Knots in Rings

THE CIRCULAR KNOTTED PROTEIN MOMORDICA COCHINCHINENSIS TRYPsin INHIBITOR-II FOLDS VIA A STABLE TWO-DISULFIDE INTERMEDIATE

Received for publication, December 16, 2005, and in revised form, January 18, 2006. Published, JBC Papers in Press, January 23, 2006, DOI 10.1074/jbc.M513399200

Maša Čemazar1, Norelle L. Daly2, Sara Hägglblad, Kai Pong Lo, Ernie Yulyaningsih, and David J. Craik3

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

The aim of this work was to elucidate the oxidative folding mechanism of the macrocyclic cystine knot protein MCoTI-II. We aimed to investigate how the six-cysteine residues distributed on the circular backbone of the reduced unfolded peptide recognize their correct partner and join up to form a complex cystine-knotted topology. To answer this question, we studied the oxidative folding of the naturally occurring peptide using a range of spectroscopic methods. For both oxidative folding and reductive unfolding, the same disulfide intermediate species was prevalent and was characterized to be a native-like two-disulfide intermediate in which the Cys1-Cys18 disulfide bond was absent. Overall, the folding pathway of this head-to-tail cyclized protein was found to be similar to that of linear cystine knot proteins from the squash family of trypsin inhibitors. However, the pathway differs in an important way from that of the cyclotide kalata B1, in that the equivalent two-disulfide intermediate in that case is not a direct precursor of the native protein. The size of the embedded ring within the cystine knot motif appears to play a crucial role in the folding pathway. Larger rings contribute to the independence of disulfides and favor an on-pathway native-like intermediate that has a smaller energy barrier to cross to form the native fold. The fact that macrocyclic proteins are readily able to fold to a complex knotted structure in vitro in the absence of chaperones makes them suitable as protein engineering scaffolds that have remarkable stability.

Although it is widely accepted that the folding of proteins is governed exclusively by their amino acid sequence (1), the prediction of the three-dimensional structure of a biologically active protein from its primary sequence remains an unsolved challenge. The importance of protein folding is highlighted by the causes of debilitating diseases such as Alzheimer disease, cystic fibrosis, and Creutzfeld-Jakob disease (2, 3), which are attributed to the loss of biological functions of specific proteins due to their inability to either fold or remain correctly folded. An understanding of protein folding should provide a greater opportunity to devise novel approaches for the treatment of these and other protein misfolding diseases.

In proteins containing cysteine residues, the oxidative formation of native disulfide bonds is an integral part of the folding process. In such proteins, conformational folding is coupled with disulfide formation in the process of oxidative folding. Because disulfide bonds play key roles in the stabilization of three-dimensional structures in vivo but their formation in a cellular environment is poorly understood, the study of oxidative folding in vitro offers a valuable tool to understand the complex pathways by which native disulfide formation takes place. In particular, the study of oxidative folding is facilitated by the ability to isolate discrete (disulfide-bonded) intermediates, something that cannot be achieved in studies of conformational folding of non-disulfide-containing proteins.

Model studies of oxidative folding of cysteine-rich proteins, such as bovine pancreatic trypsin inhibitor (4, 5) and ribonuclease A (6, 7) have established a basis for understanding this process. Most oxidative folding studies reported to date have involved the isolation of stable intermediates by liquid chromatographic purification of acid-quenched folding reactions followed by structural elucidation and disulfide connectivity analysis. The oxidative folding pathways analyzed so far can be classified in terms of the number and types (i.e. native or non-native) of disulfide bonds present in the intermediate species. Interestingly, different combinations of intermediates have been found for different proteins, and it is not clear how the primary amino acid sequence or a particular three-dimensional fold relates to a particular type of oxidative folding pathway. Thus there is a need for detailed studies on specific protein classes, and in this paper, we report on a class where the final disulfide network is particularly interesting in that it forms a “knot-ted” structure.

