Isolation, Sequencing, and Structure−Activity Relationships of Cyclotides

David C. Ireland, Richard J. Clark, Norelle L. Daly, and David J. Craik*

Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia

Received January 21, 2010

Cyclotides are a topologically fascinating family of miniproteins discovered over the past decade that have expanded the diversity of plant-derived natural products. They are approximately 30 amino acids in size and occur in plants of the Violaceae, Rubiaceae, and Cucurbitaceae families. Despite their proteinaceous composition, cyclotides behave in much the same way as many nonpeptidic natural products in that they are resistant to degradation by enzymes or heat and can be extracted from plants using methanol. Their stability arises, in large part, due to their characteristic cyclic cystine knot (CCK) structural motif. Cystine knots are present in a variety of proteins of insect, plant, and animal origin, comprising a ring formed by two disulfide bonds and their connecting backbone segments that is threaded by a third disulfide bond. In cyclotides, the cystine knot is uniquely embedded within a head-to-tail cyclized peptide backbone, leading to the ultrastable CCK structural motif. Apart from the six absolutely conserved cysteine residues, the majority of amino acids in the six backbone loops of cyclotides are tolerant to variation. It has been predicted that the family might include up to 50 000 members; although, so far, sequences for only 140 have been reported. Cyclotides exhibit a variety of biological activities, including insecticidal, nematocidal, molluscicidal, antimicrobial, antibarnacle, anti-HIV, and antitumor activities. Due to their diverse activities and common structural core from which variable loops protrude, cyclotides can be thought of as combinatorial peptide templates capable of displaying a variety of amino acid sequences. They have thus attracted interest in drug design as well as in crop protection applications.

Introduction

Traditionally, natural products chemists have focused on the study of organic molecules isolated from terrestrial or marine organisms, with plants being a particularly rich source of novel molecules. Herein we review an exciting family of plant-derived molecules called cyclotides,[1] which are different from traditionally studied natural products because they are proteinaceous. Although cyclotides are composed purely of amino acids and are translated from genes like conventional proteins, their very compact and stable structures mean they behave like organic molecules; for example, they survive severe heat treatment and are resistant to enzyme digestion.[2]

Cyclotides are distinguished from conventional proteins by their head-to-tail cyclized peptide backbone and a knotted arrangement of three disulfide bonds that link their six conserved Cys residues. Their cystine knot motif comprises an embedded ring in the structure formed by two of the disulfide bonds and their connecting backbone segments that is threaded by the third disulfide bond. Figure 1 shows the structure of the prototypic cyclotide kalata B1 and highlights the cystine knot and macrocyclic backbone made up of a ring of 29 amino acids. The combination of these two structural elements (cystine knot and cyclic backbone) is referred to as a cyclic cystine knot (CCK) motif[1] and is largely responsible for the exceptional stability of cyclotides.[2−4]

Cyclotides have been formally identified as a class of proteins for about a decade, but the discovery of their prototypic member, kalata B1, began in the 1960−1970s, when women in the Congo were reported to use a medicinal tea made from the plant *Oldenlandia affinis* (Figure 1) to accelerate childbirth.[5−8] In 1973, Gran and co-workers[9,10] isolated and partially characterized kalata B1, an active uterotonic agent of the plant, identifying it as a peptide of approximately 30 amino acids. However, it was 25 years later before the full sequence, macrocyclic nature, and cystine knot were elucidated.[11]

Several other macrocyclic peptides of similar size and cysteine content were discovered during bioassay-guided screening studies reported in the mid to late 1990s,[12−15] and in 1999 the term "cyclotides" was proposed[1] to characterize what was becoming

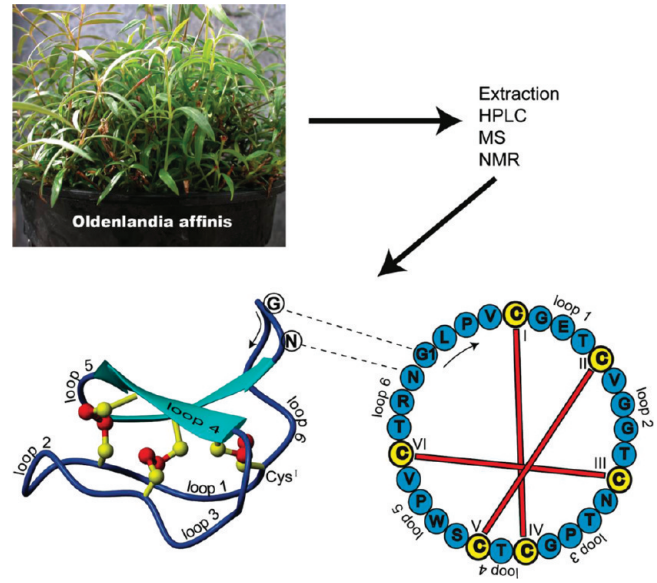


Figure 1. Plant source (*Oldenlandia affinis*), structure (PDB ID 1nb1), and sequence[11] of the prototypic cyclotide kalata B1. Cyclotides have a well-defined three-dimensional structure due to their interlocking disulfide bonds and cyclic peptide backbone. Backbone loops and selected residues are labeled on the structure. The amino acid sequence is indicated (single letter representation) and is numbered based on the presumed processing point (G1) from the precursor protein. The dashed lines indicate the ligation of terminal Gly and Asn residues to form the cyclic backbone.

apparent as a new family of proteins. Numerous members of the cyclotide family have since been discovered, with more than 140 sequences now published and documented in a database dedicated to cyclic proteins called CyBase[16] (http://www.cybase.au). Although cyclotides have been reported only in species of the Rubiaceae, Violaceae, and Cucurbitaceae plant families to date, their distribution does not seem limited by geography or habitat, with cyclotides occurring in plants from alpine, desert, rainforest, and urban environments (Figure 2). On the basis of hit rates in the species so far

* To whom correspondence should be addressed. Tel: 61-7-3346-2019. Fax: 61-7-3346-2029. E-mail: d.craik@imb.uq.edu.au.

10.1021/np1000413 © 2010 American Chemical Society and American Society of Pharmacognosy
Published on Web 08/18/2010

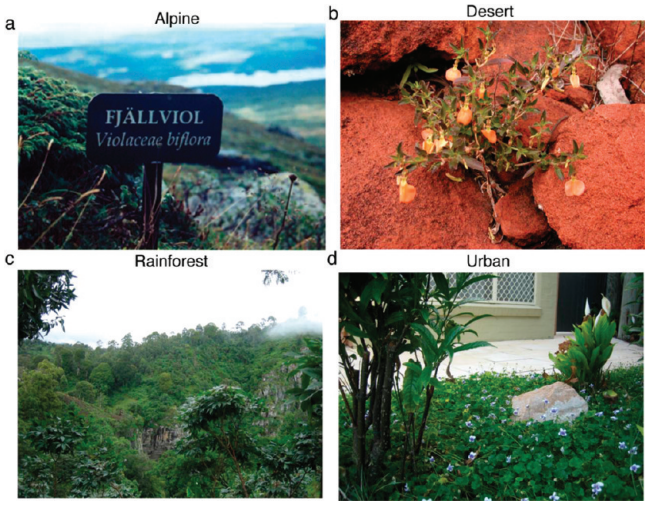


Figure 2. Example habitats of plants found to contain cyclotides. Cyclotides are found in a range of geographies and climatic conditions: (a) Alps in Sweden; (b) desert in the Australian Outback; (c) rainforest in Hawaii; (d) backyard garden in Brisbane, Australia.

examined, it has been predicted that thousands more members of the family await discovery.[17,18] The cyclotides reported so far have a range of bioactivities not limited to the uterotonic activity of kalata B1. They include hemolytic activity,[13] anti-HIV activity,[12,19−23] neurotensin antagonism,[14] antimicrobial activity,[24] cytotoxic activity,[25,26] and antifouling activity.[27] Their natural function appears to be as plant defense agents, judging by their activity in inhibiting the growth and development of insects,[28,29] nematodes,[30−32] and molluscs.[33]

As indicated in Figure 1, cyclotide structures contain six backbone loops between the conserved Cys residues, and different degrees of sequence diversity have been seen so far in the different loops. Loops 1 and 4, which make up part of the embedded ring of the cystine knot, are highly conserved in both size and residue type, whereas the other loops are more variable. For example, so far, 39 different amino acid sequences have been seen in loop 2 and 71 sequences in loop 3. Loop 6 is the site where ligation of the termini of the precursor protein occurs to produce the circular backbone.[28,34] The cystine knot of cyclotides occupies their molecular core, with the consequence that most other amino acid side chains, including hydrophobic residues that would occupy the protein core in most proteins, are exposed on the molecular surface. These residues typically cluster in a surface-exposed hydrophobic patch, which influences the biophysical characteristics of cyclotides and is thought to be involved in eliciting their various bioactivities by promoting membrane binding.[35−39]

Cyclotides fall into two main structural subfamilies, referred to as the Möbius and bracelet subfamiles.[1] Möbius cyclotides contain a *cis*-proline residue in loop 5 that induces a local 180° twist in the peptide bond angle, whereas bracelet cyclotides, the more common of the two, are composed of a ring of *trans* peptide bonds linking their amino acids. Although the presence or absence of the *cis*-proline residue is the defining feature of the two subfamilies, there also tends to be high sequence homologies within subfamilies but less so between them. A third subfamily of cyclotides, the trypsin inhibitor subfamily,[40−43] contains just two members, namely, *Momordica cochinchinensis* trypsin inhibitors I and II (MCoTI-I and MCoTI-II), which are quite different from other cyclotides in sequence and are also referred to as cyclic knottins.[43]

Previous reviews have reported on various aspects of the discovery, sequences, synthesis, mode of action, and potential applications of cyclotides.[20,22,43−55] A recent review also examined their role in plant defense.[56] Here our aim is to provide a comprehensive description of the current understanding of cyclotides from a perspective of interest to natural products chemists. Thus, as well as providing a description of cyclotides, we include text boxes that outline key practical aspects of the isolation, chemical synthesis, and structure characterization of cyclotides. We wish to emphasize the idea that cyclotides can be regarded as a natural combinatorial template; the CCK motif is essentially an ultrastable protein scaffold on which a diverse range of sequences is displayed. The framework is amenable to chemical synthesis, and so there exists the possibility of grafting non-native bioactive sequences onto the framework, making it a valuable tool in protein engineering and drug design applications.[22,44,57,58]

Isolation and Elucidation of Cyclotides. To date, the primary method for discovering and characterizing new cyclotides has been via isolation, purification, and characterization at the peptide, rather than the nucleic acid level, but discovery methods based on nucleic acid sequences are also now being applied.[59−62] Generally, isolating and purifying peptides from plants can be complicated by the propensity of peptides to degrade during exposure to solvents, heat, and enzymes. Fortunately, the CCK motif that defines the cyclotide family engenders them with a number of unique chemicophysical properties, enabling them to resist degradation[63] and distinguishing them from other plant proteins. These unique properties, such as resistance to chemical, thermal, and enzymatic degradation, and other characteristics such as their typical masses of 2.5−4 kDa and late elution times on RP-HPLC, can be exploited when isolating cyclotides from plant material.[18] Methods that utilize these defining characteristics have enabled researchers to extract cyclotides from plants with high yields, even up to 1−2 g of cyclotides per kg of wet plant material.[7,39,64] Box 1 provides a practical description of cyclotide extraction, isolation, and purification processes typically used in our laboratory and describes how to deduce their amino acid sequence.[1,39,65−67]

Box 1. Extracting, Isolating, and Purifying Cyclotides

Step 1: Chemical extraction of peptides from plant material. Harvest plant material and wash with water to remove dirt. If desired, partition the plant into different parts/organs (flowers, leaves, stems, and roots) and weigh. Grind plant material in methanol using a blender or freeze with liquid nitrogen and grind using a mortar and pestle. Immerse the plant material in a dichloromethane−methanol mixture (1:1, v/v) and leave overnight at room temperature (22 °C).

Step 2: Isolating peptides from plant material. Transfer the dichloromethane−methanol mixture to a separation flask and separate the layers after the addition of water. Collect the aqueous layer and concentrate on a rotary evaporator. Load the concentrated aqueous layer onto a C_{18} column and wash with increasing concentrations (20%, 80%, 100%) of solvent B (acetonitrile−water−trifluoroacetic acid, 9:1:0.005, v/v/v) to elute the peptides and remove any impurities. Concentrate the 80% solvent B fraction on a rotary evaporator and lyophilize for LC/MS (step 3) and purification (step 4).

Step 3: LC/MS analysis. Dissolve a sample of the dried and purified extract in 100% solvent A (water−trifluoroacetic acid, 10:0.005, v/v). Inject a 100 μL aliquot onto the LC/MS and elute with a linear gradient from 0% to 80% (acetonitrile−0.1% formic acid, v/v) at 200 μL per min. Obtain mass spectra in positive-ion mode over a range *m*/*z* 500−2000. Although it is possible that some cyclotides might elute before 25% acetonitrile−0.1% formic acid (v/v), the general hydrophobicity of cyclotides, one of the characteristics that expedites their separation from residual plant material, typically means that they elute later than 30% and before 70% acetonitrile−0.1% formic acid (v/v). Identify peaks with cyclotide-like masses and retention times and use as a guide for large-scale purification.

