Circular proteins – no end in sight

Manuela Trabi and David J. Craik

Circular proteins are a recently discovered phenomenon. They presumably evolved to confer advantages over ancestral linear proteins while maintaining the intrinsic biological functions of those proteins. In general, these advantages include a reduced sensitivity to proteolytic cleavage and enhanced stability. In one remarkable family of circular proteins, the cyclotides, the cyclic backbone is additionally braced by a knotted arrangement of disulfide bonds that confers additional stability and topological complexity upon the family. This article describes the discovery, structure, function and biosynthesis of the currently known circular proteins. The discovery of naturally occurring circular proteins in the past few years has been complemented by new chemical and biochemical methods to make synthetic circular proteins; these are also briefly described.

Manuela Trabi
David J. Craik*
Institute for Molecular Bioscience, University of Queensland, Brisbane, QLD 4072, Australia.
*e-mail: d.craik@imb.uq.edu.au

'Conventional' proteins are linear chains of amino acids that fold into a three-dimensional shape that defines their biological function. However, a chain is only as strong as its weakest link and it can be argued that linear proteins are, in fact, missing a link that might otherwise join their termini. This has significant implications because the termini of a peptide chain are often flexible and represent target points for the attack of proteolytic enzymes. Conceptually, there is no reason why this bond need be missing – the N and C termini are perfectly suited to chemical linkage in a peptide bond, just like the peptide bonds connecting all the intervening amino acids. In recent years, we and others have discovered several interesting proteins in which this linkage, to form a circular backbone, occurs.

Naturally occurring circular proteins have only been recognized and characterized over the past few years, but their history can be traced to native medicine observations three decades ago. In the early 1970s, Gran reported that the uterotonic (i.e. causes uterine contractions) properties of an extract from the African plant *Oldenlandia affinis* were associated with a 29-amino acid peptide named kalata B1 [1–3]. It was another 25 years before it was discovered that the peptide is actually backbone cyclized, has a well-defined three-dimensional structure [4] and is a gene product [5]. The discovery of numerous other circular proteins over the past few years (Table 1) in various organisms, including bacteria, plants and a mammal, redefines our classical image of proteins as simply linear chains of amino acids. It now seems that some organisms have, indeed, found ways of creating the missing link in protein chains.

The recently discovered circular proteins that are the subject of this article can be distinguished from small cyclic peptides such as cyclosporin, which have been known for some time. The latter tend to be <12 amino acids in size, often contain modified amino acids and are usually not true gene products, instead being synthesized in microorganisms by multifunctional enzyme complexes. By contrast, the circular molecules discussed here are true 'proteins' whose sequence is encoded by DNA and which adopt well-defined three-dimensional structures. As well as describing the discovery of these naturally occurring circular proteins, we will briefly describe complementary developments relating to synthetic circular proteins.

Naturally occurring circular proteins

The diversity of structures of naturally occurring circular proteins is summarized in Fig. 1. They range in size from 14 to 70 amino acids, all show well-defined three-dimensional structures, and all have been discovered only over the past decade. A common characteristic of the naturally occurring circular proteins seems to be their involvement in host defence.

Circular proteins from microorganisms

The largest naturally occurring circular protein currently known is bacteriocin AS-48, a highly basic 70-amino acid protein isolated from *Enterococcus faecalis* S-48 [6]. The function of this protein is to protect the producing strain against other bacteria and it does this by forming pores in the cytoplasmic membrane of sensitive cells. The three-dimensional structure of this protein (Fig. 1) consists of a globular arrangement of five α helices connected by five short turn regions that enclose a compact hydrophobic core [7]. The head-to-tail ligation that produces the circular backbone occurs within one of the helices and appears to confer considerable stability on the cyclized molecule evidenced, for example, by its thermal denaturation temperature of 93°C.

A second circular protein of bacterial origin is microcin J25 (MccJ25), a highly hydrophobic 21-residue peptide excreted by *Escherichia coli* [8]. Microcins are a miscellaneous group of low molecular mass antibiotic peptides produced by diverse strains of Enterobacteriaceae, but MccJ25 is the only one known to be circular. It is unrelated in sequence to the other microcins but shares a similar activity, being mainly active against bacterial species phylogenetically related to the producing strain. The antibiotic activity of MccJ25 derives from its ability to interfere with cell division [9]. The circular backbone is crucial for activity, as a linear form of the protein, produced by cleaving MccJ25 with the enzyme thermolysin, was inactive against three tested *E. coli* strains and 40 times less active than circular MccJ25 against *Salmonella newport* [8]. Although MccJ25 is less than one-third the size of bacteriocin A-48, it shares the characteristic of having a compact globular

 0968-0004/02/$ – see front matter PII: S0968-0004(02)02057-1

Table 1. Origin and characteristics of naturally occurring and synthetic circular peptide and proteins