Specifically, the current study is directed at understanding the oxidative folding process for an intriguing class of cystine knot proteins that also have a macrocyclic backbone (8, 9). In cystine knot proteins (10, 11), two disulfide bonds and their connecting backbone segments form an embedded ring that is penetrated by a third disulfide bond, as highlighted in Fig. 1. Oxidative folding studies of cystine knots, such as hirudin (12), tick anti-coagulant protein (12), Amaranthus caudatus α-amylase inhibitor (13, 14), and potato carboxypeptidase inhibitor (15) reported a highly heterogeneous mix of one-, two-, and three-disulfide intermediates, among which some contained non-native disulfide bonds. Another cystine knot, the trypsin-specific inhibitor from the squash plant Ecbalium elaterium (EETI-II) (16) forms a predominant oxidative folding intermediate that is only partially oxidized and contains only native disulfide bonds. This kind of folding pathway has also been detected for other cysteine-rich proteins that are not cystine knots, for example epi-

1 Australian Research Council postdoctoral fellow.
2 A National Health and Medical Research Council industry fellow.
3 Australian Research Council professorial fellow. To whom correspondence should be addressed: Institute for Molecular Bioscience, University of Queensland, Brisbane 4072, Queensland, Australia. Tel.: 61-7-3346-2019; Fax: 61-7-3346-2029; E-mail: d.craik@imb.uq.edu.au.

© 2006 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.
dermal growth factor (17), hen egg white lysozyme (18), and insulin-like growth factor I (19).

The class of proteins we have examined here are the cyclotides (20, 21), which are characterized by a cystine knot that is embedded within a macrocyclic backbone, defining a motif referred to as the cyclic cystine knot (CCK). As might be imagined, the combination of a macrocyclic backbone and a knotted cross-bracing motif makes the cyclotides extremely stable, and they maintain structure and biological activity after extremes of thermal, proteolytic and chemical exposure (22, 23).

As well as being exceptionally stable, the cyclotides are functionally diverse and display a wide range of activities, including antimicrobial activity, enzyme inhibition, and anti-HIV activity as well as insecticidal properties, making them attractive candidates for both drug design and agricultural applications, both in their native forms and as molecular scaffolds for the incorporation of novel bioactivities (24, 25).

Given the topological complexity of the CCK motif, it is of great interest to determine how it is formed; i.e. what is the order in which the three interlocked disulfide bonds are made? A preliminary study on the prototypic cyclotide kalata B1 showed that this uterotonic peptide folds to its native disulfide connectivity via a pathway that involves accumulation of a two-disulfide native-like intermediate species that surprisingly is not a direct precursor of the native species (26–28). Given this surprising result, it was of interest to examine oxidative folding in another subclass of CCK protein. In the current study, we examine the oxidative folding of MCoTI-II (Momordica cochinchinensis trypsin inhibitor II). As the name suggests, this molecule is a trypsin inhibitor isolated from seeds of the tropical fruit M. cochinchinensis. It was first isolated in 2000 (29), and its structure was independently determined soon after by two groups (30, 31). Fig. 1 shows the sequence and structure of MCoTI-II and highlights the disulfide bonds arranged in a CCK motif.

MCoTI-II belongs to the trypsin inhibitory subfamily of cyclotides and is quite different in sequence to the other two subfamilies, the Möbius subfamily of which kalata B1 is a member, and the bracelet subfamily (28, 29). Indeed, it is more homologous to a family of acyclic trypsin inhibitors from plants in the Cucurbitaceae family known as squash protease inhibitors. These include EETI-II (32, 33) and Cucurbita maxima trypsin inhibitor-1 (34, 35) for which structures have been reported both free in solution and bound to the proteases they inhibit.

The structure of MCoTI-II is similar to its linear counterparts apart from the presence of a linker joining what would otherwise have been the termini of the protein and forming loop 6 of the macrocyclic backbone (see Fig. 1 for a definition of the loops in cyclotides). The cystine knot is present both in the linear and the cyclic versions of these trypsin inhibitors and is a major stabilizing factor of them. The oxidative folding of EETI-II is characterized by the presence of a native-like two-disulfide intermediate species that is a direct precursor of the native species (33). In the present study, the availability of functionally related cyclic and linear counterparts of a cystine-knotted protein provides an opportunity to examine the contribution of the cyclic backbone to the oxidative folding process of cystine knots.

EXPERIMENTAL PROCEDURES

Isolation of MCoTI-II—MCoTI-II was extracted from the dormant seeds of M. cochinchinensis as described previously (29) and purified using RP-HPLC. Masses were analyzed using an electrospray ionization time-of-flight Micromass LCT mass spectrometer.

Oxidative Folding and Reductive Unfolding—Oxidative folding was performed in 0.1 M NH₄OAc, pH 8.5, 0.1 mg/ml MCoTI-II at 25 °C, containing varying concentrations of reduced glutathione from 1 to 5 mM to mimic the biochemical environment of oxidative folding in vivo. For reductive unfolding, purified native or IIa species (0.1 mg/ml) was dissolved in 0.1 M NH₄HCO₃, pH 8.0, containing 75 mM DTT at 25 °C. Aliquots were removed at various time intervals and quenched with 4% trifluoroacetic acid. Samples were analyzed using a Phenomenex RP-HPLC (250 × 10 mm, 4 μm) C18 column.

