Various mechanisms in cyclopeptide production from precursors synthesized independently of non-ribosomal peptide synthetases

Wenyan Xu1,2,3, Liling Li1, Liangcheng Du2*, and Ninghua Tan1*

1State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, China
2Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
3Graduate University of Chinese Academy of Sciences, Beijing 100049, China

*Correspondence address. Tel: +86-871-5223800; Fax: +86-871-5223800; E-mail: nhtan@mail.kib.ac.cn (N.T)/Tel: +1-402-472-2799; Fax: +1-402-472-2798; E-mail: ldu@unlserve.unl.edu (L.D.)

An increasing number of cyclopeptides have been discovered as products of ribosomal synthetic pathway. The biosynthetic study of these cyclopeptides has revealed interesting new mechanisms for cyclization. This review highlighted the recent discoveries in cyclization mechanisms for cyclopeptides synthesized independently of non-ribosomal peptide synthetases, including endopeptidase-catalyzed cyclization, intein-mediated cyclization, and peptide synthetase-catalyzed cyclization. This information may help to design hybrid ribosomal and non-ribosomal biosynthetic systems to produce novel cyclopeptides with various bioactivities.

**Keywords** ribosomal peptide; cyclopeptides; non-ribosomal peptide synthetase; cyclization mechanism

Received: April 5, 2011        Accepted: May 19, 2011

Introduction

Cyclopeptides or cyclic peptides represent a family of fantastic compounds with complex structures and various bioactivities, such as antibiotic gramicidin S [1], antitumor RA-VII [2], uterotonic polypeptide kalata B1 [3], antibacterial albonoursin [4,5], and cytotoxic patellamide D [6]. The head-to-tail cyclized backbone protects cyclopeptides against proteolytic cleavage [7]. Most cyclopeptides are products of non-ribosomal peptide synthetases (NRPSs) [8]. These peptides are cyclized by thioesterase (TE) domains located in C-terminal of NRPSs [9]. The cyclization occurs by nucleophilic attacking of nucleophile (e.g. amine group, hydroxyl group, or thiol group) against carbonyl group of ester bond linking to the TE domain. More and more cyclopeptides have been confirmed to be ribosomal products or at least independent of NRPSs, including some cyclopeptides that were previously considered as non-ribosomal peptides (NRPs) [10–12], such as microviridin [12].

Till date, pathways both of NRPS and ribosomal have been discussed [11,13] including the cyclization mechanisms in the NRPS pathway [14]. Some genetic information of these ribosomally synthesized cyclopeptides has been discussed, such as gene clusters of cyanobactins [15] and cDNA sequences of cyclotides [16]. However, no review has specifically discussed the cyclization mechanisms of peptides synthesized independently of NRPSs, although several intriguing mechanisms have been reported for the cyclization of these cyclopeptides, e.g. endopeptidase-catalyzed cyclization of cyclotides [17] and cyanobactins [18], artificial cyclization with a permuted intein in cyanobacterium [19] and peptide synthetase-catalyzed cyclization of albonoursin [20]. These mechanisms of enzymatic cyclization in the ribosomal biosynthetic pathway have extended the knowledge of enzymatic reactions and brought the ribosomal and non-ribosomal pathways together. To modify the ribosomal system is easier than modifying the non-ribosomal system, because the NRPSs are too large to be processed by gene operation. This knowledge will help to develop new technologies in combinatorial biosynthesis and bioengineering to produce novel bioactive compounds.

Endopeptidase-catalyzed Cyclization

Endopeptidases present a family of enzymes, which catalyze the hydrolysis of the peptide bond (or breaking the peptide bond in other words). Now some evidences show that the enzymes of this family also catalyze the transpeptidation by forming a peptide bond, which includes cyclization. Cyclotides and cyanobactins are good examples of endopeptidase-catalyzed cyclization.

Cyclotides belong to plant cyclopeptides type VIII [21], which is a family of mini disulfide-rich peptides derived...
by plants with ≈30 amino acids and contains a unique protein motif cyclic-cystine-knot. This motif, which includes three disulfide bonds, together with cyclic backbone makes cyclotides exceptionally stable. The first cyclotide kalata B1 was discovered in the African traditional herb Oldenlandia affinis that had been used as uterotonic medicine [22], which showed that cyclotides could be used in clinic safely. Till 2009, various cyclotides with activities of hemolysis, anti-HIV, antimicrobe, cytotoxin, and insecticide have been reported [23]. A study confirmed that the linear analogues lacked bioactivity even if the N-terminals were blocked by the acetyl group [24]. So the cyclic backbone is very important to cyclotides’ bioactivities.

