Chemical Synthesis and Folding Pathways of Large Cyclic Polypeptides: Studies of the Cystine Knot Polypeptide Kalata B1†

Norelle L. Daly, Steve Love, Paul F. Alewood, and David J. Craik*

Centre for Drug Design and Development, University of Queensland, Queensland 4072, Australia

Received March 17, 1999; Revised Manuscript Received May 24, 1999

ABSTRACT: Kalata B1 is a member of a new family of polypeptides, isolated from plants, which have a cystine knot structure embedded within an amide-cyclized backbone. This family of molecules are the largest known cyclic peptides, and thus, the mechanism of synthesis and folding is of great interest. To provide information about both these phenomena, we have synthesized kalata B1 using two distinct strategies. In the first, oxidation of the cysteine residues of a linear precursor peptide to form the correct disulfide bonds results in folding of the three-dimensional structure and preorganization of the termini in close proximity for subsequent cyclization. The second approach involved cyclization prior to oxidation. In the first method, the correctly folded peptide was produced only in the presence of partially hydrophobic solvent conditions. These conditions are presumably required to stabilize the surface-exposed hydrophobic residues. However, in the synthesis involving cyclization prior to oxidation, the cyclic reduced peptide folded to a significant degree in the absence of hydrophobic solvents and even more efficiently in the presence of hydrophobic solvents. Cyclization clearly has a major effect on the folding pathway and facilitates formation of the correctly disulfide-bonded form in aqueous solution. In addition to facilitating folding to a compact stable structure, cyclization has an important effect on biological activity as assessed by hemolytic activity.

Kalata B1 is the founding member of a new family of large cyclic polypeptides which have a disulfide-stabilized core and an unusual knotted structure. The uterotonic activity of kalata B1, which is isolated from the African plant *Oldenlandia affinis*, was first described in the early 1970s (*1*, *2*) based on anecdotal reports of the use of the plant by the women of the Lulula tribe to accelerate childbirth. Kalata B1 was purified from extracts of the plant and identified as an active uterotonic agent; however, it was not until 1995 that the primary structure of the peptide was confirmed (*3*). The peptide contains 29 amino acids, three disulfide bonds, and the unusual feature, for a peptide this large, of a cyclized backbone. Over the past few years, several other cyclic peptides with similar sizes and cysteine residue spacings have been reported, and together, they apparently form a new family of macrocyclic peptides. The currently known sequences are shown in Table 1. Prior to the discovery of these molecules, the largest known cyclic peptides contained fewer than 15 amino acids.

Most of these macrocyclic peptides have been found in plants from the Rubiaceae family, of which *O. affinis* is a member. Circulin A and B were isolated from *Chassalia parvifolia* (*4*), and cyclopsychotride A (*5*) was isolated from *Psychotria longipes*. These peptides all have a cyclic backbone, and their sequences are somewhat homologous with that of kalata B1 (Table 1). Examples also occur in the Violaceae family, with the first reported member being violapeptide I isolated from *Viola arvensis* Murray (*6*). More recently, varv peptide A, which only differs from kalata B1 by a threonine to serine substitution (Table 1), was also isolated from *V. arvensis* Murray (*7*). Thus, it appears these large cyclic peptides may be quite widespread in plants.

Table 1: Sequence Alignment of Cyclic Peptides[a]

1 5 10 15 20 25	
CGETCVGGT**C**NTP..**GC**T**C**SWP**VC**TR**N**G.L**P**V	Kalata B1
CGESCVWIP**C**ISAAL**GC**S**C**KNK**VC**YR**N**G.I**P**.	Circulin A
CGESCVFIP**C**ISTLL**GC**S**C**KNK**VC**YR**N**GVI**P**.	Circulin B
CGESCVFIP**C**VTALL**GC**S**C**KSK**VC**YK**N**.SI**P**.	Cyclopsychotride A
CGETCVGGT**C**NTP..**GC**S**C**SRP**VC**TX**N**G.L**P**V	Violapeptide I
CGETCVGGT**C**NTP..**GC**S**C**SWP**VC**TR**N**G.L**P**V	Varvpeptide A

[a] Kalata B1 was isolated from the African plant *O. affinis DC* (*3*, *8*) and varv peptide A from *V. arvensis* Murray (Violaceae) (*7*). Circulins were isolated from the African tropical tree *C. parvifolia* (Rubiaceae) (*4*). Cyclopsychotride A was isolated from the South American tropical plant *P. longipes* Muell. Arg. (Rubiaceae) (*5*) and violapeptide I from *V. arvensis* (Violaceae) (*6*).

This family of peptides is of particular interest because the individual members exhibit a diverse range of biological activities. The circulins exhibit anti-HIV activity (*4*); cyclopsychotride A inhibits neurotensin binding to cell membranes (*5*), and kalata B1 exhibits uterotonic activity (*3*, *8*). In the case of the circulins and cyclopsychotride A, the biological activities were originally identified on the basis of screening assays. Violapeptide I was found in a search for saponins in the methanol fraction of plant extracts using

† This work was supported in part by a grant from the Australian Research Council. D.J.C. is an ARC Senior Fellow.

* To whom correspondence should be addressed. Fax: +61-7-3365-2487. Phone: +61-7-3365-4945. E-mail: d.craik@mailbox.uq.edu.au.

10.1021/bi990605b CCC: $18.00
Published on Web 07/21/1999

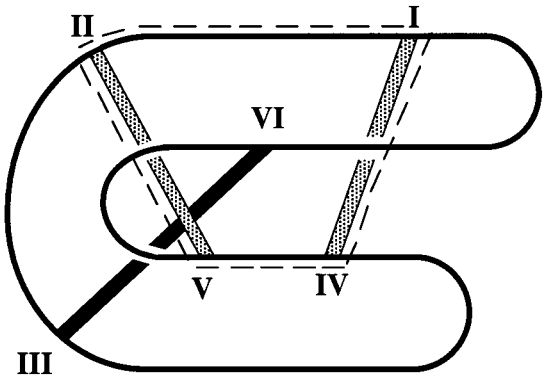

FIGURE 1: Schematic representation of the structural motif present in kalata B1 and circulin A highlighting the cyclic nature and knotted topology of the disulfide bonds. The disulfide bond which is shown in black represents the bond which threads through the other two disulfide bonds in the cystine knot motif. The starting point for numbering the residues is arbitrary and is taken from Table 1.

a hemolytic activity assay. Saponins were not found; however, violapeptide I exhibited hemolytic activity (*6*). Varv peptide A was not isolated on the basis of biological activity assays but rather in a study of isolation procedures from plants (*7*). As mentioned previously, anecdotal reports from native medicine studies led to the discovery of kalata B1 (*1*, *2*).

