Isolation and Characterization of Peptides from *Momordica cochinchinensis* Seeds

Lai Y. Chan, Conan K. L. Wang, Jodie M. Major, Kathryn P. Greenwood, Richard J. Lewis, David J. Craik, and Norelle L. Daly*

The University of Queensland, Institute for Molecular Bioscience, Brisbane QLD 4072, Australia

The plant *Momordica cochinchinensis* has traditionally been used in Chinese medicine to treat a variety of illnesses. A range of bioactive molecules have been isolated from this plant, including peptides, which are the focus of this study. Here we report the isolation and characterization of two novel peptides, MCoCC-1 and MCoCC-2, containing 33 and 32 amino acids, respectively, which are toxic against three cancer cell lines. The two peptides are highly homologous to one another, but show no sequence similarity to known peptides. Elucidation of the three-dimensional structure of MCoCC-1 suggests the presence of a cystine knot motif, also found in a family of trypsin inhibitor peptides from this plant. However, unlike its structural counterparts, MCoCC-1 does not inhibit trypsin. MCoCC-1 has a well-defined structure, characterized mainly by a triple-stranded antiparallel β-sheet, but unlike the majority of cystine knot proteins MCoCC-1 contains a disordered loop presumably as a result of flexibility in a localized region of the molecule. Of the cell lines tested, MCoCC-1 is the most toxic against a human melanoma cell line (MM96L) and is nonhemolytic to human erythrocytes. The role of these peptides within the plant remains to be determined.

Plants have traditionally been rich sources of molecules having applications as drug leads. Although previous research on plant natural products has focused on small molecules, recent attention has turned toward peptides from plants. Peptides hold great promise as drug leads, but their susceptibility to proteolytic breakdown and typical lack of oral bioavailability are major potential limitations. A solution to this problem may be found in plants that produce exceptionally stable peptides. This study focuses on *Momordica cochinchinensis*, a plant that belongs to the Cucurbitaceae family. The flesh of the fruit is widely used in Vietnam to color rice,[1] and the seeds of the fruit have been widely used as a Chinese medicine, also known as Mubiezhi, to treat conditions such as hemorrhoids and hemangiomas.[2] Previous studies have also found cytotoxic compounds in it and other plants from the same genus.[3–5] For example, analysis of an aqueous extract of *M. cochinchinensis* revealed a 34 kDa protein, which suppressed proliferation of tumor cells.[1] In addition, *M. charantia* contains triterpene glycosides, which also show anticancer activity.[6] Given the anticancer components already found in the *Momordica* genus, we were interested in the discovery of novel peptidic molecules with potential anticancer applications.

Several trypsin inhibitors have been isolated from *M. cochinchinensis,* including MCoTI-I, MCoTI-II, and MCoTI-III.[7–9] Unusually, MCoTI-I and MCoTI-II are head-to-tail cyclized peptides, whereas MCoTI-III is an acyclic peptide. The structure of MCoTI-II has been determined[10,11] and contains a cystine knot motif,[12,13] which comprises an embedded ring formed by two disulfide bonds and their connecting backbone sequences and a third disulfide bond that threads the ring. This motif has previously been shown to be responsible for structural stability in plant-derived cyclic peptides known as cyclotides.[14,15] In both MCoTI-II and the cyclotides the backbone segments between successive cysteine residues are referred to as loops.[14] Two of them (loops 1 and 4) form the embedded ring of the cystine knot, and the others (loops 2, 3, 5, and 6) protrude from the cystine knot core. Figure 1 shows the sequences of three *M. cochinchinensis* trypsin inhibitors[7] together with images of the fruit and seeds of the plant. On the basis of the limited studies on this plant and the interesting peptides already found, the aim of the current study was to isolate novel peptides from *M. cochinchinensis* seed extracts and to test their cytotoxicity
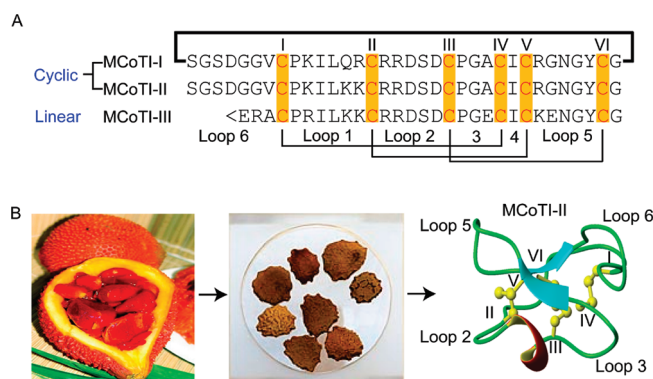

Figure 1. Sequences of previously characterized peptides from *M. cochinchinensis* and photographs of the plant's fruit and seeds. (A) Sequences are shown for the cyclic peptides MCoTI-I and MCoTI-II and the acyclic MCoTI-III. Conserved cysteine residues are highlighted, and the cystine connectivity is labeled with I–IV, II–V, and III–VI. The loop sequences between the cysteines are numbered,[14] and a thick black line joining the N and C termini indicates a cyclic backbone. In the MCoTI-III sequence, the pyro-Glu residue at the N-terminal is represented as '<'. This residue can be determined after treatment with pyroglutamyl aminopeptidase.[7] (B) From left to right: halved fruit from *M. cochinchinensis*; the seeds; and the MCoTI-II structure (PDB ID: 1IB9) showing the antiparallel β-sheet with arrows, the short 3_{10} helix as a ribbon, the disulfide bonds in yellow ball and stick format, and several labeled loops.

on cancer cell lines. Two novel peptides with toxicity against cancer cell lines were discovered and structurally characterized.