Carboxyamidomethylation—Purified IIa was dissolved in 0.5 M Tris acetate, pH 8.0, containing 2 mM Na₂EDTA and an excess of iodoacetamide. The reaction was quenched after 1 min with 0.5 mM sodium citrate, pH 3.0. The reaction mixture was purified on a semipreparative RP-HPLC (250 × 10 mm, 4 μm) C18 column.

Enzymatic Digestion and Nanospray MS-MS Sequencing—The disulfide connectivities of intermediate IIa were determined via trypsin and/or α-chymotrypsin digestion of the reduced carboxyamidomethylated analogue of IIa. After purification to >98%, this species was incubated at 37 °C in 0.1 M NH₄HCO₃, pH 8.0, with a protease:protein ratio of 1:20. The digestions were quenched after 2 h by the addition of an equal volume of 0.5% formic acid and desalted using Ziptips (Millipore). The fragments resulting from the digestion were examined first by matrix-assisted laser desorption ionization time-of-flight mass spectrometry followed by sequencing by nanospray MS-MS on a QStar mass spectrometer. A capillary voltage of 900 V was applied, and spectra were acquired between m/z 60–2000 for both time-of-flight spectra and product ion spectra. The collision energy for peptide fragmentation was varied between 10 and 50 V, depending on the size and charge of the ion. MS-MS spectra were examined and sequenced based on the presence of both b and y series of ions present (N- and C-terminal fragments).

NMR Experiments—Samples for NMR spectroscopy were prepared by dissolving the protein in 90% H₂O, 10% D₂O, and 0.1% trifluoroacetic acid to final concentrations of 0.5–0.7 mM. The trifluoroacetic acid was used to maintain a constant pH of 2 to prevent the oxidation and/or reshuffling of cysteine residues. Spectra were recorded on Bruker ARX 500 or Bruker ARX 600 NMR spectrometers at a temperature of 290 K. For TOCSY experiments, the mixing time was 80 ms and for NOESY, the mixing time was 200 ms. Slowly exchanging amide protons were identified by recording a series of one-dimensional and TOCSY spectra at 290 K over 20 h immediately after dissolution of a sample in 100% D₂O and 0.1% trifluoroacetic acid. Two-dimensional spectra were collected with 4096 data points in the f2 dimension and 512 increments in the f1 and processed using Topspin (Bruker) software. The f1 dimension was zero-filled to 2048 real data points, with the f1 and f2 dimensions multiplied by a sine-squared function prior to Fourier transformation. One-dimensional spectra were recorded with 32,768 points over a 12-ppm spectral width.

RESULTS

The main aim of this work was to elucidate the oxidative folding mechanism of the macrocyclic knotted peptide MCoTI-II. Put simply, how do the six-cysteine residues distributed on the circular backbone of the reduced unfolded peptide recognize their correct partner and join up to form a complex knotted topology? To answer this question, which is summarized in Fig. 1, we isolated a sample of the naturally occurring peptide and studied its folding/unfolding using a range of spectroscopic methods. The native peptide was extracted from the seeds of M. cochinchinensis as described previously (29) and purified over several rounds of preparative and semipreparative RP-HPLC. The identity and purity of the isolated peptide were verified by mass spectrometry.
Oxidative Folding and Reductive Unfolding of MCoTI-II—The oxidative folding pathway of MCoTI-II was examined at the level of individual disulfide species using acid quench RP-HPLC. First, oxidative folding of cyclic reduced species was studied in the presence of different concentrations of reduced glutathione to determine its role in the mechanism and rate of the folding process. Fig. 2 shows RP-HPLC traces that monitor the oxidative folding of the reduced peptide at pH 8.5 in 0.1 M ammonium acetate buffer with 2 mM reduced glutathione. The chromatograms show that the oxidative folding pathway is homogeneous, and there are only two species that are detected in significant amounts, namely the native species (N) and a folding intermediate (IIa).

The native species accumulates to >90% of the overall protein content, indicating highly efficient disulfide formation. We also examined the oxidative folding at a higher concentration of reduced glutathione (5 mM) and found that the mechanism did not change by increasing the reducing potential of the folding buffer. Increasing the reducing potential increases the rate of formation of the native species but not the overall yield.