The three-dimensional structure of kalata B1 has been determined using two-dimensional NMR spectroscopy (*3*), and we recently also determined the three-dimensional structure of circulin A (*9*). The structures of the two polypeptides are very similar, consisting of a triple-stranded β-sheet connected by turns, and a sulfur core derived from the six closely packed cysteine residues (*3*). The presence of an exceptionally well-defined structure with a tightly packed core suggests that these polypeptides may be regarded as "mini-proteins". Analyses of the structures generated from the NMR data suggest the most likely connectivity of the three disulfide bonds is 1−15, 5−17, and 10−22. A fascinating feature of the structure is that in this arrangement of disulfides, the loop formed by two of the bonds and their connected peptide backbone is "threaded" by the third disulfide bond, leading to a motif which is termed a cystine knot (*10*−*12*). In kalata B1, the added feature of a cyclic backbone in this motif has led us to refer to this as a cyclic cystine knot structure. The topology of this framework is shown in Figure 1. The dotted line represents the core of the knot through which the darkly shaded disulfide bond is threaded.

The function of these cyclic mini-proteins within plants has not been established, but it most likely involves a defense mechanism. The means by which they are produced in vivo is also unknown. This presumably involves the production of a linear precursor protein followed by cleavage and/or cyclization. Such cyclization would presumably be assisted by an enzyme, although this has yet to be established. Because of the novelty of this class of molecules and their cyclic nature, we have explored methods of their synthesis and folding in vitro.

In this report, we describe the chemical synthesis and folding of the macrocyclic peptide kalata B1 using two strategies. In the first, oxidation of the cysteine residues of a linear precursor protein to form the correct disulfide bonds results in the folding of the three-dimensional structure and preorganization of the termini in close proximity. Cyclization is then achieved via conventional peptide bond chemistry. The second approach involves cyclization prior to oxidation. The cyclization reaction is achieved using a C-terminal thioester which reacts with an N-terminal cysteine residue to produce a native peptide bond. The cysteine residues are then oxidized to form the three disulfide bonds. Both methods are shown to yield product that is identical with the native peptide. Important differences in the folding pathways were observed between the two synthetic strategies, and in particular, a dramatic effect of cyclization on folding is apparent.

MATERIALS AND METHODS

Peptide Synthesis. Linear kalata B1 was assembled using manual solid-phase peptide synthesis with Boc chemistry on a 0.5 mmol scale. The C-terminal glycine was attached to the resin via a PAM[1] linker (Applied Biosystems, Foster City, CA), and amino acids were added to the resin using HBTU with in situ neutralization (*13*). Cleavage of the peptide from the resin was achieved using hydrogen fluoride with *p*-cresol and *p*-thiocresol as scavengers [9:0.5:0.5 (v/v) HF:cresol:thiocresol]. The reaction was allowed to proceed at −5 to 0 °C for 1 h; HF was removed under vacuum and the peptide precipitated with ether. Following cleavage, the peptide was dissolved in 50% acetonitrile containing 0.1% TFA and lyophilized. The crude, reduced peptide was purified using preparative RP-HPLC on a Vydac C_{18} column. Gradients of 0.1% aqueous TFA and 90% acetonitrile/0.09% TFA were employed with a flow rate of 8 mL/min, and the eluant was monitored at 230 nm. These conditions were used in the subsequent purification steps. Mass analysis was performed on a Sciex (Thornhill, ON) triple-quadrupole mass spectrometer using electrospray sample ionization.

A second synthesis was performed using automated solid-phase peptide synthesis on a Boc-Gly-SCH_2CH_2CO Gly-PAM resin. The three-step reaction involved in synthesizing the linker is the procedure published by Camarero et al. (*14*). The linker was attached to the Gly-PAM resin by treating the resin with bromopropanoic acid and DIC for 30 min, washing with DMF, and then treating the resin with 10% thioacetic acid and 10% DIEA in DMF for 2 × 20 min. The resin was again washed with DMF and treated with 10% β-mercaptoethanol and 10% DIEA in DMF for 2 × 20 min. After a final wash with DMF, Boc-glycine was coupled to the resin using HBTU and DIEA. The formyl group was not removed from the tryptophan with ethanolamine prior to HF cleavage. A ratio of *p*-cresol and *p*-thiocresol different

[1] Abbreviations: DIEA, *N*,*N*-diisopropylethylamine; DMF, dimethylformamide; BOP, benzotriazol-1-yloxy tris(dimethylamino)phosphoniumhexafluorophosphate; HF, hydrofluoric acid; PAM, 4-(oxymethyl)-phenylacetamidomethyl; HBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium; TFA, trifluoroacetic acid; RP-HPLC, reversed-phase high-performance liquid chromatography; DIC, diisopropyl carbodiimide; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; NMR, nuclear magnetic resonance; TOCSY, two-dimensional total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; WATERGATE, water suppression by gradient-tailored excitation; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate; CD, circular dichroism; PBS, phosphate-buffered saline; SDS, sodium dodecyl sulfate.

from that used in the previous synthesis was used during the HF cleavage (800 μL of *p*-cresol and 200 μL of *p*-thiocresol). The crude peptide was solubilized and purified as in the previous synthesis.

Disulfide Formation. Oxidation reactions were originally performed by dissolving the reduced peptide in either 0.1 M ammonium bicarbonate (pH 8.5) or 0.33 M ammonium acetate and 0.5 M guanadinium chloride (pH 7.8) at concentrations of 0.1−0.5 mg/mL (by weight). Reduced glutathione (1 mM) was added to the oxidation buffers to allow disulfide interchange, and reaction mixtures were left at either 4 °C or room temperature for 1−5 days. Subsequent oxidations were carried out in the presence of organic solvents. Reduced peptide [0.1−0.5 mg/mL (by weight)] was dissolved in 0.1 M ammonium bicarbonate (pH 8.5), 50% organic solvent (v/v), and 1 mM reduced glutathione. The organic solvents that were used were methanol, ethanol, 2-propanol, and acetonitrile. Reaction mixtures were left at room temperature for 1−2 days. All oxidations were analyzed by RP-HPLC on a Vydac C_{18} column with monitoring at 214 nm and by mass spectrometry using electrospray sample ionization. Prior to purification of the oxidized peptide using preparative RP-HPLC, the pH of the oxidation reaction mixture was lowered with TFA and then the mixture diluted with 0.1% TFA and loaded onto the preparative reversed-phase column for purification.

Cyclization. In the first synthesis, oxidized open kalata B1 was cyclized (∼1 mg/mL) using a 1−10-fold molar excess of HBTU or BOP with an excess of DIEA in DMF. Aliquots were taken out of the reaction mixture, quenched with 0.1% TFA, and analyzed by RP-HPLC. HBTU and BOP gave similar results, and HBTU was used in the preparative reactions which were quenched after 30 s and the cyclic material purified using semipreparative RP-HPLC on a Vydac C_{18} column. Approximately 2% of the correctly folded material was obtained from the crude peptide using this strategy.