Results and Discussion

Peptides were extracted from homogenized seeds of *M. cochinchinensis* in NaOAc buffer, and the crude extract was freeze-dried. An analytical RP-HPLC trace of the crude extract is shown in Figure 2A. The extract was fractionated using preparative RP-HPLC and analyzed using EIMS. Two novel peptides were isolated and named MCoCC-1, for *M. cochinchinensis* with cytotoxicity activity on cancer cell lines-1, and MCoCC-2. Both peptides eluted at 29–30 min using a 2% gradient on an analytical RP-HPLC. MCoCC-1 and MCoCC-2 had masses of 3287 and 3169 Da, respectively

* To whom correspondence should be addressed. Tel: +61-7-3346 2021. Fax: +61-7-3346 2101. E-mail: n.daly@uq.edu.au.
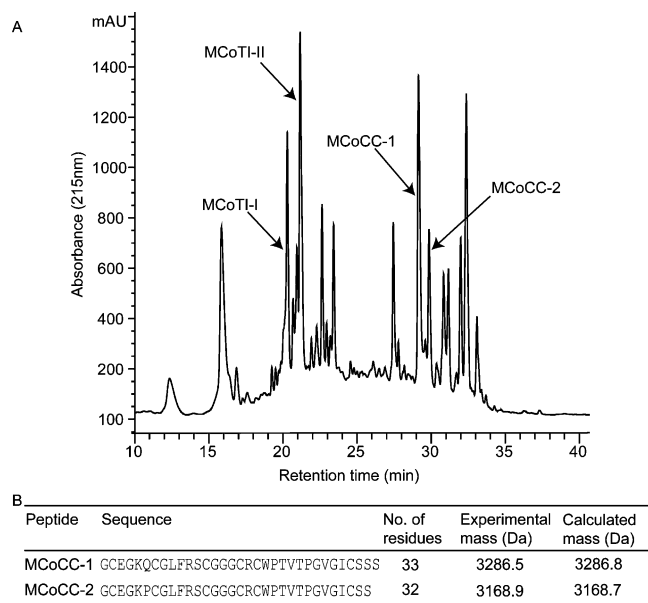
Figure 2. RP-HPLC elution profile of *M. cochinchinensis* seed extract and the sequences of MCoCC-1 and MCoCC-2. (A) The elution profile is an analytical RP-HPLC trace using a 2% gradient with a flow rate of 0.3 mL/min and demonstrates the close elution times of MCoCC-1 and MCoCC-2 as expected from their high sequence homology. The elution profile also shows the previously characterized cyclic peptides, MCoTI-I and MCoTI-II. (B) Summary table of the sequences and molecular weights of MCoCC-1 and MCoCC-2.
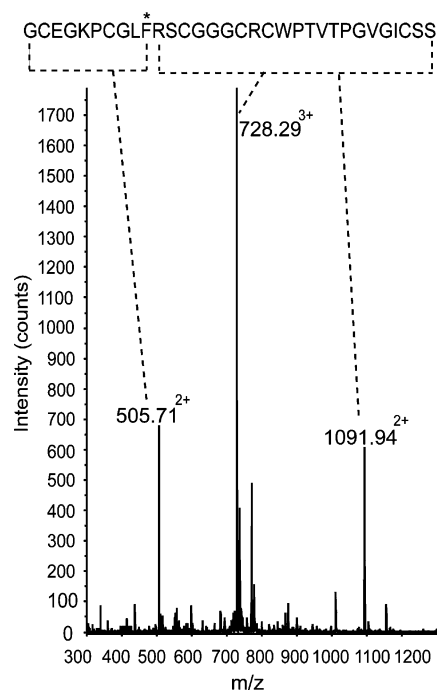

Figure 3. Nanospray MS data after chymotrypsin digestion of the MCoCC-2 peptide. The precursor ions obtained following cleavage with chymotrypsin are shown. MS/MS fragmentation on selected precursor ions was used to deduce the amino acid sequence. 505.71^{2+} is the precursor ion that resulted after the cleavage following a phenylalanine residue (marked with an asterisk) and was used to deduce the first half of the sequence (GCEGKPCGLF). Both 728.29^{3+} and 1091.94^{2+} correspond to precursor ions for the second half of the sequence (RSCGGGCRCWPTVTPGVGICSS).

(Figure 2B). MCoCC-1 was more abundant, and consequently easier to purify, than MCoCC-2 and, thus, was the major focus of the functional and structural experiments performed in this study.

To determine whether these peptides contained disulfide bonds, they were treated with the reducing agent tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and a mass increase of 6 Da was observed, suggesting that three disulfide bonds had been reduced. Both peptides were reduced and digested with a combination of trypsin and chymotrypsin, followed by sequencing by tandem mass spectrometry. Figure 3 shows an example nanospray MS spectrum of an MCoCC-2 fragment after it was digested by chymotrypsin. The sequences of both peptides were derived from analysis of the MS fragmentation patterns (see Supporting Information, Table 1S), and their amino acid compositions were verified using high-sensitivity amino acid analysis results (see Supporting Information, Table 2S). It should be noted that MCoCC-1 and MCoCC-2 were not cleaved at the single Leu residue in the sequence during chymotrypsin digestion for the nanospray analysis, and NMR analysis was required to determine the sequence position of the Ile and Leu residues. In particular, DQF-COSY spectra, which only have up to three-bond correlations, were used to distinguish between the two residues. Although leucine and isoleucine are equivalent in molecular weight, different couplings are observed in the COSY spectrum as a result of the different side chain configurations.

MCoCC-1 and MCoCC-2 comprise 33 and 32 amino acid residues, respectively, and differ by only two residues (Figure 2B). Neither peptide is homologous to any peptide found in the SWISS PROT database. Aside from having six cysteine residues, they share no sequence similarity and have different loop spacings from known *M. cochinchinensis* peptides, i.e., MCoTI-I, MCoTI-II, and MCoTI-III, as shown in Figure 1A. Thus, they represent a novel class of peptides.

