Published in final edited form as:

Angew Chem Int Ed Engl. 2018 May 22; 57(21): 6150–6154. doi:10.1002/anie.201801299.

Cyclic and Lasso Peptides: Sequence Determination, Topology Analysis, and Rotaxane Formation

Prof. Dr. Monika Raj,

Department of Chemistry and Biochemistry, Auburn University, Auburn, AL 36830 (USA),
mzr0068@auburn.edu

Hader E. Elashal,

Department of Chemistry and Biochemistry, Seton Hall University, South Orange, NJ 07079 (USA)

Heidi E. Elashal, and

Department of Chemistry and Biochemistry, Seton Hall University, South Orange, NJ 07079 (USA)

Ryan D. Cohen

Department of Process Research and Development, Merck & Co., Inc., Rahway, NJ 07065 (USA)

Dr. Chuhan Zong¹ and Prof. Dr. A. James Link²

¹Department of Chemistry and Princeton University, Princeton, NJ 08544 (USA)

²Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544 (USA)

Abstract

A broadly-applicable chemical cleavage methodology to facilitate MS/MS sequencing was developed for macrocyclic and lasso peptides, which holds promise as exciting new therapeutics. Existing methods such as Edman degradation, CNBr cleavage, and enzymatic digestion are either limited in scope or completely fail in cleavage of constrained non-ribosomal peptides. Importantly, the new method was utilized for synthesizing a unique peptide-based rotaxane (both cycle and thread) from the lasso peptide, benenodin-1 C5.

Entry for the Table of Contents

Site-selective chemical cleavage approach for sequencing of cyclic peptides and for synthesis of a unique peptide-based rotaxane (both cycle and thread) from the lasso peptide, benenodin-1 C5.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

lasso peptide; rotaxane; cyclic peptide; selective chemical cleavage; sequencing

Natural product discovery has been at the root of modern drug design, and it continues to play a vital role as there is renewed focus on orally-available peptide therapeutics¹. In particular, macrocyclic peptides often exhibit desirable characteristics such as high receptor-binding activity and specificity, metabolic stability, and cell permeability^{1,2}. Examples of uniquely-structured bioactive peptides with exceptional stability are cyclosporine A (an immunosuppressant whose structure is rich in N-methylation)¹, microcin J25 (an antibacterial lasso peptide)³, and kalata B1 (a cyclotide with insecticidal activity)¹. However, discovery efforts have been hindered in part by sequencing difficulties^{1,4,5}.

Macrocyclic peptides are difficult to characterize due to their high resistance to proteases and incompatibility with Edman degradation^{6,7}. Moreover, tandem mass spectroscopy, which is rapid and efficient for linear peptide sequencing⁸, becomes considerably more laborious and error-prone for macrocyclic peptide analysis⁹. Under collision-induced dissociation (CID), indiscriminate fragmentation may occur at any peptide bond within the macrocycle to generate a family of complex degenerate ions⁹. Meanwhile, NMR spectroscopy and single crystal x-ray diffraction (scXRD) allow for unambiguous sequence determination but are limited by sample requirements and the need to grow a single crystal of sufficient quality, respectively.

Site-specific linearization of cyclic peptides can greatly facilitate the analysis and sequencing process in the absence of commonly employed methods. Kodadek, among others, has demonstrated that site-specific linearization can be done chemo-selectively at Met residues¹⁰; however, Met is the second least abundant naturally-occurring amino acid.¹¹ Due to the rare natural presence of Met and the need for special reaction conditions¹², we sought to design a similar linearization methodology that is applicable to more commonly occurring residues.

Herein, we report a simple chemical sequencing method that overcomes challenges of existing techniques for highly stable macrocyclic and lasso peptides. This site-selective cleavage methodology is applicable to peptides containing with Ser, Thr, Cys, or Glu residues, due to their ability to form a cyclic urethane derived moiety that is susceptible to cleavage. These residues are present in approximately 60% of naturally-occurring macrocyclic peptides and 66% of lasso peptides as determined by a survey of NORINE¹³, a cyclic peptide database, and a comprehensive list of all lasso peptides discovered to date³. This method is also applicable to several less common amino acids found in non-ribosomal peptides containing reactive β - or γ -hydroxyl side chains, such as β -hydroxyl-Phe and homoserine (Hse).

