over 6024 Hz with 4096 complex data points in F_2 and 512 increments in the F_1 dimension. $^3J_{NH\text{-}H\alpha}$ coupling constants were measured from a one-dimensional spectrum or from the DQF-COSY spectrum.

Spectra were processed on a Silicon Graphics Indigo work station using XWINNMR (Bruker) software. The t_1 dimension was zero-filled to 2048 real data points, and 90° phase-shifted sine bell window functions were applied prior to Fourier transformation. Chemical shifts were referenced internal sodium 2,2-dimethyl-2-silapentane-5-sulponate.

Structure Calculations—Preliminary structures of [Ala^{1,15}]kalata B1 were calculated using a torsion angle simulated annealing protocol within the program DYANA (30). Final structures were calculated using simulated annealing and energy minimization protocols within CNS version 1.0 (31). The starting structures were generated using random (ϕ, ψ) dihedral angles and energy-minimized to produce structures with the correct local geometry. A set of 50 structures was generated by a torsion angle simulated-annealing protocol (32, 33). This protocol involves 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 500 steps of Powell minimization. Structures consistent with restraints were subjected to further molecular dynamics and energy minimization in a water shell, as described by Linge and Nilges (34). The refinement in explicit water involves the following steps. First, heating to 500 K via steps of 100 K, each comprising 50 steps of 0.005 ps of Cartesian dynamics. Second, 2500 steps of 0.005 ps of Cartesian dynamics at 500 K before a cooling phase where the temperature is lowered in steps of 100 K, each comprising 2500 steps of 0.005 ps of Cartesian dynamics. Finally, the structures were minimized with 2000 steps of Powell minimization. Structures were analyzed using PROMOTIF (35) and PROCHECK-NMR (36).

RESULTS

The cyclotides provide a unique opportunity to examine the folding of the cystine knot motif in the presence of a constrained backbone. In the current study we chose the prototypic cyclotide kalata B1 to probe oxidative refolding and reductive unfolding. We have previously shown that reduced cyclic kalata B1 can be very efficiently refolded into the native conformation (6) under appropriate conditions. Refolding was performed in 50% isopropyl alcohol (v/v) in 0.1 M ammonium bicarbonate, pH 8.5 and 1 mM reduced glutathione at room temperature over-

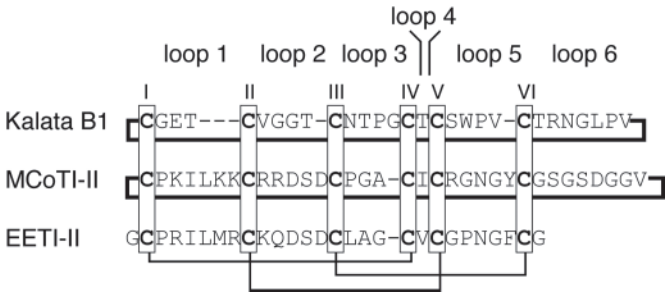


FIG. 1. **Amino acid sequences of kalata B1, MCoTI-II, and EETI-II.** The first two are cyclic cystine knot peptides while the latter is a linear cystine knot. The disulfide connectivity of all three is similar. Loops in the peptide backbone between successive cysteines are numbered for reference in the text.

night, and essentially only one product was obtained (6). In the current study the oxidative folding reaction was monitored by removing aliquots at selected time points and analyzing them with RP-HPLC. Fig. 2*a* shows the RP-HPLC traces for a time-course carried out over 10 h for kalata B1. In contrast to what is typically observed for the refolding of most peptides native oxidized kalata B1 elutes significantly later than the reduced peptide. This is most likely because of the patch of hydrophobic residues that becomes surface exposed upon correct folding (6), which provides a characteristic marker of the native folded compound.

In addition to the late eluting native peak, there are several peaks that elute with retention times close to that of the reduced cyclic peptide. LCMS confirmed that these peaks correspond to partially oxidized forms but peak overlap prevented their definitive characterization. Between the two extremes of the reduced and native forms there is an additional sharp, well-resolved peak that is detectable minutes after starting the refolding reaction (intermediate IIa in Fig. 2*a*). Its relatively late retention time suggests a conformation not dissimilar from the native form.

To determine if the presence of this well resolved folding intermediate was dependent on the refolding conditions several oxidation reactions were performed under various solution conditions (results not shown). Oxidation of reduced, cyclic kalata B1 in 0.1 M ammonium bicarbonate alone produces a small amount of the well-resolved intermediate and native peptide within 15 min. The percentage of the native product is markedly increased in the presence of reduced glutathione. Inclusion of guanidine hydrochloride in the buffer prevented folding into the native conformation and there was also no evidence for the folding intermediate. These results suggest that reduced glutathione aids refolding by allowing disulfide shuffling, while denaturing agents prevent native-like interactions that are required for folding to occur.

Isolation and Identification of the Major Folding Intermediate—Folding intermediates for model disulfide-rich peptides such as BPTI and lysozyme have been extensively studied (37, 38). It has been shown that quenching oxidative folding reactions with acid does not result in significant intramolecular disulfide rearrangement even after delays of several hours (38). In accordance with these findings, we used an acid quench method in our studies to prevent disulfide rearrangement.

Large scale oxidative refolding reactions were used to purify the folding intermediate of kalata B1. Reactions were left for ~15 min, quenched with acid and then immediately purified using semi-preparative RP-HPLC. The early eluting intermediates were not purified to homogeneity and the main focus was on the well resolved intermediate. This intermediate was purified, and an accurate molecular mass (2894 ± 0.1 Da) was determined using electrospray mass spectrometry. The molecular mass was consistent with a peptide with two disulfide bonds intact and one bond reduced.

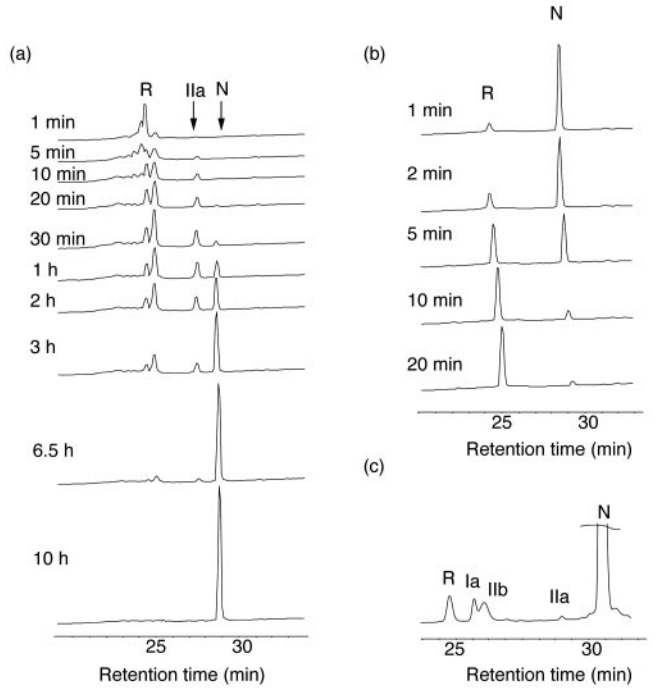


FIG. 2. **Oxidative refolding and reductive unfolding of kalata B1.** *a*, HPLC traces of the refolding of reduced, cyclic kalata B1. Oxidation was performed in 50% isopropyl alcohol, 0.1 M ammonium bicarbonate (pH 8.5) with 1 mM reduced glutathione at room temperature. Aliquots were removed at various intervals and diluted with 0.1% trifluoroacetic acid and analyzed by reverse phase HPLC. The time-points are labeled on the HPLC traces. The retention times of the (*R*) reduced, (IIa) major folding intermediate and the (*N*) native peptide are labeled. *b*, HPLC traces of the reductive unfolding of kalata B1 using 5 mM DTT in 0.1 M ammonium bicarbonate, pH 8.5. The time points are labeled on the *traces*. *c*, reductive unfolding of kalata B1 using 0.3 M TCEP in 0.1 M phosphate buffer pH 4.5. HPLC trace of a reaction quenched after 10 min.