Determining the connectivity of disulfide bonds in disulfide-rich peptides can be complicated by the presence of tightly folded structures and the proximity of multiple disulfide bonds within these molecules.[16] Nevertheless, selective reduction of the disulfide bonds followed by derivatization of the cysteine residues and sequence analysis has been successfully used to determine the disulfide connectivity of a range of peptides.[17] This approach was tried for MCoCC-1, but the selective reduction procedure failed to yield partially reduced fragments, and only the fully oxidized and fully reduced species were observed in the RP-HPLC traces (results not shown). Given the lack of partially reduced intermediates observed during the reduction of MCoCC-1, an NMR approach was used in an attempt to determine the disulfide connectivity. Although such an approach may also be complicated by the tightly folded structures of disulfide-rich peptides,[18] it has proved successful in a number of cases.[19,20] TOCSY and NOESY spectra were recorded and showed well-dispersed peaks, as well as a large number of NOE cross-peaks, which is indicative of a well-defined structure (see Supporting Information, Figure 3S). NMR assignments for MCoCC-1 were made using established techniques,[21] and the complete assignment of chemical shifts is shown in Table 4S (Supporting Information). The only complication was that the amide proton signals of Leu9 and Ser12 were not observed in the NMR spectra, suggesting that both residues may reside in a conformationally flexible region of the molecule and the signals are consequently broadened beyond detection.

Analysis of the NMR αH secondary shifts provided a first indication of the secondary structure of MCoCC-1. In general several consecutive positive αH secondary shifts[22] indicate the presence of β-sheet structures. Conversely, stretches of negative secondary shifts indicate the presence of helical structures. The αH secondary shifts of MCoCC-1 are shown in Figure 4A, and on the basis of the significant number of positive values it can be deduced that MCoCC-1 has a β-strand structure in the C-terminal region.

Three-dimensional structures were calculated using 195 distance restraints and 13 dihedral angle restraints derived from the NMR data. Two approaches were used in an attempt to determine the
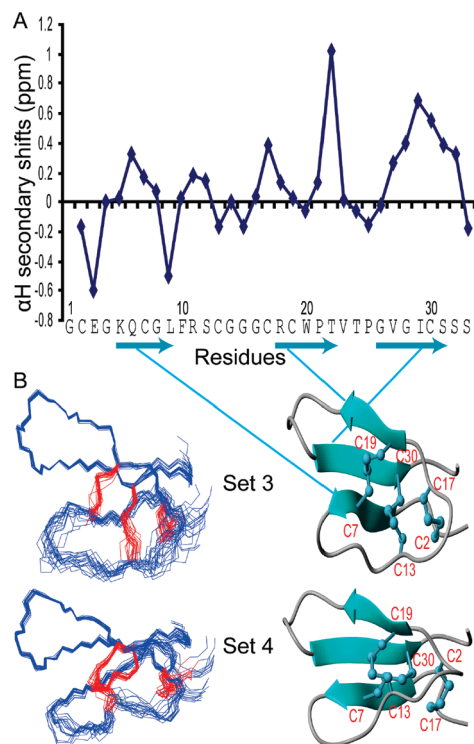
Figure 4. αH secondary shifts of MCoCC-1 determined from ^{1}H NMR spectra measured at 290 K and structures of MCoCC-1 calculated with two alternative disulfide bond connectivities. (A) The αH secondary shifts were calculated by subtracting the random coil ^{1}H NMR chemical shifts of Wishart et al.[22] from the measured αH chemical shifts. A cyan arrow below the stretch of consecutive amino acids having positive αH secondary shifts near the C-terminus supports the presence of a β-strand in this region. (B) A superposition of the 20 lowest energy structures for the two sets of disulfide connectivities is shown on the left with the disulfide bonds in red. The lowest energy structure is shown on the right with the disulfide bonds shown in ball and stick format. Cysteine residues are labeled in red: set 3 (Cys7−Cys19; Cys13−Cys30; Cys2−Cys17); set 4 (Cys2−Cys17; Cys7−Cys30; Cys13−Cys19). The diagram was created in MolMol.[38]

Table 1. Mean S−S Distances (Å) in the MCoCC-1 Structures Calculated without Disulfide Bond Restraints

bond (Cys−Cys pair)	mean distance[a] (Å)	order[b]
2−7	11.76	13
2−13	5.64	4
2−17	5.64	5
2−19	13.48	15
2−30	7.89	8
7−13	10.66	12
7−17	9.18	10
7−19	3.35	1
7−30	4.27	2
13−17	8.73	9
13−19	12.89	14
13−30	7.79	7
17−19	9.97	11
17−30	5.39	3
19−30	6.03	6

[a] Mean distance between sulfur atoms of the indicated residues calculated from the 20 lowest energy structures. [b] Rank order of mean distances.

disulfide bond connectivity. In the first approach, initial structures were calculated without restraints on any disulfide bonds and the distances between the sulfur atoms for each possible pair of the cysteine residues were measured to determine the most likely disulfide bonds based on proximity. The results of this analysis are shown in Table 1. The most probable disulfide bond was found to be between Cys7 and Cys19. The next most probable bond was between Cys7 and Cys30, but this is inconsistent with the first disulfide bond, making it difficult to determine the complete connectivity using this approach. The limitation of this simple approach was previously seen in the structural analysis of kalata B1, a 29-residue plant-derived cyclic peptide found in *Oldenlandia affinis*, also with three disulfide bonds.[14,23] A more sophisticated approach was thus undertaken. This involved calculating structures with different possible disulfide bond connectivities and comparing the energies of these structures to determine which connectivity best satisfied the experimental restraints.[18] From the 15 possible combinations of three-disulfide bond connectivities, only four of these combinations had few experimental restraint violations. In these four sets of connectivities, the Cys7−Cys19 connectivity was present in three. The results from the analysis are given in Table 2. An overlay of the 20 lowest energy structures and the lowest energy structure for sets 3 and 4, the two sets with the lowest experimental restraint violations, are illustrated in Figure 4B. Despite the differences in disulfide connectivity the overall folds

of all four sets are very similar, indicating that different connectivities can be accommodated without significant disruption to the structure. However, set 3 is the most likely disulfide bond connectivity, as this structure had the lowest NOE, dihedral, and overall energies when compared to energies obtained without disulfide connectivities. Encouragingly, it includes the Cys7−Cys19 connectivity deduced as being present from the simple distance protocol noted above.