Previously, it has been shown that MCoTI-II is prone to α-β-aspartyl isomerization, which introduces an extra CH2 group into the peptide backbone at the Asp31-Gly32 linkage in loop 6. The isomerization reaction is accelerated by alkaline pH, and the β-aspartyl isomer of MCoTI-II is separately detectable, eluting before the α-isomer on RP-HPLC. Fig. 2A shows that, during refolding of the α-aspartyl isomer, small amounts of the β-aspartyl isomer are obtained, whereas in Fig. 2B, the complementary reaction is evident; from purified reduced β-aspartyl isomer, small amounts of the oxidized α-aspartyl are obtained. The two isomers have been separated on RP-HPLC, and their structures have been studied via two-dimensional NMR (30). We found here that the β-aspartyl isomer of MCoTI-II follows the same oxidative folding mechanism as the α-aspartyl isomer and has a similar rate constant. This shows that small structural perturbations in loop 6 are not crucial for the folding pathway.

The reductive unfolding of MCoTI-II was found to be pH-dependent. MCoTI-II was highly resistant to reduction at pH 3 with 20 mM tris(2-
carboxyethyl) phosphine hydrochloride and did not produce a significant amount of any of the refolding intermediates or the reduced peptide over 8 h. However, at pH 8.0, reduction with DTT easily produced >90% of fully reduced species within 30 min (Fig. 2C). Similar to the process of oxidative folding, the reductive unfolding pathway features the native (N), fully reduced (R), and the intermediate species (IIa) detected previously and a small number of other intermediate species. It is therefore clear that MCoTI-II is not only oxidized via this intermediate species but is also reduced via this species. The intermediate species was confirmed to be the same in both the oxidative and the reductive process by RP-HPLC retention time, electrospray ionization mass spectrometry, and later by its disulfide content. It was confirmed to be a two-disulfide species by liquid chromatography-mass spectrometry. Similarly, two other minor intermediates, IIb and IIc, appearing on the reductive unfolding pathway (see Fig. 2C) were found to be two-disulfide species. A mass spectrometric confirmation of the disulfide content of the various species detected in oxidative folding and reductive unfolding is given in Table 1.

**Disulfide Bond Connectivity of IIa**—It was important to determine the disulfide connectivity of the main two-disulfide intermediate to understand the detailed mechanism of disulfide formation in MCoTI-II. A sample of IIa was isolated by partially reducing native MCoTI-II for 4 min (Fig. 3A), as it was seen in the reductive unfolding studies that, at this time point, the protein exists primarily as the IIa species. The intermediate was purified by RP-HPLC (Fig. 3B) and subjected to carboxamidomethylation with iodoacetamide. The carboxamidomethylated species, referred to as 2CM-IIa, was isolated using RP-HPLC with >99% purity (Fig. 3C). Electrospray ionization mass spectrometry of purified IIa and 2CM-IIa confirmed that IIa is a two-disulfide intermediate (Table 1). Hence, it remained to determine the identity of the missing disulfide bond. This was done by fully reducing 2CM-IIa to a species with four thiol groups (Fig. 3D) and then partially digesting it with trypsin and/or α-chymotrypsin for 2 h before sequencing the fragments using tandem mass spectrometry. It was confirmed that IIa is missing the Cys1–Cys18 bond (Table 2).

**Oxidative Folding of MCoTI-II**

![FIGURE 3. Preparation of 2CM 4SH IIa.](image)

**TABLE 2**

| Fragment sequence | Theoretical mass, MW<sub>θ</sub> | Experimental mass, MW<sub>θ</sub> | MW<sub>θ</sub> – MW<sub>θ</sub> | CM cysteines (from MS/MS) |
|------------------|----------------------------------|----------------------------------|-----------------------------|--------------------------|
| Cys<sup>14</sup>-Lys<sup>5</sup> | 2255.86                          | 2373.87                          | 118.01                      | 2 (Cys<sup>1</sup>, Cys<sup>44</sup>) |
| Asp<sup>11</sup>-Arg<sup>21</sup> | 1136.40                          | 1195.42                          | 59.02                       | 1 (Cys<sup>1</sup>) |
| Asp<sup>15</sup>-Lys<sup>9</sup> | 2572.95                          | 2690.98                          | 118.03                      | 2 (Cys<sup>1</sup>, Cys<sup>44</sup>) |
| Arg<sup>10</sup>-Lys<sup>6</sup> | 2729.05                          | 2847.08                          | 118.03                      | 2 (Cys<sup>1</sup>, Cys<sup>44</sup>) |
| Cys<sup>8</sup>-Lys<sup>6</sup> | 2987.15                          | 3106.16                          | 119.03                      | 2 (Cys<sup>1</sup>, Cys<sup>44</sup>) |
shows RP-HPLC traces of the oxidative folding process from purified IIa under the same conditions as pure reduced species was folded in Fig. 2A. There are no other species present in significant amounts on this pathway, and IIa directly interconverts to the native species. Furthermore, we found that IIa readily interconverts to the native species even in the absence of oxidative conditions and at low pH. After two-dimensional NMR studies were performed, it became clear that IIa does not have prolonged stability even at pH 2, as after 52 h in the NMR tube, 25% had interconverted to native MCoTI-II (data not shown). This finding has double significance. First, IIa under these conditions is stable enough to perform the two-dimensional NMR experiments needed for spectral assignment (Fig. 4), which take ~24 h. Second, it is clear that, even at low pH where the rate of disulfide formation is usually very low, the native species is readily formed from IIa, and no other species were present when this sample was analyzed using RP-HPLC. This finding points to the fact that IIa converts directly into the native species.