Reductive Unfolding of Kalata B1—The reductive unfolding of kalata B1 was examined using both DTT and TCEP as reducing agents in the presence and absence of isopropyl alcohol. Isopropyl alcohol did not appear to influence the reactions. The reactions performed at various concentrations of DTT (5–100 mM) at pH 8.5 essentially show only reduced and fully oxidized species with no significant accumulation of any partially reduced species (Fig. 2*b*). However, lowering the pH to 4.5 and using TCEP (5–300 mM) as the reductant revealed the presence of minor components along the unfolding pathway (Fig. 2*c*). LCMS analysis allowed determination of the mass of the two most abundant intermediates shown in Fig. 2*c*, which correspond to a one-disulfide (Ia) and a two-disulfide (IIb) species. A very minor peak was consistently observed in the reductive unfolding at the retention time observed for intermediate IIa. Isolation of this peak from RP-HPLC and mass spectrometry confirmed that the mass was consistent with intermediate IIa.

Structural Characteristics of the Major Oxidative Folding Intermediate of Kalata B1—To further characterize the intermediate IIa the peak was isolated from RP-HPLC and examined using NMR spectroscopy. A TOCSY spectrum was recorded immediately after dissolution in 10% ^2H_2O/90% H_2O and revealed a very well dispersed spectrum consistent with a single solution conformation. A NOESY spectrum recorded later had lower signal intensity than that expected based on the signal intensity of the TOCSY spectrum. This appears to be a result of the sample aggregating over time and was confirmed

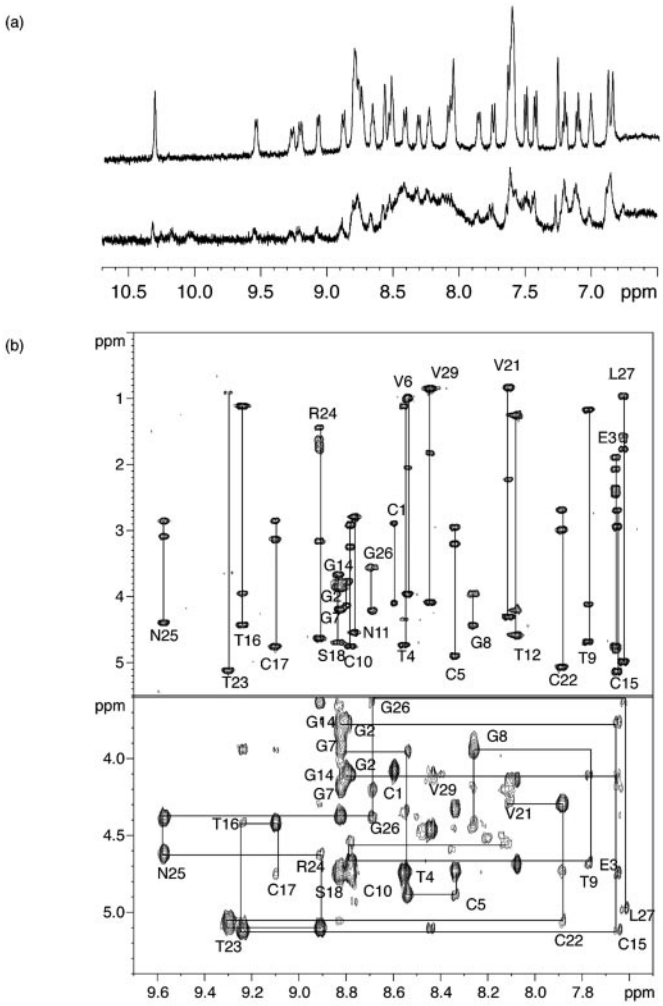


FIG. 3. **NMR spectra of the major oxidative folding intermediate of kalata B1, IIa.** *a*, one-dimensional ^{1}H NMR spectra of IIa recorded immediately after dissolution into 10% ^2H_2O/90% H_2O (*upper trace*) and after 24 h (*lower trace*). *b,* TOCSY spectrum of the folding intermediate recorded at 298 K on a 500 MHz Bruker ARX spectrometer (*upper panel*). NOESY spectrum of the folding intermediate recorded at 298 K with a mixing time of 250 ms (*lower panel*).

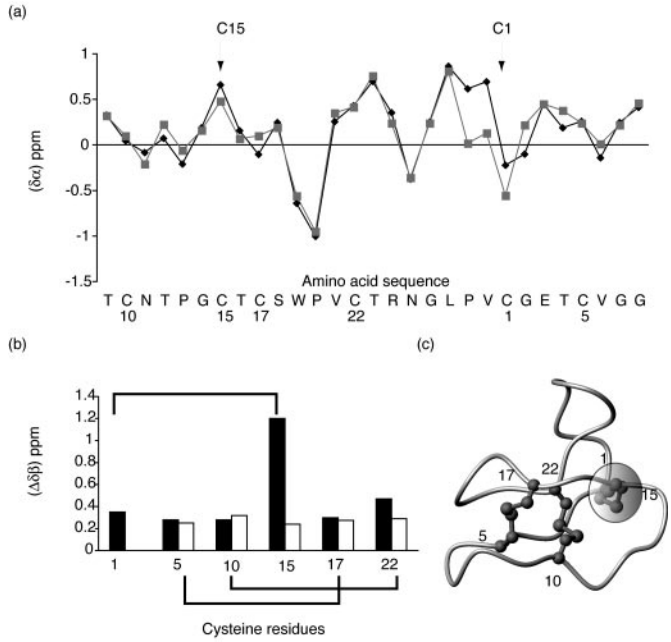


FIG. 4. **Chemical shift data for folding intermediate IIa.** *a*, comparison of αH chemical shifts of native kalata B1 and the major oxidative folding intermediate, IIa. The values plotted represent the difference between the αH chemical shift of either the native peptide or folding intermediate and the random coil chemical shifts for each residue, which were taken from Wishart *et al.* (51). The values for the native peptide are shown with *black diamonds,* and those of the intermediate are shown with *gray squares*. For clarity, the peptide sequence is circularly permuted from that shown in Fig. 1. *b*, comparison of the cysteine β-proton chemical shift differences (Δδβ) for native kalata B1 and IIa. The chemical shift difference between the βH for each cysteine residue is plotted for both peptides. The values for the native peptide are shown with *black bars* and those for the intermediate are shown with *white bars*. The disulfide connectivities are indicated on the diagram. *c,* three-dimensional structure of kalata B1 highlighting the reduced disulfide bond in IIa.

by re-recording additional one-dimensional spectra. A one-dimensional spectrum recorded immediately after dissolution of the sample and one recorded 24 h later highlights the changes that occur (Fig. 3*a*).