The 20 lowest energy structures of set 3 are in good agreement with the experimental data, showing no dihedral angle violation exceeding 3° and no distance violation more than 0.3 Å (Supporting Information, Table 5S). Analysis of the secondary structure was done with PROMOTIF[24] and revealed that an antiparallel β-sheet is the major element of secondary structure, comprising strands involving residues 5−8, 19−21, and 27−30. In addition, there is a type II β-turn at residues 24−27. These structural findings are consistent with the αH secondary shift analysis, which suggested that the C-terminal region of MCoCC-1 adopts an extended β-strand structure. The disulfide bonds of set 3 (Cys2−Cys17; Cys7−Cys19; Cys13−Cys30) form a cystine knot motif, and the loop between Cys7−Cys13 is disordered, as shown in Figure 4B. The disorder is probably due to flexibility in this loop. Interestingly, the N- and C-termini of the structure are in close proximity to each other.

Although MCoCC-1 has no sequence similarity to previously characterized peptides, the cystine knot motif seen in it is prevalent in a wide range of organisms. Indeed, a recent paper noted that the cystine knot[25] topology is found in nearly 40% of known Cys-rich peptide domains and is the most commonly observed structural motif. It is likely that this motif is prevalent because it confers significant stability to the fold. Despite the stability conferred by the cystine knot, one loop of MCoCC-1 (between residues Cys7 and Cys13) is quite disordered, most likely as a result of structural flexibility in this region. Although this is unusual, it is not unprecedented, and, for example, other cystine knot molecules such as kalata B8[26] and MCoTI-II[10,11] have individual loops that are disordered.

Biological assays were conducted to characterize the two novel peptides. MCoCC-1 was tested in all assays, but MCoCC-2 was tested only in the cytotoxicity assay because of limited amounts of material. Cytotoxicity assays were conducted to examine the effect of the peptides against four human cell lines, including human colorectal adenocarcinoma (HT29), human lung carcinoma cells (A549), human melanoma cell line (MM96L), and a noncancerous control neonatal foreskin fibroblast cell line (NFF). An *M. cochinchinensis* seed extract was also tested in this assay for comparison with the purified peptides. The highest level of cytotoxicity found for MCoCC-1 was against the MM96L cell line, which showed 57% cell survival in the presence of 2 μM MCoCC-1. By contrast, the control NFF cell line exhibited 89% cell survival at this concentra-
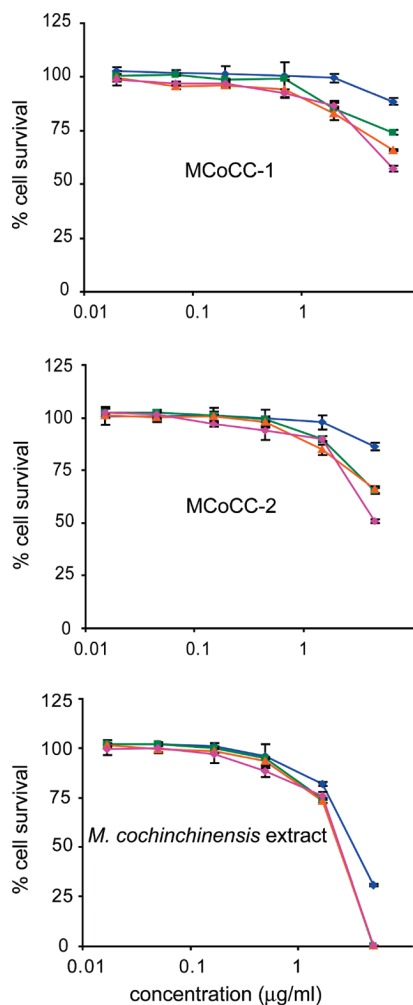
Figure 5. Cytotoxicity of purified MCoCC-1, MCoCC-2, and a *M. cochinchinensis* seed extract. Three cancer cell lines were used in the assays: a human colorectal adenocarcinoma cell line (HT29; green); a human lung carcinoma cell line (A549; orange); and a human melanoma cell line (MM96L; pink). Neonatal foreskin fibroblast cells were used as a noncancerous control (NFF; blue).

tion. Similarly, the highest and lowest cytotoxicity observed for MCoCC-2 treatments at 1.3 μM were 51% cell survival for the MM96L cell line and 86% cell survival for the NFF cell line, respectively, as shown in Figure 5. These findings highlighted that both peptides not only have similar sequences but also have similar cytotoxicity on MM96L and NFF cell lines. The *M. cochinchinensis* seed extract had potent cytotoxicity at 5 μg/mL on all cancer cell lines, with 0% cell survival for the MM96L cell line and 31% cell survival for the control NFF cell line. Thus, the *M. cochinchinensis* seed extract has more potent cytotoxicity on cancer cell lines than the isolated peptides, indicating that there are probably additional cytotoxic components that remain to be identified in the extract. Although the cytotoxic activity for both MCoCC-1 and MCoCC-2 was not significantly different between all cancer cell lines, encouragingly, both peptides caused approximately 50% cell death on the MM96L cell line.

As other disulfide-rich peptides isolated from *M. cochinchinensis* are trypsin inhibitors,[7] MCoCC-1 was tested for trypsin inhibitory activity. The sequence of MCoCC-1 is significantly different from the previously characterized trypsin inhibitors. However, a lysine residue, present in loop 1 of the trypsin inhibitors MCoTI-I, MCoTI-II, and MCoTI-III, which forms the primary trypsin recognition site, is also present in MCoCC-1, albeit located in a slightly different position. Despite the presence of this Lys in loop 1, no trypsin inhibitory effect was observed for MCoCC-1 within the 0.04−0.5 μM range.