Circular peptide and/or protein	Source or synthetic method	Size[a]	Comments	Refs
Naturally occurring				
SFTI-1	*Helianthus annuus*	14 aa	Potent trypsin inhibitor	[30]
RTD-1	*Macaca mulatta*	18 aa	Antibiotic defensin from primate leukocytes	[32]
Microcin J25	*Escherichia coli*	21 aa	Antibacterial; linear analogue less active	[8]
Cyclotide family	Rubiaceae and Violaceae spp.	28–37 aa	~45 proteins known; various bioactivities	[4,11–14, 16,17,46]
MCoTI-I and II	*Momordica cochinchinensis*	34 aa	Seed-derived trypsin inhibitor	[27]
Bacteriocin AS-48	*Enterococcus faecalis*	70 aa	Hydrophobic antibacterial protein	[6]
Synthetic				
Pseudostellarin F	Split intein mechanism	8 aa	Synthetic yields rival the ones from natural sources	[47]
Protegrins	Thia zip reaction	18 aa	Antibiotic; decrease in haemolytic activity	[48]
Tachyplesins	Thia zip reaction	18 aa	Antibiotic; changes in membranolytic selectivity	[49]
Conotoxin/cyclotide chimera	Chemical crosslinking[b]	26 aa	Potent calcium channel antagonist	[46]
SH3 domain	Native chemical ligation	57 aa	Improved ligand-binding properties	[50]
Bovine pancreatic trypsin inhibitor	Chemical crosslinking[c]	58 aa	First example of synthetic circular protein	[51]
YAP WW domain	Native chemical ligation	5 kDa	Slightly higher ligand-binding affinity	[52]
Thioredoxin	Two intein (TWIN) system	135 aa	80% yield, multimerization as side reaction	[53]
Dihydrofolate reductase	Split intein mechanism	21 kDa	Enhanced thermostability, protease resistance	[47]
Green fluorescent protein	Split intein mechanism	244 aa	High stability against chemical denaturation and exopeptidases	[41]
β-Lactamase	Intein-mediated protein ligation	30 kDa	Enhanced thermostability, protease resistance	[42]
Maltose-binding protein	Two intein (TWIN) system	395 aa	50% yield, multimerization as side reaction	[53]

[a]Number of amino acids or molecular weight in kDa.
[b]Treatment with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).
[c]Treatment with N,N′-diisopropylethylamine (DIEA).

structure (Fig. 1). However, in contrast to the helical globule of bacteriocin A-48, MccJ25 consists of a distorted antiparallel β sheet that is twisted and folded back onto itself [10]. In this sense, it is similar to the series of plant-derived circular proteins described in the subsequent text.

Circular proteins from plants

Comprising >45 members, the cyclotides [11] are by far the largest group of circular proteins discovered to date. Members of this group contain 28–37 amino acids and some representative sequences are given in Fig. 2. Their initial discoveries in the mid-1990s were made based on either screening programs or native medicine applications, both techniques involving a diverse range of biological activities. For example, the circulins A and B were found in the course of screening for anti-HIV activity [12], cyclopsychotride A was discovered on the basis of its inhibition of neurotensin binding [13], and kalata B1 [1,4] was identified by virtue of its uterotonic activity. The plants from which these molecules were derived were all tropical species from the Rubiaceae family; however, a similar macrocyclic peptide, Viola peptide 1, was reported at around the same time in a plant from the Violaceae family [14].

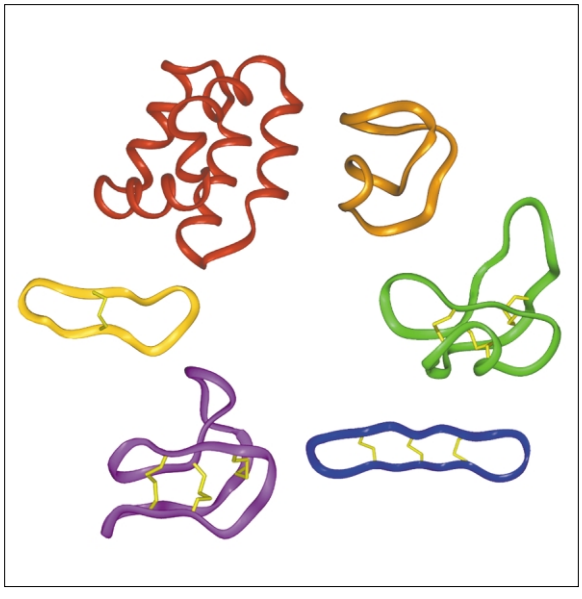

Fig. 1. Three-dimensional structures of naturally occurring proteins. Clockwise from the upper left the proteins are: bacteriocin AS-48 (1E68) from *Enterococcus faecalis*, microcin J25 (1HG6) from *Escherichia coli*, MCoTI-II (1HA9, 1IB9) from bitter melon seeds, RTD-1 (1HVZ) from the leukocytes of Rhesus macaques, kalata B1 (1KAL) from several plants of the Rubiaceae and Violaceae plant families, and SFTI-1 (1SFI, 1JBL) from the seeds of the common sunflower. Disulfide bonds are shown in yellow. PDB access codes are given in parentheses after the name of the protein.

The common features of the known molecules in these initial reports included their size of ~30 amino acids, a circular backbone, six conserved cysteine residues that formed three disulfide bonds, and a plant origin. Believing that additional examples probably existed, we set out to find other circular proteins, this time not based on bioassay-directed screening, but instead on an extraction protocol designed to find proteins of similar size and properties. This was a fruitful effort: 16 new peptides were reported by our group [11], and 8 by the group of Claeson [15] and Göransson [16] in 1998 and 1999. At this time, it became clear that these macrocyclic peptides were part of a large family and we suggested that they be referred to as the cyclotides [11]. Additional cyclotides have been reported in the past year, including circulins from *Chassalia parvifolia* [17], cycloviolins A–D from *Leonia cymosa* [18], palicourein from