In the second synthetic method, the reduced thioester peptide (∼1 mg/mL) was cyclized in 0.1 M sodium phosphate (pH 7.4), with a 6-fold excess of TCEP at room temperature. Aliquots were taken out and quenched with 0.1% TFA. All linear material was cyclized within 30 min. A significant proportion of the resulting cyclic peptide had the formyl group removed. Following oxidation of this cyclic peptide, the yield (from the crude peptide) was approximately 7%.

Comparison of Native and Synthetic Kalata. To compare the retention times of synthetic and native kalata B1, the peptides were combined at a ratio of 1:2 and run isocratically at 34% solvent B (0.09% TFA in 90% acetonitrile) on RP-HPLC using a Vydac C_{18} analytical column with monitoring at 214 nm.

CD Spectroscopy. CD spectra were recorded on a Jasco J-710 spectropolarimeter (Tokyo, Japan) over the wavelength range of 260−190 nm using a 1.0 mm path length cell, a bandwidth of 1.0 nm, a response time of 2 s, a scan rate of 20 nm min^{-1}, and averaging over two scans. Spectra were recorded on samples of native and synthetic kalata and a sample of native circulin B at 50 μM. The concentration was measured using the absorbance at 280 nm and an ϵ value calculated from the amino acid composition (*15*).

NMR Spectroscopy. The linear oxidized analogue of kalata B1 was dissolved in 10% D_2O at a concentration of approximately 2.5 mM at pH 3.5. For the comparison of synthetic and native kalata B1, the peptides were dissolved at a concentration of 0.2 mM in 10% D_2O at pH 3.25 and 3.4, respectively. ^{1}H NMR spectra were recorded at 300 and 303 K on Bruker ARX 500 and DRX 750 spectrometers. TOCSY spectra (*16*) were recorded using a MLEV-17 spin lock sequence (*17*) with a mixing time of 80 ms, and NOESY spectra (*18*) were recorded with a 200 ms mixing time. The water signal was suppressed in TOCSY and NOESY spectra using a modified WATERGATE sequence (*19*) in which two gradient pulses with a 2 ms duration and 6 G cm^{-1} strength were applied on either side of a binomial 3-9-19 pulse. Two-dimensional results were generally acquired over 6024 Hz and collected into 4096 data points with 512 t_1 increments of 32−64 scans. Spectra were processed on a Silicon Graphics Indy workstation using UXNMR (Bruker) software. Generally, both dimensions were multiplied by a sine-bell function shifted by 90°, and a polynomial baseline correction was applied to selected regions.

Hemolytic Activity Assay. The protocol used for the hemolytic assays is from Terras et al. (*20*). Human erythrocytes (blood group O) were washed several times with phosphate-buffered saline (PBS) and centrifuged at 3500 rpm until a clear supernatant was obtained. A 1% solution of the red blood cells in PBS was used in the assays which were performed by adding 20 μL of the test compound to 80 μL of the 1% suspension of red blood cells. Samples were assayed at a concentration of 1−70 μM. A 1% solution of SDS was used as a control for 100% hemolysis (*20*). The mixtures were left at room temperature for 1 h and then centrifuged at 3500 rpm for 1 min. The supernatant (20 μL) was diluted with 580 μL of milli Q water and the absorbance measured at 415 nm.

RESULTS

The family of molecules exemplified by kalata B1 represents the only known example of proteins with cyclic backbones. The mechanism of synthesis and folding of this unusual family of molecules is thus of great interest. Two distinct strategies were applied to the synthesis of kalata B1, and both provided unique information about folding. The first strategy took advantage of the compact three-dimensional structure formed upon oxidation of the cysteine residues in the linear peptide. The ability to fold the linear peptide with native disulfide bonds allowed cyclization to be achieved using conventional peptide bond chemistry because of the extremely close proximity of the termini. In the absence of such preorganization of the reactive termini, head-to-tail oligomerization would normally be favored. In the second strategy, the alternative approach was carried out in that cyclization was performed prior to oxidation. Cyclization was achieved using a C-terminal thioester which reacted with an N-terminal cysteine to form a cyclic thioester which then underwent an S,N acyl transfer to produce a native peptide bond. This strategy revealed a dramatic effect of cyclization on folding; i.e., cyclization increases the propensity for correct folding. A schematic representation of the two synthetic strategies is given in Figure 2. The linear precursors differ for the two methods; however, both approaches result in fully oxidized cyclic polypeptides.

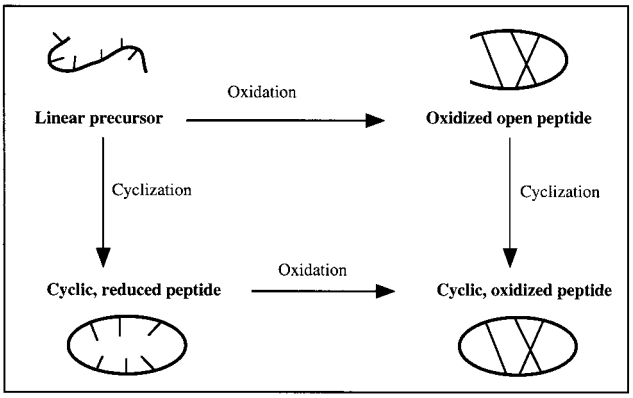

FIGURE 2: Representation of the steps involved in the two synthetic strategies used in the synthesis of kalata B1. One protocol involved oxidizing the peptide prior to cyclization, and the other involved cyclizing the peptide and then oxidizing the reduced cyclic peptide. Both protocols produced fully oxidized cyclic peptide.

Because the native material is amide-cyclized and because the gene encoding kalata B1 has not yet been identified, it is not obvious where the N- and C-termini of the native precursor molecule might be. Thus, there are in principle 29 possible linear precursors from which an in vitro synthesis could commence. We have chosen two different starting sites for the synthesis on the basis of the two different synthetic strategies as discussed below.

Strategy 1: Oxidation Followed by Cyclization. In this approach, oxidation of a linear precursor is achieved prior to cyclization, but the choice of "end groups" for the cyclization reaction is critical in the design of the linear precursor. In particular, the possibility of racemization is an important factor to consider when performing peptide bond cyclization reactions in solution. Fortuitously, kalata B1 contains a Gly-Gly sequence (residues 7 and 8, Table 1) which provided a cyclization site without the prospect of racemization. Further, this Gly-Gly sequence is involved in a β-turn and not in the major structural element of kalata B1, a triple-stranded β-sheet, and thus, we considered that initial notional disconnection of the bond between these residues was unlikely to significantly affect the folding of the peptide. A schematic representation of the protocol followed to obtain the cyclic polypeptide is given in Figure 3. This highlights the position of the (now) terminal residues and the fact that correct oxidation and folding of the molecule necessarily result in their spatial proximity. The linear oxidized peptide which has not been cyclized is referred to throughout the text as open kalata B1. Open kalata B1 is cyclized to form synthetic kalata B1.