Hemolytic assays were conducted to examine the effect of MCoCC-1 on human erythrocytes, as hemolytic peptides are not appropriate candidates for drug design. The percentage of hemolysis associated with the incubation of red blood cells with MCoCC-1 was examined and compared to melittin, a highly hemolytic peptide from honey bee venom.[27] The extent of hemolysis caused by melittin increased sharply between 2.5 and 20 μM, as expected for this positive control, whereas no hemolysis was observed for MCoCC-1 up to 50 μM.

The lack of toxicity on erythrocytes and lack of trypsin inhibitory activity pose the question of the natural function of MCoCC-1. There are several examples of plant-derived peptides of similar size (i.e., ~30 amino acids) displaying insecticidal activity, including the cyclotides kalata B1[28] and kalata B2[29] and plant defensins.[30] Although MCoCC-1 has a different sequence than these other peptides, its plant origin and disulfide-rich structure in common with the cyclotides meant that it was of interest to test it for insecticidal activity. MCoCC-1 was tested against the insect species *Helicoverpa armigera,* which, along with other species in the *Helicoverpa* genus, is a major pest in cotton and other commercial crops, such as corn and sorghum. *H. armigera* larvae were tested in feeding trials starting at the late second to early third instar developmental stage because more visible effects from changes in diet can be observed at this stage. Larvae were divided into three groups and given different diets: a control diet, a diet containing an *O. affinis* extract, and a diet containing MCoCC-1. The control diet consisted of soy flour and wheat germ without the addition of peptide.[29] *O. affinis* extract was added to a second diet as a positive control, since it has previously been shown to display insecticidal activity.[28] Analysis of the weights of larvae over a feeding period of 96 h revealed that larvae fed with a diet containing MCoCC-1 gained more weight than larvae fed with the *O. affinis* extract diet. In contrast, larvae fed with the control diet and larvae fed with the diet containing MCoCC-1 showed similar results, indicating that MCoCC-1 does not have an insecticidal effect.

A comparison of the surfaces of MCoCC-1 and kalata B1 is given in Figure 6 and helps to explain why no insecticidal effect is seen for MCoCC-1. The surface-exposed hydrophobic residues of kalata B1 have been shown to be involved in membrane binding,[31] and disruption of the midgut membranes of insects fed kalata B1 is the presumed mechanism of insecticidal action.[32] Given the different locations of the hydrophobic patches in the MCoCC-1 and kalata B1, it seems unlikely that they have a similar mode of membrane binding.

In summary, we have discovered two novel peptides from *M. cochinchinensis* seeds. They have no sequence homology to any previously characterized peptides and display cytotoxicity effects on cancer cell lines. MCoCC-1 likely contains a cystine knot motif, a common structural feature that imparts considerable stability to protein folds. MCoCC-1 is not hemolytic and, unlike peptides

Table 2. Structural Statistics of MCoCC-1 Structures Calculated with Alternative Disulfide Restraints

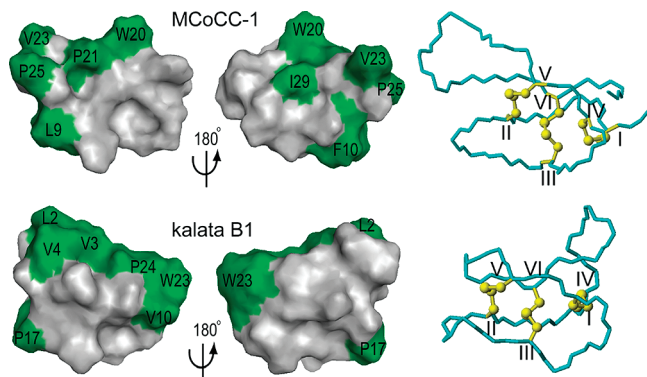| disulfide connectivities	NOE energy (kJ mol^{-1})	dihedral energy (kJ mol^{-1})	overall energy (kJ mol^{-1})
no disulfide connectivities	14.44 ± 2.61	1.08 ± 0.73	−1159.58 ± 19.07
2−13, 7−19, 17−30 (set 1)	13.43 ± 4.34	2.51 ± 0.91	−1116.88 ± 6.64
2−30, 7−19, 13−17 (set 2)	20.08 ± 5.55	1.66 ± 1.14	−1110.57 ± 23.61
2−17, 7−19, 13−30 (set 3)	12.00 ± 3.22	0.69 ± 0.28	−1186.40 ± 14.25
2−17, 7−30, 13−19 (set 4)	14.16 ± 3.05	1.45 ± 0.55	−1168.04 ± 9.66

Figure 6. Comparison of surface features of MCoCC-1 and the insecticidal plant peptide kalata B1. Hydrophobic residues are shown in green, and selected amino acids are labeled. The ball and stick diagrams on the right highlight the common orientation of the peptides. Kalata B1 has been shown to involve membrane binding via a surface-exposed hydrophobic patch in its mechanism of insecticidal activity.[32] The different molecular surface of MCoCC-1, despite a similar overall fold, explains its lack of insecticidal activity.

previously characterized from *M. cochinchinensis* seeds, does not have trypsin inhibitory activity. Overall, this study has expanded knowledge on *M. cochinchinensis* peptides and contributes to the field of cystine knot peptides.