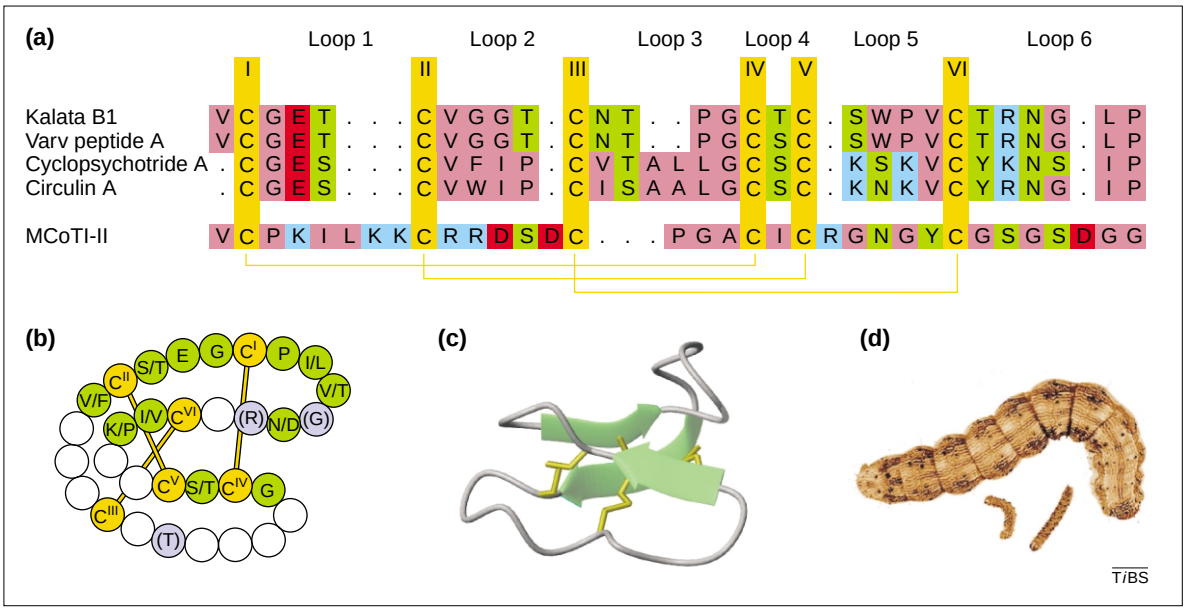

Fig. 2. Features of plant cyclotides. (a) Representative sequences with the disulfide-bonding pattern highlighted in yellow. The cysteine residues are numbered with Roman numbers, the backbone loops between cysteine residues with Arabic digits. The residues are coloured according to their properties: hydrophobic, pink; hydrophilic, green; basic, blue; acidic, red. Although the cyclotides and the trypsin inhibitor from *Momordica cochinchinensis* (MCoTI-II) share the same general disulfide framework, they have different cysteine spacing and little sequence similarity. Loop 1 of MCoTI-II contains the trypsin inhibitory sequence. (b) Topology of the cyclic cystine knot (CCK) motif and summary of the conserved and variable residues of all the known cyclotides (except palicourein). The disulfide-bonding pattern is shown in yellow with the cysteine residues numbered. Residues are indicated by their one letter code. Conserved residues are shown in green (where two highly homologous residues are interchangeable both are shown). Residues with moderate sequence conservation are represented in purple, with the most common amino acid given in parentheses. Variable residues are represented by blank circles. This illustration shows that the residues involved in the CCK (bold circles) are highly conserved. (c) Solution structure of kalata B1 showing the cystine knot and associated triple stranded β sheet. (d) Insecticidal activity of kalata B1 demonstrated by effect on growth and development of *Helicoverpa punctigera* larvae. After 16 days, the average weight of kalata-fed larvae (below) was 3 mg, whereas animals from the control group (represented by the larger caterpillar) weighed 284 mg on average. None of the kalata-fed larvae had progressed past the first instar stage of development, whereas most of the larvae on the control diet reached fifth instar [5].

Palicourea condensata [19], and hypa A from *Hybanthus parviflorus* [20]. Again, all these plants come from the Violaceae or Rubiaceae families.

Kalata B1 was the first cyclotide to have its structure determined [4], leading to the finding that it contains a 'cystine knot' motif in which a small embedded ring formed by two disulfide bonds and their connecting peptide backbone segments is penetrated by a third disulfide bond. Similar cystine knot motifs have been observed in other proteins, including small toxins [21] and growth factors [22] (although, in the case of growth factors, the topology of the knot is different). However, cystine knot motifs are only combined with a circular peptide backbone in the cyclotides. The combination of these two features – cystine knot and circular backbone – defines the so-called cyclic cystine knot (CCK) motif [11] (Fig. 2b). A small triple stranded β sheet is also associated with this motif (Fig. 2c). Structures of circulin A [23] and cycloviolacin O1 [11] also show this consensus motif and it is anticipated that, based on their sequence homologies, the CCK motif will be common to all cyclotides.

The knotted disulfide arrangement and circular backbone of the CCK motif renders the cyclotides extremely stable. They are resistant to enzymatic hydrolysis and thermal denaturation as shown, for example, by the retention of biological activity after boiling in native medicine applications [24]. Studies of acyclic permutants [25] of kalata B1 showed that although open chain analogues maintain the basic three-dimensional structure, they are intrinsically less stable and have weaker internal hydrogen bonds than the circular parent proteins. This finding suggests that cyclization plays a pivotal role in stabilizing rather than defining the three-dimensional structure.

The diverse range of biological activities displayed by the cyclotides gives them potential as lead compounds for the development of pharmaceutical agents. In addition to the anti-HIV and uterotonic activities already mentioned, several cyclotides have antibacterial or antifungal activity [26]. Although it might be speculated that the natural function of the cyclotides is related to their antimicrobial activity, we believe that their true role in plants is as a defence against insect predation. This is based on the finding that kalata B1 is a potent inhibitor of the growth and development of *Helicoverpa* species [5] (Fig. 2d). It will be interesting to see whether other cyclotides also have insecticidal activity.

The family of circular, disulfide-rich proteins from plants was expanded recently with the report [27] of two trypsin inhibitors, MCoTI-I and MCoTI-II, from the seeds of *Momordica cochinchinensis*, a vine plant from the Cucurbitaceae family. The structure of MCoTI-II has recently been determined [28,29] and contains the CCK motif, so the peptide can tentatively be regarded as a new member of the cyclotide family.

