Available online at www.sciencedirect.com

SciVerse ScienceDirect

ELSEVIER

Current Opinion in Chemical Biology

Cyclotide biosynthesis

David J Craik and Uru Malik

Cyclotides are bioactive macrocyclic peptides from plants that are characterized by their exceptional stability and potential applications as protein engineering or drug design frameworks. Their stability arises from their unique cyclic cystine knot structure, which combines a head-to-tail cyclic peptide backbone with three conserved disulfide bonds having a knotted topology. Cyclotides are ribosomally synthesized by plants and expressed in a wide range of tissues, including leaves, flowers, stems and roots. Here we describe recent studies that have examined the biosynthesis of cyclotides and in particular the mechanism associated with post-translational backbone cyclization.

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

Corresponding author: Craik, David J (d.craik@imb.uq.edu.au)

Current Opinion in Chemical Biology 2013, **17**:546–554

This review comes from a themed issue on **Mechanisms**

Edited by **Hung-wen Liu** and **Tadhg Begley**

For a complete overview see the Issue and the Editorial

Available online 25th June 2013

http://dx.doi.org/10.1016/j.cbpa.2013.05.033

Introduction

Cyclotides [1•] are head-to-tail cyclic mini-proteins of approximately 30 amino acids that incorporate a cystine knot motif built from their six conserved cysteine residues. They occur in a range of plant families, including the Rubiaceae (coffee), Violaceae (violet), Fabaceae (legume), Solanaceae (nightshade) and Cucurbitaceae (cucurbit). Dozens to hundreds of different cyclotides occur in leaves, flowers and other tissues of individual plants [2], where their natural function is thought to be host defense agents [3,4]. They were originally discovered on the basis of the pharmaceutically interesting bioactivities of extracts from a few medicinal Rubiaceae and Violaceae plants, but the number of known cyclotides has dramatically increased over the last few years since focused discovery efforts have been directed at them [5•,6•,7•,8•,9,10]. Cyclotides have attracted interest because of their unique topology, exceptional stability [11], and potential applications as protein engineering or drug design frameworks [12–16]. It has also been reported recently that they might have immunosuppressive properties [17].

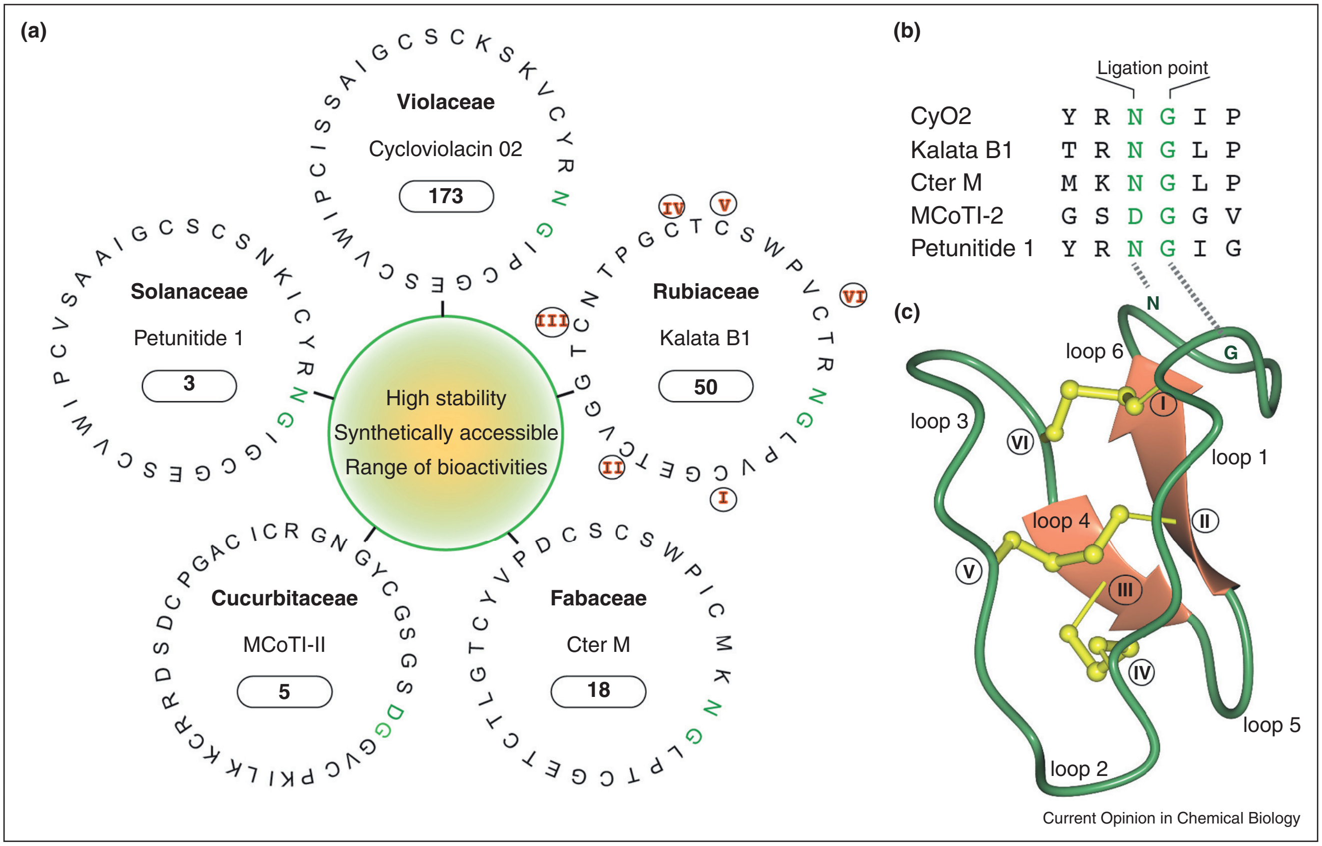
Cyclotides are just one class of a growing number of circular proteins found in nature [5•,18•,19,20,21••,22]. Until a decade or so ago, most known naturally occurring cyclic peptides were small (<12 amino acids) and thought to be non-ribosomally synthesized, as exemplified by the well-known fungal cyclic peptide, cyclosporine [23]. However, it is now clear that all kingdoms of life produce ribosomally synthesized cyclic peptides [24]. They range in size from around six to 80 amino acids [25] and so far little is known about their biosynthesis, other than that they are gene-encoded, expressed as precursor proteins and post-translationally processed to produce a cyclic backbone. Cyclotides are the largest family of circular proteins, with more than 250 sequences reported so far [26]. In this article we review progress in delineating their biosynthesis, focusing on articles published over the last two years. Other recent reviews provide background on their discovery, characterization and applications [5•,27–32], and CyBase [26,33] is a regularly updated database where information on cyclotides and other circular proteins can be found.