Experimental Section

General Experimental Procedures. UV spectra were measured on a Varian Cary 50 UV visible spectrophotometer. 1D and 2D NMR spectra were recorded on both Bruker Avance 500 and 600 MHz spectrometers. Data were processed using TOPSIN (Bruker). Mass spectrometry data were obtained using a Micromass LCT mass spectrometer. MALDI-TOFMS analysis was carried out using a Voyager DE-STR mass spectrometer (Applied Biosystems). Data were collected between 500 and 4000 Da. Peptide sequencing was conducted using nanospray MS/MS on a QStar mass spectrometer. Ion spray voltage was applied at 900 V, and collision energy between 10 and 50 V was applied for peptide fragmentation. Data processing and acquisition were done using the Analyst software program with spectra acquired at m/z 400−2000 for TOF spectra and m/z 60−2000 for product ion spectra. RP-HPLC was carried out using an Agilent 1100 series with a UV detector at variable wavelengths of 215, 254, and 280 nm and Waters 600 system controller with a Waters 484 tunable absorbance detector at 215 nm.

Extraction of Peptides from the *M. cochinchinensis* Seed Extract. The frozen ripe fruit of *M. cochinchinensis* was purchased from a Vietnamese store in Melbourne, Australia. The fruit was thawed, and the seeds were removed from the pulp and homogenized in the presence of ice-cold NaOAc (20 mM, pH 5.0).[10] The mixture was blended and stirred overnight at 4 °C, before centrifugation for 30 min at 5343*g*. The soluble fraction was stirred with an equal volume of ice-cold acetone for 1 h, and the precipitate was removed using a Whatman 541 filter paper. The acetone was evaporated using a rotary evaporator, and the final extract was freeze-dried.

Purification of Peptides from the *M. cochinchinensis* Seed Extract. The peptides were purified to >95% purity from the dried *M. cochinchinensis* seed extract by a series of RP-HPLC purifications on Phenomenex C18 columns. Gradients of 1% min⁻¹ of 0−60% solvent B (90% MeCN in 0.045% TFA in H_2O) from solvent A (aqueous 0.05% TFA in H_2O) were employed, and the eluant was monitored at 215 nm. The final purity of the peptides was examined by analytical RP-HPLC on a Phenomenex Jupiter 5u C18 300 Å 150 × 2.0 mm column and the mass assessed by electrospray mass spectrometry.

Sequencing of the Peptides Using Nanospray Mass Spectrometry. The peptides from the *M. cochinchinensis* seed extract were sequenced using nanospray MS-MS on a QStar mass spectrometer. Initially, disulfide bonds of the peptides were reduced with TCEP. These peptides were then digested with endoproteases, trypsin and chymotrypsin. The digestions were quenched after 2 h by adding an equal volume of 0.5%

HCO_2H and desalted using Ziptips (Millipore). Precursor ions from the digested peptide were selected and sequenced via nanospray MS-MS.[33] A voltage of 900 V was applied, and spectra were acquired in the range m/z 60−2000 for both product ion and time-of-flight spectra. Collision energy was applied between 10 and 50 V for ion fragmentation. The composition of the peptide was then confirmed with high-sensitivity amino acid analysis, which was conducted by the Australian Proteome Analysis Facility Ltd. Peptide samples were hydrolyzed for 24 h with 6 N HCl at 110 °C, and amino acids were analyzed using the Waters AccQTag chemistry. This analysis was performed in duplicate.

NMR Spectroscopy. Samples for ¹H NMR measurements contained ∼2 mg of peptide in 90% H_2O/10% D_2O (v/v) at ∼pH 5.5. D_2O (99.9%) was obtained from Cambridge Isotope Laboratories, Woburn, MA. Spectra were recorded at 290 K on both Bruker Avance 500 and Bruker Avance 600 MHz spectrometers. One-dimensional ¹H spectra were acquired, followed by two-dimensional spectra, which included TOCSY, NOESY, and DQF-COSY. Spectra were analyzed using the program SPARKY.[34] The sequential assignment procedure pioneered by Wüthrich[21] was initially used to sequence-specifically assign the amino acids, using TOCSY and NOESY spectra. DQF-COSY spectra were used to confirm the side chain assignments for several residues.

Structure Calculations. The three-dimensional structure was calculated by deriving distance and angle restraints from the NOESY and DQF-COSY spectra, respectively. A family of structures that are consistent with the experimental restraints was calculated using the programs CYANA[35] and CNS.[36] Preliminary structures were calculated using CYANA and used to resolve ambiguities in the NOE assignments. Once a complete set of input restraints (distance and dihedral angle restraints) was determined, structures were calculated with a simulated annealing protocol within the program CNS. A set of 50 structures was calculated, and the 20 lowest energy structures were selected for further analysis. Structures were analyzed using the programs PROCHECK[37] and PROMOTIF[24] to generate statistical analyses, such as a Ramachandran plot, from the structure calculations. The programs MolMol[38] and PyMol[39] were used to display the structural ensembles and surfaces of the proteins, respectively.

Partial Reduction of MCoCC-1 to Determine the Connectivity of the Disulfide Bond. MCoCC-1 (∼100 μg) peptide was dissolved with 100 μL of citrate buffer (0.2 M, pH 3). Freshly prepared tris(2-carboxyethyl)phosphine hydrochloride (10 μL, 20 mM) was added to 10 μL of peptide solution and then incubated at 37 °C for 30 min in a steam bath. Samples were analyzed on analytical RP-HPLC using a 2% gradient with 0.3 mL/min flow rate on a Phenomenex Jupiter 5u C18 300 Å 150 × 2.0 mm column. The success of reduction of MCoCC-1 was determined by a comparison of its oxidized and reduced elution profiles. The elution profiles were overlaid to observe the presence of any intermediates.