Step 4: Purifying peptide extracts. For preparative RP-HPLC, dissolve the individual fractions acquired during step 2 in 100% solvent A. Filter and load filtrate onto a C_{18} column, elute with a linear gradient from 0% to 80% solvent B, and collect individual peaks. Analyze via mass spectrometry and retain samples containing peptides with appropriate masses (2.5−4 kDa). Repeat purification as necessary by varying column, elution, and flow rates to ensure extracts are at least 95% pure. Concentrate samples on a rotary evaporator and lyophilize for subsequent peptide reduction and sequence determination. To separate peptides with similar retention times under the above conditions, a tetraethylenepentamine buffer has proven useful.[39]

Step 5: Reduction of peptides and MALDI-MS analysis. Add 1 μL of 0.1 M tris(2-carboxyethyl)phosphine (TCEP) to approximately 6 nmol of peptide in 20 μL of 0.1 M ammonium bicarbonate (pH 8.0), and incubate at 65 °C for 10 min. Confirm chemical reduction (addition of 6 Da) by MALDI-TOF-MS,[68] after desalting using Ziptips (Millipore) involving several washing steps, then elution in 10 μL of 80% acetonitrile (v/v) containing 0.5% formic acid in water. Mix the desalted samples in a 1:1 ratio (v/v) with matrix consisting of a saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile containing 0.5% aqueous formic acid.

Step 6: Enzymatic digestion of reduced peptides. Add trypsin, endoGlu-C, or a combination of both to the reduced peptide to give a final peptide-to-enzyme ratio of 50:1 (w/v). Allow the trypsin incubation to proceed for 1 h and the endoGlu-C digestion for 3 h, and for the combined digestion add the trypsin for 1 h followed by endoGlu-C for a further 3 h. Quench the digestions by the addition of an equal volume of 0.5% formic acid(aq) and desalt the samples using Ziptips (Millipore). Store the samples at 4 °C before analysis.

Step 7: Nanospray MS/MS sequencing of peptides. Examine the fragments resulting from the digestion first by MALDI-TOFMS, followed by sequencing by nanospray MS/MS. Examine the MS/MS spectra for each peptide and sequence based on the presence of both b- and y-series of ions (N- and C-terminal fragments). Use amino acid analysis to confirm sequence composition for all peptides. Because MS/MS alone cannot distinguish Ile from Leu, identify and confirm the position of these residues by a chymotrypsin digest, which cleaves preferentially after Leu, using the same conditions as for trypsin digestion, then sequence. Edman sequencing can be used as an alternative to MS/MS sequencing.

Characteristics of Cyclotides. Table 1 presents the sequences and known bioactivities of cyclotides reported to date. Although a cursory glance at the table might suggest extensive variations in the composition and size of cyclotide loops, a closer analysis reveals many sequence homologies that are important for understanding the chemicophysical properties and bioactivities of cyclotides.

The six Cys residues are absolutely conserved in all cyclotides and, in combination with the cyclic backbone, are an integral part of the CCK motif. The cystine knot is formed when the disulfide bonds $Cys^{I}-Cys^{IV}$ and $Cys^{II}-Cys^{V}$ and their interconnecting backbone segments form a ring that is penetrated by the $Cys^{III}-Cys^{VI}$ disulfide bond.[1,69] This knotted disulfide connectivity was first proposed for cyclotides in the mid 1990s[11,70] on the basis of NMR studies. An alternative (and subsequently shown to be incorrect) disulfide connectivity[71] was proposed in 2002, but the knotted connectivity has since been confirmed from chemical and additional NMR studies.[3,4,71] The structures of 16 cyclotides have now been determined using NMR methods, and each has confirmed the knotted disulfide connectivity.[36,39,66,72−81] Recently, the X-ray structure of the cyclotide varv F was reported and confirmed the presence of the cystine knot in this peptide.[82] The solution structure and micelle-bound structure of varv F were reported in the same paper and are essentially identical to the crystal structure.

The formation of the cystine knot in cyclotides in vivo does not depend on the presence of a cyclic backbone.[66] Thus, violacin A, a relatively rare example of a linear cyclotide derivative recently discovered in the plant *Viola odorata*, was shown by NMR to exhibit a cystine knot motif, which locks it into a typical cyclotide fold, despite having a noncyclic backbone. This observation is consistent with the fact that cystine knots are not only found in cyclotides but are also present in a wide range of proteins from insects, plants, and animals, including toxins, inhibitory proteins, and growth factors.[83−86]

Although only the Cys residues are directly involved in maintaining the characteristic CCK fold, several other highly conserved residues are thought to provide additional stability.[3] Throughout known cyclotides, loop 1 is highly conserved; the Glu residue in position 2 of this three-residue loop is almost invariant, being present in 139 of the 140 currently known Möbius and bracelet cyclotides and being replaced by Asp in the other example (kalata B12). Similarly, Gly (or Ala) and Ser (or Thr) residues in positions 1 and 3 of this loop, respectively, are highly conserved. NMR analysis of a range of cyclotides has revealed that the conserved residues in this loop play an important role in stabilizing cyclotide structures through hydrogen bonding with residues from loops 3 and 5.[3]

Loops 2−6 also have highly conserved residues that are thought to contribute to the stability and fold of the CCK motif. These residues include a hydroxy-bearing amino acid in loop 4 that is involved in side chain−side chain hydrogen bonding with loop 6; a hydroxy-bearing residue in loop 3 that acts as a hydrogen bond donor to loop 1; a Gly residue with a positive ϕ-angle in the final position of loop 3; a basic and a Pro residue in the penultimate position in loop 5 of bracelet and Möbius cyclotides respectively; and an Asn (or occasionally aspartic acid) residue at the putative cyclization site[28,34,66,87,88] in loop 6. It is of interest to note that not only are a number of specific residues highly conserved but the backbone and side chain angles are as well.[3] Unlike other cystine knot proteins that show significant sequence and loop size variations,[85,89] the core of cyclotides is structurally conserved.

There are a number of common sequence patterns that exist within the two main subfamilies of cyclotides. For example, loop 5 of many bracelet cyclotides comprises a Lys-Ser-Lys-Val sequence, whereas this loop is more typically Ser-Trp-Pro-Val in Möbius cyclotides. Sequence homologies specific to individual subfamilies were, in the past, thought to reflect regions of importance specific to the stability of members of the subfamily. However, the recent discovery of the "chimeric" cyclotides kalata B8,[74] kalata B9,[90] palicourein,[72] and tricyclon A,[77] which contain both Möbius and bracelet cyclotide features,[23,74] suggests that sequence homologies within the Möbius and bracelet subfamilies might not be as distinct as previously thought.

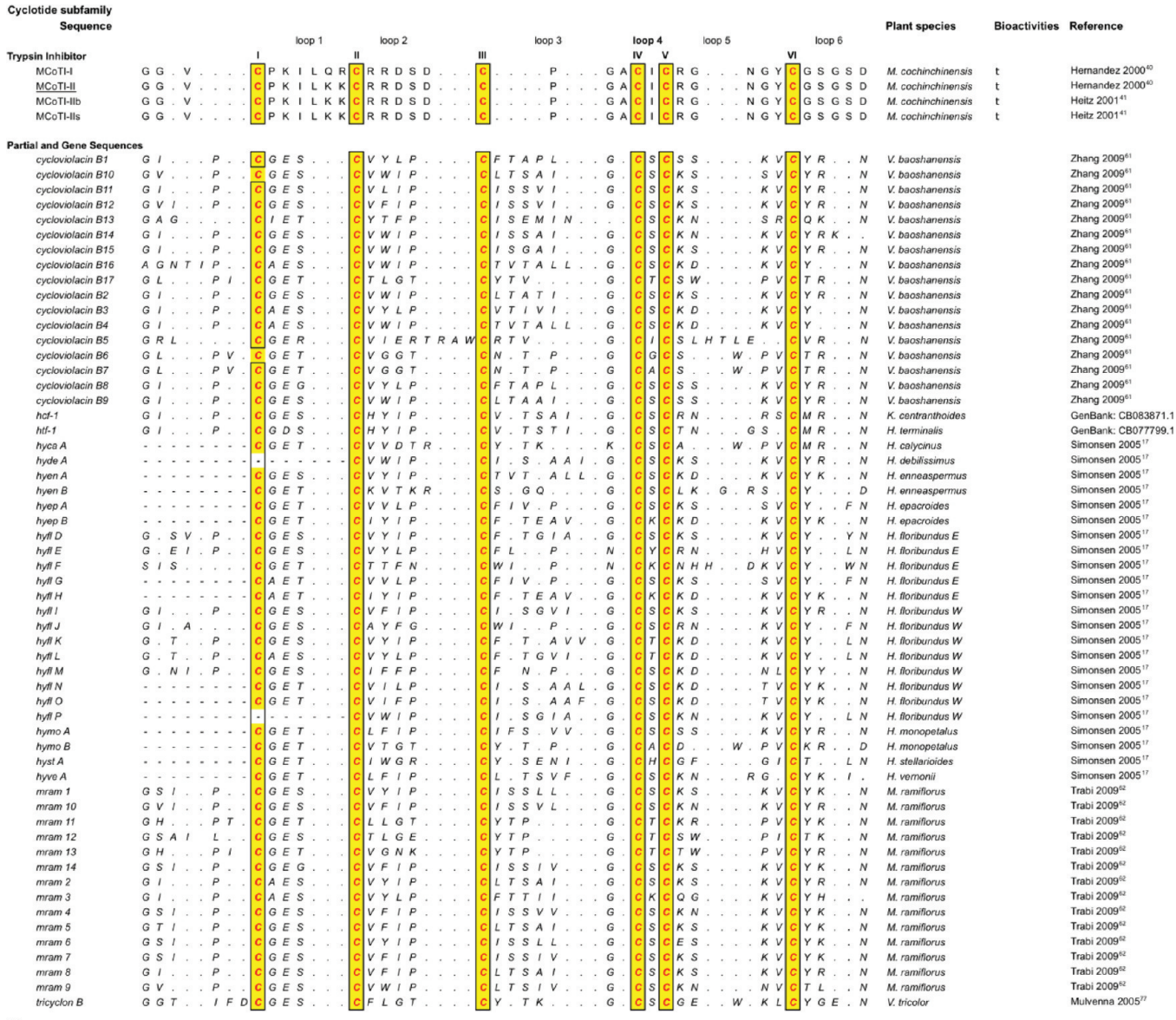
Although cyclotides are resistant to a range of degradative environments, recently, the Trp of loop 5 has been shown to be susceptible to oxidation,[90] yielding hydroxytryptophan and/or oxindolylalanine rings in kalatas B1, B2, and B10 on standing in solution for prolonged periods. It was suggested that exposure to sunlight might have caused this degradation because no degradation was observed in control cyclotides that were kept in the dark. Degradation of these cyclotides resulted in loss of hemolytic activity, suggesting that the Trp in loop 5 plays an important role in this bioactivity.

Three-Dimensional Structures of Cyclotides. Cyclotides are amenable to structure determination with NMR spectroscopy due

Table 1. Sequences, Sources, and Bioactivities of Known Cyclotides and Their Naturally Occurring Derivatives