Due to difficulties in sequencing macrocycles by MS/MS, our method first converts them to their linear counterparts. The strategy involves reaction of nucleophilic side chains (e.g., Ser, Thr, or Cys) with *N,N*-disuccinimidyl carbonate (DSC), Hünig's base (DIEA), and a catalytic amount of *N,N*-dimethylaminopyridine (DMAP) in *N,N*-dimethylformamide

(DMF) to generate five-membered ring moieties (e.g., an oxazolidinone in the case of Ser), which are susceptible to N-terminal peptide bond hydrolysis under mild aqueous conditions (Figure 1). In a previous study, conversions from 70 to >99% were observed for linear, model peptides with up to 12 residues.¹⁴ To test the methodology on cyclic peptides, reaction conditions were first evaluated on a macrocyclic, model peptide, *cyc*-Ser-Phe-Arg-Tyr-Ala-Glu, and complete conversion to its linear analog was observed in 48 h (Figures S1 and S2-Supporting information).

The method was then tested on polycarponin C, a mono-N-methylated cyclic peptide with anti-cancer activity¹⁵. Polycarponin C, which contains Phe, was incubated with chymotrypsin to confirm its resistance to enzymatic digestion, and no detectable cleavage was observed after 48 h (Figure 2a). MS/MS of the unmodified polycarponin C was then performed, and the spectrum was exceedingly complex and difficult to assign even with the structure being known due to several expected isobaric fragments (Figure 2b). We then proceeded to selectively hydrolyze the peptide backbone at the N-terminus of Thr using our method (Figures S3-Supporting information). Notably, scission proceeded smoothly even in the presence of a neighboring Pro residue, unlike enzymatic digestion¹⁶. The MS/MS spectrum of the linear counterpart was considerably simpler to interpret (Figure 2c). Nearly complete sequence coverage was observed except for the N-terminal Thr (i.e., missing b_1 ion). Nevertheless, the N-terminal residue can be easily discerned since only a small set of amino acids may be modified.

Next, we investigated a highly stable tri-N-methylated somatostatin analog (Figure 3)¹⁷. Again, after trypsin or chymotrypsin incubation for 48 h, the peptide remained intact. Direct MS/MS analysis as expected was troublesome due to complications previously noted. Conversely, our method completely cleaved the peptide to its linear analog following Thr activation and hydrolysis (Figures S4-Supporting information). Thus, both examples demonstrate the method's efficiency for macrocyclic peptides containing multiple N-methyl groups.

To demonstrate applicability with less common amino acids that are prevalent in non-ribosomally produced peptides isolated from nature, the model macrocyclic peptide, *cyc*-Gly-Tyr-Val-Hse-Phe-Leu-Ala containing homoserine (Hse), was synthesized. Due to Hse's hydroxyl side chain and its potential to form a six-membered oxazinanone ring with the amide nitrogen, we hypothesized that this residue would also be compatible with our chemical methodology. As expected, the Hse residue cleaved at its N-terminus after the modification, thereby opening the macrocycle to its linear counterpart (Figure S5-Supporting information).

We then further explored the method's applicability for lasso peptides, which are known for their diverse bioactivities (e.g., antimicrobial, antitumorigenic, and antiretroviral) and unique topologies consisting of a macrolactam ring, loop, and tail components^{3,18}. Sequencing this distinct class of peptides, including their linear loop and tail regions, is problematic due to their high resistance towards commonly utilized digestive enzymes. While isopeptidases are able to selectively cleave the isopeptide bond and thereby open the ring to a linear peptide, they are specific to the respective organism from which the lasso peptide was isolated.

Moreover, lasso peptide isopeptidases are found only in a subset of lasso peptide gene clusters¹⁹⁻²¹. Therefore, there is a considerable need for a chemical cleavage method since no general procedure currently exists.

The proteolytic stability of lasso peptide, lariatin A, was verified before testing the chemical cleavage methodology. Lariatin A is an antimicrobial peptide containing an Arg residue in its macrocyclic ring²². Since Arg is a substrate for trypsin, lariatin A was incubated overnight to investigate the extent, if any, of cleavage. As expected, no observable cleavage was detected. Lariatin A was then incubated with a mixture of trypsin and chymotrypsin since it contains the chymotrypsin substrate, Trp. After 48 hours, lariatin A remained intact (Figure S6-Supporting information).

We then applied our chemical cleavage method to sequence lariatin A because it contains two Ser residues, one in the ring and the other in the loop region (Figure 4). Overnight reaction with DSC was expected to modify only Ser-2 in the ring because the one in the loop should be less accessible. To our surprise, the major product was modification of Ser-13, and after hydrolysis, this generated a branched, cyclized peptide (**A**) and a short, linear peptide (**B**). A small amount of the expected branched, ring-opened product (**C**) was also observed; however, unexpectedly Lys17 was converted into a cycloocta-urea moiety. Previously on short model peptides, modification of the Lys side chain amino group was observed after treatment with DSC to the corresponding carbamic acid that subsequently would regenerate Lys under the hydrolytic conditions¹⁴. It was surmised that entropic difficulties negated formation of a cyclic ring moiety for Lys. However, in this case, Lariatin A's molecular topology may facilitate intramolecular ring formation.