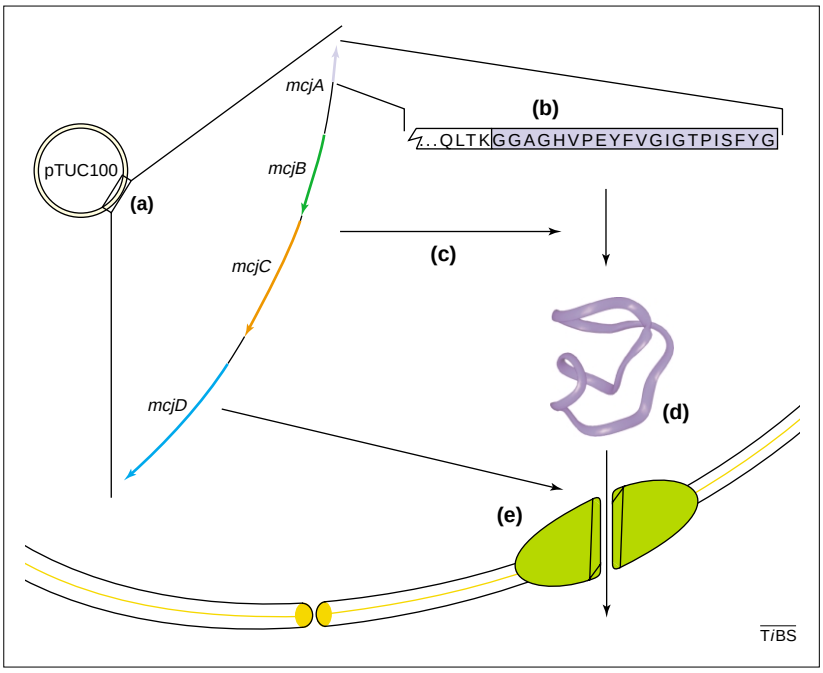

Fig. 3. Gene arrangements and biosynthesis of microcin J25. (a) A 4.8-kb region of the 50-kb plasmid pTUC100 carries the four genes *mcjABCD* necessary and sufficient for production of, and immunity against, the circular peptide antibiotic MccJ25. *mcjA* encodes a propeptide containing the amino acid sequence of mature MccJ25 (b) plus a 37-amino acid leader peptide (which differs significantly from traditional secretion sequences). (c) The products of both genes *mcjB* and *mcjC* are necessary to process the linear precursor McjA into a mature, circular protein (d), involving cleaving off the leader sequence and cyclizing the peptide backbone. Finally, *mcjD* encodes a putative inner membrane ABC transporter (e), which is not involved in MccJ25 production and maturation, but is required for secretion of the mature, circular protein. Additionally, *mcjD* confers cellular immunity to exogenous MccJ25.

However, despite this topological similarity and the six conserved cysteine residues, there are substantial sequence differences with the previously known cyclotides (Fig. 2a). Furthermore, the two MCoTI peptides are the only CCK family members that inhibit trypsin and they differ from the other cyclotides in that a linear homologue co-exists with circular forms in the same tissue. The linear form, MCoTI-III, is present only in minor amounts and is homologous, rather than identical, in sequence with the circular forms, so it is not necessarily a precursor of the circular forms. These linear and circular forms have high sequence similarities with members of the well-known squash trypsin inhibitor family, suggesting that they might be better classified in that group rather than with the cyclotides. Interestingly, the head-to-tail linker region of MCoTI-II is flexible, making entropic factors an unlikely reason why backbone cyclization evolved. Rather, resistance to proteolysis seems to be the driving force for cyclization [28,29] of these peptides.

In 1999, another circular trypsin inhibitor was discovered in plants, this time in sunflower seeds, and it was therefore named sunflower trypsin inhibitor 1 (SFTI-1). Comprising just 14 amino acids, SFTI-1 shows both sequence and conformational similarity with the Bowman–Birk inhibitors, a family of small serine proteinase inhibitors found in the seeds of legumes and several other plants [30]. So far, SFTI-1 is the most potent Bowman–Birk inhibitor, with a sub-nanomolar K_i value. The crystal structure of SFTI-1 in complex with trypsin revealed that the inhibitor comprises two antiparallel β strands stabilized by a disulfide bond and connected by an extended loop at the reactive site end and by a hairpin turn at the other end [30]. Korsinczky *et al.* [31] compared the solution structure and inhibitory activity of native SFTI-1 with that of an acyclic analogue having the peptide backbone broken at the hairpin end. The three-dimensional structures of the two molecules proved to be almost identical to each other and to the crystal structure of SFTI-1 bound to trypsin, indicating that the circular nature of SFTI-1 has probably evolved to increase the *in vivo* lifetime rather than to confer conformational stability onto the active loop region; as with the Momordica peptides, enhanced proteolytic stability seems to be a major driving force for cyclization.

Circular peptides from mammals

Rhesus theta defensin-1 (RTD-1) [32] is a new type of defensin (small, disulfide-rich peptides involved in host defence) from the leukocytes of rhesus macaques. It comprises just 18 amino acids, including six cysteines and five arginines, with the backbone cyclized through a peptide bond. The solution structure of RTD-1 consists of two β strands connected by two tight turns. Like the cyclotides, RTD-1 contains three disulfide bonds which, in contrast to those in the cyclotides, are organized in a ladder-like arrangement (Fig. 1). Despite the small size of RTD-1 and the constraints imposed by the circular backbone and the three disulfide bonds, the peptide appears to be relatively flexible, showing that the laddered tricystine arrangement is not as efficient as the knotted arrangement of the cyclotides in stabilizing structure [33].