Despite the instability of the sample there was still sufficient information gained from the TOCSY/NOESY spectra to assign the peaks in IIa (Fig. 3*b*). A comparison of the αH chemical shifts of the intermediate with the native peptide is shown in Fig. 4*a* and reveals that IIa has a native-like structure. The only significant differences occur near residue Cys^1, suggesting that the disulfide bond involving this residue is perturbed (absent) in the intermediate.

An analysis of the βH chemical shift differences between the native kalata B1 and IIa further confirmed the identity of the reduced cysteine residues. The separation of the β-proton chemical shifts (Δδβ) for four of the six cysteine residues is very similar in the native peptide and in IIa (Fig. 4*b*). However, Cys^1 and Cys^{15} are both affected significantly. In the native peptide the Δδβ values for Cys^1 and Cys^{15} are 0.35 and 1.2 ppm, respectively, but in IIa they fall to 0.0 and 0.3 ppm, which suggests an altered conformation for these two residues. Furthermore, the direction of the change, *i.e.* toward a decreased chemical shift separation suggests a less structured local conformation. Thus, it appears that IIa contains a native-like structure with the two native disulfide bonds Cys^5–Cys^{17} and Cys^{10}–Cys^{22} intact, but with Cys^1 and Cys^{15} in the reduced state. The location of the reduced disulfide bond is highlighted in Fig. 4*c*.

Analysis of the NMR sample of IIa by RP-HPLC after 24 h in solution revealed a trace very similar to that seen in Fig. 2*a* at 10 min. This suggests that the intermediate does not directly convert to the native product with time but equilibrates via the earlier eluting forms. Refolding reactions on the isolated IIa were performed to further examine the folding properties of kalata B1. Dissolving the purified intermediate in the oxidation buffer used to fully refold kalata B1 (0.1 M ammonium bicarbonate, pH 8.5, 50% isopropyl alcohol (v/v) and 1 mM reduced glutathione) (6), resulted in a small percentage of the earlier eluting peaks. The intermediate was initially retained but a significant amount of the native peptide was formed over a 6-h period. However, when the buffer was changed to 0.5 M GnHCl in 0.05 M potassium phosphate, pH 7.8, the majority of the intermediate converted to the early eluting forms and over a 6-h period only a small percentage of native peptide was formed. Both of these reactions were performed in the presence of reduced glutathione. When the purified intermediate was placed in 0.1 M ammonium bicarbonate alone, similar profiles to the reaction containing denaturant were observed, where the majority of the intermediate converted to the early eluting forms and very little native peptide was formed over the time the reaction was monitored (6 h).

When IIa is incubated without isopropyl alcohol in the pres-

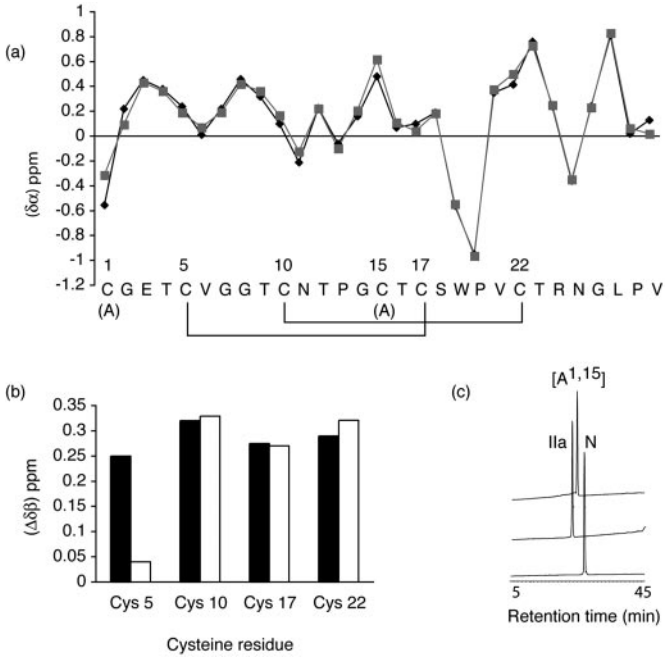


FIG. 5. **Comparison of data for IIa and [Ala^{1,15}]kalata B1.** *a*, comparison of αH chemical shifts of IIa and [Ala^{1,15}]kalata B1. The values plotted represent the difference between the αH chemical shift of either IIa or [Ala^{1,15}]kalata B1 and the random coil chemical shifts for each residue taken from Wishart *et al.* (51). The values for IIa are shown with black diamonds and those for [Ala^{1,15}]kalata B1 are shown with *gray squares*. *b*, comparison of the cysteine β-proton chemical shift differences (Δδβ) for IIa and [Ala^{1,15}]kalata B1. The chemical shift difference between the βH for each cysteine residue is plotted for both peptides. The values for the IIa are shown with *black bars* and those for [Ala^{1,15}]kalata B1 are shown with *white bars*. *c*, HPLC traces of the native kalata B1 (*N*), IIa and [Ala^{1,15}]kalata B1 run on a C_{18} RP-HPLC column with a gradient of 0–80% 0.045% trifluoroacetic acid in 90% acetonitrile over 40 min.

TABLE I
Structural statistics for the 20 lowest energy structures of [Ala1, 15]kalata B1

Experimental restraints	
Sequential NOEs	124
Medium range NOEs	52
Long range NOEs	145
Hydrogen bonds	20
Dihedral angles	19
Mean RMSDs from experimental restraints[c]	
NOE distances (Å)	0.03 ± 0.0018
Dihedral angles (°)	0.25 ± 0.09
Mean RMSDs from idealized covalent geometry	
Bonds (Å)	0.0040 ± 0.0002
Angles (°)	0.48 ± 0.02
Impropers (°)	0.39 ± 0.02
Mean energies (kcal mol^{-1})	
E_{NOE}^{c}	15.99 ± 2.0
E_{cdih}^{c}	0.09 ± 0.05
E_{bond}	5.98 ± 0.66
E_{angle}	24.78 ± 1.59
E_{improper}	4.58 ± 0.48
E_{vdw}	−76.1 ± 4.49
Atomic RMSDs (Å)[b]	
Backbone atoms	0.42 ± 0.15
Heavy atoms	0.90 ± 0.21
Backbone atoms (residues 3–28)	0.37 ± 0.13
Heavy atoms (residues 3–28)	0.89 ± 0.21
Ramachandran statistics (%)[c]	
Residues in most favored regions	81.8
Residues in additional allowed regions	18.2

[a] The values in the table are the mean ± S.D.
[b] Atomic RMSDs are the pairwise RMS difference for the family of structures.
[c] Procheck NMR was used to calculate the Ramachandran statistics.

ence of reduced glutathione, once again the majority of the intermediate is converted into the early eluting forms but then native peptide is present within an hour. In contrast, when isopropyl alcohol is included with no reduced glutathione, the intermediate persists for longer and does not appear to convert to the native conformation, suggesting that disulfide reshuffling is required to attain the native conformation.