Reductive unfolding from purified IIa features some other minor intermediates that have been previously seen on the reductive unfolding pathway of native MCoTI-II (Fig. 6B). Overall, IIa has a similar mechanism of reductive unfolding as the native species.

Thermal Stability of the Cystine Knot in MCoTI-II—One-dimensional NMR spectra of native MCoTI-II and 2CM-IIa were recorded at a range of temperatures to compare the thermal stability of the peptides. Solutions of both peptides were heated from 293 to 353 K and cooled back to 293 K, and their NMR spectra were recorded at 20-degree intervals. The spectra of native MCoTI-II at 293 K before and after heating to 353 K are the same (supplemental data), showing that the molecular scaffold is extremely stable to heating and its three-dimensional structure is preserved. The equivalent spectra for the thermal cycling of 2CM-IIa show that the amide resonances at intermediate temperatures are broader and less dispersed than at 293 K, although again there was recovery back to a native fold after heating. Overall, the cyclotide sca-
fold that has a missing disulfide bond appears to be somewhat more flexible and accesses multiple conformations on heating.

Rigidity/Amide Exchange of the CCK Motif—Hydrogen/deuterium exchange experiments were performed on both native MCoTI-II and 2CM-II. Freeze-dried samples of both species were dissolved in D2O, and the disappearance of peaks in the amide region was monitored by one- and two-dimensional 1H-NMR spectroscopy. A representative set of spectra for each species is shown in Fig. 7. Remarkably, the spectrum for the native species after 13 days shows at least 10 resonances that have not exchanged, whereas the corresponding spectrum of the capped intermediate shows only traces of one or two overlapped resonances. This indicates that the hydrogen-bonding network in the two-disulfide intermediate is not as strong as in the native species. However, the two-disulfide species still adopts a native-like fold, as can be seen from the spectra after 3 h of dissolution in D2O, where the difference between the two peptides is not as striking.

To examine the amide exchange in more detail, we determined the rate constants for the exchange of all of the amide protons in both species by integrating the cross-peaks in TOCSY spectra and plotting their volumes against time. The results are mapped on surface representations of the native protein in Fig. 8. The regions highlighted in green have rate constants <0.001 min⁻¹ and are the most resistant to hydrogen exchange. It is clear that, with the exception of a few residues, these regions are in the same locations of the protein backbone in both species. The missing disulfide Cys¹-Cys¹₈ does not disrupt the overall hydrogen-bonding network of the scaffold, but both cysteine residues that are involved in making this disulfide bond exchange faster with the solvent in the two-disulfide species. Indeed there is a general loosening of the hydrogen-bonding network, as is most apparent from an increase in the number of amides that are in moderate exchange in native MCoTI-II (Fig. 8, light magenta) but are in fast exchange in the intermediate (magenta), such as Leu⁵ and Arg¹⁰. Overall, the two-disulfide intermediate appears to have a similar structure but a weakened hydrogen-bonding network relative to the native form.

DISCUSSION

In the past, the oxidative folding pathways of single-domain three-disulfide proteins have generally been characterized in terms of the number of species present on the pathway and their disulfide content (12, 14, 15, 17). In the current study, we focused on a new class of protein that is characterized by a unique structural motif involving a head-to-tail macrocyclic backbone and a cystine knot. The results show that the oxidative folding pathway of MCoTI-II has only one major intermediate present (IIα). The disulfides and three-dimensional structure of this intermediate are native-like, and it appears to be the immediate precursor of the native species. It is remarkable that the complex knotted structure of the cyclotides is able to form so efficiently. This efficiency occurs despite, or even perhaps because of, the presence of the cyclic backbone.