Native circular RTD-1 has an antibacterial activity threefold higher than a synthetic acyclic analogue and, in contrast with the acyclic form, maintains its staphlocidal and colicidal activity in the presence of physiological NaCl [32]. Elucidation of the three-dimensional structure of this acyclic form shows that cyclization does not change the overall structure of the peptide [33]. Therefore, the difference in antimicrobial activity cannot be explained by structural changes upon cyclization, and might, again, involve stabilization of the molecule *in vivo*. Resistance to exoproteases would be particularly advantageous in the protease-rich inflammatory milieu in which these peptides function.

Biosynthesis

So far, not much is known about the biosynthesis of circular proteins. A common feature is that they all appear to be derived from longer precursor proteins, and therefore both cleavage and cyclization steps must be involved. The gene sequences for most of the precursors are known, but putative cleaving and/or cyclizing enzymes, or other mechanisms of cyclization, have not yet been reported. However, the

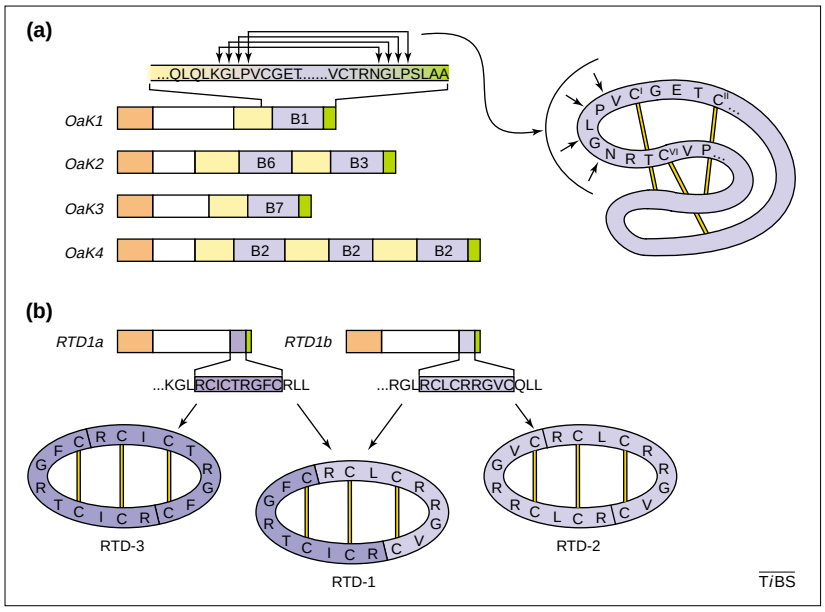

Fig. 4. Gene arrangements and posttranslational processing of the cyclotides and RTD-1. (a) The cyclotides are derived from a multigene family. Four clones, *OaK1–4*, are shown, each encoding up to three mature cyclotide domains. Each mature cyclotide sequence (i.e. kalata B1, B2, B3, B6 or B7) is flanked by two GLP sequences, resulting in four potential processing sites. The GLP sequence retained in the mature circular protein could be derived entirely from one of the two flanking elements or could contain parts of both, depending on initial cleavage. However, in the case of kalata B2, the flanking C-terminal element is SLP, suggesting a cleavage before the G and S residues. (b) Both *RTD1a* and *RTD1b* encode a prepropeptide containing half the mature RTD-1 sequence. In addition to processing of these two linear precursors, two head-to-tail ligations are necessary to produce mature, circular RTD-1. The homodimeric products of precursors *RTD1a* and *RTD1b*, peptides RTD-2 and RTD-3, are also found in Rhesus monkey leukocytes, although in smaller concentrations. Orange shading depicts an endoplasmic reticulum signal sequence, yellow indicates a precursor repeat fragment, lilac indicates a mature peptide sequence and green indicates a hydrophobic tail.

precursor sequences provide some clues about the processing steps leading to circular proteins. Interestingly, no mature linear forms that correspond exactly to the various circular peptides have been isolated from natural sources. The only example of a related linear form is a 30-amino acid trypsin inhibitor isolated from the seeds of *Momordica cochinchinensis*, which shows high sequence similarity to MCoTI-II [27]. The absence of exact linear analogues in general suggests that the cyclization process is very rapid and efficient once a mature peptide sequence is excised from its precursor.

The two circular proteins from bacteria, microcin J25 and bacteriocin AS-48, are both plasmid encoded. Interestingly, the plasmids not only contain the structural gene for the respective protein, but also several genes for proteins involved in protein maturation and export. In each case, only one of the additional proteins, namely McjD and AS-48D, thought to be responsible for protein export and cellular immunity against the respective protein, show sequence similarities with known proteins, in both cases ABC transporters [34,35]. Figure 3 summarizes the presumed role of the processing and export auxiliary proteins. Identification of the structure and function of the processing proteins represents an exciting challenge that should shed light onto the mechanism of cyclization.

Recently, genes encoding members of the cyclotide family have been isolated [5]. After screening the genomic DNA of *Oldenlandia affinis*, it was estimated that the cyclotides are derived from a multigene family with up to 12 related genes, four of which were isolated from a cDNA library for further investigation. The clones encode predicted prepropeptides with a 20-amino acid signal sequence, an N-terminal prosequence and 1–3 cyclotide domains (Fig. 4a). The mature cyclotide domains are preceded by a 22-amino acid repeat fragment and there is a small hydrophobic tail at the C terminus of the precursor protein. The exact processing site of this precursor has not been reported, but one probable cleavage site lies at a highly conserved GLP or SLP sequence that flanks both sides of the cyclotide domains. The mature circular protein retains one copy of this G/SLP motif, which could have originated entirely in one of the initial flanking elements, or contain parts from both [5]. So far, enzymes with the specificity necessary to perform this task have not been reported.