Three-dimensional Structure of [Ala^{1,15}]Kalata B1—The instability of purified IIa over time prevented a complete structural analysis. Therefore, to further structurally characterize the intermediate an analogue was synthesized with Cys^{1} and Cys^{15} replaced with alanine residues. Assignments of [Ala^{1,15}]kalata B1 were made using established techniques (39) and the ^{1}H chemical shifts are supplied as supplementary material. Chemical shifts in the amide region are well dispersed and the large number of resolved cross peaks in the NOESY spectrum allowed determination of a well defined structure. A comparison of the secondary shifts of [Ala^{1,15}]kalata B1 with IIa shown in Fig. 5*a* highlights the similarities between the two molecules. The Δδβ shifts and retention times on RP-HPLC also show the similarities (Fig. 5, *b* and *c*). A slight change in Δδβ shifts is observed for Cys^{5}, suggesting minor differences for this residue (Fig. 5*b*).

The three-dimensional structure of [Ala^{1,15}]kalata B1 was calculated with 321 distance restraints and 19 angle restraints using a simulated annealing protocol in CNS. The disulfide connectivities (Cys^{5}–Cys^{17} and Cys^{10}–Cys^{22}) were also included as restraints in the structure calculations. The resulting family of structures had good structural and energetic statistics, as shown in Table I. A stereoview and ribbon representation of the secondary structure is shown in Fig. 6 along with a comparison with native kalata B1.

Analysis of the structures with PROMOTIF (35) identified a type I β-turn between residues 5–8, a type II β-turn between residues 12–15, a type VIa1 β-turn between residues 18–21, and a type I β-turn between residues 24–27. An anti-parallel β-sheet is recognized between residues 16–18, 21–24, and 27–29, as shown in Fig. 6*b*. Residues 16–18 and 21–24 constitute the β-hairpin that is often present in cystine knot proteins. The disulfide bond between Cys^{10} and Cys^{22} is characterized as a left-handed spiral, however the disulfide bond between residues Cys^{5} to Cys^{17} did not fall into one of the established classifications of disulfide bond geometries (35).

Given the recent suggestion that kalata B1 may not contain a cystine knot (16), but rather a Cys 1–22, 5–17, 10–15 disulfide connectivity it was of interest to examine the structures of the other theoretically possible disulfide connectivities of [Ala^{1,15}]kalata B1 (*i.e.* Cys^{5}–Cys^{10}, Cys^{17}–Cys^{22}, and Cys^{5}–Cys^{22}, Cys^{10}–Cys^{17}). Structures were thus calculated in which these two alternative disulfide connectivities were included as restraints. The energies associated with the experimental constraints, provided as supplementary material, are significantly lower (more favorable) for the Cys^{5}–Cys^{17} and Cys^{10}–Cys^{22} connectivity than the two alternative connectivities. Thus a side benefit of the synthetic mutant is that it was able to provide strong evidence in favor of the originally proposed knotted topology of kalata B1. Since the mutant contains only two disulfide bonds it must adopt one of only three possible disulfide isomers (relative to the 15 isomers possible for a three-disulfide species), and thus it provides a convenient model system for confirming the native disulfide connectivity.

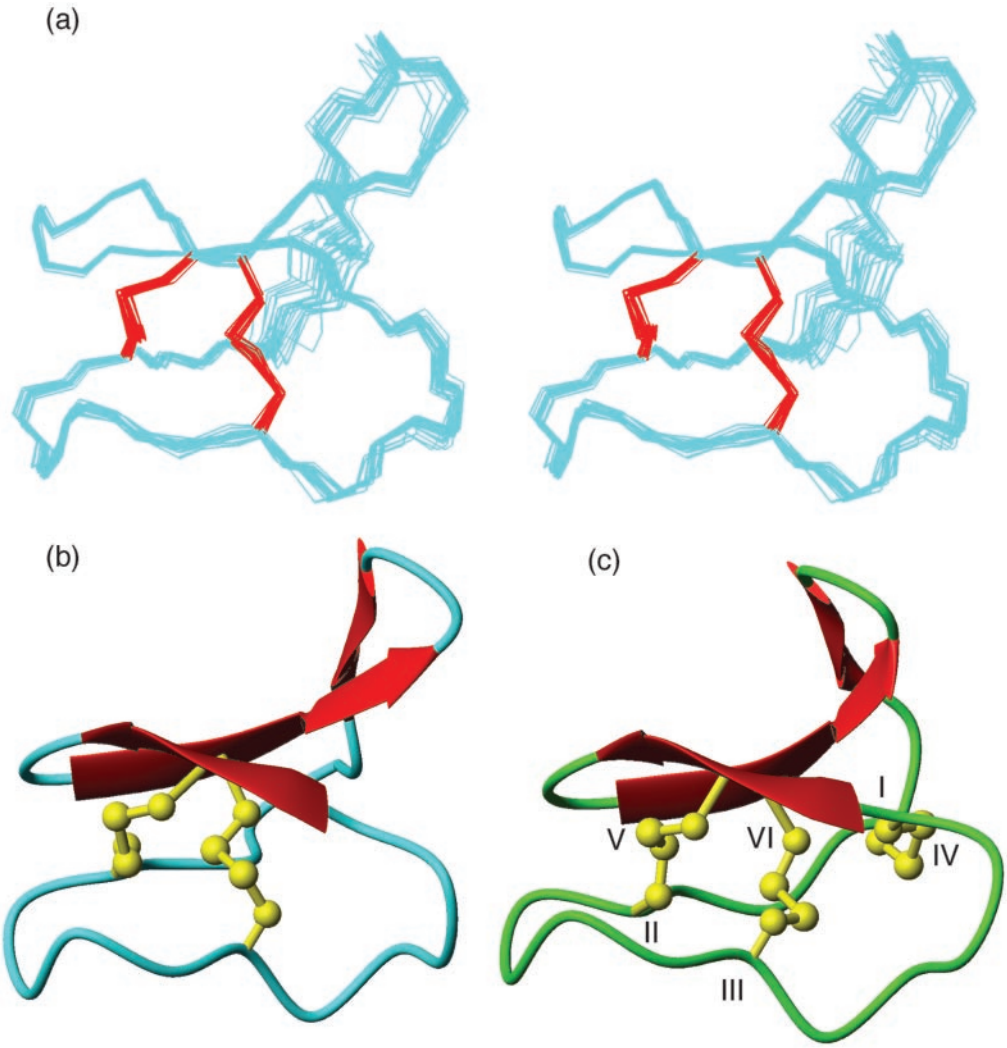


FIG. 6. **Three-dimensional structure of [Ala]kalata B1 and comparison with native kalata B1.** *a*, stereoview of a family of 20 lowest energy structures for [Ala^{1,15}]kalata B1. The structures are superimposed over the backbone atoms of residues 3–29, and the disulfide bonds are shown in *red*. *b*, ribbon representation of the secondary structure of [Ala^{1,15}]kalata B1 and *c*, native kalata B1 with the disulfide connectivities shown in *ball and stick* format. Diagrams were generated using MOLMOL (52).

DISCUSSION

The cyclotides have only recently been discovered but represent a structurally fascinating class of miniproteins with important biological functions including antimicrobial and insecticidal activity (40, 41). Little is known about the *in vivo* folding and biosynthesis of these molecules but the present study has provided an insight into the *in vitro* folding of kalata B1 and into the intricacy of the folding pathways of cystine knot proteins. We have shown that a native-like two-disulfide intermediate is present during oxidative refolding but does not accumulate to a significant extent in the reductive unfolding of kalata B1. Chemical shift analysis of the isolated intermediate, IIa, revealed a native-like three-dimensional fold and strongly suggested the presence of two-native disulfide connectivities (Cys^5–Cys^{17} and Cys^{10}–Cys^{22}). This connectivity and the native-like structure of the intermediate were confirmed by a three-dimensional structural analysis of an analogue of the intermediate, [Ala^{1,15}]kalata B1.