Our study of the reductive unfolding of MCoTI-II showed that the disulfide bonds are not stabilized in an interdependent, concerted fashion, as has been seen for some other cystine knot proteins (14, 38). Rather, reductive unfolding involves several intermediate species, among them the one observed in oxidative folding (IIα). The fact that we are able to reduce MCoTI-II partially (to form the native-like two-disulfide intermediate) means that the disulfides in this cyclotide are independently stabilized and have different degrees of solvent accessibilities. A similar observation is made when IIα is subjected to reductive unfold-
Oxidative Folding of MCoTI-II

ing; several intermediates are detected on its reductive pathway apart from the fully reduced peptide, which emphasizes that the two native disulfides in this species are also independently stabilized.

The conversion of IIa to the native peptide appears to be the rate-determining step in the oxidative folding of MCoTI-II. This intermediate is stabilized to an extent that it prevails over the formation of other intermediate species, including three-disulfide non-native intermediates that have been seen in other studies of oxidative folding (13). The interesting question is why is it stabilized kinetically with respect to the formation of native MCoTI-II? The stability of this intermediate, as judged by its relative abundance, suggests that it is of low energy, and because it folds directly to the native species, it must have a reasonably simple pathway for this conversion. However, the energy barrier of this pathway must be sufficiently high for the intermediate to be kinetically detected along with the native state. The main contributions to the energy barrier are most likely associated with the oxidation of the two thiol groups and the minor rearrangement in three-dimensional structural features to orientate the side chains bearing the thiols into immediate proximity of each other.

MCoTI-II is a member of the trypsin inhibitor subfamily of cyclotides. The oxidative folding of a cyclotide from another subfamily, kalata B1, shows some similarities to the pathway described here for MCoTI-II (33). Both oxidative folding pathways have a main intermediate that accumulates and then declines, and in both cases, it is a native-like two-disulfide species missing the I-IV disulfide of the CCK. However, the difference between these pathways is that the MCoTI-II intermediate IIa is a direct precursor of the native protein, whereas in kalata B1, the intermediate does not lead directly to native product and requires disulfide shuffling for eventual formation of the cystine knot (26–28).

Native-like intermediates that are direct precursors of the corresponding native proteins have been observed in a number of different

FIGURE 8. Mapping of amide exchange rates of MCoTI-II and its intermediate onto three-dimensional structures. Surface representation of native MCoTI-II (A) and 2CM-IIa species (B), color-coded according to the hydrogen/deuterium exchange rates extracted from two-dimensional TOCSY spectra taken after the dissolution of these species in 99.9% D2O and 0.1% trifluoroacetic acid at 290 K. The bright magenta regions have the fastest exchange rates (1000 min⁻¹ > 1000 kex > 10 min⁻¹) for the amide protons, the lighter pink have intermediate rates (10 min⁻¹ > 1000 kex > 1 min⁻¹), and the green regions have the slowest rates of exchange (1 min⁻¹ > 1000 kex), where kex is the exchange constant obtained by assuming a first order exponential decay process for hydrogen/deuterium exchange. A few residues are labeled for comparison.

FIGURE 9. Summary of the oxidative folding pathways in cyclic cystine knot and related proteins. The oxidative folding pathways of MCoTI-II, EETI-II, and kalata B1 are summarized, and a schematic representation of the size of the cystine knot rings embedded in their cores is given. MCoTI-II and EETI-II are cystine knot squash trypsin inhibitors, and they share the property that the main folding intermediate with two native disulfides is a direct precursor of the native protein. By contrast, for kalata B1, which has a cyclic cystine knot motif similar to MCoTI-II and also has a native-like intermediate, this intermediate is not a direct precursor of the native protein. The ring of the cystine knot embedded in the structures differs in the various structures, with EETI-II and MCoTI-II having a ring made up of 11 residues, whereas kalata B1 only has 8 residues in the ring. The backbone sequences of the rings are indicated, with the upper part of the ring corresponding to loop 1 and the lower part to loop 4 of the cyclotides.
cystine knots and other three- and four-disulfide proteins, including, for example, the trypsin inhibitor EETI-II, which also shares high sequence and structural similarity with MCoTI-II (16, 33). Both EETI-II and MCoTI-II have 11 residues embedded in the backbone ring that form in the core of the cystine knot, whereas kalata B1 has only eight residues because of a shorter loop 1, as illustrated in Fig. 9. The difference in ring size appears to be a major factor in the difference in folding pathways. The larger cystine knot ring for MCoTI-II means that the disulfide bonds are more separated in space and hence can adopt a more independent character in their oxidation or reduction.