The case of RTD-1 biosynthesis is even more complex. Tang *et al*. [32] found two highly similar cDNAs that each contained half the sequence of mature RTD-1. RTD1a and RTD1b each encode 76-amino acid prepropeptides that have a 20-residue signal sequence, a 44-residue prosegment, half the mature RTD-1 sequence and a QLL/RLL tail. Therefore, transformation of the two linear precursors into mature, circular RTD-1 involves processing of both prepropeptides as well as two head-to-tail ligation reactions (Fig. 4b). Recently, two further theta defensins were discovered, representing the homodimeric splicing products of RTD1a and RTD1b. The cellular abundances of RTD-1, RTD-2 (product of RTD1b) and RTD-3 (from RTD1a) were 29:1:2, indicating that the heterodimeric ligation producing RTD-1 is clearly preferred [36,37]. However, the mechanism of this cyclization and the enzyme(s) involved remain to be discovered.

In summary, there are many unanswered questions relating to the biosynthesis of circular proteins. It is clear that the mature circular protein must be first excised from a larger linear precursor protein, but whether the excision and cyclization occur together or sequentially, and whether they are enzymatic or autocatalytic processes, remains to be determined. Further insight into these questions will come as more circular proteins and their precursors are discovered. Studies on methods to produce synthetic circular proteins will also probably assist in this understanding.

Synthetic circular proteins

In parallel with the discovery of increasing numbers of naturally occurring circular proteins, the development of new synthetic techniques has made it possible to cyclize proteins whose termini are in reasonably close proximity in their native (linear) structures. Figure 5 summarizes two major approaches based on modified

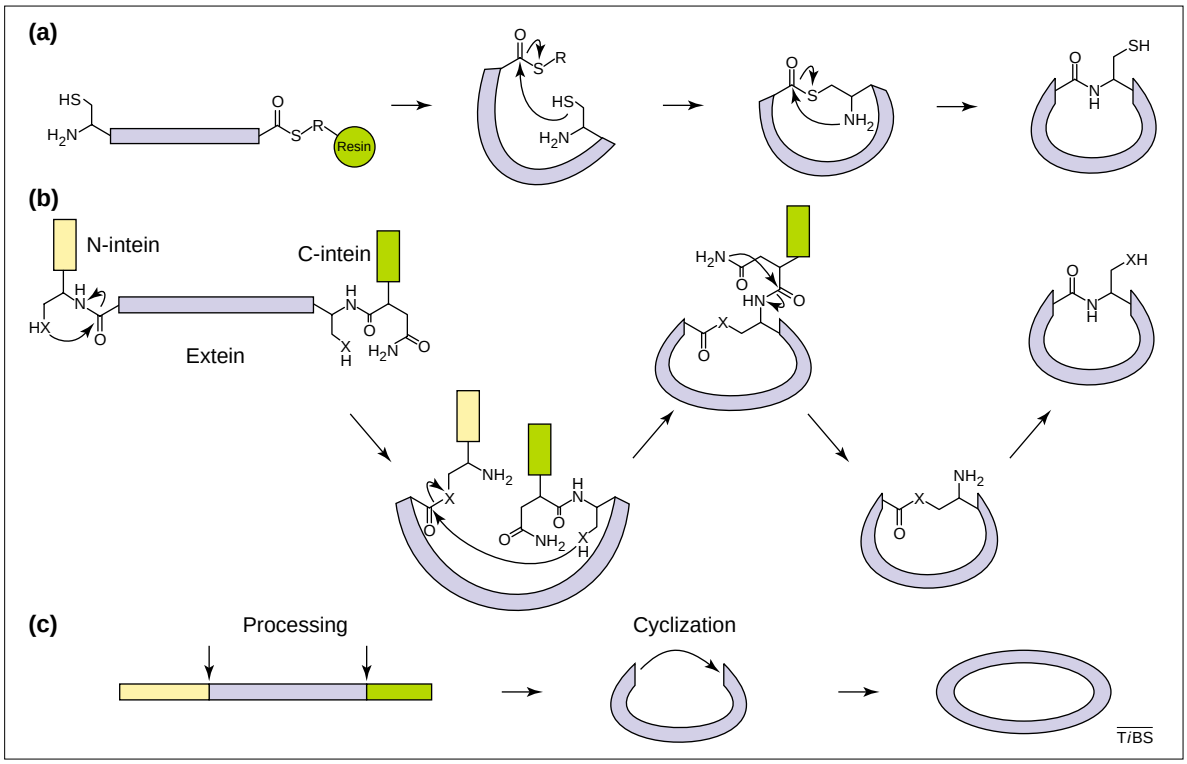

Fig. 5. Strategies for synthetic circular proteins and their parallels to the processing of some naturally occurring ones. (a) Native chemical ligation in solid phase peptide synthesis. The peptide is synthesized on resin with an N-terminal cysteine residue and a C-terminal α-thioester moiety. These two groups react to form a thioester-linked intermediate as the initial covalent product and this intermediate undergoes spontaneous, rapid intramolecular rearrangement to form a native peptide bond at the ligation site. (b) Split intein mechanism. The N-terminal splice junction is activated by an N–O or N–S acyl rearrangement forming an ester or thioester intermediate. In a second step, the ester or thioester bond is attacked by the hydroxyl or thiol group of a S/T/C residue on the C-terminal side of the extein, resulting in a branched cyclic intermediate. The branch is resolved by cyclization of a conserved C-terminal asparagine residue to form a succinimide ring, resulting in cleavage of the C-terminal splice junction. A final spontaneous O–N or S–N acyl rearrangement results in formation of a circular extein with a native peptide bond at the reaction site. (X represents the S or O atom of the C/S/T side chain.) (c) Generic model for the processing of naturally occurring circular proteins: a precursor protein is processed to release the mature sequence, which is consequently folded and cyclized in a process not yet understood.