The linear precursor (GTCNTPGCTCSWPVCTRNGLP-VCGETCVG) was assembled and cleaved from the resin with HF and subsequently purified using RP-HPLC. No oxidation of the cysteine residues was observed during the purification. The molecular mass and purity were confirmed by mass spectrometry and C_{18} analytical RP-HPLC, respectively. Various conditions were employed in the oxidation of the reduced peptide. Initial oxidations were performed with 0.1 M ammonium bicarbonate (pH 7.8, 8.5, or 9) in the presence of 1 mM reduced glutathione. Multiple product peaks were observed on RP-HPLC under these conditions, as shown in Figure 4A. Similar results were obtained at 4 °C and at room temperature (results not shown). In addition,

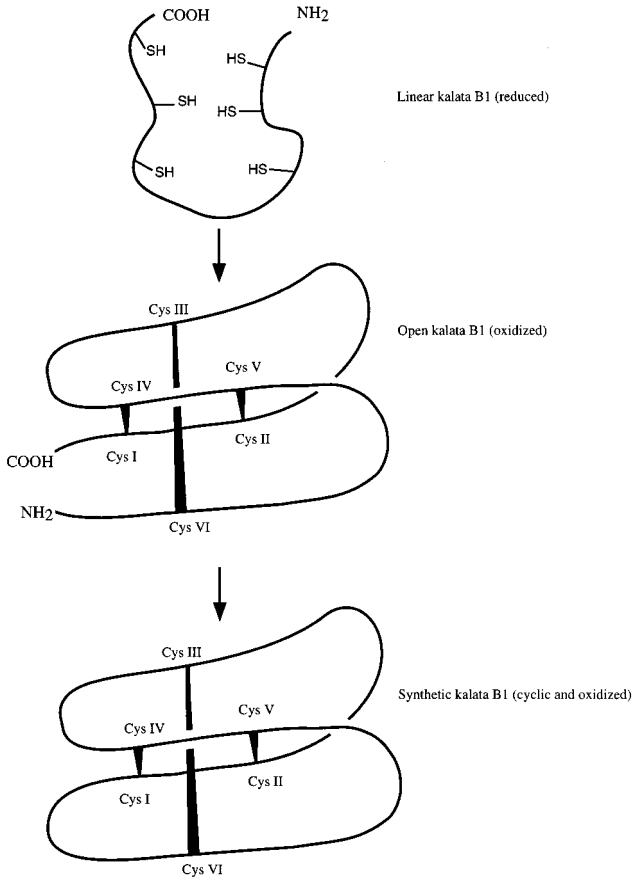

FIGURE 3: Schematic representation of the two steps involved in the synthetic strategy used in the first synthesis of kalata B1. The linear peptide was assembled and oxidized, placing the N- and C-termini in close proximity and greatly facilitating the final cyclization step.

the presence of denaturants (urea and guanidine) or varying the peptide concentration between 0.1 and 0.5 mg/mL did not significantly alter the distribution of oxidized forms, and the results were very similar to that in Figure 4A.

Given the low yield of correctly folded product in a wide range of aqueous solution conditions, it became clear that correctly folded kalata B1 would require significantly different folding conditions. Examination of the three-dimensional structure of native polypeptide showed that a large proportion of hydrophobic residues are surface-exposed. Subsequent oxidation reactions were performed in the presence of organic solvents in an attempt to stabilize the formation of these surface-exposed hydrophobic residues. Early studies were consistent with this rationale as it was reported that the solubility of kalata B1 was much higher in aqueous methanol and ethanol than in water alone (*21*). HPLC traces from oxidations [0.1 M ammonium bicarbonate (pH 8.5) and 1 mM reduced glutathione] performed in the presence of 50% (v/v) methanol, ethanol, 2-propanol, and acetonitrile are given in Figure 4. It appeared this strategy had produced the correctly folded material as a late eluting peak which was observed with a retention time much closer to that of the native peptide than to those of the other isomers present. The highest proportion of the late eluting peak was observed with 2-propanol, and there appears to be a correlation with polarity within this series of alcohols that were examined. As the polarity decreases, the proportion of

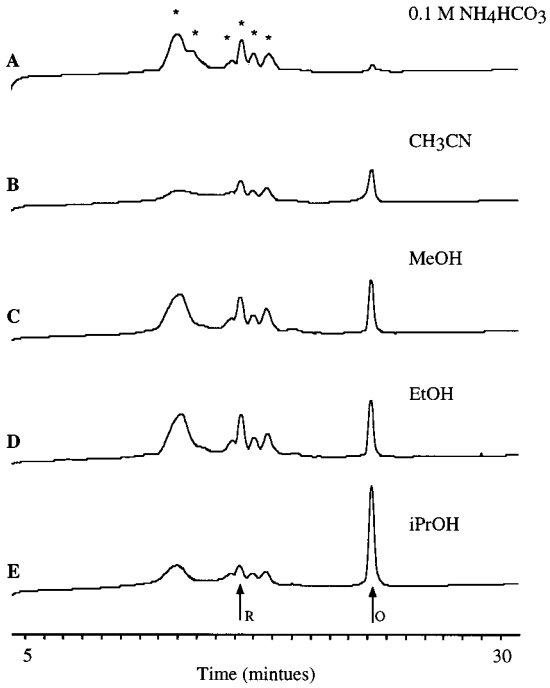

FIGURE 4: HPLC traces of oxidation reactions with the linear analogue obtained from the first synthetic strategy. Reactions were performed at room temperature in 0.1 M ammonium bicarbonate (pH 8) and 1 mM reduced glutathione (A) and in the presence of 50% acetonitrile (B), 50% methanol (C), 50% ethanol (D), and 50% 2-propanol (E). The gradient used was 20 to 50% B over the course of 30 min. The retention times of the reduced and correctly oxidized forms are denoted with arrows and marked with R and O, respectively. The earlier eluting peaks which are labeled with asterisks have molecular masses consistent with the fully oxidized peptide. The polarities of the solvents acetonitrile, methanol, ethanol, and 2-propanol are 0.65, 0.95, 0.88, and 0.82, respectively. The proportions of the correctly folded material to the misfolded material for each solvents are approximately 2% for buffer alone, 10% for acetonitrile, 9% for methanol, 13% for ethanol, and 28% for 2-propanol.

correctly folded material increases. Presumably, 2-propanol allows stabilization of the hydrophobic residues and is compatible with the hydrophilic residues. This late eluting isomer (open kalata B1) was purified using RP-HPLC and further analyzed by NMR spectroscopy prior to N-to-C cyclization in solution using HBTU. The major forms of the earlier eluting peaks from the oxidation reactions were isolated, and their masses were consistent with fully oxidized peptides; however, the retention times on RP-HPLC clearly show they are significantly more hydrophilic. Further characterization of these forms was not carried out.