In addition to providing valuable new information on the folding pathway of kalata B1, [Ala^{1,15}]kalata B1 provides strong support for the originally proposed disulfide connectivity of kalata B1. Skjeldal *et al.* (16) recently suggested that kalata B1 may have a disulfide connectivity of 1–22, 5–17, and 10–15, (based on the numbering in Fig. 1) rather than the originally proposed knotted connectivity (1–15, 5–17 and 10–22) (11). Clearly, the disulfide bonds 1–22 and 10–15 are not possible in [Ala^{1,15}]kalata B1 and given the highly native-like conformation of this analogue it appears that the originally proposed connectivity is correct. Furthermore, the calculation of alternative structures of [Ala^{1,15}]kalata B1 with the three possible two-disulfide connectivities strongly supports the knotted topology. Indeed the suggestion by Skjeldal *et al.* of a different disulfide connectivity for kalata B1 was somewhat surprising given a range of supporting evidence for the originally proposed connectivity for other members of the cyclotide family, including that from directed chemical synthesis of circulin, A and circulin B and cyclopsychotride A (7, 15) and mass spectral analysis of the circulins (42). A recent high resolution structural analysis of kalata B1 has further supported the original disulfide assignment (43).

Despite the native-like overall structure of IIa it appears that direct formation of the final disulfide bond does not occur directly from this intermediate and that a structural rearrangement may be required to form the native conformation. Refolding of the purified intermediate is greatly facilitated by reduced glutathione, suggesting that disulfide bond reshuffling is required. A predominant folding intermediate that does not readily form the native peptide has also been observed for

human epidermal growth factor (EGF), a 53-residue protein containing three-disulfide bonds. In that case a stable intermediate (EGF-II) containing two of the three native disulfide bonds was isolated along the oxidative folding pathway and shown to be a major kinetic trap in the folding. Formation of the third disulfide bond did not occur directly and it appeared that the native structure was accessed instead via three-disulfide scrambled isomers (44). It appears that a similar situation applies for kalata B1, except that the native structure is accessed by an alternative two-disulfide intermediate.

Structural analysis of [Ala^{1,15}]kalata B1 provides a possible explanation for why IIa accumulates as a kinetic trap in the oxidative refolding reactions and does not readily form the native peptide. The three-dimensional structure of [Ala^{1,15}]-kalata B1 is very similar to the native peptide but a surface analysis of the structure reveals that the side chain of Ala1 is solvent exposed (Table II) and not in close proximity to Ala15 (Fig. 7). The most significant differences between native kalata B1 and [Ala^{1,15}]kalata B1, in both chemical shifts and in three-dimensional structure, occur near Cys^1/Ala^1. In particular, the chemical shift of proline 28 αH is shifted significantly upfield compared with the native peptide. This shift is seen in both [Ala^{1,15}]kalata B1 and in the intermediate IIa, suggesting that the differences observed in this region of [Ala^{1,15}]kalata B1 are also present in IIa. Consequently, Cys1 in IIa may be distant from Cys15, thereby preventing direct oxidation of the final disulfide bond. It would appear that to lead to a productive intermediate this disulfide bond must be formed before the other two are locked in place, to avoid its constituent Cys residues being placed in an unfavorable geometric arrangement.

TABLE II
Solvent accessible surface (%) for native kalata B1 and [Ala1, 15]kalata B1

The solvent accessible surfaces are calculated from MOLMOL (52) and represented as a percentage of the surface exposure for a particular residue relative to the surface of the isolated residue. The percentages quoted are for the sulfur atoms in the cysteine residues and the methyl groups for the alanine residues.

Residue	Native kalata B1	[Ala1, 15]kalata B1
1	0	97.5
5	0	0
10	0.3	1.4
15	0	0
17	13.8	14.2
22	0	0

The accumulation of intermediates in the reductive unfolding of kalata B1 is influenced by the pH of the reaction. When the reduction was performed at pH 8.5 with DTT there was no significant accumulation of intermediates (Fig. 2*b*). However at pH 4.5, using TCEP, intermediates were observed (Fig. 2*c*). These intermediates do not accumulate to a significant extent relative to the levels observed in the oxidative refolding (Fig. 2*a*). The two most significant peaks have retention times very close to the reduced species, suggesting a non-native three-dimensional structure. However, it can be assumed that the disulfide connectivities are native, based on the low level of disulfide reshuffling that occurs at low pH (45). These peaks were shown to correspond to a one-disulfide and two-disulfide species based on LCMS analysis (Fig. 2*c*). Trace amounts of IIa were also observed in the TCEP reduction. The observation of two intermediates containing two disulfide bonds (IIa and IIb) suggests that reduction of kalata B1 may occur via multiple pathways. Clearly, IIb represents a more populated pathway than that involving IIa, and we propose here that this reflects a kinetic preference for this intermediate. Parallel unfolding pathways have previously been observed for ribonuclease A, a protein containing four disulfide bonds. Two well-populated pathways were observed in the unfolding reaction in that case (46).

The accumulation of IIa during oxidative refolding is consistent with it having both a favorable low energy conformation and a high energy barrier between it and the native peptide as schematically represented in Fig. 8. The conformational rearrangement required to form the native peptide would account for the high energy barrier for conversion to fully oxidized kalata B1. The low level of IIa in the reductive unfolding is also consistent with a high-energy barrier for the reverse process (*i.e.* reduction of native kalata B1 to IIa). The latter may reflect steric hindrance to access of reducing agents to the Cys^1–Cys^{15} disulfide bond because of the cyclic backbone. The observation of IIb in the reductive unfolding suggests that it is accessed via a lower energy barrier, however the non-native structure of IIb, based on the retention time being close to the reduced rather than native conformation, suggests that it is of higher energy