Cyclotides were confirmed as gene-coded products [25] and spliced from larger propeptides [26]. The cyclization of backbone occurs after the forming of a cyclic cystine knot in vivo [27], although chemical synthetic research shows that the cyclic backbone is preferred for the generation of cyclic compound in vitro [28,29]. The cyclization of the backbone is catalyzed by asparaginyl endopeptidase (AEP) [30,31]. There are six conserved residues (XXNGLP), which are recognized by AEPs. The reaction is initiated by the electron transferring from histidine to cysteine. Then the thio group will attack the carbonyl group of the asparagine, break the peptide bond, and link the N-terminal of propeptide to the enzyme. The amino group of glycine accepts the proton from histidine and the C-terminal germin-like protein (GLP) tripeptide leaves. Then the N-terminal propeptide folds and the first three residues, which are conserved GLPs, fit into recognizing site S1’, S2’ and S3’. The amino group of glycine residue initiates a nucleophilic attack to form the peptide bond and completes the cyclization. This research gives a novel mechanism of peptidase, in which a pair of reversing activities occurs in a single catalytic site. Peptidases were considered to catalyze the reaction of either breaking or forming a peptide bond [32]. Although the activities both of hydrolysis and transpeptidation share a similar catalytic center (catalytic triad Asp-Ser-His), the nucleophilic attacking groups are different in these two reactions, which are H₂O in hydrolysis but NH₂ in transpeptidation. If we may define the cyclization of cyclotides as intramolecular transpeptidation, this research is the first example to catalyze both of two reactions in order.

Cyanobactins are small bioactive cyclic peptides produced by cyanobacteria [13]. These cyclic compounds show a wide range of bioactivities such as anti-tumor, anti-HIV, and anti-viral [33]. Both cyclotides and cyanobactins are ribosomal backbone cyclic or N–C cyclic, while the other ribosomal cyclic peptides are side-chain cyclized. Distinguished from cyclotides, cyanobactins are cyclized by subtilisin-like protease which belongs to non-specific serine endopeptidase [18] instead of AEPs [25]. The precursor proteins encoded one or more cyanobactins, which were flanking by the putative recognition sequences [13,20,34]. According to this demonstration, the two proteases encoded by cyanobactin gene cluster were proposed to carry out different functions. One was confirmed to cleave the recognition sequence in the N-terminal, and the other to cleave the C-terminal and catalyze the cyclization [13,20].

In 2009, the cyclization mechanism of cyanobactin with PatG was reported by Lee et al. [20]. An artificial substrate of linear peptide QGGRGDWP-AYDGE could be split and cyclized into cyclo-(QGGRGDWP) by PatG in vitro. PatG is the first cyclase involved in the ribosomal pathway. This enzyme shares the same catalytic triad (Asp-Ser-His) with serine protease. PatG cut AYDGE off from the precursor and loaded QGGRGDWP into the hydroxyl group of Ser in the catalytic triad. The amine group of Gln attacked the carboxyl group as nucleophile to form the cyclic product. No cofactors, extra metals, or energy were required in the reaction. This mechanism is like TE cyclization in NRPS, except that the reaction breaks a peptide bond instead of TE bond.

The substrate selectivity of PatG was fairly flexible. Over 30 natural peptide sequences and over 60 natural products could be cyclized by PatG in vivo or in vitro, although the turnover number was as low as 1 per enzyme molecule per day. The authors also excluded the possibility of spontaneous cyclization by feeding the peptide of QGGRGDWP as the substrate [20]. Later in 2010 [18], McIntosh et al. introduced non-proteinogenic amino acids and polyketide-like linkers in the artificial substrates. Results showed that PatG generated products in lengths from 6 to 11 residues, including non-proteinogenic amino acids and polyketide-like linkers. Similar to TE domains, unnatural substrates triggered the hydrolysis of PatG. This result makes it possible to hybrid the ribosomal pathway and non-ribosomal pathway and produces more novel interesting bioactive compounds.