In the native peptide, the residues corresponding to the termini in open kalata B1 are involved in a β-turn. The disulfide bonds formed during oxidation of the linear precursor result in a three-dimensional fold of the molecule in which the termini are very close in space as determined by NMR spectroscopy. Thus, cyclization, rather than oligomerization, occurs readily, presumably because of the effectively higher local concentration of intramolecular reactive groups. Figure 5 shows that after 30 s a significant yield of cyclized, oxidized kalata B1 is present. Allowing the reaction to proceed for a longer period of time generally resulted in reduced yield, most likely due to the formation of polymeric material through side reactions involving the unprotected glutamic acid residue. Despite this, sufficient

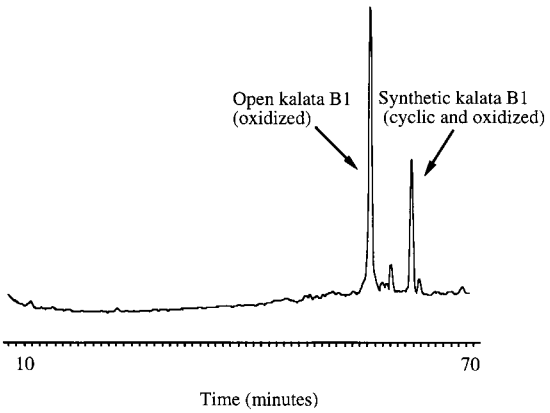

FIGURE 5: HPLC trace monitoring cyclization with HBTU of open kalata B1. The reaction was stopped after 30 s and quenched with 0.1% TFA. The gradient used was 10 to 45% B over the course of 70 min.

amounts of cyclic polypeptide were isolated for NMR analysis. Cyclization and correct folding were confirmed using mass spectrometry and HPLC coelution with the native peptide.

Strategy 2: Cyclization Followed by Oxidation. While this work was in progress, Tam and Lu (*22*, *23*) and Camarero et al. (*14*) suggested an alternative approach to the cyclization of large peptides which overcomes the potential disadvantage of side chain interactions. In the elegant procedures suggested by both groups, a thioester is formed between the N-terminal cysteine residue and the functionalized C-termini which subsequently undergoes an S,N acyl migration to form a native peptide bond. The second general strategy used in the study described here to synthesize kalata B1 is via an approach similar to those above and is given in Figure 6.

An N-terminal cysteine was required for the thioester method, and a glycine was chosen as the C-terminal residue to decrease steric hindrance. Thus, residues 14 and 15 which correspond to Gly and Cys were chosen as the C- and N-termini, respectively, for the linear peptide (i.e., CTC-SWPVCTRNGLPVCGETCVGGTCNTPG-SCH_2CH_2CO-Gly). The integrity of the thioester peptide after cleavage from the resin was confirmed using mass spectrometry and analytical HPLC. Cyclization was achieved using 0.1 M sodium phosphate (pH 7.4) with an excess of TCEP; data were taken at time points during the reaction, and the samples were quenched with 0.1% TFA and analyzed by RP-HPLC. The results are summarized in Figure 7 and show that despite the large size of the peptide the reaction occurs relatively quickly and cleanly.

Oxidation of the reduced cyclic peptide in aqueous buffer [0.1 M ammonium bicarbonate (pH 8.5) and 1 mM reduced glutathione] (Figure 8) was strikingly different from that observed for the linear peptide (Figure 4A). A very large proportion of the cyclic peptide folded into the native conformation in contrast to the linear form where negligible amounts of the correct fold were produced, indicating cyclization has a major effect on the folding pathway. The effect of organic solvent on the folding of the cyclic reduced peptide is also shown in Figure 8. Clearly, in the presence of 2-propanol, the folding efficiency is improved as the peptide folds into a single disulfide isomer. Once again, the integrity of the purified peptide was confirmed by mass spectrometry and HPLC coelution with the native peptide.

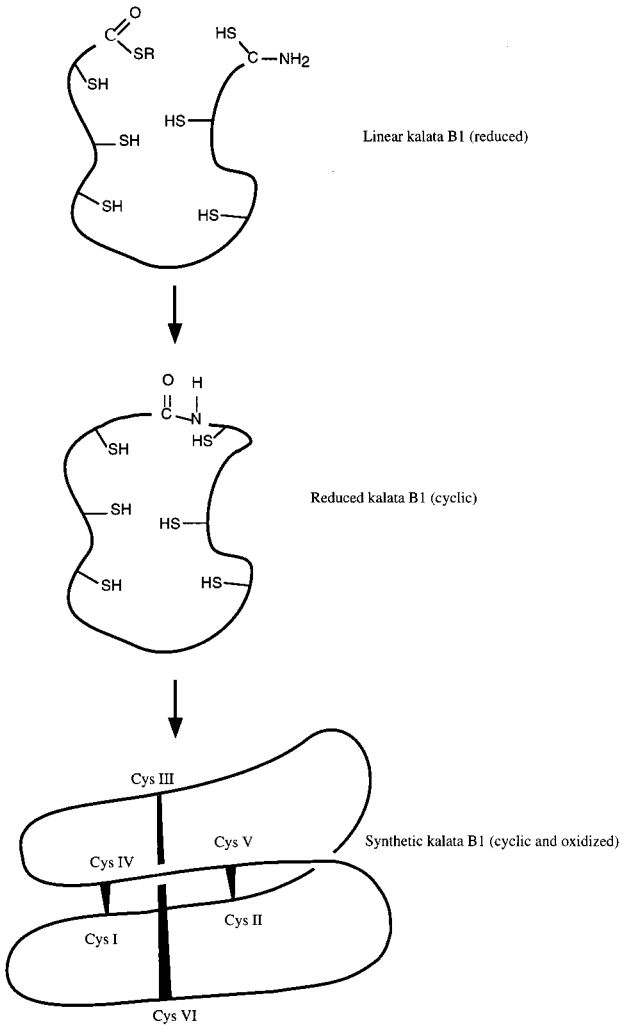

FIGURE 6: Schematic representation of the second protocol used to synthesize kalata B1. Cyclization was achieved using a C-terminal thioester which reacts with the N-terminal cysteine residue and was followed by oxidation of the cysteine residues.