Cytotoxicity Assay. The cytotoxicity of MCoCC-1 and MCoCC-2 was tested on three human tumor cell lines, namely, the human colorectal adenocarcinoma (HT29), human lung carcinoma cells (A549), the human melanoma cell line (MM96L), and a noncancerous cell line, neonatal foreskin fibroblast cells (NFF). These cell lines were chosen because they have been used in recent cancer disease studies.[40−42] All cells were cultured in RPMI 1640 supplemented with 10% FBS at 37 °C, 5% CO_2, and 99% humidity. Cells were seeded into individual wells of the 96-well tissue culture plates (5000 cells/well) and grown for 24 h prior to treatment. The peptides were dissolved in H_2O and diluted in culture medium. Cells were treated with 0.02, 0.07, 0.2, 0.7, 2.0, and 7 μg/mL and no peptide (culture media only).[43] Three days after treatment initiation, cells were washed with phosphate-buffered saline (PBS) and fixed with methylated spirits, and total protein was determined using sulforhodamine B as described in Skehan et al.[44] Cytotoxicity was calculated as a percentage of the nonpeptide-treated control cell protein, where 100% of control represented 100% cell survival. All treatments were performed in triplicate, and each experiment was performed in duplicate.

Trypsin Inhibitor Assay. A stock solution of *N*α-benzoyl-DL-arginine 4-nitroanilide (BAPNA) was prepared at 2 mg/mL and stored on ice. Immediately prior to use, a stock solution of trypsin was prepared (10 mg/mL in H_2O). The substrate (chicken trypsin inhibitor) was prepared by dissolving 1.5 mg/mL in distilled water. Four samples were tested in this assay: 1 (0.21 μM stock) and 10 μL (2.1 μM stock) of MCoCC-1, 10 μL of chicken trypsin inhibitor (a positive control), and 20 μL of trypsin solution (a negative control). Each Eppendorf tube

consisted of 20 μL of trypsin solution, 40 μL of BAPNA, variable volume of test compounds as noted above, and a buffer of 50 mM tris-HCl pH 7.4; 10 mM CaCl$_2$ was added to make up a total volume of 600 μL. BAPNA was added last after 1.5 min of incubation at room temperature. A sample that contained only trypsin solution was used to measure 100% trypsin activity. Absorbance readings were obtained at 410 nm using a CARY-50 UV−visible spectrophotometer (Varian). The assay was performed in triplicate.

Hemolytic Assay. Peptides were prepared by two serial dilutions with PBS as previously described.[16,34] Human red blood cells (RBCs) were washed with PBS and centrifuged repeatedly at 1500*g* for 1 min until a clear supernatant was obtained. Synthetic melittin was used as a positive control due to its high hemolytic activity. Twenty microliters of each peptide solution, melittin, or 1% Triton X-100 was added to 100 μL of 0.25% suspension of washed RBCs in PBS. The highest test concentration for the peptides and synthetic melittin (Sigma) solutions was 50 and 20 μM, respectively. The plate was incubated at 37 °C for 1 h and centrifuged at approximately 150*g* for 5 min using a plate centrifuge. The supernatant (100 μL) was transferred out of each well into a 96-well, flat-bottomed plate, and the absorbance was measured at 415 nm (Multiskan Ascent plate reader, Labsystems). Hemolysis was calculated as the percentage of maximum lysis (Triton X-100) and was plotted using GraphPad Prism software.

Insecticidal Assay. *H. armigera* larvae were obtained from the Queensland Department of Primary Industries (QDPI). A feeding trial adapted from one reported recently[32] was conducted for 96 h; larvae were kept at 25 °C throughout the experiment. Larvae were given one of three diets and were tested in triplicate. All diets contained mainly wheat germ, yeast, and soy flour. The control diet did not have any added peptide. Jennings et al.[28] showed that pure kalata B1 from *O. affinis* extract had insecticidal activity on larvae; hence a diet containing *O. affinis* extract was used as a positive control. The quantity of *O. affinis* extract in the diet was equivalent to a calculated content of 0.825 μmol of kalata B1. The third diet contained approximately 0.825 μmol of MCoCC-1. Initially each larva was weighed and starved for 6 h in an individual Eppendorf tube with holes. On the sixth hour each larva was transferred into a feeding container, given 300 mg of their respective diet, and allowed to feed for 96 h. Each larva was weighed after the first feeding on the 24th hour and again at 24 h intervals until the end of this feeding trial. Following this, the larvae, remaining diet, and frass were weighed and photographed.

Supporting Information Available: (1) Sequence fragments from the enzymatic digestion of peptides from *M. cochinchinensis,* (2) a comparison of amino acid composition for MCoCC-1 and MCoCC-2, (3) 2D NMR spectra of MCoCC-1, (4) chemical shifts of MCoCC-1 at 290 K, and (5) structural statistics of the MCoCC-1 structures calculated with set 3 disulfide connectivity are available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Tien, P. G.; Kayama, F.; Konishi, F.; Tamemoto, H.; Kasono, K.; Hung, N. T.; Kuroki, M.; Ishikawa, S. E.; Van, C. N.; Kawakami, M. *Int. J. Oncol.* **2005**, *26*, 881–9.

(2) Huang, B.; Ng, T. B.; Fong, W. P.; Wan, C. C.; Yeung, H. W. *Int. J. Biochem. Cell Biol.* **1999**, *31*, 707–15.

(3) Akihisa, T.; Higo, N.; Tokuda, H.; Ukiya, M.; Akazawa, H.; Tochigi, Y.; Kimura, Y.; Suzuki, T.; Nishino, H. *J. Nat. Prod.* **2007**, *70*, 1233–9.

(4) Tien, P. G.; Kayama, F.; Konishi, F.; Tamemoto, H.; Kasono, K.; Hung, N. T.; Kuroki, M.; Ishikawa, S. E.; Van, C. N.; Kawakami, M. *Int. J. Oncol.* **2005**, *26*, 881–9.

(5) Ng, T. B.; Chan, W. Y.; Yeung, H. W. *Gen. Pharmacol.* **1992**, *23*, 579–90.

(6) Akihisa, T.; Higo, N.; Tokuda, H.; Ukiya, M.; Akazawa, H.; Tochigi, Y.; Kimura, Y.; Suzuki, T.; Nishino, H. *J. Nat. Prod.* **2007**, *70*, 1233–9.