Generally, in the endopeptidase-catalyzed cyclization, the precursors of cyclopeptides are loaded to the catalyst site companying the release of C-terminal fragments. An N-to-X acyl shift transpepsterification cyclizes the backbones and yields the products (Fig. 1).

**Intein-mediated Cyclization**

Inteins are internal sequences of proteins that are self-excluded from larger precursor proteins while ligating the remaining parts collectively termed as exteins [35]. There are three types of intein-splicing, which are canonical splicing, intermolecular trans-splicing, and cyclization using permuted inteins. Canonical splicing gives a linear intein with a succinimide residue at the C-terminal. The reaction...
Peptide Synthetase-catalyzed Cyclization

Cyclodipeptides were reported as the products of peptide synthetase by ligating two amino acids together and then cyclizing, which were discovered in diverse bacterial [43], fungi [44], animals [43], and plants [19]. These compounds were believed to be formed from protein hydrolysates as by-products of protein degradation or produced by NRPSs [43]. Some were reported as the products of spontaneous cyclization in the NRPS pathway [43,45]. The isolation of albonoursin biosynthetic gene cluster [43] elucidated a
novel non-ribosomal mechanism of cyclodipeptide formation, which was totally independent of NRPSs. A peptide synthetase AlbC was present in this gene cluster, and it was believed to catalyze the formation of cyclodipeptide cyclo-(L-Phe-L-Leu). In further research \[46\], AlbC was confirmed as a synthetase of cyclodipeptide using tRNA charged amino acids as substrates (Fig. 3). The purified AlbC, which was expressed in $E. \text{coli}$ and secreted into a medium, showed the activities of ligation and cyclization to a range of amino acids $\textit{in vitro}$. The

![Figure 2 Mechanism for intein-mediated cyclization of pseudostellarin F](image)

The expressed precursor folded to form an active protein ligase (reaction 1); an N-to-S acyl shift was catalyzed to give a thioester intermediate (reaction 2) that underwent transesterification to form a lariat intermediate (reaction 3); Succinimide formation released the cyclic product as a lactone (reaction 4); the final cyclopeptides was generated by an X-to-N acyl shift (reaction 5). The illusion was modified from references [19,40].

![Figure 3 Cyclization of cyclodipeptides via cyclodipeptide synthase (CDPS)](image)

Cyclodipeptide is synthesized by CDPS using two amino acyl tRNA as substrates companying with releasing two unloaded tRNAs. The detail of catalytic mechanism is still unknown, which was marked as a question marker in the illustration.
same activity was also detected in the enzymatic assay with other related proteins. The authors named these proteins as cyclodipeptide synthases (CDPSs). CDPSs were distinguished from other enzymes using amino acyl-tRNA as a substrate to generate peptide bonds by additional cyclization activity. CDPSs use amino acyl-tRNA to generate two peptide bonds and form a cyclodipeptide, while the other enzymes only generate one peptide bond between the amino acyl group of amino acyl-tRNA donor and the N-terminal amine group of peptide or protein acceptor.

There is a gap between the in vitro activity and in vivo function, which is the specificity of substrate recognition. AlbC showed a tolerant specificity to different amino acyl-tRNA in vitro while only cyclo-(L-Phe-L-Leu) could be generated in vivo. von Dohren [47] proposed that the recognition site might locate on some special tRNA moieties. Although the mechanism of CDPSs remains to be exploited, this new class of enzymes opens a door to the utilization of tRNA charged amino acids to generate a wide range of cyclodipeptides as substrates for more modification which may find novel bioactive compounds.

Final Remark

An increasing number of natural peptides have been confirmed to be synthesized independently of NRPS especially by the ribosomal pathway [11,48]. The understanding of their biosynthetic mechanisms is important for the utility of these compounds, especially the small-molecule peptides that show the outstanding potential as anti-infectious and antitumor agents. The non-ribosomal pathway usually gives complex products while the ribosomal pathway is easier to be modified by genetic operations. The cyclases discovered independent of NRPSs might combine the advantages of both ribosomal and non-ribosomal pathways and provide more novel complex drugs or drug-leads with a simpler molecular biological technique.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (30725048, U1032602, and 91013002), the Chinese Academy of Sciences (KSCX2-YW-R-177), the State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Faculty Seed Grant from Nebraska Research Council, and National Institute of Health of USA (RR015468-01). W.X. was supported in part by China Scholarship Council.

References