'native ligation' in solid phase peptide synthesis [38–40] and on intein-based methods [41–43]. The former requires the design of a linear precursor sequence with a Cys residue at the N terminus, whereas the latter involves the autocatalytic ligation of a peptide sequence by reactive protein domains called inteins. Figure 5c highlights a parallel between these synthetic methods and the biosynthesis of naturally occurring proteins. In both cases, production of the mature circular protein involves the removal of auxiliary flanking domains from a linear precursor protein. Other cyclization methods have also been applied, including the use of oxidative conditions to form disulfide bonds, thereby correctly folding the linear precursor to pre-organize the termini in proximity before cyclization [44], and enzyme-assisted approaches [45].

Table 1 summarizes examples of synthetic circular proteins, ranging in size from 8 to 395 amino acids. Many studies were undertaken with the aim of enhancing the stability of the linear parent protein. For example, a circular backbone can be used to improve the thermodynamic stability of enzymes (e.g. used in industrial processes) or the *in vivo* stability of therapeutically useful proteins. The field is still in its infancy, but promising results are being obtained. It is noteworthy that many proteins are, in principle, amenable to cyclization. A sample of ~2000 representative proteins from the protein database shows that 31% have their termini within 20 Å and 11% are within 15 Å, a gap that could be filled with as few as 5–6 residues (L. Chiche and J. Gracy, pers. commun.).

As well as being applied to cyclize conventional linear proteins, new synthetic methodologies are also being used to make analogues of naturally occurring circular proteins such as the cyclotides. These studies are motivated by the fact that the exceptional stability of the CCK framework makes it valuable in drug design applications. It is being used as a stable scaffold onto which amino acids different from those in the native sequence can be grafted to achieve new functionalities [46].

Concluding remarks

In this review, we have described recently discovered circular proteins, noting the mystery that still surrounds their biosynthesis, and examined their diverse properties. Clearly, the additional complexity and biosynthetic resources necessary to produce such circular proteins must have evolved because of advantages conferred by cyclization. The two main advantages appear to be increased resistance to

protease digestion and improved thermodynamic stability, both of which can improve the biological activity of a protein *in vivo*. Other possible benefits of cyclization, such as orienting the active residues in a preferred conformation for binding, or diminishing the flexibility of peptide termini and thereby minimizing entropic losses on receptor binding, might occur in some cases, but do not appear to be the primary role of backbone cyclization.

With increasing discoveries of naturally occurring circular proteins and an active literature on the engineering of synthetic circular proteins, one thing is certain in relation to circular proteins – there is no end in sight.

Acknowledgements
This work was supported by grants from the Australian Research Council. We thank our colleagues cited in the references for their valuable contributions.