Cyclotide subfamily / Sequence	loop 1 (I)	loop 2 (II)	loop 3 (III)	loop 4 (IV)	loop 5 (V)	loop 6 (VI)	Plant species	Bioactivities	Reference
Bracelet									
circulin A	GIPCGESCVWIPCISAALGCSCKNKVCYRN						*C. parvifolia*	b, h, v	Gustafson 1994[12]
circulin B	GVIPCGESCVFIPCITSTLLGCSCKNKVCYRN						*C. parvifolia*	b, c, h, v	Gustafson 1994[12]
circulin C	GIPCGESCVFIPCITSVAGCSCKSKVCYRN						*C. parvifolia*	v	Gustafson 2000[19]
circulin D	KIPCGESCVWIPCVTSIFNCKCENKVCYHD						*C. parvifolia*	v	Gustafson 2000[19]
circulin E	KIPCGESCVWIPCLTSVFNCKCENKVCYHD						*C. parvifolia*	v	Gustafson 2000[19]
circulin F	AIPCGESCVWIPCISAAIGCSCKNKVCYR.						*C. parvifolia*	v	Gustafson 2000[19]
cyclopsychotride A	SIPCGESCVFIPCTVTALLGCSCKSKVCYKN						*P. longipes*	b, c, h, n	Witherup 1994[14]
cycloviolacin H1	GIPCGESCVYIPCLTSAIGCSCKSKVCYRN						*V. hederaceae*		Craik 1999[1]
cycloviolacin H2	SAIACGESCVYIPCFI.PGGCSCRNRVCY.LN						*V. hederaceae*		Chen 2005[73]
cycloviolacin H4	GPACAESCVWIPCTVTALLGCSCSNNVCY.N						*V. hederaceae*		Chen 2006[135]
cycloviolacin O1	GIPCAESCVYIPCTVTALLGCSCSNRVCY.N						*V. odorata*		Craik 1999[1]
cycloviolacin O10	GIPCGESCVYIPCLTSAVGCSCKSKVCYRN						*V. odorata*		Craik 1999[1]
cycloviolacin O11	GTLPCGESCVWIPCISAVVGCSCKSKVCYKN						*V. odorata*		Craik 1999[1]
cycloviolacin O13	GIPCGESCVWIPCISAAIGCSCKSKVCYRN						*V. odorata*	h	Ireland 2006[39]
cycloviolacin O17	GIPCGESCVWIPCISAAIGCSCKNKVCYRN						*V. odorata*		Ireland 2006[39]
cycloviolacin O18	GIPCGESCVYIPCTVTALAGCKCKSKVCY.N						*V. odorata*		Ireland 2006[39]
cycloviolacin O19	GTLPCGESCVWIPCISSVVGCSCKSKVCYKD						*V. odorata*		Ireland 2006[39]
cycloviolacin O2	GIPCGESCVWIPCISSAIGCSCKSKVCYRN						*V. odorata*	c, f, h	Craik 1999[1]
cycloviolacin O20	GIPCGESCVWIPCLTSAIGCSCKSKVCYRD						*V. odorata*		Ireland 2006[39]
cycloviolacin O25	DI.FCGETCAFIPCITHVPGTCSCKSKVCY.FN						*V. odorata*		Ireland 2006[39]
cycloviolacin O3	GIPCGESCVWIPCLTSAIGCSCKSKVCYRN						*V. odorata*		Craik 1999[1]
cycloviolacin O4	GIPCGESCVWIPCISSAIGCSCKNKVCYRN						*V. odorata*		Craik 1999[1]
cycloviolacin O5	GT.PCGESCVWIPCISSAVGCSCKNKVCYKN						*V. odorata*		Craik 1999[1]
cycloviolacin O6	GTLPCGESCVWIPCISAAVGCSCKSKVCYKN						*V. odorata*		Craik 1999[1]
cycloviolacin O7	S.IPCGESCVWIPCTITALAGCKCKSKVCY.N						*V. odorata*		Craik 1999[1]
cycloviolacin O8	GTLPCGESCVWIPCISSVVGCSCKSKVCYKN						*V. odorata*		Craik 1999[1]
cycloviolacin O9	GIPCGESCVWIPCLTSAVGCSCKSKVCYRN						*V. odorata*		Craik 1999[1]
cycloviolacin Y1	GGTIFDCGETCFLGTCYTPGCSCGNYGLCYGTN						*V. yedoensis*	h, v	Wang 2008[21]
cycloviolacin Y2	GGTIFDCGESCFLGTCYTAGCSCGNWGLCYGTN						*V. yedoensis*	h, v	Wang 2008[21]
cycloviolacin Y3	GGTIFDCGETCFLGTCYTAGCSCGNWGLCYGTN						*V. yedoensis*	h, v	Wang 2008[21]
cycloviolacin Y4	GVPCGESCVFIPCITGVIGCSCSSNVCY.LN						*V. yedoensis*	h, v	Wang 2008[21]
cycloviolacin Y5	GIPCAESCVWIPCTTALVGCSCSDKVCY.N						*V. yedoensis*	h, v	Wang 2008[21]
cycloviolin A	GVIPCGESCVFIPCISAAIGCSCKNKVCYRN						*L. cymosa*	v	Hallock 2000[130]
cycloviolin B	GTACGESCYVLPCFTVGGCTCTSSQCFKN						*L. cymosa*	v	Hallock 2000[130]
cycloviolin C	GIPCGESCVFIPCLTTVAGCSCKNKVCYRN						*L. cymosa*	v	Hallock 2000[130]
cycloviolin D	GFPCGESCVFIPCISAAIGCSCKNKVCYRN						*L. cymosa*	v	Hallock 2000[130]
hyfl A	SISCGESCVYIPCTVTALVGCTCKDKVCY.LN						*H. floribundus E*		Simonsen 2005[17]
hyfl B	GSPIQCAETCFIGKCYTEELGCTCTAFLCMK.N						*H. floribundus E*		Simonsen 2005[17]
hyfl C	GSPRQCAETCFIGKCYTEELGCTCTAFLCMK.N						*H. floribundus E*		Simonsen 2005[17]
hypa A	GIPCAESCVYIPCTITALLGCSCKNKVCY.N						*H. parviflorus*		Broussalis 2001[65]
kalata B16	GIPCAESCVYIPCTITALLGCKCQDKVCY.D						*O. affinis*		Craik 1999[1]
kalata B17	GIPCAESCVYIPCTITALLGCKCKDQVCY.N						*O. affinis*		Craik 1999[1]
kalata B5	GTPCGESCVYIPCISGVIGCSCSTDKVCY.LN						*O. affinis*		Craik 1999[1]
kalata B8	GSVLNCGETCLLGTCYTTGCTCNKYRVCTK.D						*O. affinis*	v	Daly 2006[74]
kalata B9	GSVFNCGETCVLGTCYTPGCTCNTYRVCTK.D						*O. affinis*		Plan 2007[90]
palicourein	GDPTFCGETCRVIPVCTYSAALGCTCDDRSDGLCKR.N						*P. condensata*	v	Bokesch 2001[127]
tricyclon A	GGTIFDCGESCFLGTCYTKGCSCGEWKLCYGTN						*V. tricolor, V. arvensis*		Mulvenna 2005[77]
vhl-1	SIISCGESCAMISFCFTEVIGCSCKNKVCY.LN						*V. hederaceae*	v	Chen 2005[73]
vhr1	GIPCAESCVWIPCTVTALLGCSCSNKVCY.N						*V. hederaceae*		Trabi 2004[78]
vibi E	GIPCAESCVWIPCVTALIGCGCSNKVCY.N						*V. bicolor*		Herrman 2008[139]
vibi F	GTIPCGESCVFIPCLTSALGCSCKSKVCYKN						*V. bicolor*		Herrman 2008[139]
vibi G	GTFPCGESCVFIPCLTSAIGCSCKSKVCYKN						*V. bicolor*		Herrman 2008[139]
vibi H	GLLPCAESCVYIPCLTTVIGCSCKSKVCYKN						*V. bicolor*		Herrman 2008[139]
vibi I	GIPCGESCVWIPCLTSTVGCSCKSKVCYRN						*V. bicolor*		Herrman 2008[139]
vibi J	GTFPCGESCVWIPCISKVIGCACKSKVCYKN						*V. bicolor*		Herrman 2008[139]
vibi K	GIPCGESCVWIPCLTSAVGCPCKSKVCYRN						*V. bicolor*		Herrman 2008[139]
vico A	GSIPCAESCVYIPCFTGIAGCSCKNKVCY.YN						*V. cotyledon*		Göransson 2003[140]
vico B	GSIPCAESCVYIPCITGIAGCSCKNKVCY.YN						*V. cotyledon*		Göransson 2003[140]
vitri A	GIPCGESCVWIPCITSAIGCSCKSKVCYRN						*V. tricolor*	c	Svangard 2004[25]
Möbius									
cycloviolacin H3	GLPVCGETCFGGTCNTPGCICDPWPVCTRN						*V. hederaceae*		Chen 2005[73]
**cycloviolacin O12	GLPICGETCVGGTCNTPGCSCSWPVCTRN						*V. odorata*		Craik 1999[1]
cycloviolacin O14	GSIPACGESCFKGKCYTPGCSCSKYPLCAKN						*V. odorata*	h	Ireland 2006[39]
cycloviolacin O15	GLVPCGETCFTGKCYTPGCSCSYPICKKN						*V. odorata*	h	Ireland 2006[39]
cycloviolacin O16	GLPCGETCFTGKCYTPGCSCSYPICKKIN						*V. odorata*		Ireland 2006[39]
cycloviolacin O21	GLPVCGETCVTGSCYTPGCTCSWPVCTRN						*V. odorata*		Ireland 2006[39]
cycloviolacin O22	GLPICGETCVGGTCNTPGCTCSWPVCTRN						*V. odorata*		Ireland 2006[39]
cycloviolacin O23	GLPTCGETCFGGTCNTPGCTCDSSWPICTHN						*V. odorata*		Ireland 2006[39]
cycloviolacin O24	GLPTCGETCFGGTCNTPGCTCDPWPVCTHN						*V. odorata*	h	Ireland 2006[39]
kalata B1	GLPVCGETCVGGTCNTPGCTCSWPVCTRN						*O. affinis, V. odorata*	b, h, l, u, v	Craik 1999[1]
kalata B10	GLPTCGETCFGGTCNTPGCSCSSWPICTRD						*O. affinis*		Plan, 2007[90]
kalata B11	GLPVCGETCFGGTCNTPGCSCTDPICTRD						*O. affinis*		Plan, 2007[90]
kalata B12	GSLCGDTCFVLGCNDSSSCSCNYPICVKD						*O. affinis*		Plan, 2007[90]
kalata B13	GLPVCGETCFGGTCNTPGCACDPWPVCTRD						*O. affinis*		Plan, 2007[90]
kalata B14	GLPVCGESCFGGTCNTPGCACDPWPVCTRD						*O. affinis*		Plan, 2007[90]
kalata B15	GLPVCGESCFGGSCYTPGCSCTWPICTRD						*O. affinis*		Plan, 2007[90]
kalata B2	GLPVCGETCFGGTCNTPGCSCTWPICTRD						*O. affinis*	i	Craik 1999[1]
kalata B3	GLPTCGETCFGGTCNTPGCTCDPWPICTRD						*O. affinis*		Craik 1999[1]
kalata B4	GLPVCGETCVGGTCNTPGCTCSWPVCTRD						*O. affinis*		Craik 1999[1]
kalata B6	GLPTCGETCFGGTCNTPGCSCSSWPICTRN						*O. affinis*		Jennings 2001[28]
kalata B7	GLPVCGETCTLGTCYTQGCTCSWPICKRN						*O. affinis*		Jennings 2001[28]
*kalata S	GLPVCGETCVGGTCNTPGCSCSWPVCTRN						*O. affinis*		Craik 1999[1]
*varv A	GLPVCGETCVGGTCNTPGCSCSWPVCTRN						*V. arvensis, V. odorata, V. tricolor*	c, h	Claeson 1998[141]
varv B	GLPVCGETCFGGTCNTPGCSCDPWPMCSRN						*V. arvensis*		Göransson 1999[15]
varv C	GVPICGETCVGGTCNTPGCSCSWPVCTRN						*V. arvensis*		Göransson 1999[15]
varv D	GLPICGETCVGGSCNTPGCSCSWPVCTRN						*V. arvensis*		Göransson 1999[15]
**varv E	GLPICGETCVGGTCNTPGCSCSWPVCTRN						*V. arvensis, V. tricolor*	c	Göransson 1999[15]
varv F	GVPICGETCTLGTCYTAGCSCSWPVCTRN						*V. arvensis*	c	Göransson 1999[15]
varv G	GVPVCGETCFGGTCNTPGCSCDPWPVCSRN						*V. arvensis*		Göransson 1999[15]
varv H	GLPVCGETCFGGTCNTPGCSCETWPVCSRN						*V. arvensis*		Göransson 1999[15]
vhl-2	GLPVCGETCFTGTCYTNGCTCDPWPVCTRN						*V. hederaceae*		Chen 2005[73]
vibi A	GLPVCGETCFGGTCNTPGCSCSYPICTRN						*V. biflora*		Herrman 2008[139]
vibi B	GLPVCGETCFGGTCNTPGCTCSYPICTRN						*V. biflora*		Herrman 2008[139]
vibi C	GLPVCGETCAFGSCYTPGCSCSWPVCTRN						*V. biflora*		Herrman 2008[139]
vibi D	GLPVCGETCFGGRCNTPGCTCSYPICTRN						*V. biflora*		Herrman 2008[139]
violapeptide 1	GLPVCGETCVGGTCNTPGCSCSRPVCTXN						*V. arvensis*	h	Schöpke 1993[13]
vodo M	GAPICGESCFTGKCYTVGQCSCSSWPVCTRN						*V. odorata*		Svangard 2003[67]
vodo N	GLPVCGETCTLGKCYTAGCSCSWPVCYRN						*V. odorata*		Svangard 2003[67]
Natural linear derivatives									
violacin A	SAISCGETCFKFKCYTPGRCSCSYPVCKN						*V. odorata*	h	Ireland 2006[39]

