Minireviews:
Chemical Synthesis of Circular Proteins

James P. Tam and Clarence T. T. Wong

doi: 10.1074/jbc.R111.323568 originally published online June 14, 2012

Access the most updated version of this article at doi: 10.1074/jbc.R111.323568

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Read an Author Profile for this article at
http://www.jbc.org/content/suppl/2012/08/02/R111.323568.DCAuthor_profile.html

This article cites 61 references, 14 of which can be accessed free at
http://www.jbc.org/content/287/32/27020.full.html#ref-list-1
MINIREVIEW

Chemical Synthesis of Circular Proteins

Published, JBC Papers in Press, June 14, 2012, DOI: 10.1074/jbc.R111.323568
James P. Tam and Clarence T. T. Wong
From the School of Biological Sciences, Nanyang Technological University, Singapore 637551

Circular proteins, once thought to be rare, are now commonly found in plants. Their chemical synthesis, once thought to be difficult, is now readily achievable. The enabling methodology is largely due to the advances in entropic chemical ligation to overcome the entropy barrier in coupling the N- and C-terminal ends of large peptide segments for either intermolecular ligation or intramolecular ligation in end-to-end cyclization. Key elements of an entropic chemical ligation consist of a chemoselective capture step merging the N and C termini as a covalently linked O/S-ester intermediate to permit the subsequent step of an intramolecular O/S-N acyl shift to form an amide. Many ligation methods exploit the supernucleophilicity of a thiol side chain at the N terminus for the capture reaction, which makes cysteine-rich peptides ideal candidates for the entropy-driven macrocyclization. Advances in desulfurization and modification of the thiol-containing amino acids at the ligation sites to other amino acids add extra dimensions to the entropy-driven ligation methods. This minireview describes recent advances of entropy-driven ligation to prepare circular proteins with or without a cysteiny1 side chain.

Circular proteins and their smaller versions, cyclic peptides, have a characteristic head-to-tail or end-to-end peptide backbone structure. The absence of both N and C termini in these macrocycles confers resistance to exopeptidase and heat degradation, enhances their conformational stability, and maximizes epitope display upon their circular contiguous sequences for interactions with other molecules. These advantages have provided incentives to engineer circular proteins by grafting linear bioactive peptides or epitopes into various structural scaffolds for therapeutic applications (1).

For the purpose of this minireview, we refer to a cyclic peptide of >15 amino acids as a “circular protein” or “miniprotein.” This arbitrary cutoff point has two justifications. First, many cyclic peptides of <15 amino acids are derived from nonribosomal synthesis, and the substrate specificity of a cyclase enzyme such as tyrocidine thioesterase could accommodate precursors of 6–14 amino acids in length (2). Second, circular proteins are gene-encoded and processed by specific enzymes from their linear precursors containing a signal peptide and one or more precursors (3). Cyclic peptides and circular proteins are frequently found in bacteria (microcin), fungi (cyclosporin), animals (s-defensins), and more commonly, plants (cyclotides) (4–7).

From a synthetic standpoint, chemical synthesis of circular proteins has been a formidable challenge using the traditional enthalpic activation methods, which require partially or globally protected linear precursors and a strong enthalpic activation of the C-terminal residue for the cyclization reaction (Fig. 1A). Strong activation of the C-terminal moiety is necessary to overcome the entropy barrier in the coupling reactions and often leads to epimerization of the C-terminal amino acid residue and oligomerization to dimers and trimers. Work prior to 1997 showcased the challenges associated with enthalpy-driven cyclization of peptides of <15 amino acids. In 1997, we reported the total synthesis of circular proteins of 31 amino acids, cyclotides circulin B and cyclopsychotride, and in the following year, two other cyclotides (8, 9). These studies represented a breakthrough because they were the first reports on a successful chemical synthesis of naturally occurring circular proteins using an entropy-driven ligation chemistry, which is conceptually different and operationally much simpler than the conventional enthalpy cyclization method.

Many review articles have been published over the last few years, with a majority articulating the occurrence, chemistry, and biological functions of cyclic peptides (10–13). Here, we will focus on contemporary entropy-driven ligation chemistry for the synthesis of circular proteins.

Entropy-driven Ligation Chemistry

During the 1990s, there was a paradigm shift in the synthesis of large peptides and proteins (14–18). The sea change was driven by the discovery of entropic activation in convergent peptide synthesis using unprotected peptides as building blocks to enable an efficient coupling reaction (ligation) of the N and C termini of two peptide segments (intermolecular ligation) or a single peptide segment (macrocyclization). The conceptual difference between an entropic and an enthalpic ligation is that the entropic ligation is proximity-driven to overcome the entropy barrier, merging the N and C termini as a covalent intermediate, usually as an O- or S-ester, to permit an intramolecular proximity-driven O-N or S-N acyl shift to form an amide (Fig. 1B). As such, it eliminates the risk of the C-terminal epimerization in peptide synthesis, a major side reaction that results in a side product that is difficult to remove by chromatographic methods.