Next, we investigated the applicability of our methodology for structural determination of albusnodin, a newly discovered lasso peptide containing a single Ser in its ring and two consecutive Arg residues in its linear chain²³. Reaction with DSC and subsequent hydrolysis generated the desired linear, ring-opened peptide (Figure S7-Supporting information). Our methodology was used in this context to determine the site of isopeptide bond formation due to the applicability of our method to Glu. If the isopeptide bond was forming on the Asp neighboring to Glu, then Ser and Glu would both react with DSC to afford two cleavage sites. However, this was not the case; the reaction was only observed at Ser, thereby confirming the isopeptide bond formation on the side chain of Glu²³. Conversely, no scission was observed upon treatment with carboxypeptidase and trypsin, which further emphasizes the high stability towards proteases and the difficulty of lasso peptide structural analysis even in the presence of two enzymatic cleavage sites.

Finally, the reactivity of benenodin-1 C5, a lasso peptide containing three possible cleavage sites (Ser, Thr, and Glu) in the loop region¹⁹, was investigated. The major product was Ser cleavage, which interestingly resulted in the formation of a rotaxane structure (Figure 5). Rotaxanes have an interlocked architecture consisting of a dumbbell-shaped molecule threaded through a macrocyclic ring²⁴. In the case of benenodin-1 C5, Glu and Gln serve as steric locks that prevent unthreading. We have developed the chemical approach for the synthesis of peptide-based rotaxane (both cycle and thread) for the first time. All the current methods for the synthesis of rotaxanes and catenanes require non-peptidic components due

the challenges associated with the synthesis of all-peptide based interlocked molecules.²⁵⁻²⁷ The significance of our chemical approach is that it converts a strong amide bond into an easily breakable bond in the lasso peptide scaffold thus has a potential for the synthesis of peptide based-rotaxanes and various other interlocked peptide-based molecular machines. Small amount of other benenodin-1 C5 cleavage products were also observed from Ser, Thr, and Glu residues (Figure S8-Supporting information). Cleavage at or near the steric lock unthreads the loop from the ring, while distal site cleavage generates the rotaxane structure through conservation of the threaded architecture. The advantage of these synthetic peptide-rotaxanes is that they may provide insights into natural interlocked systems.

Notably, the reaction of DSC with benenodin-1 C5 and subsequent cleavage gave rise to some unexpected results. While our method has shown compatibility with both Ser and Thr residues, it was surprisingly to observe the rotaxane product as the major product, where only Ser displayed reactivity. In an effort to further explore the difference in reactivity between Ser, Thr, and Glu, two linear peptides were synthesized OAc-FEYFSYF-NH₂ **1** and OAc-FTYFSYF-NH₂ **2**. The relative reactivity of Ser, Thr and Glu residues towards modification by DSC and hydrolysis were analyzed by using different equivalents of DSC, monitoring the reactions after regular intervals of time, and by performing the hydrolysis reaction at different temperatures and pH conditions (Supporting Information). These studies showed that serine underwent DSC modification at faster rate than threonine. Meanwhile, Ser and Glu did not display a difference in reactivity. Hydrolysis, however, occurred at a faster rate at modified Ser as compared to modified Glu. On the other hand, modified Ser and modified Thr did not display a difference in the rate of hydrolysis (Supporting Information).

In summary, a general chemical cleavage method combined with tandem mass spectrometry was developed for primary structure determination of macrocyclic peptides and those with unique topologies, such as lasso peptides. The methodology possesses broad substrate scope due to N-terminal peptide bond scission of highly abundant residues such as Ser, Thr, Cys, and Glu, as well as less common amino acids such as Hse and presumably others containing β - or γ -hydroxyl side chains, and applicability was estimated at >60% of naturally-occurring macrocyclic and lasso peptides. We were able to sequence peptides with unique structures, and thus our method provides an enabling technology platform for discovering newly isolated biologically-active peptides. Furthermore, the synthetic scheme enables interrogation of a lasso peptide's structure (e.g., how it is threaded and effectiveness of steric locks). Most importantly, we have utilized this chemical cleavage approach for synthesis of peptide-based rotaxane (both cycle and thread) for the first time from lasso peptide benenodin-1 C5. There is no other chemical method for the synthesis of entirely peptide-based interlocked rotaxane structures. Ultimately, this chemical cleavage method will provide a generic approach for synthesizing interlocking peptides that could be used to make peptide based supramolecular structures such as [2]catenanes, peptide-daisy chains and quasi-catenanes from lasso scaffolds. The work in this direction is currently underway in our laboratories.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by NIH grant GM107036 to AJL.