Table 1. Continued



Cyclotide subfamily / Sequence	Plant species	Bioactivities	Reference
Trypsin Inhibitor			
MCoTI-I	*M. cochinchinensis*	t	Hernandez 2000[40]
MCoTI-II	*M. cochinchinensis*	t	Hernandez 2000[40]
MCoTI-IIb	*M. cochinchinensis*	t	Heitz 2001[41]
MCoTI-IIs	*M. cochinchinensis*	t	Heitz 2001[41]
Partial and Gene Sequences			
cycloviolacin B1	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B10	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B11	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B12	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B13	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B14	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B15	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B16	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B17	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B2	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B3	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B4	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B5	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B6	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B7	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B8	*V. baoshanensis*		Zhang 2009[61]
cycloviolacin B9	*V. baoshanensis*		Zhang 2009[61]
hcf-1	*K. centranthoides*		GenBank: CB083871.1
htf-1	*H. terminalis*		GenBank: CB077799.1
hyca A	*H. calycinus*		Simonsen 2005[17]
hyde A	*H. debilissimus*		Simonsen 2005[17]
hyen A	*H. enneaspermus*		Simonsen 2005[17]
hyen B	*H. enneaspermus*		Simonsen 2005[17]
hyep A	*H. epacroides*		Simonsen 2005[17]
hyep B	*H. epacroides*		Simonsen 2005[17]
hyfl D	*H. floribundus E*		Simonsen 2005[17]
hyfl E	*H. floribundus E*		Simonsen 2005[17]
hyfl F	*H. floribundus E*		Simonsen 2005[17]
hyfl G	*H. floribundus E*		Simonsen 2005[17]
hyfl H	*H. floribundus E*		Simonsen 2005[17]
hyfl I	*H. floribundus W*		Simonsen 2005[17]
hyfl J	*H. floribundus W*		Simonsen 2005[17]
hyfl K	*H. floribundus W*		Simonsen 2005[17]
hyfl L	*H. floribundus W*		Simonsen 2005[17]
hyfl M	*H. floribundus W*		Simonsen 2005[17]
hyfl N	*H. floribundus W*		Simonsen 2005[17]
hyfl O	*H. floribundus W*		Simonsen 2005[17]
hyfl P	*H. floribundus W*		Simonsen 2005[17]
hymo A	*H. monopetalus*		Simonsen 2005[17]
hymo B	*H. monopetalus*		Simonsen 2005[17]
hyst A	*H. stellarioides*		Simonsen 2005[17]
hyve A	*H. vernonii*		Simonsen 2005[17]
mram 1	*M. ramiflorus*		Trabi 2009[62]
mram 10	*M. ramiflorus*		Trabi 2009[62]
mram 11	*M. ramiflorus*		Trabi 2009[62]
mram 12	*M. ramiflorus*		Trabi 2009[62]
mram 13	*M. ramiflorus*		Trabi 2009[62]
mram 14	*M. ramiflorus*		Trabi 2009[62]
mram 2	*M. ramiflorus*		Trabi 2009[62]
mram 3	*M. ramiflorus*		Trabi 2009[62]
mram 4	*M. ramiflorus*		Trabi 2009[62]
mram 5	*M. ramiflorus*		Trabi 2009[62]
mram 6	*M. ramiflorus*		Trabi 2009[62]
mram 7	*M. ramiflorus*		Trabi 2009[62]
mram 8	*M. ramiflorus*		Trabi 2009[62]
mram 9	*M. ramiflorus*		Trabi 2009[62]
tricyclon B	*V. tricolor*		Mulvenna 2005[27]

Notes:

1 Bioactivity key: b = anti-bacterial, c = cytotoxicity/anti-tumor, f = marine antifouling, h = hemolytic, i = insecticidal, n = neurotensin antagonist, t = trypsin inhibitor, v = anti-HIV, u = uterotonic

2 Partial sequences identified by PCR as partial mRNA transripts. "-" Denotes unknown residue.

3 Cyclotides denotes the structure has been resolved. PDB codes are; circulin A - 1BH4; circulin B - 2ERI; cycloviolacin O1 - 1NBJ; cycloviolacin O14 - 2GJ0; kalata B1 - 1NB1/1ZNU/1JJZ/1K48; kalata B2 - 1PT4; kalata B7- 2jwm; kalata B8 - 2B38; MCoTI-II - 1IB9/1HA9; palicourein - 1R1F; tricyclon A - 1YP8; varv peptide F- 2k7g; vhl-1 - 1ZA8; vhr1 - 1VB8; violacin A - 2FQA;

4 Full species names are as follows. *Chassalia parvifolia, Hybanthus calycinus, Hybanthus debilissimus, Hybanthus enneaspermus, Hybanthus epacroides, Hybanthus floribundus E, Hybanthus floribundus W, Hybanthus monopetalus, Hybanthus parviflorus, Hybanthus stellarioides, Hybanthus vernonii, Kadua centranthoides, Kadua terminalis, Leonia cymosa, Momordica cochinchinensis, Oldenlandia affinis, Palicourea condensata, Psychotria longipes, Viola arvensis, Viola biflora, Viola baoshanensis, Viola cotyledon, Viola hederaceae, Viola odorata, Viola tricolor*

5 * and ** indicates cyclotides with identical sequences but named differently

8 hcf-1 and htf-1 are hypothetical cyclotides derived from precursor sequence assuming similar processing points to other cyclotides

9 hyfl D-M sequences determined using a forward primer annealing to an identified conserved element within the ER signal of previously characterized cyclotide precursors from the Violaceae family.

to their small size, solubility, stability, and well-defined β-sheet structures. The structurally characterized cyclotides so far are the bracelet subfamily members, circulin A (PDB ID: 1BH4), circulin B (PDB ID: 2ERI), vhl-1 (PDB ID: 1ZA8), vhr1 (PDB ID: 1VB8), cycloviolacin O1 (PDB ID: 1NBJ, 1DF6), cycloviolacin O2 (PDB ID: 2KCG, 2KNM), kalata B8 (PDB ID: 2B38), palicourein (PDB ID: 1R1F), and tricyclon A (PDB ID: 1YP8); the Möbius subfamily members, kalata B1 (PDB ID: 1NB1/1ZNU/1JJZ/1K48), kalata B2 (PDB ID: 1PT4, 2KCH), kalata B7 (PDB ID: 2JWM), varv F (PDB ID: 2K7G, 3E4H), and cycloviolacin O14 (PDB ID: 2GJ0); the trypsin inhibitor cyclotide MCoTI-II (PDB ID: 1IB9); and the linear derivative violacin A (PDB ID: 2FQA). Box 2 describes the steps involved in determining the three-dimensional structures of cyclotides using NMR spectroscopy.

Box 2. Structure Determination of Cyclotides with NMR Spectroscopy

Step 1: Sample preparation. For NMR analysis, dissolve the cyclotide in either water−deuterated water (9:1, v/v) or water−deuterated acetonitrile−deuterated water (7:2:1, v/v/v) depending on solubility. A 1 mM solution is sufficient for high-quality spectra, but lower concentrations can be used with NMR spectrometers equipped with high-sensitivity cryo-probes. A range of pH values can be used (~3−7), and pH-induced conformational changes (if present) can be assessed by measuring changes in chemical shifts with pH.

Step 2: Record NMR spectra. Record a ^{1}H one-dimensional NMR spectrum to check on sample quality. The β-sheet structure

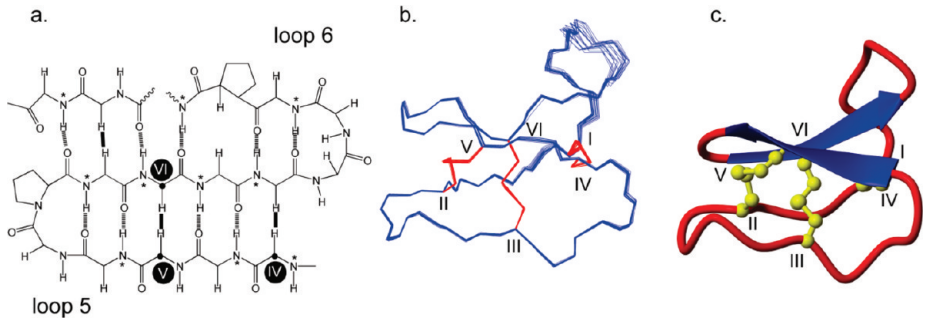


Figure 3. Structural representations of kalata B1.[3] (a) Schematic representation of the secondary structure of kalata B1. The amide protons marked by an asterisk are slowly exchanging, indicative of hydrogen bonding. (b) Superposition of 20 lowest energy NMR structures demonstrating the high precision of the structure determination. (c) Ribbon representation showing the overall fold. The β-strands are marked with arrows. Note that although kalata B1 contains a triple-stranded β-sheet, one of the strands is typically distorted [upper strand in panel (a)] and is often not formally identified by molecular drawing programs. Cysteines are labeled with Roman numerals.

of cyclotides results in well-dispersed NMR signals. Record 2D TOCSY and NOESY spectra to allow sequence-specific assignment of individual residues. Use mixing times of approximately 80 and 200 ms for the TOCSY and NOESY spectra, respectively. Spectra are usually recorded at 290 and 298 K, because different temperatures can be useful where resonances are overlapped in the amide region. After sequence-specific assignment, distance restraints can be derived from NOESY spectra. Angle restraints are derived for ϕ angles by measuring αN couplings from either the one-dimensional spectrum or a DQFCOSY spectrum. χ1 angle restraints are derived from analysis of NOESY peak intensities and measuring $\alpha\beta$ couplings from an ECOSY spectrum. Analysis of the NOESY peak intensities can be complicated by spin-diffusion, and thus a mixing time of 100 ms is generally used for this analysis to minimize this complication. Determine the slowly exchanging amide protons by dissolving freeze-dried cyclotide in deuterated water and monitoring the exchange of the amide protons with the solvent by recording one-dimensional and TOCSY spectra over time.

Step 3: Structure calculations. The distance and angle restraints are used to calculate three-dimensional structures consistent with these experimentally derived data. One program for calculating these structures is CNS, and the methods used for cyclotides have been summarized previously.[3] Analyze the initial structures for the presence of hydrogen bonds. If the hydrogen bond donor is slowly exchanging in the deuterated water exchange experiments, hydrogen bond restraints can be added to the structure calculations. Calculating structures is an iterative process that improves as additional restraints, such as these hydrogen-bonding restraints, are successively added and as any errors in the initial restraints sets are discovered and eliminated. Analyze the structures for violations of the experimental data. If there are violations greater than 0.3 Å for the distance restraints or 3° for the angle restraints, the restraints should be checked for errors and the structures recalculated. Analyze the final structures in terms of their Ramachandran and energetic statistics. These statistics provide information on the precision and quality of the structures, and examples are given in the references.[41,74,75]

As noted earlier in this review, the cystine knot and circular backbone are the most characteristic features of cyclotides. However, β-sheet secondary structure is typically associated with the cystine knot motif in other proteins,[85] and this is also the case with the cyclotides. In cyclotides, the antiparallel β-sheet is composed of a β-hairpin associated with a third distorted β-strand, as illustrated in Figure 3.[46] Minor variations in other parts of the structure are associated with the different subfamilies of cyclotides. For instance, bracelet cyclotides generally have a short helical region

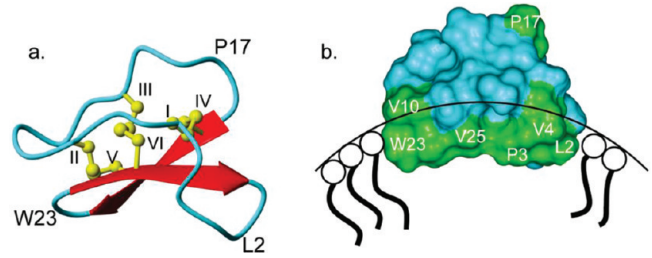


Figure 4. Representation of kalata B1 and its interaction with membranes. (a) Orientation of the backbone of kalata B1 corresponding to the space-filling model (b). The hydrophobic patch of kalata B1 (hydrophobic residues labeled in green) is responsible for interactions with membranes. Selected residues are labeled. Figure drawn based on data presented by Shenkarev et al.[37]

in loop 3, in contrast to Möbius cyclotides, which have no helical structure in this loop. MCoTI-II has a structure very similar to homologous linear squash trypsin inhibitors, with a single turn of a helix in loop 2.[41,42] Overall, the cyclotide framework can accommodate different types of charge and hydrophobic residue distributions while conserving the CCK scaffold.

In general, cyclotide structures are very well-defined, as shown for kalata B1 in Figure 3b. However, there are examples where individual loops are disordered, presumably due to flexibility. MCoTI-II[41,42] and kalata B8[74] display loops with structural disorder, albeit in different loops. For MCoTI-II, loop 6 is disordered, whereas loop 1 is disordered in kalata B8. Interestingly, the disorder in both of these cyclotides appears to be related to isomerization around an Asp-Gly sequence.

A surface-exposed patch of hydrophobic residues is another characteristic feature of Möbius and bracelet cyclotides. Apart from assisting in their isolation by causing relatively long retention times on HPLC, the hydrophobic nature of cyclotides is thought to have a role in their bioactivities and, in particular, their membrane binding.[81,91,92] For kalata B1, this surface-exposed hydrophobic patch is thought to associate with membrane surfaces, as schematically represented in Figure 4.[37]

Nomenclature of Cyclotides. Several nomenclature schemes have been used to describe cyclotides, including stylized names (e.g., circulins;[12] tricyclon A[77]), species- and tissue-specific names (vhl-1 for Viola hederaceae-1[73]), and more systematic approaches based on genus and species names.[65] Settling on a suitable naming system for cyclotides remains difficult given the occurrence of individual cyclotides in multiple plant species, high sequence homologies between many cyclotides, the large potential size of the cyclotide family, and the existence of naturally occurring and synthetic derivatives.

To illustrate some of the potential difficulties with nomenclature, in 1999 two groups[1,15] simultaneously reported the discovery of identical cyclotides from three different plant species, *O. affinis*, *V. odorata*, and *V. arvensis*. We reported the discovery of kalata S from *O. affinis*[1] (with the name derived by analogy with the original name proposed for kalata B1[7]) and cycloviolacin O12 (*V. odorata*), while Göransson and co-workers[15] reported the discovery of two cyclotides extracted from *V. arvensis*, varv A and varv E. From these publications, it became apparent that peptides kalata S and varv A and peptides cycloviolacin O12 and varv E were, respectively, identical despite being isolated from different plant species. Furthermore, the earlier reported sequence for violapeptide I,[13] for which the loop 5 and 6 regions were uncertain due to an ambiguity in the sequencing data used in the original study, is probably identical to the sequence reported for varv A.