Although the principles of entropy-driven reactions are well established in the cyclization of small molecules in organic synthesis, Kemp and co-workers are pioneers and strong advocates of entropic chemical ligation in the arena of large molecules, including peptide synthesis. In the 1980s, they demonstrated its feasibility by placing both peptide segments on rigid organic templates to facilitate an O-N acyl shift in organic solvents to form a peptide bond (14). Kemp’s template-based ligation approach was limited to general applications due to the use of a tricyclic organic template and the attendant slow rates of the
O-N acyl shift reactions mediated by a large 12-member ring intermediate. In 1994, our laboratory reported an entropy-driven chemical ligation method. Without the use of an organic template, we used an N-terminal Cys-, Ser-, or Thr-containing peptide segment with a C-terminal fragment bearing an ester aldehyde (15, 16). The ligation was performed in water, without the use of protecting groups or a coupling reagent. The key element of our approach is a chemoselective capture, forming a covalent ester between two peptide fragments to enable an intramolecular O-N acyl shift mediated by a small five-member ring intermediate to form an amide bond and, in our case, a pseudo-proline bond. In the same year, Kent and co-workers (17) reported a cysteine-based ligation with another segment bearing a C-terminal thioester and an S-N acyl shift through a five-member ring intermediate to form an amide bond. Because cysteine is regenerated at the ligation site, they called their ligation method native chemical ligation. The supernucleophilic cysteinyl thiol of an unprotected peptide permits a rapid thiol-thioester exchange at a basic pH to form a covalent thioester intermediate, and the rate of S-N acyl shift is faster than the corresponding O-N acyl shift. Both factors contribute to the effectiveness of native chemical ligation, which our laboratory confirmed shortly after Kent’s communication (18). Indeed, the use of a C-terminal thioester in the chemoselective capture step, first reported by Wieland and Schneider in 1953 (19), is an excellent choice for chemical ligation. The combined work of Kemp’s group and the early ligation studies published in the 1994 and 1995 period provided the conceptual framework of the entropic ligation approach for nearly all subsequent chemical ligation methods developed in the past 16 years (14–18). Interestingly, we found that an N-terminal His-containing segment could also be exploited for chemical ligation because of the exceptional reactivity of its nucleophilic imidazole side chain. Under acidic conditions, the imidazole-based capture of another peptide segment containing a C-terminal thioester derivative affords a covalent acyl imidazole intermediate, leading to an N-N acyl shift to form a histidine peptide bond at the ligation site (20).

Two congruent but independent developments accelerated the acceptance of entropy-driven ligation during the 1990s. Xu et al. (21) found that the intein-mediated protein splicing process could undergo four acyl shift reactions, three of which are catalyzed by the intein enzyme in the early stages. The final acyl shift involving either an O-N or S-N acyl shift to form the peptide bond between the spliced extein fragments is uncatalyzed and spontaneous. This O-N or S-N acyl shift reaction is also mediated by a five-member ring intermediate, similar to the entropic chemical ligation reactions described in the 1994–1995 articles (15–18). In 1998, Severinov and Muir (22) developed “expressed protein ligation,” a semibiosynthetic approach for protein synthesis to generate a peptide thioester using the intein-mediated splicing mechanism. The ability to generate building blocks of the N-terminal Cys-containing and C-terminal thioester-containing segments by either chemical or recombinant methods further facilitates protein synthesis of all
MINIREVIEW: Chemical Synthesis of Circular Proteins