In terms of reductive unfolding, kalata B1 has a so-called “all or none” mechanism at high pH, and having no intermediates on its reductive unfolding pathway, stabilizes its disulfides in a concerted, interdependent manner. This is likely why there is a large barrier between the native protein and the low energy, but non-productive, intermediate that accumulates during folding. This intermediate needs to unfold to a one-disulfide species to undergo oxidation to the correct disulfide pattern, reinforcing the notion about the interdependence of its disulfides. By contrast, MCoTI-II stabilizes its disulfides in an independent manner (as seen from its reductive unfolding, where one disulfide can be reduced although keeping the other intact). This is also why the energy barrier between II, of MCoTI-II and that of the native species is much lower; the disulfide bonds in this protein are independently stabilized and hence the native-like II species can proceed directly to the native state.

Cyclotides are, in general, very stable molecules, resistant to proteolytic, chemical, and thermal degradation, and their reduction at low pH (far from the pKs of cysteine residues) is not as effective as their reduction at high pH. MCoTI-II essentially does not produce any reduced protein when subjected to reduction with tris(2-carboxyethyl)phosphine hydrochloride at pH ~3 for over 8 h, whereas at pH 8, the reduced protein accumulates to ~99% in just 30 min. Interestingly, intermediates are observed upon reductive unfolding of kalata B1 at low pH, although at very low concentrations, and again this highlights the differences in how the disulfide bonds are stabilized for kalata B1 and MCoTI-II.

Oxidative folding has now been studied for representatives of all three subfamilies of cystine knots, namely the inhibitory cystine knots (12, 14, 15), the growth factor knots (36, 37), and the cyclic cystine knot (20), and the present study is the first report on the oxidative folding of a representative of the trypsin inhibitor subfamily of CCK. Some studies have reported that the cystine knot is either necessary for folding or has a role in directing the folding (36, 37), but that it is not required for the overall thermodynamic stability of the protein (37). Several reports have also emphasized that the cystine knot is a structural feature that governs the exceptional thermal (as opposed to thermodynamic) stability of proteins containing it (22, 37). The results of the current study, combined with previous reports, shows that one group of cystine knot proteins features concerted stabilization, interdependence of disulfide bonds, and the presence of non-native disulfide species on the folding pathway (14, 38). Other cystine knots, such as EETI-II (16) and MCoTI-II described in this paper, show independent stabilization of disulfides and the presence of native-like two-disulfide intermediate species.

The key features of the folding pathways of MCoTI-II, EETI-II, and kalata B1 are compared in Fig. 9. Apart from the cystine knot embedded in these three peptides, they all have a topologically equivalent two-disulfide intermediate with two (Cys^II–Cys^V and Cys^III–Cys^VI) native disulfide bonds on their oxidative folding pathways. MCoTI-II and EETI-II traverse this same intermediate upon reductive unfolding. Clearly, the major differences between these peptides are the size of the cystine knot and the presence or absence of the cyclic backbone. Given that the folding pathways and indeed the sequences of MCoTI-II and EETI-II are more similar to each other than to kalata B1, it appears that it is not the cyclic backbone that is responsible for the differences in the folding pathways. Rather, the size of the embedded ring of cystine knot appears to be crucial for determining the way the disulfide bonds are stabilized.

In summary, cyclotides from different subfamilies, MCoTI-II and kalata B1, possess a common feature of having a native-like folding intermediate. However, despite the similarities in structures of the two cyclotides, the disulfide bonds are stabilized by different mechanisms, and the size of the cystine knot appears to be important for the productivity and efficacy of the folding process. Larger rings of the cystine knot directly contribute to the independence of disulfides in the native species and favor an on-pathway-native intermediate species that has a smaller energy barrier to cross to form the native state. Hence, a larger knot in a ring means a more direct and less bumpy road to the low energy native state.

Acknowledgments—We thank Michelle Colgrave for assistance with mass spectrometry and Kathryn Greenwood for the extraction of crude protein from M. cochinchinesis seeds.