The most recent suggestion for a naming system was in 2001, when Broussalis and co-workers[65] proposed that for newly discovered cyclotides a name should be constructed as an indicative and pronounceable acronym of the Latin binomial of the plant from which the cyclotide was first isolated, followed by a letter indicating the order of appearance. Since then, alternative naming systems have been used, the number of cyclotides has more than doubled, the incidence of individual cyclotides in multiple plants species has increased, and in species such as *V. odorata* the number of cyclotides discovered has passed the number of letters available in the English alphabet. Furthermore, the emergence of the chimera cyclotides, cyclotides that have both bracelet and Möbius features, has eliminated any opportunity of naming specific to cyclotide subfamily.

With the number of novel cyclotide discoveries expected to increase dramatically as gene- and peptide-sequencing tools develop over the coming years, it is becoming increasingly clear that a systematic naming system is required. Whatever system is chosen, it must be logical and non-species-specific to reduce the opportunity for double naming and to accommodate data analysis and mining. The recently established database Cybase,[16,93] which holds all published cyclotide sequences with related bioactivities, structures, references, and other related links, potentially could be expanded to manage naming.

Chemical and Biological Synthesis of Cyclotides. Isolating low-abundance cyclotides from host plants often provides insufficient material to permit a complete structural and bioactivity analysis. Fortunately, the fact that cyclotides are only ~30 amino acids in length makes them amenable to chemical synthesis using solid-phase methodologies.[94] In addition, chemical synthesis can be used to probe the structure/activity properties of cyclotides by making artificial derivatives (point mutants, acyclic permutants, and chimeras) and also opens the possibility of developing engineered cyclotides as molecular scaffolds for drug design.[95]

The most widely used methodology for making cyclotides utilizes solid-phase peptide synthesis to assemble the peptide chain, followed by backbone cyclization via an intramolecular native chemical ligation reaction (see Box 3 and Figure 5).[96−103] The ligation reaction requires the linear peptide precursor to have both an N-terminal Cys and a C-terminal thioester. The need for a C-terminal thioester for cyclization means that currently the use of BOC chemistry is the preferred method, as the thioester can be incorporated during chain assembly; however, recently, two FMOC synthetic strategies for the production of cyclotides have been reported. The trypsin inhibitor cyclotide MCoTI-II was synthesized using a sulfonamide-based linker to generate the required thioester for the cyclization step,[104] and the synthesis of cycloviolacin O2 was achieved by producing the thioester in solution after cleavage of the peptide chain from the solid support.[101] It should be noted that the production of thioesters in solution may give rise to epimerization at the C-terminal, but this can be avoided by using Gly as the C-terminal residue.

Box 3. Cyclotide Chemical Synthesis Methodology

To date, the predominant method for the solid-phase peptide synthesis of cyclotides utilizes BOC in situ neutralization chemistry,[112] which is therefore the focus of this protocol. Recently, the synthesis of cyclotides using FMOC chemistry has been described. These studies included the synthesis of MCoTI-II utilizing a sulfonamide-based linker[101,104] and cycloviolacin O2 via the generation of the required thioester after peptide cleavage.[96]

Step 1: Design the sequence of the linear precursor based on the desired cyclotide sequence. The intramolecular native chemical ligation reaction requires an N-terminal Cys residue, leading to the potential of six possible linear precursors. It is best to select a precursor where the C-terminal residue is sterically small, as this has been shown to be kinetically more favorable.[113] For example, when synthesizing kalata B1 and related molecules, we design the break in the peptide chain at the Gly-Cys sequence at the end of loop 3, producing a linear molecule with an N-terminal Cys residue and a C-terminal Gly. Although this is the preferred linear precursor, we also routinely produce cyclotides from linear precursors generated by breaking the peptide chain at any of the other five Cys residues.

Step 2: To synthesize peptides with a C-terminal thioester for native chemical ligation, we generally begin with PAM resin preloaded with either phenylalanine or Gly. However, in principle, any preloaded amino acid could be used.

Step 3: Generate the thioester linker by coupling *S*-trityl-β-mercaptopropionic acid using standard amino acid coupling conditions.[112] Remove the trityl group by treating the resin with triisopropylsilane−water−trifluoroacetic acid (9.6:0.2:0.2, v/v/v) (100 mL of this solution per 0.5−1 mmol of resin, ~10 × 1 min treatments).

Step 4: Couple the C-terminal amino acid of the peptide to the free thiol using standard coupling procedures.[112] Continue and complete the remainder of the sequence.

Step 5: To cleave the peptide chain from the resin, use hydrofluoric acid with *p*-cresol as a scavenger (HF−*p*-cresol, 9:1), precipitate with diethyl ether, and collect by filtration. Dissolve the peptide in an aqueous acetonitrile buffer (e.g., 50% acetonitrile in H_2O with 0.5% trifluoroacetic acid) and freeze-dry.

Step 6: Purify the crude peptide mixture from the HF cleavage by RP-HPLC. Pool and freeze-dry the fractions containing the desired product (as determined by ESIMS analysis) for peptide cyclization and oxidative folding.

Step 7: Cyclization and disulfide formation is generally performed in a one-pot reaction. The primary buffer requirement for cyclization and oxidation is that the pH is basic (~8.0). Typically, 0.1 M ammonium bicarbonate (pH 8.0−8.5) is used either on its own or with additional additives such as reduced and oxidized glutathione and 2-propanol. An initial screen is performed on a 100 μg scale (in 100 μL of buffer) with a variety of conditions, for example:

a. 0.1 M ammonium bicarbonate (pH 8)
b. 25% 2-propanol in 0.1 M ammonium bicarbonate
c. 50% 2-propanol in 0.1 M ammonium bicarbonate
d. Condition (a) plus 2 mM reduced glutathione and 0.4 mM oxidized glutathione
e. Condition (b) plus 2 mM reduced glutathione and 0.4 mM oxidized glutathione
f. Condition (c) plus 2 mM reduced glutathione and 0.4 mM oxidized glutathione

Step 8: Leave the reaction mixture overnight at room temperature and then purify and analyze by RP-HPLC and MS or LC-MS (see Box 1). A major sharp and late-eluting peak usually indicates efficient folding of the cyclotide, with misfolded peptides typically eluting earlier than correctly folded peptides.

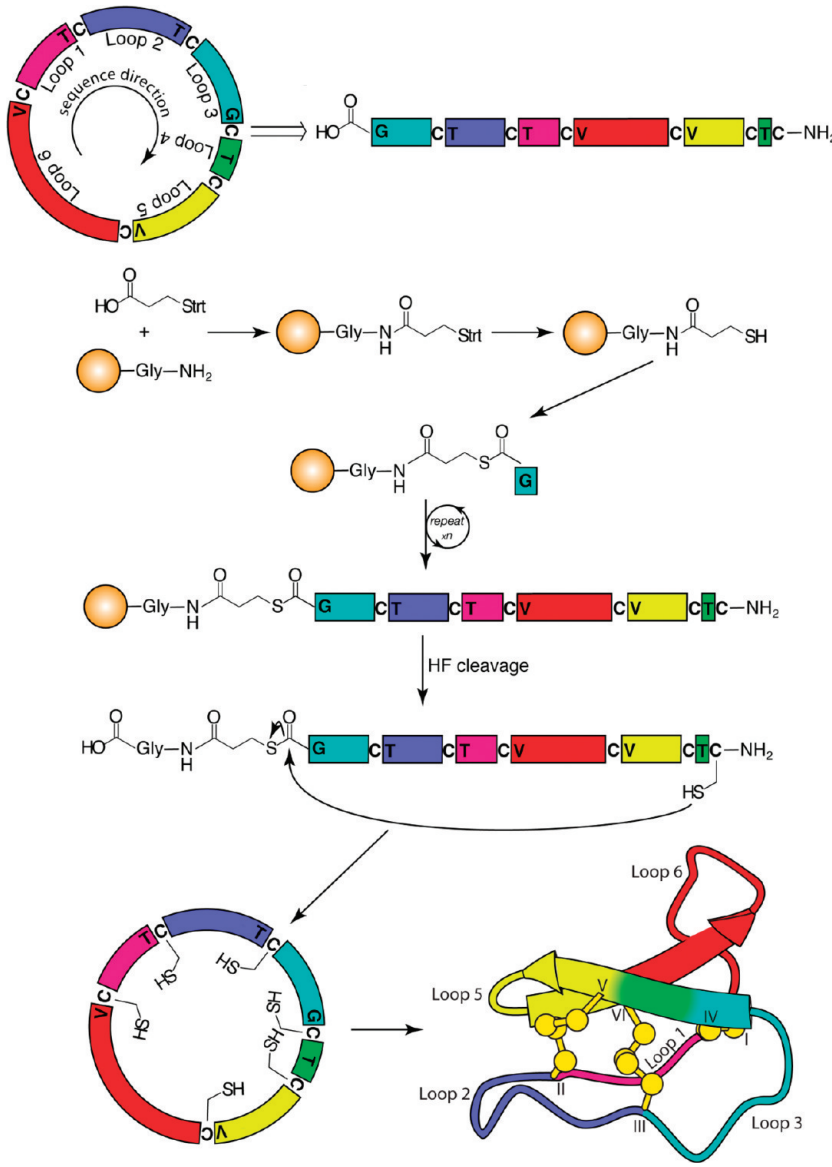


Figure 5. Cyclotide chemical synthesis methodology. Cyclotides are synthesized using a solid-phase peptide synthesis approach to assemble the peptide chain followed by backbone cyclization via an intramolecular native chemical ligation reaction. The upper panel shows the deconvolution of a cyclotide structure into a corresponding linear peptide. Reflecting the order of synthesis, the sequence is drawn from the C-terminus to the N-terminus. Selected residues at one end of each of the backbone loops are marked, along with the Cys residues that demarcate the loop. The lower panels show successive steps in the synthesis process, eventually leading to thioester-mediated ligation and folding of the assembled cyclotide.

Step 9: Once a suitable cyclization/oxidation condition has been established, the reaction can be scaled-up by maintaining the total peptide concentration at approximately 0.3 to 0.5 mg/mL.

After cleavage of the peptide from the solid-phase resin, cyclization of the linear precursor and formation of the disulfide bonds are performed in one pot. It has been proposed that the intramolecular ligation reaction occurs via a thia-zip mechanism followed by a final S to N-acyl shift to generate the native amide bond between the termini.[105] As the ligation reaction is performed at basic pH, the disulfide bonds are also formed during the reaction.

A problem that can be encountered during the chemical synthesis of cyclotides is formation of non-native disulfide bonds (step 8, Box 3). Although the buffer used permits the shuffling of disulfide bonds, it is common for non-native disulfide bonds to form preferentially, particularly for bracelet cyclotides, thereby reducing the yield of the correctly folded cyclotide. In an effort to understand this process and improve the conditions used for disulfide bond formation, recent studies have investigated disulfide folding pathways and intermediates. It was shown that for the oxidative folding pathways of both MCoTI-II[106] and the related cystine knot protein EETI-II,[107] the main folding intermediate having two native disulfides is a direct precursor of the native protein. For both MCoTI-II and EETI-II this intermediate is detected in both the oxidative folding and reductive unfolding pathways. In contrast, for kalata B1, the same intermediate is not a direct precursor of the native protein.[108] Although the intermediate contains a native-like structure with two native disulfide bonds, it does not accumulate during the reductive unfolding process. Despite the topological and structural similarities of this intermediate with the native kalata B1, it is a kinetic trap that converts to the correctly folded kalata B1 over time. In the synthesis of cycloviolacin O2, a bracelet cyclotide, a three non-native disulfide species was found to form preferentially in the typical buffers used for folding cyclotides. After extensive screening of folding conditions, a buffer system containing shuffling agents, an organic cosolvent, and a detergent was found, which

resulted in the formation of the native peptide.[101] It appears that different cyclotides fold via different folding pathways, with the potential for non-native intermediates to act as inhibitors along the folding pathway; thus additional studies are required to understand all of the factors involved. A recent systematic study of hybrids of bracelet and Möbius cyclotides was helpful in defining specific residues that assist in folding.[109]

The presence of six Cys residues within a cyclotide sequence means there are six possible linear precursors that could produce an identical cyclotide on cyclization. This redundancy is advantageous because it can be utilized in a combinatorial approach to produce a series of related cyclotides.[58,110] For example, if a series of cyclotides were to be made that were identical except for differences in loop 5, then the peptide chain can be "broken" on the N-terminal side of the Cys residue preceding this loop. The majority of the sequence can then be assembled using solid-phase synthesis, and when loop 5 is reached, the resin can be split and each analogue completed individually. In practice the different linear precursors of kalata B1 cyclize with different efficiencies,[111] and this should be taken into account in designing syntheses.

The development of new synthetic methodologies for producing cyclotides, in combination with the characterization of new naturally occurring variants, will assist in further understanding structure−activity relationships of cyclotides. In addition, efficient methods for the production of the cyclotides will be required for their development as potential drug leads, and chemical synthesis provides one way of achieving this large-scale synthesis.