Figure 1 shows selected cyclotide sequences, highlights some properties that make them of interest, and summarizes their prevalence among the five major plant families in which they have been discovered so far. It also shows the structure of the prototypic cyclotide, kalata B1, and gives an overview of the standard nomenclature used to define the six conserved Cys residues (I–VI) and six backbone loops (1–6) between them. Figure 1 also highlights the location of the ligation point where the proto-termini are joined to form the circular backbone, and shows the cystine knotted structure at the core of cyclotides, linking Cys^{I}–Cys^{IV}, Cys^{II}–Cys^{V} and Cys^{III}–Cys^{VI}. Cyclotides have been divided into three subfamilies, the Möbius, bracelet and trypsin inhibitor subfamilies, on the basis of structures and activities [32].

Cyclotide precursor sequences

The first indication that cyclotides are ribosomally synthesized came when the precursors for several cyclotides, including kalata B1, B2, B6 and B7, were deduced from a cDNA library prepared from *Oldenlandia affinis* (Rubiaceae) leaves [34•]. This African plant was the first species in which cyclotides were discovered, on the basis of their role as uterotonic agents in a medicinal tea used to accelerate childbirth [35]. Since then a large number of cyclotides and their precursor sequences have been reported from Rubiaceae and Violaceae plants [8•,36,37] and some common features of the gene architectures have become apparent. The most recently discovered cyclotide precursors are from the Solanaceae [7•], Fabaceae

Figure 1



Cyclotide distribution in the plant kingdom, sequences and structures. **(a)** Representative examples of cyclotide sequences from each of the five major plant families expressing cyclotides are shown; the total number of cyclotides reported in each family is in a box within each circular sequence. **(b)** Cyclotides have highly conserved Gly and Asx (Asn or Asp) residues at their proto N-termini and C-termini, respectively; a selection of sequences flanking the joined termini is shown to illustrate the conserved residues flanking the biosynthetic ligation site that forms the cyclic backbone. **(c)** The structure of the prototypic cyclotide kalata B1 is shown to illustrate the locations of the six conserved Cys residues (I–VI) and backbone loops (1–6).

[38••,39••], and Cucurbitaceae plant families [40•], some species of which are widely used in human and animal diets. The discovery of cyclotide genes in plants from these families places additional prominence on cyclotides; in particular, suggesting the possibility that humans might have routinely ingested cyclotides, thus providing confidence that they might have a good safety profile in pharmaceutical applications. Furthermore, the new precursor protein architectures suggest diversity in the way cyclotides are biosynthesized.

Figure 2 shows representative structures of the precursor proteins of cyclotides from various plant families. The precursors typically incorporate an endoplasmic reticulum (ER) signal sequence, which is followed by an N-terminal propeptide, part of which is often classified as an N-terminal repeat (see below). This is followed by the cyclotide domain and a C-terminal pro-peptide. In some cases elements of this modular structure are repeated, namely the cyclotide-encoding domain and its flanking N-terminal and C-terminal regions, referred to as the N-terminal repeat (NTR) and C-terminal repeat (CTR) respectively. For example, the first cyclotide gene discovered, *Oak1* (*O. affinis* kalata B1) encodes a single copy of kalata B1, whereas the *Oak4* gene encodes three copies of kalata B2. Interestingly, different genes can produce the same cyclotide. For example *Oak1* and *Oak10* both produce kalata B1. Multiple genes are present in a given plant, and combined with the multidomain nature of some of the genes, this accounts for the large number of cyclotides expressed by an individual plant. Generally there is little overlap in cyclotide content between different species, but in some cases the same cyclotide is produced in different species. Kalata B1, for example, occurs in several plant species in addition to *O. affinis*, including *Viola odorata*, whose Voc1 precursor is shown in Figure 2.

As is apparent from the precursor sequences in Figure 2, a conserved feature of all cyclotides, besides the six Cys residues, is an Asn or Asp residue at the cyclotide proto C-terminus. Also highly conserved is a small amino acid