M.R. thanks Gary E. Martin for helpful suggestions.

References

- [1]. Henninot A , Collins JC , Nuss JM , J. Med. Chem 2017, DOI: 10.1021/acs.jmedchem.7b00318.
- [2]. Giordanetto F , Kihlberg J , J. Med. Chem 2014, 57, 278.24044773
- [3]. Hegemann JD , Zimmermann M , Xie X , Marahiel MA , Acc. Chem. Res 2015, 48, 1909.26079760
- [4]. Butler MS , J. Nat. Prod 2004, 67, 2141.15620274
- [5]. Liang X , Girard A , Biron E , ACS Comb. Sci 2013, 15 , 535.24016186
- [6]. Millward SW , Fiacco S , Austin RJ , Roberts RW , ACS Chem. Biol 2007, 2, 625.17894440
- [7]. Edman P , Begg G , Eur. J. Biochem 1967, 1, 80.6059350
- [8]. Medzihradsky KF , Chalkley RJ , Mass Spectrom. Rev 2015, 34, 43.25667941
- [9]. Ngoka LCM , Gross ML , J. Am. Soc. Mass Spectrom 1999, 10, 732.10439511
- [10]. Simpson LS , Kodadek T , Tetrahedron Lett 2012, 53, 2341.22736872
- [11]. King JL , Jukes TH , Science 1967, 164, 788.
- [12]. Kaiser R , Metzka L , Anal. Biochem 1999, 266, 1.9887207
- [13]. Caboche S , Pupin M , Leclère V , Fontaine A , Jacques P , Kucherov G , Nucleic Acids Res 2008, 36, 326.
- [14]. Elashal HE , Raj M , Chem. Commun 2016, 52, 6304.
- [15]. Shinde NV , Dhake AS , Havalavalaval KP , Orient. J. Chem 2016, 32, 515.
- [16]. Milovi NM , Kostić NM , J. Am. Chem. Soc 2003, 125, 781.12526679
- [17]. Chatterjee J , Laufer B , Kessler H , Nat. Protoc 2012, 7, 432.22322216
- [18]. Maksimov MO , Pan SJ , Link AJ , Nat. Prod. Rep 2012, 29, 996.22833149
- [19]. Zong C , Wu MJ , Qin JZ , Link AJ , J. Am. Chem. Soc 2017, 139, 10403.28696674
- [20]. Chekan JR , Koos JD , Zong C , Maksimov MO , Link AJ , Nair SK , J. Am. Chem. Soc 2016, 138, 16452.27998080
- [21]. Maksimov MO , Link AJ , J. Am. Chem. Soc 2013, 135, 12038.23862624
- [22]. Iwatsuki M , Tomoda H , Uchida R , Gouda H , Hirono S , Omura S , J. Am. Chem. Soc 2006, 128, 7486.16756302
- [23]. Zong C , Cheung-Lee W , Elashal HE , Raj M , Link AJ , Chem. Commun 2018, 54, 1339.
- [24]. Xue M , Yang Y , Chi X , Yan X , Huang F , Chem. Rev 2015, 115, 7398.25734835
- [25]. Aucagne V , Leigh DA , Lock JS , Thomson AR , J. Am. Chem. Soc 2006, 128, 1784.16464065
- [26]. Bruns CJ , Liu H , Francis MB , J. Am. Chem. Soc 2016, 138, 15307.27933926
- [27]. Fernando IR , Frasconi M , Wu Y , Liu W-G , Wasielewski MR , Goddard WA , Stoddart JF , J. Am. Chem. Soc 2016, 138, 10214.27398609

Figure 1.
Macrocyclic peptide sequencing strategy.

Figure 2. Sequencing polycarponin C. a) HRMS spectrum after 48 h incubation with chymotrypsin showing proteolytic resistance, b) complex MS/MS spectrum of polycarponin C, and c) simplified MS/MS spectrum of its ring-opened analog.

Figure 3.

Sequencing data for a somatostatin analog. A) HRMS spectrum after protease incubation showing no detectable cleavage, b) MS/MS spectrum of cyclized peptide, and c) MS/MS spectrum of ring-opened peptide.

Figure 4.

Sequencing data for lasso peptide, lariatin A. MS/MS spectra of the fully ring-opened peptide, C, and cleavage product, A, from reaction at Ser-13, which is present at the steric lock region.

Figure 5.
LC-MS of a rotaxane obtained from the modification/cleavage at Ser-9 of benendonin-1 C5.