References

1 Gran, L. (1970) An oxytocic principle found in *Oldenlandia affinis* DC. An indigenous, Congolese drug "Kalata-Kalata" used to accelerate delivery. *Meddelelser fra Norsk Farmaceutisk Selskap* 32, 173–180
2 Gran, L. (1973) On the effect of a polypeptide isolated from "Kalata-Kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta Pharmacol. Toxicol.* 33, 400–408
3 Sletten, K. and Gran, L. (1973) Some molecular properties of kalatapeptide B-1. A uterotonic polypeptide isolated from *Oldenlandia affinis* DC. *Meddelelser fra Norsk Farmaceutisk Selskap* 7–8, 69–82
4 Saether, O. *et al.* (1995) Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide Kalata B1. *Biochemistry* 34, 4147–4158
5 Jennings, C. *et al.* (2001) Biosynthesis and insecticidal properties of plant cyclotides – cyclic knotted proteins from *O. affinis*. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10614–10619
6 Martínez-Bueno, M. *et al.* (1994) Determination of the gene sequence and the molecular structure of the enterococcal peptide antibiotic AS-48. *J. Bacteriol.* 176, 6334–6339
7 González, C. *et al.* (2000) Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11221–11226
8 Blond, A. *et al.* (1999) The cyclic structure of microcin J25, a 21-residue peptide antibiotic from *Escherichia coli*. *Eur. J. Biochem.* 259, 747–755
9 Salomón, R.A. and Farías, R.N. (1992) Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. *J. Bacteriol.* 174, 7428–7435
10 Blond, A. *et al.* (2001) Solution structure of microcin J25, the single macrocyclic antimicrobial peptide from *Escherichia coli*. *Eur. J. Biochem.* 268, 2124–2133
11 Craik, D.J. *et al.* (1999) Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.* 294, 1327–1336
12 Gustafson, K.R. *et al.* (1994) Circulins A and B: novel HIV-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*. *J. Am. Chem. Soc.* 116, 9337–9338
13 Witherup, K.M. *et al.* (1994) Cyclopsychotride A, a biologically active, 31-residue cyclic peptide isolated from *Psychotria longipes*. *J. Nat. Prod.* 57, 1619–1625
14 Schöpke, T. *et al.* (1993) Hämolytisch aktive komponenten aus *Viola tricolor* L. und *Viola arvensis* Murray. *Sci. Pharm.* 61, 145–153
15 Claeson, P. *et al.* (1998) Fractionation protocol for the isolation of polypeptides from plant biomass. *J. Nat. Prod.* 61, 77–81
16 Göransson, U. *et al.* (1999) Seven novel macrocyclic polypeptides from *Viola arvensis*. *J. Nat. Prod.* 62, 283–286
17 Gustafson, K.R. *et al.* (2000) New circulin macrocyclic polypeptides from *Chassalia parvifolia*. *J. Nat. Prod.* 63, 176–178
18 Hallock, Y.F. *et al.* (2000) Cycloviolins A–D, anti-HIV macrocyclic peptides from *Leonia cymosa*. *J. Org. Chem.* 65, 124–128
19 Bokesch, H.R. *et al.* (2001) A novel anti-HIV macrocyclic peptide from *Palicourea condensata*. *J. Nat. Prod.* 64, 249–250
20 Broussalis, A.M. *et al.* (2001) First cyclotide from *Hybanthus* (Violaceae). *Phytochemistry* 58, 47–51
21 Pallaghy, P.K. *et al.* (1994) A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides. *Protein Sci.* 3, 1833–1839
22 McDonald, N.Q. and Hendrickson, W.A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell* 73, 421–424
23 Daly, N.L. *et al.* (1999) Solution structure by NMR of Circulin A: a macrocyclic knotted peptide having anti-HIV activity. *J. Mol. Biol.* 285, 333–345
24 Gran, L. *et al.* (2000) *Oldenlandia affinis* (R&S) DC. A plant containing uteroactive peptides used in African traditional medicine. *J. Ethnopharmacol.* 70, 197–203
25 Daly, N.L. and Craik, D.J. (2000) Acyclic permutants of naturally occurring cyclic proteins: characterization of cystine knot and β-sheet formation in the macrocyclic polypeptide Kalata B1. *J. Biol. Chem.* 275, 19068–19075
26 Tam, J.P. *et al.* (1999) An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8913–8918
27 Hernandez, J-F. *et al.* (2000) Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry* 39, 5722–5730
28 Felizmenio-Quimio, M.E. *et al.* (2001) Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J. Biol. Chem.* 276, 22875–22882
29 Heitz, A. *et al.* (2001) Solution structure of the squash trypsin inhibitor MCoTI-II. A new family for cyclic knottins. *Biochemistry* 40, 7973–7983
30 Luckett, S. *et al.* (1999) High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.* 290, 525–533
31 Korsinczky, M.L.J. *et al.* (2001) Solution structures by ^{1}H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J. Mol. Biol.* 311, 579–591
32 Tang, Y-Q. *et al.* (1999) A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated α-defensins. *Science* 286, 498–502
33 Trabi, M. *et al.* (2001) Three-dimensional structure of RTD-1, a cyclic antimicrobial defensin from rhesus macaque leukocytes. *Biochemistry* 40, 4211–4221
34 Solbiati, J.O. *et al.* (1999) Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. *J. Bacteriol.* 181, 2659–2662
35 Martínez-Bueno, M. *et al.* (1998) Analysis of the gene cluster involved in production and immunity of the peptide antibiotic AS-48 in *Enterococcus faecalis*. *Mol. Microbiol.* 27, 347–358
36 Tran, D. *et al.* Homodimeric theta defensins from *Rhesus macaque* leukocytes: isolation, synthesis, anti-microbial activities and bacterial binding properties of the cyclic peptides. *J. Biol. Chem.* (in press)
37 Leonova, L. *et al.* (2001) Circular minidefensins and posttranslational generation of molecular diversity. *J. Leukocyte Biol.* 70, 461–464
38 Camarero, J.A. and Muir, T.W. (1997) Chemoselective backbone cyclization of unprotected peptides. *Chem. Commun.* 15, 1369–1370
39 Dawson, P.E. *et al.* (1994) Synthesis of proteins by native chemical ligation. *Science* 266, 776–779
40 Tam, J.P. and Lu, Y-A. (1998) A biomimetic strategy in the synthesis and fragmentation of cyclic protein. *Protein Sci.* 7, 1583–1592
41 Iwai, H. *et al.* (2001) Cyclic green fluorescent protein produced *in vivo* using an artificially split PI-*Pfu*I intein from *Pyrococcus furiosus*. *J. Biol. Chem.* 276, 16548–16554
42 Iwai, H. and Plückthun, A. (1999) Circular β-lactamase: stability enhancement by cyclizing the backbone. *FEBS Lett.* 459, 166–172
43 Perler, F.B. and Adam, E. (2000) Protein splicing and its applications. *Curr. Opin. Biotechnol.* 11, 377–383
44 Daly, N.L. *et al.* (1999) Chemical synthesis and folding pathway of large cyclic polypeptides: studies of the cystine knot polypeptide Kalata B1. *Biochemistry* 38, 10606–10614
45 Jackson, D.Y. *et al.* (1995) Enzymatic cyclization of linear peptide esters using subtiligase. *J. Am. Chem. Soc.* 117, 819–820
46 Craik, D.J. (2001) Plant cyclotides: circular, knotted peptide toxins. *Toxicon* 39, 1809–1813
47 Scott, C.P. *et al.* (1999) Production of cyclic peptides and proteins *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13638–13643
48 Tam, J.P. *et al.* (2000) Membranolytic selectivity of cysteine-stabilized cyclic protegrins. *Eur. J. Biochem.* 267, 3289–3300
49 Tam, J.P. *et al.* (2000) Marked increase in membranolytic selectivity of novel cyclic tachyplesins constrained with an antiparallel two-β strand cystine knot framework. *Biochem. Biophys. Res. Commun.* 267, 783–790
50 Camarero, J.A. *et al.* (2001) Rescuing a destabilized protein fold through backbone cyclization. *J. Mol. Biol.* 308, 1045–1062
51 Goldenberg, D.P. and Creighton, T.E. (1983) Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 165, 407–413
52 Camarero, J.A. *et al.* (1998) Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization. *Angew. Chem., Int. Ed. Engl.* 37, 347–349
53 Evans, T.C., Jr *et al.* (2000) Protein *trans*-splicing and cyclization by a naturally split intein from the dnaE gene of *Synechocystis* species PCC6803. *J. Biol. Chem.* 275, 9091–9094