Another method for synthetically producing cyclotides based on an intramolecular version of native chemical ligation using chemically engineered inteins for the bacterial expression of cyclotides has been reported recently.[111] Inteins are internal protein domains in precursor proteins that are post-translationally processed to join flanking external domains, termed exteins, via a native peptide bond.[114] The use of inteins in bacterial expression of cyclotides is necessary because although bacteria contain the machinery for the generation of certain classes of naturally occurring circular proteins, they do not naturally produce cyclotide-like proteins. This method can be used to generate combinatorial libraries and might become the method of choice for drug screening studies, as it is amenable to the production of a large number of cyclotide mutants.[115−118] In yet another approach, Leatherbarrow and colleagues recently reported a chemicoenzymatic method for the synthesis of MCoTI-II and engineered variants.[119,120] This approach uses solid-phase peptide synthesis to make a linear precursor that is subsequently cyclized in vitro using trypsin. Precedent for the approach of using a proteolytic enzyme such as trypsin as a peptide ligase was earlier established using the cyclic sunflower trypsin inhibitor SFTI-1 as a model system.[121] Cultured plant cells also show great promise as production factories for cyclotides.[122,123] As far as the last two methods are concerned, the chemicoenzymatic approach, which uses an enzyme to achieve cyclization, can be considered as a mimic in an in vitro system of the natural biosynthesis pathway, and the use of cultured cells takes this mimicry a stage closer to in planta biosynthesis, which is described below.

Biosynthesis of Cyclotides. Plants are a rich source of small cyclic peptides,[124] with the majority of earlier known cyclic peptides produced via nonribosomal biosynthetic pathways. In contrast, cyclotides are gene-encoded products generated via processing of a ribosomally synthesized precursor protein.[17,28,34,61,62,77] Figure 6 illustrates the generic configuration of cyclotide precursor proteins, which comprise an endoplasmic reticulum (ER) signal sequence, a Pro-region, a highly conserved region known as the N-terminal repeat (NTR), the mature cyclotide domain, and finally, a short C-terminal tail. The combined NTR−cyclotide domain contains either one cyclotide sequence, as in the case of the Oak1 precursor, or multiple copies separated by additional NTR sequences, as seen for the Oak2 and Oak4 precursors (see Figure 6 legend for explanation of Oak nomenclature). In precursor proteins containing multiple cyclotide domains, either the sequences can be identical, as is the case for Oak4, which encodes three copies of kalata B2, or they can be different, as in Oak2, which contains sequences corresponding to kalatas B3 and B6.[28]

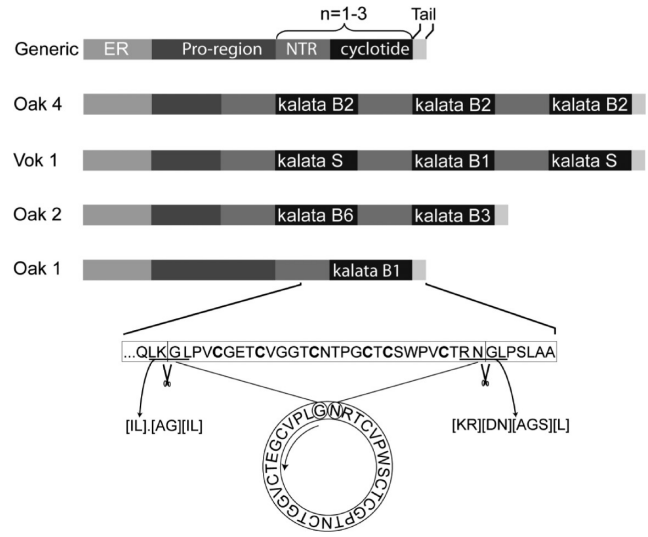


Figure 6. Representation of cyclotide precursor structures. The generic precursor contains an ER signal, a Pro-region, an NTR (N-terminal repeat), and a cyclotide sequence domain, followed by a short tail. Individual cyclotide genes encode between one and three repeats of the NTR and cyclotide domain to form multiple cyclotides from a single precursor. The processing is highlighted for kalata B1, where the upstream and downstream GLP sequences mark processing sites. The consensus sequences found across multiple cyclotide precursors at each processing site are shown in square brackets, with each bracket representing the range of residues found with two amino acids on either side of the processing sites. The "*Oak*" nomenclature introduced by Jennings et al.[28] for cyclotide genes refers to *O. affinis* kalata *n*, where "*n*" is the order of discovery of the gene.

Within the generic structure of the precursor protein there are several highly conserved motifs that are thought to be important for the processing of mature cyclotides. Although the NTR sequence is highly conserved in precursor proteins within a species, significant variation is observed between species. Despite this variation, NTR sequences are predicted to share a common helical structure,[34,77] as shown by NMR studies of the isolated NTR sequences from Oak2 (from *O. affinis*) and Vok1 (from *V. odorata*).[34] It has been suggested that the amphipathic helical nature of the NTR might assist in directing the correct folding of the cyclotide domain via an interaction with the hydrophobic surface of the correctly folded product.[34] It must be emphasized though that at this stage this is only speculation and the molecular details of the folding process have not yet been elucidated.

All cyclotide precursor proteins contain a short C-terminal sequence containing several hydrophobic residues. This motif has been referred to as the C-terminal region (CTR) or the C-terminal propeptide (CTPP). Although the role of this motif is unclear, vacuole targeting and a recognition site for the enzyme(s) involved in processing are possibilities.[49]

The biosynthetic pathway that leads from the precursor protein to the mature cyclotide is still not fully understood, although some insights have been obtained by examination of the conserved residues around the putative processing sites (Figure 6). At the C-terminal portion of the cyclotide sequence the processing site is thought to be at the C-terminal side of a highly conserved Asn (or Asp) residue: the possible target of an asparaginyl endoproteinase (AEP).[28,34,49] AEPs are widespread in plants, and there is precedent for their involvement in the maturation of a circularly permuted

protein; specifically, a plant AEP has been shown to facilitate cleavage and transpeptidation reactions in the formation of concanavalin A in *Canavalia ensiformis* (Jack bean).[125] In addition to the conserved Asn residue, there are several other residues flanking the C-terminal processing site that are conserved. Figure 6 shows a common sequence motif consisting of a positively charged residue preceding the conserved Asn, an amino acid with a small side chain at the position after the Asn, followed by a conserved Leu. At the predicted N-terminal processing site there is a conserved motif comprising two hydrophobic residues (Leu or Ile) flanking the cleavage site and a Gly (or Ala) as the C-terminal residue of the cleaved peptide bond. These conserved motifs are suggestive of recognition sites for enzymes involved in cleaving the mature cyclotide domain from the precursor protein and might also have a role in the backbone cyclization reaction. Two recent studies involving the use of transgenic plants transiently expressing cyclotide precursor proteins have confirmed the vital role of the Asn residue and the involvement of an AEP.[87,88]

A recent study examined the oxidative folding of cyclotides and found that the process is aided by a plant protein disulfide isomerase (PDI).[126] OaPDI (a PDI found in *O. affinis*) and the Oak1 precursor were shown to coexpress on a transcriptional level and to have a strong biomolecular interaction in vitro, which together suggested an in vivo interaction of both proteins. OaPDI also significantly enhanced the yield of correctly disulfide-bonded species in vitro under physiological conditions. It was thus concluded that the interaction between OaPDI and the Oak1 precursor was a potentially important step in the production of correctly folded cyclotides.

Clearly, the further elucidation of the biosynthetic mechanism remains one of the most challenging and important areas in the field of cyclotide research. If the biosynthetic pathway and the factors affecting associated processes can be understood, then potential applications of cyclotides can be expanded. For example, engineered cyclotide drug leads could be produced in plants, which could significantly lower manufacturing costs compared with traditional peptide synthesis. In addition, the natural insecticidal properties of cyclotides could be harnessed by producing transgenic cyclotide-producing crop cultivars.[56]

Bioactivities of Cyclotides. As mentioned in the Introduction, naturally occurring cyclotides have a range of biological activities, including uterotonic,[7] antimicrobial,[24] anti-HIV,[12,19,21,23,73,74,127−130] antitumor,[25,26] hemolytic,[13,14,24,39,66] insecticidal,[28,75] neurotensin antagonism,[14] antifouling,[27] molluscicidal,[33] and anthelminthic activities.[30−32] Table 2 summarizes the potency data for these activities for a range of cyclotides. It is probable that insecticidal activity is the native function of cyclotides in plants.[28,29,75] Cyclotides have potential applications in both the pharmaceutical and agricultural industries due to these activities being present in a very stable peptide framework; consequently there is much interest in determining their mechanisms of action.[35,36,110]

It has been proposed that membrane interactions might be involved in the various biological activities of cyclotides. Supporting this proposal, it has been shown, through NMR[36,37] and surface plasmon resonance experiments,[38] that cyclotides do indeed interact with membranes. Furthermore, electron microscopy studies of the midguts of cyclotide-fed lepidopteran larva have clearly demonstrated disruption of the gut, confirming membrane binding as a mode of action in target insects.[29] Simonsen et al.[110] proposed a mechanism for membrane binding that involves self-association of cyclotides, a phenomenon that, earlier, had been shown to occur for kalata B2 in the solution state.[131] Electrophysiological studies[91] and a Lys-scan of kalata B1 also support this suggestion.[92]

Although the precise details of how membrane binding is involved in the mechanism of action of cyclotides remain unknown, certain aspects of the structural features of cyclotides that are important for activity are beginning to emerge. Anti-HIV activity is the most extensively studied activity, and analysis of acyclic permutants of kalata B1 (i.e., linear analogues synthesized with openings in the various loops in the cyclotide backbone) has shown that the dynamics of the backbone might have a role.[129,132] Acyclic permutants with breaks in any of loop 2, 3, 4, or 5 did not show any anti-HIV activity, despite the fact that they were structurally similar to the native peptide. Measurements of amide exchange rates of acyclic permutants showed them to be more flexible than the parent cyclotide.[129,133,134] Regardless of whether these acyclic permutants lose anti-HIV activity due to decreased stability after linearization of the peptide backbone or due to an intrinsic effect of flexibility, these studies show that a cyclic backbone is critical for anti-HIV activity.

Initially it appeared that trends in the anti-HIV activity of cyclotides had parallels with trends in hemolytic activity; that is, cyclotides with high anti-HIV activity tended to have high hemolytic activity. However, this is not always the case. For instance, one of the most hydrophilic cyclotides, kalata B8, is devoid of hemolytic activity but still has anti-HIV activity.[74] Furthermore, the presence of a CCK motif alone is not sufficient to produce anti-HIV or hemolytic activity. Tricyclon A, isolated from *V. tricolor*, has neither hemolytic nor anti-HIV activity despite maintaining the characteristic CCK structural motif of the cyclotides. The hydrophobic patch on the surface of cyclotides does, however, appear to be important for activity. The hydrophobic patch of tricyclon A is not as prominent as that observed in other cyclotides, and this has been suggested to account for its lack of hemolytic activity.[77] Previously, hemolytic activity was reported to correlate with hydrophobicity.[135] In kalata B8 the degree of surface-exposed hydrophobic residues is even less pronounced than in tricyclon A, consistent with the lack of hemolytic activity for kalata B8. Studies on the antitumor activity of cyclotides have also provided some insight into functionally important residues. A study on cyclotides isolated from *V. arvensis* and *V. odorata* revealed that cycloviolacin O2 was a more potent cytotoxic agent than varv A or varv F.[25] Cycloviolacin O2 is more cationic than varv A or varv F, and it was suggested that its extra positive charge might be responsible for its increased potency. However, a recent study has shown that the conserved Glu has a much greater influence on activity than the positively charged residues. Methylation of the Glu resulted in a 48-fold decrease in cytotoxicity activity, whereas modification of the Arg or Lys residue only decreased activity by 3-fold.[136] The cytotoxic activity profiles of these various cyclotides differ from anticancer drugs currently in use, and differences were also found in the activity profiles for varv A and varv F compared with cycloviolacin O2.[25] At present, the reasons for these differences are not fully understood, but the fact that cyclotides have different profiles of cytotoxic activity from current anticancer drugs is a promising indicator that they might have niche applications in cancer therapy.