REFERENCES
1. Anfinsen, C. B. (1973) Science 181, 223–230
2. Dobson, C. M. (2004) Semin. Cell Dev. Biol. 15, 3–16
3. Stefani, M. (2004) Biochim. Biophys. Acta 1739, 5–25
4. Creighton, T. (1974) J. Mol. Biol. 87, 579–602
5. Weissman, J. S., and Kim, P. S. (1991) Science 253, 1386–1393
6. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) Biochemistry 39, 4207–4216
7. Wedemeyer, W. J., Xu, X., Welker, E., and Scheraga, H. A. (2002) Biochemistry 41, 4843–1491
8. Mulvonna, J. P., Wang, C., and Craik, D. J. (2006) Nucleic Acids Res. 34, D192–D194
9. Trabi, M., and Craik, D. J. (2002) Trends Biochem. Sci. 27, 132–138
10. Pallaghy, P. K., Nielsen, K. J., Craik, D. J., and Norton, R. S. (1994) Protein Sci. 3, 1833–1839
11. Craik, D. J., Daly, N. L., and Waine, C. (2001) Toxicol 39, 43–60
12. Chang, J. Y., Li, L., and Bulychev, A. (2000) J. Biol. Chem. 275, 8287–8289
13. Cemazar, M., Zahariev, S., Lopez, J. I., Carugo, O., Jones, J. A., Hore, P. J., and Pongor, S. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 5754–5759
14. Cemazar, M., Zahariev, S., Pongor, S., and Hore, P. J. (2004) J. Biol. Chem. 279, 16697–16705
15. Chang, J. Y., Li, L., Canals, F., and Aviles, F. X. (2000) J. Biol. Chem. 275, 14205–14211
16. Le Nguyen, D., Heitz, A., Chiche, L., el Hajji, M., and Castro, B. (1993) Protein Sci. 2, 165–174
17. Chang, J. Y., Li, L., and Lai, P. H. (2001) J. Biol. Chem. 276, 4845–4852
18. van den Berg, B., Chung, E. W., Robinson, C. V., and Dobson, C. M. (1999) J. Mol. Biol. 290, 781–796
19. Miller, J. A., Narhi, L. O., Hua, Q. X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) Biochemistry 32, 5203–5213
20. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) J. Mol. Biol. 294, 1327–1336
21. Craik, D. J., Daly, N. L., Mulvonna, J., Pongor, M. R., and Trabi, M. (2004) Curr. Protein Pept. Sci. 5, 297–315
22. Colgrave, M. L., and Craik, D. J. (2004) Biochemistry 43, 5965–5975
23. Grant, L., Sandberg, F., and Sletten, K. (2000) J. Ethnopharmacol. 6314–6322
24. Clark, R. J., Daly, N. L., and Craik, D. J. (2006) Biochem. J. 34, 85–93
25. Craik, D., Cemazar, M., and Daly, N. L. (2006) Curr. Opin. Drug Discovery Dev., in press
26. Craik, D. J., and Daly, N. L. (2005) Protein Pept. Lett. 12, 147–152
27. Daly, N. L., Clark, R. J., and Craik, D. J. (2003) J. Biol. Chem. 278, 6314–6322
28. Daly, N. L., Clark, R. J., Goransson, U., and Craik, D. J. (2003) Lett. Pept. Sci. 10, 523–531
29. Hernandez, J. F., Gagnon, J., Chiche, L., Nguyen, T. M., Andreiu, J. P., Heitz, A., Trinh Hong, T., Pharm, T. T., and Le Nguyen, D. (2000) Biochemistry 39, 5722–5730
30. Felizmenio-Quimio, M. E., Daly, N. L., and Craik, D. J. (2001) J. Biol. Chem. 276, 22875–22882
31. Heitz, A., Hernandez, J. F., Gagnon, J., Hong, T. T., Pharm, T. T., Nguyen, T. M.,
Oxidative Folding of MCoTI-II

Le-Nguyen, D., and Chiche, L. (2001) *Biochemistry* **40**, 7973–7983
32. Kratzner, R., Debreczeni, J. E., Pape, T., Schneider, T. R., Wentzel, A., Kolmar, H., Sheldrick, G. M., and Uson, I. (2005) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 1255–1262
33. Heitz, A., Chiche, L., Le-Nguyen, D., and Castro, B. (1995) *Eur. J. Biochem.* **233**, 837–846
34. Thaimattam, R., Tykarska, E., Bierzynski, A., Sheldrick, G. M., and Jaskolski, M. (2002) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 1448–1461
35. Helland, R., Berglund, G. I., Otlewski, J., Apostoluk, W., Andersen, O. A., Willassen, N. P., and Smalas, A. O. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 139–148
36. Darling, R. J., Ruddon, R. W., Perini, F., and Bedows, E. (2000) *J. Biol. Chem.* **275**, 15413–15421
37. Muller, Y. A., Heiring, C., Misselwitz, R., Welfle, K., and Welfle, H. (2002) *J. Biol. Chem.* **277**, 43410–43416
38. Chang, J. Y., Canals, F., Schindler, P., Querol, E., and Aviles, F. X. (1994) *J. Biol. Chem.* **269**, 22087–22094