TABLE 1

Summary of entropic ligation methods

<table>
<thead>
<tr>
<th>Starting materials</th>
<th>N-terminal</th>
<th>C-terminal</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site-based ligation methods</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Pseudo-Pro</td>
<td>Cys/Ser/Thr</td>
<td>Ester aldehyde</td>
<td>15, 16, 55</td>
</tr>
<tr>
<td>2. Cys</td>
<td>Cys</td>
<td>Thioester</td>
<td>17, 18</td>
</tr>
<tr>
<td>3. His</td>
<td>His</td>
<td>Perthioester</td>
<td>20</td>
</tr>
<tr>
<td>4. Met</td>
<td>Hcy</td>
<td>Thioester</td>
<td>35</td>
</tr>
<tr>
<td>5. Ala</td>
<td>Cys</td>
<td>Thioester</td>
<td>38</td>
</tr>
<tr>
<td>6. Phe</td>
<td>6-Mercaptob-Phe</td>
<td>Thioester</td>
<td>40</td>
</tr>
<tr>
<td>7. Val</td>
<td>Penicillamine</td>
<td>Thioester</td>
<td>41</td>
</tr>
<tr>
<td>8. Val</td>
<td>γ-Mercapto-Val</td>
<td>Thioester</td>
<td>42</td>
</tr>
<tr>
<td>9. Ser/Asp</td>
<td>Thiourea Ser/Asp</td>
<td>Thioester</td>
<td>56</td>
</tr>
<tr>
<td>10. Pseudo-Lys</td>
<td>Cys</td>
<td>Thioester</td>
<td>57</td>
</tr>
<tr>
<td>11. Gly</td>
<td>N'-Auxiliary Gly</td>
<td>Thioester</td>
<td>58</td>
</tr>
<tr>
<td>12. Lys</td>
<td>γ-Mercapto-Lys</td>
<td>Thioester</td>
<td>36</td>
</tr>
<tr>
<td>13. Ser/Thr</td>
<td>Ser/Thr</td>
<td>Ester salicylaldehyde</td>
<td>59</td>
</tr>
</tbody>
</table>

| Other ligation methods | | | |
| 14. Ag-assisted Ser/Thr | Ser/Thr/Asn/Gly | Thioester | 28 |
| 15. Staudinger | Azido amino acids | Phosphanyl thioester | 60, 61 |
| 16. Subtiligase | Met/Ala/Lys/Arg | Glycolated phenyl ester | 44 |

- Thiazolidine/oxazolidine ligation without (Method 1) or with (Method 13) a subsequent chemical manipulation.
- Native chemical ligation.
- S-Alkylation after ligation.
- Desulfurization after ligation.
- Removal of thiosugar (Method 9), O-auxiliary (Method 13), and N-auxiliary (Method 11) group after ligation.
- Deprotection step required after ligation.

shapes and sizes by the entropic chemical ligation approach for various biological studies (23).

Chemical Ligation and Macrocyclization Based on Entropic Ligation

With the development of entropy-driven ligation, the synthesis of circular proteins became approachable in the late 1990s. This was timely, as many cysteine-rich peptides consisting of 30–60 amino acids had been identified from natural sources, and their pharmacological evaluations were limited by suitable methods of chemical synthesis (24, 25). Their characteristic signature of a multiple-cysteine structure, which is a liability in conventional chemical synthesis, becomes an asset in entropic chemical ligation methods. An example is the family of cyclotides, which contain 28–37 residues, six cysteines, and a circular peptide backbone. For cyclotides, there are six possible X-Cys ligation sites, and the choice is often governed by choosing the least hindered amino acid to serve as the C-terminal thioester. Thus, the order preference for X is Gly > Ala > Leu, Phe, Ser > Thr, Val, Ile >> Pro. The C-terminal thioesters of Asp and Glu are unstable and seldom used for ligation reactions.

The thiol-thioester exchange reaction in the capture reaction of a thiol-based ligation also plays an essential role in facilitating the entropy-driven macrocyclization of cysteine-rich peptides (Fig. 1C). We found that the internal cysteines enhance the cyclization rates of cyclotides kalata B1, cyclosporictide, and circulin A and B, which contain six cysteines (26). The cyclization rate was fast under denaturing conditions of 6 M guanidine HCl. In the enthalpic activation, the cyclization is performed at high dilution to prevent oligomerization. In contrast, we observed that there was no dimer or oligomer formation even at peptide concentrations up to 0.5 mM. We proposed that the rate enhancement and lack of oligomerization of the Cys-rich cyclotides could be attributed to the thia-zip cyclization mechanism (26, 27). Thia-zip cyclization involves two reactions: the reversible thiol-thioester exchanges through intramolecular transthioesterifications as thiolactones and the irreversible S-N acyl migration of the N-terminal thiolactone to the lactam (amide). The facile and reversible intramolecular transthioesterifications lead to the ring expansion as discrete thiolactone intermediates, pulling the two ends successively into close proximity as the end-to-end N-terminal thiolactone to permit an S-N acyl shift via a five-member ring to form an end-to-end circular protein. As a result, the thia-zip-assisted cyclization, through small discrete intermediates, is entropy-driven and more efficient than the corresponding one-step end-to-end cyclization. Along the same line of reasoning, we demonstrated the use of silver ions in assisting cyclization between the N-terminal Ser/Thr/Asn/Gly and a C-terminal thioester. In this approach, the silver ion captures and brings two ends into close proximity for an acyl transfer to occur. Furthermore, the silver ion forms a complex with the thiol group on the thioester, making it a better leaving group in the acyl transfer reaction (Table 1, Method 14) (28).