Concluding Remarks. Recent studies on cyclotides have provided important insights into their structure, bioactivities, biosynthesis, and functional role in plants. These fascinating natural products behave like organic molecules in terms of stability and biophysical properties, but are proteinaceous in nature. They have a unique knotted and cyclic structure that gives them their exceptional stability and are of interest for their potential applications in agriculture and human therapeutics. Aside from their potential applications, their unique structure is of fundamental interest from a topological perspective. This review has highlighted the topological idiosyncrasies of cyclotides and has explored how they contribute to the physicochemical behavior of cyclotides. The diversity, abundance, and broad geographic distribution of cyclotide-containing plants make cyclotides accessible to natural products chemists throughout the world. We hope that the studies and methodologies described in this review will inspire and enable other scientists to begin working on this fascinating class of molecules. The scope for cyclotide research is large, including isolation and sequencing novel bioactive cyclotides and structure/function, bio-

Table 2. Bioactivities of Cyclotides[a]

activity	test organism	compound	potency (μM)	reference
antifouling	*Balanus improvisus*	cycloviolacin O2	0.025[b]	Göransson et al., 2004[48]
anti-HIV	10 HIV strains	circulins A−F	(0.04−0.26)	Gustafson et al., 1994[12] and 2000[19]
	HIV-I	cycloviolns A−D	(0.13)	Hallock et al., 2000[130]
	CEM-SS cells	kalata B1	(0.14)	Daly et al., 2004[129]
		kalata B8	(2.5)	Daly et al., 2006[74]
		palicourein	(0.1)	Bokesch et al., 2001[127]
		vhl-1	(0.87)	Chen et al., 2005[73]
antimicrobial	*Candida kefyr*	kalata B1	21.4[c]	Tam et al., 1999;[105] Tam et al., 2001[137]
		circulin A	18.6[c]	
		circulin B	29.0[c]	
		cyclopsychotride A	14.0[c]	
	Candida tropicalis	circulin A	19.4[c]	
		cyclopsychotride A	56.5[c]	
	Escherichia coli	circulin B	0.41[c]	
		cyclopsychotride A	1.55[c]	
	Klebsiella oxytoca	kalata B1	54.8[c]	
		circulin B	8.2[c]	
		cyclopsychotride A	5.8[c]	
	Pseudomonas aeruginosa	circulin B	25.5[c]	
		cyclopsychotride A	13.5[c]	
	Proteus vulgaris	circulin A	54.6[c]	
		circulin B	6.8[c]	
		cyclopsychotride A	13.2[c]	
	Staphylococcus aureus	kalata B1	0.26[c]	
		circulin A	0.19[c]	
		circulin B	13.5[c]	
		cyclopsychotride A	39.0[c]	
cytotoxic/antitumor	10 HIV strains	circulins A−F	0.5	Gustafson et al., 1994[12] and 2000[19]
	HIV-I	cycloviolins A−D	0.56	Gustafson et al., 2004[20]
	CEM-SS cells	kalata B1	11	Daly et al., 1999[103]
		kalata B8	3.5	Daly et al., 2006[74]
		palicourein	1.5	Bokesch et al., 2001[127]
	RPMI-8226/(s, Dox40 and Lr-5), U-937(GTB and Vcr), ACHN, CCRF-	varv A, varv F cycloviolacin O2	2.73−4.97 2.63−7.49 0.11−0.26	Lindholm et al., 2002[25]
	(CEM, and CEM/VM-1), NCI-H(69 and 69 AR)	vitri A	0.6−1	Svangard et al., 2004[26]
		varv E	4	
	mouse R1 fibroblast	circulin B	820	Tam et al., 1999[105]
		cyclopsychotride A	1850	
hemolytic	human RBC	circulin A	1020[d]	Tam et al., 1999[24]
		circulin B	550[d]	
		cyclopsychotride A	405[d]	Witherup et al., 1994[14]
		cycloviolacin O2	36	Ireland et al., 2006[39]
		cycloviolacin O13	11	
		cycloviolacin O14	25[e]	
		cycloviolacin O15		
		cycloviolacin O24	25[f]	
		kalata B1	50−1510[d]	Daly et al., 1999;[103] Tam et al., 1999;[105] Barry et al., 2003;[132] Simonsen et al., 2004[134]
		varv A		Ireland et al., 2006[39]
		violacin A		Ireland et al., 2006[66]
		violapeptide 1		Schopke et al., 1993[13]
insecticidal	*Helicoverpa punctigera*	kalata B1	0.825[d]	Jennings et al., 2001[28]
	Helicoverpa armigera	kalata B2		Jennings et al., 2005[75]
neurotensin antagonist	human HT-29 cells	cyclopsychotride A	3	Witherup et al., 1994[14]
trypsin inhibitory	trypsin	MCo-TI-I	0.02[g]	Avrutina et al., 2005[138]
		MCo-TI-II	0.03[g]	
uterotonic	rat, rabbit, and human uteri	kalata B1	10−20[h]	Gran, 1973[7]

[a] Note: Unless otherwise specified, values written as potency are IC_{50}, whereas values inside parentheses are EC_{50}. [b] Concentration required to inhibit ~80% barnacle settlement (barnacle inhibition). [c] MIC (minimum inhibitory concentration in μM). [d] HD_{50} (hemolytic dosage for 50% of the erythrocytes in μM). [e] Indicates the required concentration to cause ~11% hemolysis. [f] Indicates the required concentration to cause ~75% hemolysis. [g] nM. [h] μg/mL.

synthetic, and synthetic studies. Resolving the biosynthesis of cyclotides remains the largest and most important challenge, as it is intimately connected to understanding their evolution and natural role in plants.

Acknowledgment. Work in our laboratory on cyclotides is supported by grants from the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC). D.C.I. is a Dora Lush NHMRC Fellow, N.L.D. is a Queensland Smart State Fellow, R.J.C. is a NHMRC Biomedical CDA Fellow, and D.J.C. is a NHMRC Professorial Fellow. We thank our colleagues and collaborators listed in the references for their invaluable contributions to cyclotide research.

References and Notes

(1) Craik, D. J.; Daly, N. L.; Bond, T.; Waine, C. *J. Mol. Biol.* **1999**, *294*, 1327–1336.

(2) Colgrave, M. L.; Craik, D. J. *Biochemistry* **2004**, *43*, 5965–5975.
(3) Rosengren, K. J.; Daly, N. L.; Plan, M. R.; Waine, C.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 8606–8616.
(4) Göransson, U.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 48188–48196.
(5) Sandberg, F. *Cah. Moboke* **1965**, *27*, 5–49.
(6) Gran, L. *Medd. Nor. Farm. Selsk.* **1970**, *12*, 173–180.
(7) Gran, L. *Lloydia* **1973**, *36*, 174–178.
(8) Gran, L. *Acta Pharmacol. Toxicol.* **1973**, *33*, 400–408.
(9) Gran, L. *Lloydia* **1973**, *36*, 207–208.
(10) Sletten, K.; Gran, L. *Medd. Nor. Farm. Selsk.* **1973**, *7*−*8*, 69–82.
(11) Saether, O.; Craik, D. J.; Campbell, I. D.; Sletten, K.; Juul, J.; Norman, D. G. *Biochemistry* **1995**, *34*, 4147–4158.
(12) 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.; Boyd, M. R. *J. Am. Chem. Soc.* **1994**, *116*, 9337–9338.
(13) Schöpke, T.; Hasan Agha, M. I.; Kraft, R.; Otto, A.; Hiller, K. *Sci. Pharm.* **1993**, *61*, 145–153.
(14) Witherup, K. M.; Bogusky, M. J.; Anderson, P. S.; Ramjit, H.; Ransom, R. W.; Wood, T.; Sardana, M. *J. Nat. Prod.* **1994**, *57*, 1619–1625.
(15) Göransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. *J. Nat. Prod.* **1999**, *62*, 283–286.
(16) Wang, C. K.; Kaas, Q.; Chiche, L.; Craik, D. J. *Nucleic Acids Res.* **2008**, *36*, D206–D210.
(17) Simonsen, S. M.; Sando, L.; Ireland, D. C.; Colgrave, M. L.; Bharathi, R.; Göransson, U.; Craik, D. J. *Plant Cell* **2005**, *17*, 3176–3189.
(18) Gruber, C. W.; Elliott, A. G.; Ireland, D. C.; Delprete, P. G.; Dessein, S.; Goransson, U.; Trabi, M.; Wang, C. K.; Kinghorn, A. B.; Robbrecht, E.; Craik, D. J. *Plant Cell* **2008**, *20*, 2471–2483.
(19) Gustafson, K. R.; Walton, L. K.; Sowder, R. C., II; Johnson, D. G.; Pannell, L. K.; Cardellina, J. H., II; Boyd, M. R. *J. Nat. Prod.* **2000**, *63*, 176–178.
(20) Gustafson, K. R.; McKee, T. C.; Bokesch, H. R. *Curr. Protein Pept. Sci.* **2004**, *5*, 331–340.
(21) Wang, C. K.; Colgrave, M. L.; Gustafson, K. R.; Ireland, D. C.; Goransson, U.; Craik, D. J. *J. Nat. Prod.* **2008**, *71*, 47–52.
(22) Craik, D. J.; Čemažar, M.; Wang, C. K.; Daly, N. L. *Biopolymers* **2006**, *84*, 250–266.
(23) Ireland, D. C.; Wang, C. K.; Wilson, J. A.; Gustafson, K. R.; Craik, D. J. *Biopolymers* **2008**, *90*, 51–60.
(24) Tam, J. P.; Lu, Y. A.; Yang, J. L.; Chiu, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 8913–8918.
(25) Lindholm, P.; Göransson, U.; Johansson, S.; Claeson, P.; Gulbo, J.; Larsson, R.; Bohlin, L.; Backlund, A. *Mol. Cancer Ther.* **2002**, *1*, 365–369.
(26) Svangard, E.; Göransson, U.; Hocaoglu, Z.; Gullbo, J.; Larsson, R.; Claeson, P.; Bohlin, L. *J. Nat. Prod.* **2004**, *67*, 144–147.
(27) Göransson, U.; Sjogren, M.; Svangard, E.; Claeson, P.; Bohlin, L. *J. Nat. Prod.* **2004**, *67*, 1287–1290.
(28) Jennings, C.; West, J.; Waine, C.; Craik, D.; Anderson, M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10614–10619.
(29) Barbeta, B. L.; Marshall, A. T.; Gillon, A. D.; Craik, D. J.; Anderson, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 1221–1225.
(30) Colgrave, M. L.; Kotze, A. C.; Kopp, S.; McCarthy, J. S.; Coleman, G. T.; Craik, D. J. *Acta Trop.* **2009**, *109*, 163–166.
(31) Colgrave, M. L.; Kotze, A. C.; Ireland, D. C.; Wang, C. K.; Craik, D. J. *ChemBioChem* **2008**, *9*, 1939–1945.
(32) Colgrave, M. L.; Kotze, A. C.; Huang, Y. H.; O'Grady, J.; Simonsen, S. M.; Craik, D. J. *Biochemistry* **2008**, *47*, 5581–5589.
(33) Plan, M. R. R.; Saska, I.; Gagauan, A. G.; Craik, D. J. *J. Agric. Food Chem.* **2008**, *56*, 5237–5241.
(34) Dutton, J. L.; Renda, R. F.; Waine, C.; Clark, R. J.; Daly, N. L.; Jennings, C. V.; Anderson, M. A.; Craik, D. J. *J. Biol. Chem.* **2004**, *279*, 46858–46867.
(35) Svangard, E.; Burman, R.; Gunasekera, S.; Lovborg, H.; Gullbo, J.; Goransson, U. *J. Nat. Prod.* **2007**, *70*, 643–647.
(36) Shenkarev, Z. O.; Nadezhdin, K. D.; Lyukmanova, E. N.; Sobol, V. A.; Skjeldal, L.; Arseniev, A. S. *J. Inorg. Biochem.* **2008**, *102*, 1246–1256.
(37) Shenkarev, Z. O.; Nadezhdin, K. D.; Sobol, V. A.; Sobol, A. G.; Skjeldal, L.; Arseniev, A. S. *FEBS J.* **2006**, *273*, 2658–2672.
(38) Kamimori, H.; Hall, K.; Craik, D. J.; Aguilar, M. I. *Anal. Biochem.* **2005**, *337*, 149–153.
(39) Ireland, D. C.; Colgrave, M. L.; Craik, D. J. *Biochem. J.* **2006**, *400*, 1–12.
(40) 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*, 5722–5730.
(41) Heitz, A.; Hernandez, J. F.; Gagnon, J.; Hong, T. T.; Pham, T. T.; Nguyen, T. M.; Le-Nguyen, D.; Chiche, L. *Biochemistry* **2001**, *40*, 7973–7983.
(42) Felizmenio-Quimio, M. E.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2001**, *276*, 22875–22882.
(43) Chiche, L.; Heitz, A.; Gelly, J. C.; Gracy, J.; Chau, P. T.; Ha, P. T.; Hernandez, J. F.; Le-Nguyen, D. *Curr. Protein Pept. Sci.* **2004**, *5*, 341–349.
(44) Craik, D. J.; Čemažar, M.; Daly, N. L. *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 251–260.
(45) Craik, D. J.; Daly, N. L. *Protein Pept. Lett.* **2005**, *12*, 147–152.
(46) Craik, D. J.; Daly, N. L.; Mulvenna, J.; Plan, M. R.; Trabi, M. *Curr. Protein Pept. Sci.* **2004**, *5*, 297–315.
(47) Craik, D. J.; Simonsen, S.; Daly, N. L. *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 251–260.
(48) Göransson, U.; Svangard, E.; Claeson, P.; Bohlin, L. *Curr. Protein Pept. Sci.* **2004**, *5*, 317–329.
(49) Gunasekera, S.; Daly, N. L.; Anderson, M. A.; Craik, D. J. *IUBMB Life* **2006**, *58*, 515–524.
(50) Trabi, M.; Craik, D. J. *Trends Biochem. Sci.* **2002**, *27*, 132–138.
(51) Craik, D. J.; Anderson, M.; Daly, N. L. *Today's Life Sci.* **2003**, 28–32.
(52) Craik, D. J.; Daly, N. L.; Saska, I.; Trabi, M.; Rosengren, K. J. *J. Bacteriol.* **2003**, *185*, 4011–4021.
(53) Pelegrini, P. B.; Quirino, B. F.; Franco, O. L. *Peptides* **2007**, *28*, 1475–1481.
(54) Craik, D. J.; Mylne, J. S.; Daly, N. L. *Cell. Mol. Life Sci.* **2010**, *67*, 9–16.
(55) O'Keefe, B. R. *J. Nat. Prod.* **2001**, *64*, 1373–1381.
(56) Craik, D. J. *Trends Plant Sci.* **2009**, *14*, 328–335.
(57) Christmann, A.; Walter, K.; Wentzel, A.; Krätzner, R.; Kolmar, H. *Protein Eng.* **1999**, *12*, 797–806.
(58) Clark, R. J.; Daly, N. L.; Craik, D. J. *Biochem. J.* **2006**, *394*, 85–93.
(59) Sando, L.; Bharathi, R.; Craik, D. J. *Biopolymers* **2005**, *80*, 514.
(60) Mulvenna, J. P.; Mylne, J. S.; Bharathi, R.; Burton, R. A.; Shirley, N. J.; Fincher, G. B.; Anderson, M. A.; Craik, D. J. *Plant Cell* **2006**, *18*, 2134–2144.
(61) Zhang, J.; Liao, B.; Craik, D. J.; Li, J. T.; Hu, M.; Shu, W. S. *Gene* **2009**, *431*, 23–32.
(62) Trabi, M.; Mylne, J. S.; Sando, L.; Craik, D. J. *Org. Biomol. Chem.* **2009**, *7*, 2378–2388.
(63) Čemažar, M.; Craik, D. J. *Int. J. Peptide Res. Ther.* **2006**, *12*, 253–260.
(64) Seydel, P.; Dornenburg, H. *Plant Cell Tissue Organ Cult.* **2006**, *85*, 247–255.
(65) Broussalis, A. M.; Göransson, U.; Coussio, J. D.; Ferraro, G.; Martino, V.; Claeson, P. *Phytochemistry* **2001**, *58*, 47–51.
(66) Ireland, D. C.; Colgrave, M. L.; Nguyencong, P.; Daly, N. L.; Craik, D. J. *J. Mol. Biol.* **2006**, *357*, 1522–1535.
(67) Svangard, E.; Göransson, U.; Smith, D.; Verma, C.; Backlund, A.; Bohlin, L.; Claeson, P. *Phytochemistry* **2003**, *64*, 135–142.
(68) Colgrave, M. L.; Jones, A.; Craik, D. J. *J. Chromatogr. A* **2005**, *1091*, 187–193.
(69) Craik, D. J. *Toxicon* **2001**, *39*, 1809–1813.
(70) Derua, R.; Gustafson, K. R.; Pannell, L. K. *Biochem. Biophys. Res. Commun.* **1996**, *228*, 632–638.
(71) Skjeldal, L.; Gran, L.; Sletten, K.; Volkman, B. F. *Arch. Biochem. Biophys.* **2002**, *399*, 142–148.
(72) Barry, D. G.; Daly, N. L.; Bokesch, H. R.; Gustafson, K. R.; Craik, D. J. *Structure* **2004**, *12*, 85–94.
(73) Chen, B.; Colgrave, M. L.; Daly, N. L.; Rosengren, K. J.; Gustafson, K. R.; Craik, D. J. *J. Biol. Chem.* **2005**, *280*, 22395–22405.
(74) Daly, N. L.; Clark, R. J.; Plan, M. R.; Craik, D. J. *Biochem. J.* **2006**, *393*, 619–626.
(75) 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*, 851–860.
(76) Koltay, A.; Daly, N. L.; Gustafson, K. R.; Craik, D. J. *Int. J. Pept. Res. Ther.* **2005**, *11*, 99–106.
(77) Mulvenna, J. P.; Sando, L.; Craik, D. J. *Structure* **2005**, *13*, 691–701.
(78) Trabi, M.; Craik, D. J. *Plant Cell* **2004**, *16*, 2204–2216.
(79) Nair, S. S.; Romanuka, J.; Billeter, M.; Skjeldal, L.; Emmett, M. R.; Nilsson, C. L.; Marshall, A. G. *Biochim. Biophys. Acta* **2006**, *1764*, 1568–1576.
(80) Goransson, U.; Herrmann, A.; Burman, R.; Haugaard-Jonsson, L. M.; Rosengren, K. J. *ChemBioChem* **2009**, *10*, 2354–2360.
(81) Wang, C. K.; Colgrave, M. L.; Ireland, D. C.; Kaas, Q.; Craik, D. J. *Biochem. J.* **2009**, *97*, 1471–1481.
(82) Wang, C. K.; Hu, S. H.; Martin, J. L.; Sjogren, T.; Hajdu, J.; Bohlin, L.; Claeson, P.; Goransson, U.; Rosengren, K. J.; Tang, J.; Tan, N. H.; Craik, D. J. *J. Biol. Chem.* **2009**, *284*, 10672–10683.
(83) Pallaghy, P. K.; Nielsen, K. J.; Craik, D. J.; Norton, R. S. *Protein Sci.* **1994**, *3*, 1833–1839.
(84) Isaacs, N. W. *Curr. Opin. Struct. Biol.* **1995**, *5*, 391–395.
(85) Craik, D. J.; Daly, N. L.; Waine, C. *Toxicon* **2001**, *39*, 43–60.