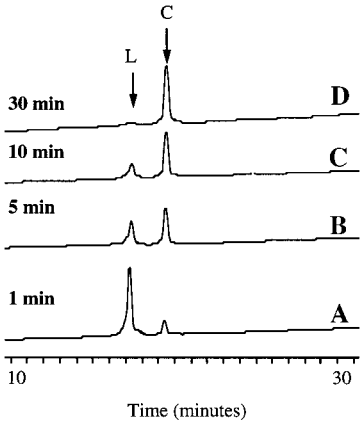

FIGURE 7: HPLC traces of the cyclization of the thioester of kalata B1. Aliquots were taken out of the reaction mixture at 1 (A), 5 (B), 10 (C), and 30 min (D), quenched with 0.1% TFA, and analyzed using a gradient of 20 to 70% B over the course of 50 min. The base-labile formyl protecting group on the tryptophan side chain, which was maintained during HF cleavage, was removed to a significant extent (~60%) during the cyclization reaction; the protected and deprotected peptides coelute under these conditions.

Coelution of Native and Synthetic Kalata. Synthetic kalata B1 derived from each of the above methods was separately mixed at a 2:1 ratio with native kalata B1 and analyzed with RP-HPLC run isocratically at 34% solvent B (0.09% TFA in 90% acetonitrile/water). In both cases, a single symmetrical peak was obtained (results not shown), providing very strong evidence that the correct isomer was obtained. The synthetic products were further examined by CD and NMR spectroscopy.

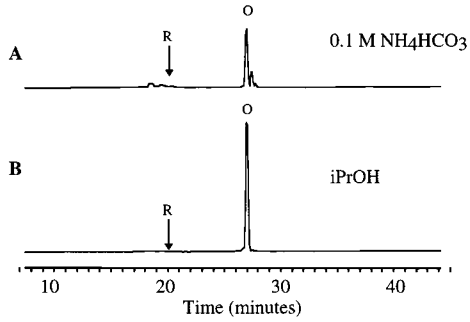

FIGURE 8: HPLC traces comparing the oxidation of cyclic reduced kalata B1 with and without organic solvent. Reactions were performed at room temperature in 0.1 M ammonium bicarbonate (pH 8) and 1 mM reduced glutathione in the absence (A) and presence (B) of 2-propanol. The gradient used was 20 to 60% B over the course of 40 min. The retention times of the reduced thioester peptide and cyclic reduced peptide are labeled with R and C, respectively, and denoted with arrows. The peptide used in these oxidation reactions was from previously oxidized and cyclic peptide which was reduced and reoxidized in the presence and absence of organic solvent.

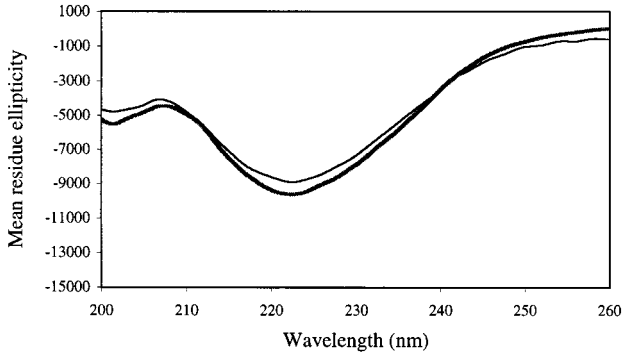

FIGURE 9: CD spectra of synthetic and native kalata B1 shown as thick and thin black lines, respectively. Spectra were recorded at a concentration of 50 μM on a Jasco J-710 spectropolarimeter.

CD Spectroscopy. A comparison of the CD spectra of native and synthetic kalata B1 is given in Figure 9. The spectra are essentially the same, within experimental error, confirming that the same elements of secondary structure are present in both peptides. The CD spectra are nonclassical in that they do not correspond to particular types of secondary structure and thus indicate the presence of irregular secondary structure. This interpretation is consistent with the three-dimensional structure established from NMR studies (*3*) which contains a series of β-turns and loops and a distorted triple-stranded β-sheet.

NMR Spectroscopy. The NMR signals of the open, oxidized form of kalata B1 are well-dispersed in the amide region, and the individual amino acid spin systems were readily identified. These were assigned to specific amino acids in the sequence using two-dimensional TOCSY and NOESY experiments (*24*). ^{1}H chemical shifts for backbone and side chain protons are provided as Supporting Information. An $\alpha H_i - \alpha H_{i+1}$ NOE is present between Trp 19 and

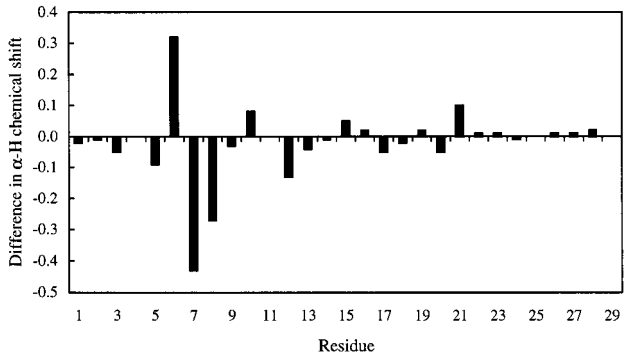

FIGURE 10: Comparison of the difference in α-proton chemical shifts of synthetic kalata B1 and the open analogue. Chemical shifts were measured relative to internal DSS on spectra recorded on a Bruker ARX 500 MHz spectrometer at 300 K. Samples were approximately 1 mM at pH 3.5.

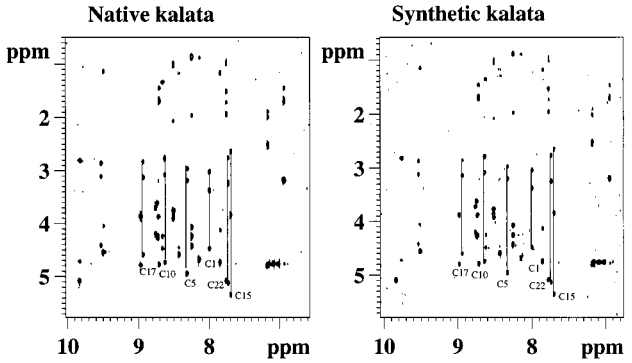

FIGURE 11: Amide region of TOCSY spectra of native and synthetic kalata B1. Spectra were recorded on a Bruker Avance 750 MHz spectrometer on samples at a concentration of 0.2 mM at pH 3.2 for the synthetic material and at pH 3.4 for the native material.