Given the advances of chemical ligation, the synthesis of circular proteins, both designed and naturally occurring, became routinely achievable. Naturally occurring circular proteins, including cycloviolacin, O2, hedyotide B1, and rhesus θ-defensin-1, have been successfully prepared (6, 29, 30). Several groups, including Craik and Camarero, have extensively applied cysteine-based ligation to prepare cyclotides and their analogs for biochemical and biophysical studies (31–33). For designed circular proteins, Yu et al. (34) employed cysteine ligation to introduce an end-to-end cyclic backbone into an α-defensin. The circularized defensin leads to improved stability against exopeptidases, increased conformational stability, and enhanced tolerance to high salt sensitivity in antimicrobial assays.

Cysteine-free Ligation and Macrocyclization

The occurrence of cysteine (1.7%) in proteins is lower than the other amino acids, and its absence is common in naturally occurring cyclic peptides. Furthermore, there is also a need to
increase the flexibility of entropic ligation methods that use non-cysteinyl N-terminal amino acids. Over the past 16 years, a suite of ligation methods has emerged to meet these needs, and they are summarized in Table 1. Many employ a combination strategy of entropic ligation and a follow-up chemical modification step. They include the use of N-terminal cysteine mimetics (Table 1, Methods 4–8 and 12) or thiol auxiliary groups on or near the N-terminal amino acids (Methods 9 and 11). These methods generally adhere to the principle of the thiol-based ligation but use various forms of N-terminal thiols to facilitate the capture reactions by thiol-thioester exchange reactions with C-terminal esters and then an S–N acyl shift to form an amide. The resulting thiolate side chain or auxiliary group at the ligation site is followed by an additional step, either a chemical modification or removal to give rise to a different amino acid (Table 1).

An example of combining thiol-based ligation and chemical modification is methionine ligation. It employs homocysteine (Hcy)\(^2\) to replace Cys as the N-terminal amino acid (35). The S-alkylation of Hcy by a methylating agent at the ligation site smoothly converts it to methionine (Table 1, Method 3). The combination of using an N-terminal Cys or Hcy for ligation followed by an alkylation reaction has been extended to pseudo-Lys, pseudo-Asp, and pseudo-Glu (Table 1, Method 10) (36, 37).

A promising approach is the combination of thiol-based chemical ligation and desulfurization of the resulting thiol. Yan and Dawson (38) exploited cysteine-based ligation in the synthesis of a cyclic peptide and then converted the resulting cysteinyl residue to alanine through a metal-based desulfurization by Raney nickel or palladium/Al\(_2\)O\(_3\). Metal-based desulfurization of small proteins may not be practical, as it often requires a large excess of metal, which produces side reactions and irreversible peptide loss due to aggregation and absorption onto the metal surfaces. To address these limitations, Wan and Danishefsky (39) developed a nonmetal desulfurization of cysteine via reaction with a free radical using tris(2-carboxyethyl)phosphine, a thiol, and the radical initiator 2,2’-azobis(2-(imidazolin-2-yl)propane) dihydrochloride in water. Desulfurization was extended to N-terminal \(\beta\)-mercaptophenylalanine, \(\beta\)-mercaptovaline (penicillamine) and \(\gamma\)-mercaptovaline, and \(\gamma\)-mercaptoisine to afford Phe, Val, and Lys, respectively (Table 1, Methods 6–8 and 12) (36, 40–42). A limitation of this approach and the N-terminal auxiliary approach is that few such building blocks (except for Cys to Ala and penicillamine to Val) are commercially available and thus may require considerable synthetic expertise for their preparation.

Chemo-enzymatic Macrocyclization

The use of enzymes for peptide coupling reactions through reverse proteolysis was described by Bergmann and Fruton over 70 years ago (43). Over the past few decades, enzymatic coupling of peptide segments has been explored as an alternative to enthalpic ligation (44, 45). Jackson et al. (46) reported the use of subtiligase in peptide cyclization in 1995 (Table 1, Method 16). The linear peptide precursor was synthesized as a C-terminal glycolate phenylalanilamide ester for subtiligase recognition, which results in end-to-end backbone cyclization (46). Another chemo-enzymatic method was reported by Tan et al. (47), who exploited subtiligase to catalyze the hydrolysis or aminolysis of a peptide glycolate ester substrate as a hydrothiolease in the synthesis of peptide thyoacids for peptide ligation. In the circular trypsin inhibitor family, trypsin can play the role as a cyclase to cyclize a linear synthetic precursor to a circular protein such as the sunflower trypsin inhibitor (48) and the cyclic trypsin inhibitor MCoT1-I (49).