(86) McDonald, N. Q.; Hendrickson, W. A. *Cell* **1993**, *73*, 421–424.
(87) Gillon, A. D.; Saska, I.; Jennings, C. V.; Guarino, R. F.; Craik, D. J.; Anderson, M. A. *Plant J.* **2008**, *53*, 505–515.
(88) Saska, I.; Gillon, A. D.; Hatsugai, N.; Dietzgen, R. G.; Hara-Nishimura, I.; Anderson, M. A.; Craik, D. J. *J. Biol. Chem.* **2007**, *282*, 29721–29728.
(89) Norton, R. S.; Pallaghy, P. K. *Toxicon* **1998**, *36*, 1573–1583.
(90) Plan, M. R.; Goransson, U.; Clark, R. J.; Daly, N. L.; Colgrave, M. L.; Craik, D. J. *ChemBioChem* **2007**, *8*, 1001–1011.
(91) Huang, Y. H.; Colgrave, M. L.; Daly, N. L.; Keleshian, A.; Martinac, B.; Craik, D. J. *J. Biol. Chem.* **2009**, *284*, 20699–20707.
(92) Huang, Y. H.; Colgrave, M. L.; Clark, R. J.; Kotze, A. C.; Craik, D. J. *J. Biol. Chem.* **2010**, *285*, 10797–10805.
(93) Mulvenna, J. P.; Wang, C.; Craik, D. J. *Nucleic Acids Res.* **2006**, *34*, D192–D194.
(94) Marglin, A.; Merrifield, R. B. *Annu. Rev. Biochem.* **1970**, *39*, 841–866.
(95) Daly, N. L.; Craik, D. J. *Future Med. Chem.* **2009**, *1*, 1613–1622.
(96) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776–779.
(97) Tam, J. P.; Lu, Y.-A. *Tetrahedron Lett.* **1997**, *38*, 5599–5602.
(98) Camarero, J.; Muir, T. *Chem. Commun.* **1997**, 1369–1370.
(99) Camarero, J. J, P.; Muir, T. *Angew. Chem., Int. Ed.* **1998**, *37*, 347–349.
(100) Tam, J. P.; Lu, Y.-A. *Protein Sci.* **1998**, *7*, 1583–1592.
(101) Leta Aboye, T.; Clark, R. J.; Craik, D. J.; Goransson, U. *ChemBioChem* **2008**, *9*, 103–113.
(102) Čemažar, M.; Craik, D. J. *J. Pept. Sci.* **2008**, *14*, 683–689.
(103) Daly, N. L.; Love, S.; Alewood, P. F.; Craik, D. J. *Biochemistry* **1999**, *38*, 10606–10614.
(104) Thongyoo, P.; Tate, E. W.; Leatherbarrow, R. J. *Chem. Commun.* **2006**, 2848–2850.
(105) Tam, J. P.; Lu, Y.-A.; Yu, Q. *J. Am. Chem. Soc.* **1999**, *121*, 4316–4324.
(106) Čemažar, M.; Daly, N. L.; Haggblad, S.; Lo, K. P.; Yulyaningsih, E.; Craik, D. J. *J. Biol. Chem.* **2006**, *281*, 8224–8232.
(107) Le-Nguyen, D.; Heitz, A.; Chiche, L.; El Hajji, M.; Castro, B. *Protein Sci.* **1993**, *2*, 165–174.
(108) Daly, N. L.; Clark, R. J.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 6314–6322.
(109) Gunasekera, S.; Daly, N.; Clark, R.; Craik, D. J. *Antioxid. Redox Signaling* **2009**, *11*, 971–980.
(110) Simonsen, S. M.; Sando, L.; Rosengren, K. J.; Wang, C. K.; Colgrave, M. L.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2008**, *283*, 9805–9813.
(111) Kimura, R. H.; Tran, A. T.; Camarero, J. A. *Angew. Chem., Int. Ed.* **2006**, *45*, 973–976.
(112) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180–193.
(113) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 10068–10073.
(114) Xu, M. Q.; Perler, F. B. *EMBO J.* **1996**, *15*, 5146–5153.
(115) Camarero, J. A.; Kimura, R. H.; Woo, Y. H.; Cantor, J.; Steenblock, E. *Chim. Oggi* **2007**, *25*, 20–23.
(116) Camarero, J. A.; Kimura, R. H.; Woo, Y. H.; Shekhtman, A.; Cantor, J. *ChemBioChem* **2007**, *8*, 1363–1366.
(117) Camarero, J. A. *Adv. Drug Delivery Rev.* **2009**, *61*, 897–898.
(118) Austin, J.; Wang, W.; Puttamadappa, S.; Shekhtman, A.; Camarero, J. A. *ChemBioChem* **2009**, *10*, 2663–2670.
(119) Thongyoo, P.; Bonomelli, C.; Leatherbarrow, R. J.; Tate, E. W. *J. Med. Chem.* **2009**, *52*, 6197–6200.
(120) Thongyoo, P.; Roque-Rosell, N.; Leatherbarrow, R. J.; Tate, E. W. *Org. Biomol. Chem.* **2008**, *6*, 1462–1470.
(121) Marx, U. C.; Korsinczky, M. L.; Schirra, H. J.; Jones, A.; Condie, B.; Otvos, L., Jr.; Craik, D. J. *J. Biol. Chem.* **2003**, *278*, 21782–21789.
(122) Seydel, P.; Gruber, C. W.; Craik, D. J.; Dornenburg, H. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 275–284.
(123) Dornenburg, H. *Biotechnol. Lett.* **2008**, *30*, 1311–1321.
(124) Tan, N. H.; Zhou, J. *Chem. Rev.* **2006**, *106*, 840–895.
(125) Sheldon, P. S.; Keen, J. N.; Bowles, D. J. *Biochem. J.* **1996**, *320*, 865–870.
(126) Gruber, C. W.; Cemazar, M.; Clark, R. J.; Horibe, T.; Renda, R. F.; Anderson, M. A.; Craik, D. J. *J. Biol. Chem.* **2007**, *282*, 20435–20446.
(127) Bokesch, H. R.; Pannell, L. K.; Cochran, P. K.; Sowder, R. C., 2nd; McKee, T. C.; Boyd, M. R. *J. Nat. Prod.* **2001**, *64*, 249–250.
(128) Daly, N. L.; Koltay, A.; Gustafson, K. R.; Boyd, M. R.; Casas-Finet, J. R.; Craik, D. J. *J. Mol. Biol.* **1999**, *285*, 333–345.
(129) Daly, N. L.; Gustafson, K. R.; Craik, D. J. *FEBS Lett.* **2004**, *574*, 69–72.
(130) Hallock, Y. F.; Sowder, R. C., II; Pannell, L. K.; Hughes, C. B.; Johnson, D. G.; Gulakowski, R.; Cardellina, J. H., II; Boyd, M. R. *J. Org. Chem.* **2000**, *65*, 124–128.
(131) Nourse, A.; Trabi, M.; Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2004**, *279*, 562–570.
(132) Barry, D. G.; Daly, N. L.; Clark, R. J.; Sando, L.; Craik, D. J. *Biochemistry* **2003**, *42*, 6688–6695.
(133) Daly, N. L.; Craik, D. J. *J. Biol. Chem.* **2000**, *275*, 19068–19075.
(134) Simonsen, S. M.; Daly, N. L.; Craik, D. J. *FEBS Lett.* **2004**, *577*, 399–402.
(135) Chen, B.; Colgrave, M. L.; Wang, C.; Craik, D. J. *J. Nat. Prod.* **2006**, *69*, 23–28.
(136) Herrmann, A.; Svangard, E.; Claeson, P.; Gullbo, J.; Bohlin, L.; Göransson, U. *Cell. Mol. Life Sci.* **2006**, *63*, 235–245.
(137) Tam, J. P.; Lu, Y. A.; Yang, J. L.; Yu, Q. Cyclic cystine-knot β-stranded antimicrobial peptides. In *Development of Novel Antimicrobial Agents: Emerging Strategies*; Lohner, K., Ed.; Horizon Scientific Press: Wymondham, UK, 2001; pp 215−240.
(138) Avrutina, O.; Schmoldt, H. U.; Gabrijelcic-Geiger, D.; Le Nguyen, D.; Sommerhoff, C. P.; Diederichsen, U.; Kolmar, H. *Biol. Chem.* **2005**, *386*, 1301–1306.
(139) Herrmann, A.; Burman, R.; Mylne, J. S.; Karlsson, G.; Gullbo, J.; Craik, D. J.; Clark, R. J.; Goransson, U. *Phytochemistry* **2008**, *69*, 939–952.
(140) Göransson, U.; Broussalis, A. M.; Claeson, P. *Anal. Biochem.* **2003**, *318*, 107–117.
(141) Claeson, P.; Göransson, U.; Johansson, S.; Luijendijk, T.; Bohlin, L. *J. Nat. Prod.* **1998**, *61*, 77–81.

NP1000413