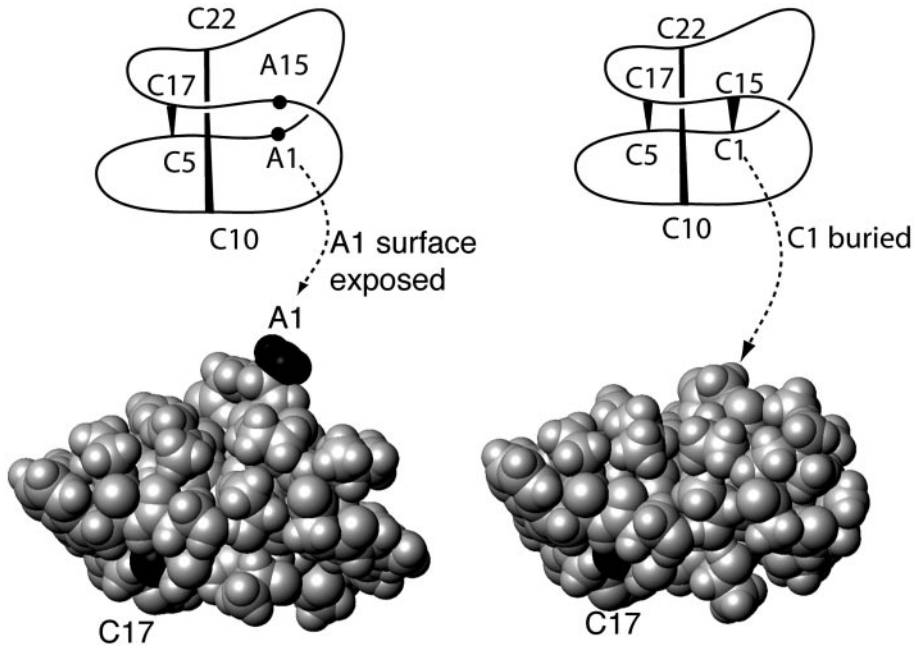


FIG. 7. **CPK representation of [Ala^{1,15}]kalata B1 (*left*) and native kalata B1 (*right*).** The alanine methyl groups and all cysteine sulfur atoms are shown in *black*. Ala1 in [Ala^{1,15}]kalata B1 is highly surface exposed, and the sidechain of Ala^{15} is buried. The only cysteine sulfur atom significantly surface exposed is from Cys^{17} in both [Ala^{1,15}]kalata B1 and native kalata B1. The exposure of this residue is consistent with the Cys^5–Cys^{17} disulfide bond being the first to be reduced. Schematic representations of the backbone structures are shown *above* the CPK diagrams to highlight the disulfide bonds present.

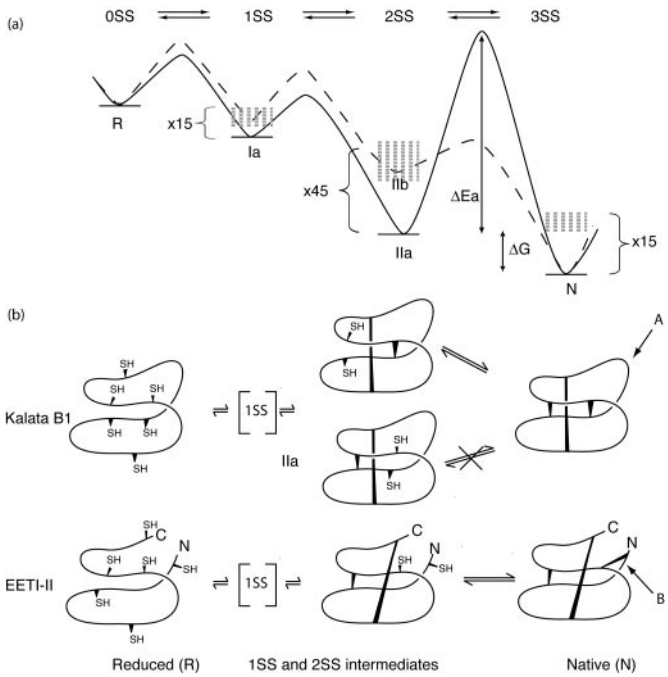


FIG. 8. **Schematic representation of the folding pathways of kalata B1 and EETI-II.** *a*, the folding intermediates identified in the current study (Ia, IIa, IIb) are represented on the diagram with arbitrary relative energies. The activation pathway to proceed from IIa to the native form is shown with a *solid line* and has a high barrier (ΔE_a) to represent the lack of conversion of IIa directly to the native peptide. However the free energy difference between IIa and the native state, ΔG, is small as IIa is native-like. The pathways involving IIb are shown with a *dashed line*. IIb is assumed to have a higher relative energy than IIa based on its non-native-like structure; however, the energy barrier for conversion to native kalata B1 is lower, accounting for the observation of IIb in the reductive unfolding studies. In a three-disulfide containing protein there are three possible one-disulfide species, 45 two-disulfide species, and 15 three-disulfide species; these species are schematically represented as *dashed lines*. *b*, *schematic illustration* of some of the proposed intermediates in the folding of kalata B1 and EETI-II. The stable two-disulfide intermediate of kalata B1 (IIa) is only produced to a significant extent along the oxidative refolding pathway in contrast to EETI-II. *Arrow A* highlights the cyclic backbone that masks the Cys(I–IV) disulfide bond in kalata, while *arrow B* shows its more exposed nature in EETI-II.

than IIa. At this stage the direct precursor to kalata B1 is uncertain, however the observation of IIb in the reductive unfolding implies that it is a likely candidate. Further, it is likely that the Cys^5–Cys^{17} bond is reduced in this intermediate but it was not possible to isolate sufficient quantities to definitively confirm this. However, a consideration of the degree of surface exposure of the Cys residues in native kalata B1 strongly supports this suggestion. The percentage surface exposure of the sulfur atoms in native kalata B1 and [$Ala^{1,15}$]kalata B1 is given in Table II. Cys^{17} is the only cysteine residue to have a significant degree of surface exposure, as is further highlighted in Fig. 7. Thus, the disulfide bond involving Cys^{17} (*i.e.* CysII-CysV in the generic cystine knot numbering scheme of Fig. 1), is likely to be the most susceptible to reduction and to be the bond broken in IIb.

Comparison of the folding intermediates characterized for kalata B1 and EETI-II reveals that they are topologically equivalent as both lack the I-IV disulfide bond and contain II-V and III-VI disulfide bonds. On this basis it might be thought that the folding pathways are similar. However, when the rate of conversion to the native peptide and the reductive unfolding of kalata B1 and EETI-II are considered it is apparent that the pathways are clearly different. These pathways are schematically represented in Fig. 8*b*. The two-disulfide intermediate of kalata B1, IIa, does not appear to be the immediate precursor of the native peptide and is only weakly detected during reductive unfolding. In contrast, the predominant intermediate in the folding of EETI-II was proposed to be a direct precursor of the native form, based on the absence of any other intermediates observed in the NMR sample, which showed the presence of both intermediate and native forms after incubation for several weeks (22). Furthermore, the EETI-II intermediate was also present during reductive unfolding (22). Thus, despite the presence of topologically equivalent intermediates in the oxidative refolding of kalata B1 and EETI-II there appears to be alternative folding pathways for these cystine knot proteins.

The alternative folding pathways observed for kalata B1 and EETI-II are likely to be related to structural differences between the molecules. The most obvious differences are the cyclic backbone and the size of the ring comprising the cystine knot. Kalata B1 contains eight residues in the knot whereas EETI-II contains twelve. The restraints of the cyclic backbone and the smaller ring size may increase the stability of intermediate IIa and disfavor its conversion to the native conformation without rearrangement of the two already formed disulfide bonds. Similarly, the susceptibility of individual disulfide bonds to reduction may be affected by the added constraints of the cyclic backbone in the cyclotides. In particular the linker region that cyclizes kalata B1 is in close proximity to the Cys^1–Cys^{15} disulfide bond (*arrow A* in Fig. 8*b*). By contrast, the corresponding bond is much more exposed in EETI-II (*arrow B* in Fig. 8*b*). In summary, the cyclic backbone and compact structure of kalata B1 appears to have a significant influence on how the disulfide bonds are stabilized.

Previous studies have highlighted the diversity of disulfide folding pathways of disulfide-rich proteins, but a general correlation between the reductive unfolding and oxidative refolding pathways of proteins has been suggested (47). Reduction of disulfide bonds in a simultaneous manner is often associated with a high degree of heterogeneity of folded intermediates, whereas sequential reduction of the disulfide bonds is associated with predominant intermediates with native-like structures. The results from the reductive unfolding of kalata B1 at high pH with DTT suggest that the disulfide bonds are reduced in a simultaneous manner under these conditions. However, the reduction at lower pH with TCEP suggests a sequential pathway with the accumulation of both a two-disulfide and one-disulfide intermediate. The sequential pathway is consistent with the general correlation mentioned above as kalata B1 has a predominant intermediate during oxidative refolding. However, the current study has emphasized the differences in folding pathways between linear and cyclic cystine knot peptides even when common intermediates are observed.