Oxidative Folding of Cysteine-rich Peptides

Because many circular proteins are cysteine-rich, the oxidative folding step in forming the correct connectivity of disulfide bonds is another key step in their synthesis. However, forming the native disulfide bonds under oxidative folding conditions remains empirical with unpredictable outcomes. Therefore, the chemoselective method was introduced for the synthesis of cyclotides in 1997 (8). With a pair of hydrofluoric acid-stable cysteine-protecting groups such as the acetamidomethyl group, it greatly reduces the number of disulfide isomers formed. This method had been used in the synthesis of many cysteine-rich peptides such as engineered protegrins and defensins. A variation of the chemoselective method was reported by Alewood and co-workers (50) in their work on conotoxins, in which they replaced a pair of cysteines with selenocysteine. As selenium is much more reactive than sulfur, the diselenide bond forms almost spontaneously to confer its selectivity over cystines.

In addition to chemoselective methods, global oxidative folding has also been employed in the folding of cysteine-rich peptides. This reaction involves concurrent reduction (S–S breakage) and oxidation (S–S formation) in the presence of both reducing and oxidizing agents, forming a redox pair that can reshuffle and rearrange the disulfide pairs into the native disulfide bonds and presumably one of the stable conformational states. The reaction is mediated by the thiol-thiol exchange reaction that takes place at basic pH to generate active thiols for the S–S exchange reactions. This method involves fewer synthetic steps than the chemoselective method.

The global oxidative folding of cysteine-rich peptides generally requires high dilution of peptides in a suitable combination of solvents to prevent aggregation and precipitation. It also requires maintaining a proper redox potential such as GSH and GSSG in a basic buffer to minimize the accumulation of dead end products (51–53). Interestingly, the use of a combination of aqueous and hydrophobic organic solvents appears to work well for the oxidative folding of cyclotides, which display an unusual hydrophobic side chain arrangement due to the bulky interior disulfide core. A hydrophobic organic solvent may prevent aggregation during the folding process, as the hydrophobic side chains are being externalized to the solvent surface in forming the cystine knot. Such a rationale was used successfully in the oxidative folding of the Möbius cyclotide kalata B1 by Daly et al. (54), who used an aqueous hydrophobic solvent combination containing 50% 2-propanol. Under such a condition, they found that much higher yields of cyclic oxidized cyclotides were obtained by allowing the backbone cyclization to proceed.

\(^2\)The abbreviation used is: Hcy, homocysteine.
prior to the folding compared with the reverse order of folding first and then oxidation. In general, this order of synthetic operation, cyclization first and then oxidation, has been employed for the synthesis of circular proteins irrespective of the disulfide-forming strategy. However, the use of the solvent combination containing 50% 2-propanol in the oxidative folding of the bracelet family of cyclotides was found to be unsuccessful. For the successful oxidative folding of the bracelet cyclotide cycloviolacin O2, Aboye et al. (29) reported the use of a high percentage of Me3SO and the detergent Brij 35 in a solvent combination to fold it with a reasonable yield. In a systematic study, various parameters for the oxidative folding of a bracelet cyclotide were examined, and the key parameter was found to be the concentration of an organic co-solvent. It appears that the more hydrophobic cyclotides may require a strong hydrophobic solvent combination to fold the circular protein successfully. Hedyotide B1, one of the most hydrophobic bracelet cyclotides, requires a combination of solvents containing 70% 2-propanol for its optimal folding (30).

Conclusions

The development of entropic ligation and macrocyclization protocols based on the thiol capture of an N-terminal cysteine enables the chemical synthesis of many circular proteins in recent years. The combination of ligation and chemical modifications would likely extend entropic coupling methods to include most of the N-terminal aliphatic amino acids as ligation sites in the macrocyclization of circular proteins. Thus, the availability of a suite of ligation methods would permit the chemical synthesis of circular proteins of any size for biochemical, biophysical, and pharmacological studies in the laboratory and for evaluation of biologics as potential drug candidates.

References

MINIREVIEW: Chemical Synthesis of Circular Proteins

cyclized protein with improved biological activity. J. Am. Chem. Soc. 121, 5597–5598
51. Gunasekera, S., Daly, N. L., Clark, R. J., and Craik, D. J. (2009) Dissecting the oxidative folding of circular cystine knot miniproteins. Antioxid. Redox Signal. 11, 971–980