Pro 20 (numbering system used for kalata B1), indicating that the intervening peptide bond is in a cis conformation. The other two prolines have NOEs indicative of trans conformations. A comparison of the chemical shifts of the open analogue and synthetic kalata B1 is given in Figure 10. The αH chemical shifts of open kalata B1 follow very closely those of the native peptide with the exception of those of residues 5−8 which are either very close to or involved in the ligation point. Given the extremely sensitive dependence of chemical shifts on the local three-dimensional environment, the similarity between the open analogue and synthetic kalata B1 indicates that the purified isomer does contain the native fold and thus the native disulfide bonding pattern. Further, the three-dimensional structure is not greatly perturbed by the break at residues 7 and 8.

Following cyclization of the open analogue, NMR spectra were recorded for the product, cyclic kalata B1. Resonance assignments were obtained as described for the open analogue. The spectra of the native and synthetic peptides are essentially identical (Figure 11). The chemical shifts of native kalata B1 are within 0.05 ppm of those for synthetic kalata B1 with the exception of those of Asn 11, Thr 16, and Asn 25 which differ by only 0.1 ppm. The pH of the samples was slightly different (pH 3.4 for native kalata B1 and pH 3.2 for synthetic kalata B1), and this readily accounts for the small difference in the amide chemical shift of these residues. Asn 11 is particularly sensitive to changes in pH; however, the significance of this is unknown. A comparison of the αH shifts of the synthetic and native peptides, which are in general less affected by changes in pH, emphasizes the identical nature of the two peptides. As seen in the open analogue, a cis proline is present in synthetic and native kalata B1 in position 20.

Table 2: Hemolytic Activity of Cyclic Peptides and a Linear Analogue[a]

peptide	approximate HD_{50} (μM)
native kalata B1	50
synthetic kalata B1	50
circulin B	30
open kalata B1 (linear peptide)	no activity detected at 70 μM

[a] HD_{50} is the concentration at half-maximal lysis of erythrocytes.

Hemolytic Activity. As one of the earliest discovered peptides from this family, violapeptide I, was reported to have hemolytic activity (*6*), but other members have not been tested it was of interest to probe for hemolytic activity in native and synthetic kalata B1, and in the open analogue. The effects of these analogues on human erythrocytes were thus examined using the procedure described by Terras et al. (*20*). Native and synthetic kalata B1 exhibited identical activities, with 50% hemolysis occurring at 50 μM kalata B1. No hemolysis occurred with the open analogue at a concentration of 70 μM. Circulin B was also examined in the hemolytic assay to provide a comparison with another member of the cyclic cystine knot family. Approximately 50% hemolysis occurred with 30 μM circulin B, making it slightly more potent than kalata B1. These results are presented in Table 2. The fact that moderate hemolytic activity has now been seen for several members of the cyclic cystine knot family suggests that it may be a common characteristic of the family.

DISCUSSION

Insights into the folding of the cyclic cystine knot class of molecules were obtained from the two synthetic strategies applied to kalata B1. The correctly disulfide-bonded form of open kalata B1 was not a favored product of oxidation reactions carried out in the absence of organic solvent. To allow formation of the correctly folded form in significant yield, a hydrophobic environment was required, presumably because this helps to stabilize the hydrophobic residues which are exposed on folding. In general, it is considered that protein folding is accompanied by burial of hydrophobic residues, and this is consistent with the results of folding linear kalata B1 in aqueous solution. This is shown by the shorter retention times, in general, of the folded products on RP-HPLC compared to that of the reduced form, indicating the products are more hydrophilic. However, in the presence of organic solvent, the driving force of the folding of the linear peptide no longer appears to be burial of hydrophobic residues as the major form produced in organic solvent (i.e., 50% 2-propanol) elutes significantly later than the reduced material on RP-HPLC, suggesting the molecule is now much more hydrophobic. The major form was isolated and shown by NMR spectroscopy to have a fold similar to that of the native peptide which has a cystine knot core and surface-exposed hydrophobic residues. Thus, it appears likely that in this case burial of the cysteine residues drives the folding. In support of this proposal, disulfide bond formation has been found to be an important

factor in the folding of other small cysteine-rich peptides, such as the ω-conotoxin MVIIA (*25*). However, in the case of MVIIA and other ω-conotoxins, oxidation generally results in a structure which is less hydrophobic than the reduced form, in stark contrast to kalata B1. In a further contrast, the α-conotoxin MII, like kalata B1, becomes more hydrophobic upon oxidation; however, unlike kalata B1, it still folds readily in aqueous buffers (*26*). The above observations suggest that the balance between disulfide-driven folding and hydrophobic core-driven folding is finely poised and can vary depending on the size and amino acid content of disulfide-rich peptides.

The folding efficiency of cyclic reduced kalata B1 was much greater than that of the linear form. Prior cyclization of the molecule overcame the difficulties associated with folding in the absence of organic solvent, and for the cyclic precursor, a large proportion of the correctly folded material was obtained from oxidation in aqueous buffer. Folding of the cyclic material occurs even more efficiently in 2-propanol, suggesting that a range of folding conditions is tolerated for cyclic kalata B1. Cyclization appears to allow the disulfides to drive the folding at the expense of placing the hydrophobic residues on the surface and, thus, provides an interesting mechanism by which folding is controlled.

It is interesting that oxidation of circulin B in the presence of reduced and oxidized glutathione produced relatively low yields of correctly folded material (2.5%), and selective protection of the cysteine residues was required (*23*). Although the cysteine frameworks of kalata B1 and circulin B are similar, the remainder of the sequences differ significantly and this apparently leads to different folding efficiencies.

The misfolded products of kalata B1 detected by HPLC under various oxidation conditions most likely arise from non-native disulfide bond connectivities. In the original structure determination of kalata B1 (*3*), the disulfide connectivities were deduced from NOE data, although there still remained a degree of ambiguity among two or three possible S−S connectivity patterns. In the course of that initial structural study, families of structures were calculated for each of the 15 possible disulfide connectivities and many of these structural families were found to be very similar to each other. The disulfide bonds form a core which is very closely packed, and the structure calculations suggested that, in principle, different disulfide connectivities could easily be accommodated within the same global fold. However, the experimental data in the study presented here suggest that misfolded derivatives do not necessarily reach this equilibrium fold with a tightly packed disulfide core. The misfolded forms of open kalata B1 all behave very differently on HPLC with respect to the correctly folded form in that they are significantly less hydrophobic. This change in hydrophobic character suggests that the misfolded forms have structures quite different from the native form. It appears that the misconnection of S−S bonds results in the trapping of structures in misfolded intermediates which fail to reach the theoretical equilibrium states predicted by the calculations with a disulfide core and global fold not too different from those of the native state but with different disulfide connectivities.