The folding pathway of the cystine knot-containing glycoprotein hormone α-subunit has recently been examined using transiently transfected cells labeled with [^{35}S]cysteine (48). Glycoprotein hormone α-subunit is a member of the growth factor cystine knot family that is topologically distinct from the cyclotides and EETI-II, which belong to the inhibitor cystine knot family (49, 50). Although the connectivities are identical, the penetrating disulfide bond is I-IV in the growth factor cystine knots instead of III-VI in the inhibitor cystine knots (50). The study by Darling *et al.* (48) indicated that folding proceeds through only one detectable intermediate that was determined to be a one-disulfide species containing the Cys^{II} to Cys^{V} disulfide bond. Through a series of experiments involving analogues containing Ala residues in place of selected cysteine residues it was established that the order of disulfide bond formation was Cys^{II}–Cys^{V}, Cys^{I}–Cys^{IV}, followed by Cys^{III}–Cys^{VI}. The presence of only one detectable disulfide intermedi-

ate containing one disulfide bond is in contrast to kalata B1 and EETI-II and suggests that parallels cannot be drawn between the folding of the two different families of cystine knots.

In conclusion, a native-like intermediate accumulates in the oxidative refolding of kalata B1 but is only a trace component in the reductive unfolding. This intermediate appears to be a kinetic trap and not the immediate precursor to the native structure. Despite the topological and structural similarities in the intermediates characterized for kalata B1 and the related cystine knot peptide EETI-II there appears to be alternative pathways in the folding of these peptides. The cyclic backbone of kalata B1 has a significant influence on the folding and in particular on the unfolding pathway. Furthermore, besides differences in the folding of cyclic and linear cystine knot peptides, it appears that inhibitor and growth factor cystine knots follow different folding pathways.

Acknowledgment—We thank Ulf Goransson for helpful comments.

REFERENCES

1. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294,** 1327–1336
2. Craik, D. J. (2001) *Toxicon* **39,** 1809–1813
3. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27,** 132–138
4. Gran, L. (1973) *Acta Pharmacol. Toxicol.* **33,** 400–408
5. 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
6. Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) *Biochemistry* **38,** 10606–10614
7. Tam, J. P., Lu, Y. A., Yang, J. L., and Chiu, K. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8913–8918
8. Gustafson, K. R., Sowder II, R. C., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina II, J. H., McMahon, J. B., Buckheit Jr., R. W., Pannell, L. K., and Boyd, M. R. (1994) *J. Am. Chem. Soc.* **116,** 9337–9338
9. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10614–10619
10. 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
11. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) *Biochemistry* **34,** 4147–4158
12. 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
13. Felizmenio-Quimio, M. E., Daly, N. L., and Craik, D. J. (2001) *J. Biol. Chem.* **276,** 22875–22882
14. Heitz, A., Hernandez, J. F., Gagnon, J., Hong, T. T., Pham, T. T., Nguyen, T. M., Le-Nguyen, D., and Chiche, L. (2001) *Biochemistry* **40,** 7973–7983
15. Tam, J. P., Lu, Y. A., Yang, J. L., and Yu, Q. (2001) *Development of Novel Antimicrobial Agents: Emerging Strategies*, pp. 215–240, Horizon Scientific Press, Wymondham, UK
16. Skjeldal, L., Gran, L., Sletten, K., and Volkman, B. F. (2002) *Arch. Biochem. Biophys.* **399,** 142–148
17. Favel, A., Mattras, H., Coletti-Previero, M. A., Zwilling, R., Robinson, E. A., and Castro, B. (1989) *Int. J. Pept. Protein Res.* **33,** 202–208
18. Wieczorek, M., Otlewski, J., Cook, J., Parks, K., Leluk, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T., and Laskowski, M., Jr. (1985) *Biochem. Biophys. Res. Commun.* **126,** 646–652
19. Heitz, A., Chiche, L., Le-Nguyen, D., and Castro, B. (1989) *Biochemistry* **28,** 2392–2398
20. Bode, W., Greyling, H. J., Huber, R., Otlewski, J., and Wilusz, T. (1989) *FEBS Lett.* **242,** 285–292
21. Daly, N. L., and Craik, D. J. (2000) *J. Biol. Chem.* **275,** 19068–19075
22. Le-Nguyen, D., Heitz, A., Chiche, L., El Hajji, M., and Castro, B. (1993) *Protein Sci.* **2,** 165–174
23. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113,** 967–974
24. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson.* **53,** 521–528
25. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson.* **65,** 355–360
26. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 479–485
27. Griesinger, C., Sørensen, O. W., and Ernst, R. R. (1987) *J. Magn. Reson.* **75,** 474–492
28. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys.* **71,** 4546–4553
29. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2,** 661–665
30. Guntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
31. Brünger, A. T., Adams, P. D., and Rice, L. M. (1997) *Structure* **5,** 325–336
32. Rice, L. M., and Brünger, A. T. (1994) *Proteins* **19,** 277–290
33. Stein, E. G., Rice, L. M., and Brünger, A. T. (1997) *J. Magn. Reson.* **124,** 154–164
34. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
35. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci.* **5,** 212–220
36. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
37. Weissman, J. S., and Kim, P. S. (1991) *Science* **253,** 1386–1393
38. van den Berg, B., Chung, E. W., Robinson, C. V., Mateo, P. L., and Dobson, C. M. (1999) *EMBO J.* **18,** 4794–4803
39. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York
40. Craik, D. J., Simonsen, S., and Daly, N. L. (2002) *Curr. Opin. Drug Discov. Devel.* **5,** 251–260
41. Craik, D. J., Anderson, M. A., Barry, D. G., Clark, R. J., Daly, N. L., Jennings, C. V., and Mulvenna, J. (2002) *Let. Pep. Sci.* **8,** 119–128
42. Derua, R., Gustafson, K. R., and Pannell, L. K. (1996) *Biochem. Biophys. Res. Commun.* **228,** 632–638
43. Rosengren, K. J., Daly, N. L., Plan, M. R., Waine, C., and Craik, D. J. (December 12, 2002) *J. Biol. Chem.* 10.1074/jbcm211147200
44. Chang, J.-Y., Li, L., and Lai, P.-H. (2001) *J. Biol. Chem.* **276,** 4845–4852
45. English, B. P., Welker, E., Narayan, M., and Scheraga, H. A. (2002) *J. Am. Chem. Soc.* **124,** 4995–4999
46. Li, Y. J., Rothwarf, D. M., and Scheraga, H. A. (1995) *Nat. Struct. Biol.* **2,** 489–494
47. Chang, J. Y., Li, L., and Bulychev, A. (2000) *J. Biol. Chem.* **275,** 8287–8289
48. Darling, R. J., Wilken, J. A., Ruddon, R. W., and Bedows, E. (2001) *Biochemistry* **40,** 577–585
49. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) *Protein Sci.* **3,** 1833–1839
50. Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon* **39,** 43–60
51. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81
52. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.* **14,** 51-55, 29–32