(7) Hernandez, J. F.; Gagnon, J.; Chiche, L.; Nguyen, T. M.; Andrieu, J. P.; Heitz, A.; Trinh Hong, T.; Pham, T. T.; Le Nguyen, D. *Biochemistry* **2000**, *39*, 5722–30.

(8) Tsoi, A. Y.; Ng, T. B.; Fong, W. P. *J. Pept. Sci.* **2005**, *11*, 665–8.

(9) Wong, R. C.; Fong, W. P.; Ng, T. B. *Peptides* **2004**, *25*, 163–9.

(10) Felizmenio-Quimio, M. E.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2001**, *276*, 22875–82.

(11) Heitz, A.; Hernandez, J. F.; Gagnon, J.; Hong, T. T.; Pham, T. T.; Nguyen, T. M.; Le-Nguyen, D.; Chiche, L. *Biochemistry* **2001**, *40*, 7973–83.

(12) Pallaghy, P. K.; Nielsen, K. J.; Craik, D. J.; Norton, R. S. *Protein Sci.* **1994**, *3*, 1833–9.

(13) Craik, D. J.; Daly, N. L.; Waine, C. *Toxicon* **2001**, *39*, 43–60.

(14) Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. *J. Mol. Biol.* **1999**, *294*, 1327–36.

(15) Colgrave, M. L.; Craik, D. J. *Biochemistry* **2004**, *43*, 5965–75.

(16) Göransson, U.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 48188–96.

(17) Daly, N. L.; Clark, R. J.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 6314–22.

(18) Daly, N. L.; Koltay, A.; Gustafson, K. R.; Boyd, M. R.; Casas-Finet, J. R.; Craik, D. J. *J. Mol. Biol.* **1999**, *285*, 333–45.

(19) Walewska, A.; Skalicky, J. J.; Davis, D. R.; Zhang, M. M.; Lopez-Vera, E.; Watkins, M.; Han, T. S.; Yoshikami, D.; Olivera, B. M.; Bulaj, G. *J. Am. Chem. Soc.* **2008**, *130*, 14280–6.

(20) Rosengren, K. J.; Daly, N. L.; Plan, M. R.; Waine, C.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 8606–16.

(21) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, 1986.

(22) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67–81.

(23) Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. *Biochemistry* **1995**, *34*, 4147–58.

(24) Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477–86.

(25) Cheek, S.; Krishna, S. S.; Grishin, N. V. *J. Mol. Biol.* **2006**, *359*, 215–37.

(26) Daly, N. L.; Clark, R. J.; Plan, M. R.; Craik, D. J. *Biochem. J.* **2006**, *393*, 619–26.

(27) Chen, B.; Colgrave, M. L.; Wang, C.; Craik, D. J. *J. Nat. Prod.* **2006**, *69*, 23–8.

(28) Jennings, C.; West, J.; Waine, C.; Craik, D.; Anderson, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10614–9.

(29) Jennings, C. V.; Rosengren, K. J.; Daly, N. L.; Plan, M.; Stevens, J.; Scanlon, M. J.; Waine, C.; Norman, D. G.; Anderson, M. A.; Craik, D. J. *Biochemistry* **2005**, *44*, 851–60.

(30) Liu, Y. J.; Cheng, C. S.; Lai, S. M.; Hsu, M. P.; Chen, C. S.; Lyu, P. C. *Proteins* **2006**, *63*, 777–86.

(31) Shenkarev, Z. O.; Nadezhdin, K. D.; Sobol, V. A.; Sobol, A. G.; Skjeldal, L.; Arseniev, A. S. *FEBS J.* **2006**, *273*, 2658–72.

(32) Barbeta, B. L.; Marshall, A. T.; Gillon, A. D.; Craik, D. J.; Anderson, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1221–5.

(33) Wattenberg, A.; Organ, A. J.; Schneider, K.; Tyldesley, R.; Bordoli, R.; Bateman, R. H. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 772–83.

(34) Kneller, D.; Kuntz, I. *J. Cell. Biochem.* **1993**, 254.

(35) Ikeya, T.; Terauchi, T.; Guntert, P.; Kainosho, M. *Magn. Reson. Chem.* **2006**, *44*, S152–7.

(36) Brünger, A. T.; Adams, P. D.; Rice, L. M. *Structure* **1997**, *5*, 325–36.

(37) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1996**, *5*, 212–20.

(38) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graph.* **1996**, *14*, 29–32.

(39) DeLano, W. *The PyMOL Molecular Graphics System*; DeLano Scientific San Carlos: CA, 2002.

(40) Glenn, M. P.; Kahnberg, P.; Boyle, G. M.; Hansford, K. A.; Hans, D.; Martyn, A. C.; Parsons, P. G.; Fairlie, D. P. *J. Med. Chem.* **2004**, *47*, 2984–94.

(41) Murillo, G.; Hirschelman, W. H.; Ito, A.; Moriarty, R. M.; Kinghorn, A. D.; Pezzuto, J. M.; Mehta, R. G. *Nutr. Cancer* **2007**, *57*, 28–37.

(42) Sato, H.; Fukumoto, K.; Hada, S.; Hagiwara, H.; Fujimoto, E.; Negishi, E.; Ueno, K.; Yano, T. *Cancer Chemother. Pharmacol.* **2007**, *60*, 449–57.

(43) Zhang, H.; Major, J. M.; Lewis, R. J.; Capon, R. J. *Org. Biomol. Chem.* **2008**, *6*, 3811–5.

(44) Skehan, P.; Storeng, R.; Scudiero, D.; Monks, A.; McMahon, J.; Vistica, D.; Warren, J. T.; Bokesch, H.; Kenney, S.; Boyd, M. R. *J. Natl. Cancer Inst.* **1990**, *82*, 1107–12.