The initial ambiguity in determining the disulfide bonding pattern for native kalata B1 arose because classical chemical methods failed due to the very tightly packed disulfide core and resistance of the oxidized peptide to enzymatic cleavage (*3*). Subsequent studies have however supported the original NMR-based deduction of S−S connectivities. Calculation of the circulin A structure also involved the use of NMR data in deducing the disulfide connectivities (*9*), and these were found to be equivalent to those in kalata B1. A chemical analysis of the disulfide connectivities of circulin A has also been carried out using acid hydrolysis and mass spectrometry and led to the same conclusion (*27*). This result has also been confirmed by synthetic studies carried out with cyclopsychotride A using cyclization via a thioester combined with a partial selective oxidation (*22*, *23*).

In addition to examining the role of disulfide bonds in the folding of kalata B1 and related molecules, we also examined the role of proline residues. Proline residues have been implicated in producing trapped non-native disulfide-bonded peptides (*28*) because cis−trans isomerization is relatively slow (*29*). For example, the folding of decorsin, a 39-residue RGD protein, was relatively difficult compared to that of hirudin HVI and charybdotoxin (*28*), with non-native species present in the oxidation of decorsin in contrast to the other two peptides. It was postulated that this difference may be ascribed to the presence of a large proportion of prolines in decorsin compared to that in hirudin and charybdotoxin. Interestingly, kalata B1 folds very efficiently in the presence of 2-propanol despite containing three prolines, one of which is in the cis conformation.

In this study, synthetic and native kalata B1 and circulin B have been shown to have relatively weak hemolytic activity. The open analogue of kalata B1 did not display this activity at a concentration of 70 μM. The cytolytic activity may be a result of the cyclic molecules forming pores in membranes. In general, peptides and proteins which form pores in membranes either have amphipathic helices which are long enough to span a membrane or oligomerize and consist mainly of β-sheet structure (*30*). The main structural feature of kalata B1 is a triple-stranded β-sheet with a solvent-exposed hydrophobic face which is formed upon disulfide bond formation as discussed above. This hydrophobic face of the molecule may promote oligomerization in the presence of membranes, although aggregation does not appear to occur in aqueous solution. Many pore-forming peptides have been shown to contain cationic residues which bind to the negatively charged lipids (*31*). Kalata B1 only contains one arginine residue and no lysines; by contrast, circulin B contains one arginine and a KNK sequence which is also present in circulin A and cyclopsychotride A. If the peptides act by forming pores, the presence of two extra positively charged residues in circulin B may explain the difference in hemolytic activity between kalata B1 and circulin B. Other possible mechanisms of action include the formation of a transient hexagonal phase in the lipids leading to membrane disruption or by forming "holes" which allow water and ion transit leading to cells bursting.

As discussed above, the open analogue of kalata B1 retains the basic three-dimensional structure of the native peptide, suggesting that either the cyclic nature of the peptide, and the subsequent loss of a positive and negative charge, or the β-turn which has been disrupted may be integral in membrane disruption. Further analogues are required to assess these possibilities.

CONCLUSIONS

Dramatic differences in folding are observed between the linear and cyclic forms of kalata B1. Entropic effects may play a role in the improved folding efficiency of the cyclic peptide as entropic losses upon folding are presumably less for the cyclic than the linear peptide. Further, the cyclic nature of the molecule may restrict the mobility of the peptide in a way which allows the native conformation to form more readily and may represent an important mechanism of directing folding.

The observation of cytolytic activity for kalata B1 and circulin B supports the proposal that the natural function of these molecules may involve a defense mechanism for the plants. The membrane disruption of these peptides may explain the anti-HIV activity which is observed for the circulin peptides (*4*).

ACKNOWLEDGMENT

We thank Dr. K. Gustafson for providing a sample of circulin B and Clement Waine for purifying native kalata B1.

SUPPORTING INFORMATION AVAILABLE

^{1}H chemical shifts of open kalata B1. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gran, L. (1970) *Medd. Nor. Farm. Selsk. 12*, 173−180.
2. Gran, L. (1972) *Medd. Nor. Farm. Selsk. 34*, 125−135.
3. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) *Biochemistry 34*, 4147−4158.
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) *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) *J. Nat. Prod. 57*, 1619−1625.
6. Schöpke, T., Hasan Agha, M. I., Kraft, R., Otto, A., and Hiller, K. (1993) *Sci. Pharm. 61*, 145−153.
7. Claeson, P., Göransson, U., Johansson, S., Luijendijk, T., and Bohlin, L. (1998) *J. Nat. Prod. 61*, 77−81.
8. Gran, L. (1973) *Lloydia 36*, 207−208.
9. 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.
10. McDonald, N. Q., and Hendrickson, W. A. (1993) *Cell 73*, 421−424.
11. Isaacs, N. W. (1995) *Curr. Opin. Struct. Biol. 5*, 391−395.
12. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) *Protein Sci. 3*, 1833−1839.
13. Schnölzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. H. (1992) *Int. J. Pept. Protein Res. 40*, 180−193.
14. Camarero, J. A., Cotton, G. J., Adeva, A., and Muir, T. W. (1998) *J. Pept. Res. 51*, 303−316.
15. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
16. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
17. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
18. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.
19. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
20. Terras, F. R. G., Schoofs, H. M. E., De Bolle, M. F. C., Van Leuven, F., Rees, S. B., Vanderleyden, J., Cammue, B. P. A., and Broekaert, W. F. (1992) *J. Biol. Chem. 267*, 15301−15309.
21. Sletten, K., and Gran, L. (1973) *Medd. Nor. Farm. Selsk. 7−8*, 69−82.
22. Tam, J. P., and Lu, Y.-A. (1997) *Tetrahedron Lett. 38*, 5599−5602.
23. Tam, J. P., and Lu, Y.-A. (1998) *Protein Sci. 7*, 1583−1592.
24. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
25. Price-Carter, M., Gray, W. R., and Goldenberg, D. P. (1996) *Biochemistry 35*, 15537−15546.
26. Hill, J. M., Oomen, C. J., Miranda, L. P., Bingham, J. P., Alewood, P. F., and Craik, D. J. (1998) *Biochemistry 37*, 15621−15630.
27. Derua, R., Gustafson, K. R., and Pannell, L. K. (1996) *Biochem. Biophys. Res. Commun. 228*, 632−638.
28. Polverino de Laureto, P., Scaramella, E., De Filippis, V., Marin, O., Gabriella Doni, M., and Fontana, A. (1998) *Protein Sci. 7*, 433−444.
29. Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953−4963.
30. Lesieur, C., Vecsey-Semjen, B., Abrami, L., Fivaz, M., and van der Goot, F. G. (1997) *Mol. Membr. Biol. 14*, 45−64.
31. Chen, D., Kini, R. M., Yuen, R., and Khoo, H. E. (1997) *Biochem. J. 325*, 685−691.

BI990605B