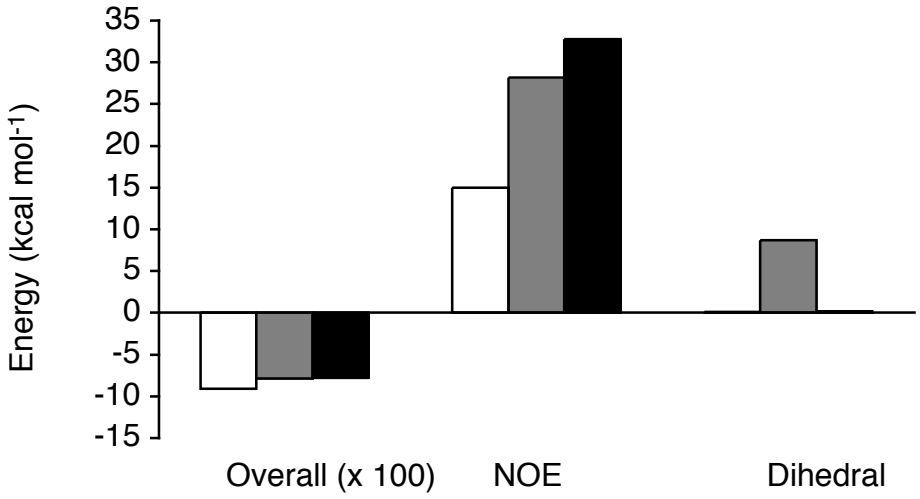
Supplemental data

Chemical shifts of [A^{1,15}]kalata B1 at pH 3.8, 290 K

2	8.578	0.000	HN	1
4	4.037	0.003	HA	1
5	1.338	0.000	QB	1
12	8.737	0.000	HN	2
14	4.063	0.000	HA1	2
15	3.737	0.003	HA2	2
19	7.729	0.000	HN	3
21	4.722	0.000	HA	3
23	2.055	0.000	HB2	3
24	1.847	0.020	HB3	3
27	2.443	0.000	HG2	3
28	2.356	0.000	HG3	3
34	8.547	0.000	HN	4
36	4.714	0.000	HA	4
38	4.294	0.016	HB	4
39	1.106	0.001	QG2	4
47	8.286	0.000	HN	5
49	4.839	0.006	HA	5
53	3.111	0.000	QB	5
57	8.604	0.000	HN	6
59	4.022	0.000	HA	6
61	2.013	0.000	HB	6
72	0.956	0.014	QQG	6
75	8.972	0.000	HN	7
77	4.163	0.000	HA1	7
78	3.865	0.000	HA2	7
82	8.324	0.000	HN	8
84	4.390	0.000	HA1	8
85	3.955	0.000	HA2	8
89	7.855	0.000	HN	9
91	4.712	0.000	HA	9
93	4.158	0.001	HB	9
94	1.173	0.000	QG2	9
102	8.793	0.000	HN	10
104	4.817	0.000	HA	10
106	2.837	0.000	HB2	10
107	3.242	0.000	HB3	10
112	8.775	0.000	HN	11
114	4.627	0.000	HA	11
118	2.767	0.002	QB	11
126	7.819	0.014	HN	12

128 4.574 0.002 HA 12
130 4.021 0.020 HB 12
131 1.209 0.003 QG2 12
141 4.342 0.000 HA 13
143 1.915 0.000 HB2 13
144 2.334 0.000 HB3 13
147 2.136 0.000 HG2 13
148 2.008 0.000 HG3 13
150 4.108 0.000 HD2 13
151 3.703 0.000 HD3 13
155 8.891 0.000 HN 14
157 4.176 0.000 HA1 14
158 3.717 0.006 HA2 14
162 7.768 0.000 HN 15
164 4.968 0.000 HA 15
165 1.269 0.022 QB 15
172 9.373 0.000 HN 16
174 4.461 0.000 HA 16
176 3.997 0.022 HB 16
177 1.124 0.000 QG2 16
185 9.064 0.000 HN 17
187 4.691 0.000 HA 17
189 2.871 0.000 HB2 17
190 3.143 0.000 HB3 17
195 8.890 0.000 HN 18
197 4.685 0.007 HA 18
199 3.849 0.000 HB2 18
200 3.677 0.012 HB3 18
205 8.195 0.013 HN 19
207 4.155 0.016 HA 19
211 3.255 0.018 QB 19
218 7.306 0.022 HD1 19
219 7.502 0.014 HE3 19
222 10.367 0.001 HE1 19
225 7.562 0.023 HZ2 19
231 3.475 0.000 HA 20
233 1.703 0.000 HB2 20
234 0.150 0.005 HB3 20
237 1.499 0.022 HG2 20
238 1.407 0.026 HG3 20
240 3.342 0.019 HD2 20
241 3.175 0.021 HD3 20
245 8.108 0.014 HN 21
247 4.326 0.020 HA 21
249 2.251 0.024 HB 21
260 0.808 0.007 QQG 21

263 7.928 0.000 HN 22
265 5.150 0.019 HA 22
267 2.916 0.000 HB2 22
268 2.582 0.000 HB3 22
273 9.384 0.000 HN 23
275 5.078 0.000 HA 23
277 3.643 0.018 HB 23
278 0.913 0.017 QG2 23
286 9.010 0.000 HN 24
288 4.630 0.000 HA 24
290 1.714 0.008 HB2 24
291 1.374 0.001 HB3 24
294 1.497 0.026 HG2 24
295 1.438 0.028 HG3 24
298 3.145 0.002 HD2 24
299 3.039 0.002 HD3 24
302 7.222 0.002 HE 24
312 9.637 0.000 HN 25
314 4.401 0.000 HA 25
316 3.075 0.000 HB2 25
317 2.839 0.000 HB3 25
326 8.728 0.000 HN 26
328 4.202 0.000 HA1 26
329 3.555 0.000 HA2 26
333 7.636 0.000 HN 27
337 1.751 0.026 HB2 27
338 1.602 0.016 HB3 27
341 1.608 0.056 HG 27
352 0.955 0.000 QQD 27
357 4.504 0.000 HA 28
359 1.800 0.000 HB2 28
360 2.195 0.000 HB3 28
363 2.139 0.000 HG2 28
364 1.950 0.000 HG3 28
366 3.915 0.000 HD2 28
367 3.666 0.017 HD3 28
371 8.615 0.001 HN 29
373 3.970 0.002 HA 29
375 1.811 0.020 HB 29
376 0.827 0.006 QG1 29
377 0.764 0.021 QG2 29



Relative energies of structures calculated with the three alternative disulfide connectivities of [Ala^{1,15}]kalata B1. The overall, NOE and dihedral energies are plotted with the white bars representing the values for the native-like connectivity (Cys 5-17, 10-22), the grey bars representing the Cys 5-22, 10-17 connectivity, and the black bars representing the Cys 5-10, 17-22 connectivity. In all cases the native connectivity (Cys5-17, Cys 10-22) has the most favorable energy terms.

Disulfide Folding Pathways of Cystine Knot Proteins: TYING THE KNOT WITHIN THE CIRCULAR BACKBONE OF THE CYCLOTIDES

Norelle L. Daly, Richard J. Clark and David J. Craik

J. Biol. Chem. 2003, 278:6314-6322.
doi: 10.1074/jbc.M210492200 originally published online December 12, 2002