Check for updates

Peptide Sequencing

Deutsche Ausgabe: DOI: 10.1002/ange.201801299
Internationale Ausgabe: DOI: 10.1002/anie.201801299

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

*Hader E. Elashal, Ryan D. Cohen, Heidi E. Elashal, Chuhan Zong, A. James Link, and Monika Raj**

Abstract: *A broadly applicable chemical cleavage methodology to facilitate MS/MS sequencing was developed for macrocyclic and lasso peptides, which hold 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 nonribosomal peptides. Importantly, the new method was utilized for synthesizing a unique peptide-based rotaxane (both cyclic and threaded) from the lasso peptide, benenodin-1 ΔC5.*

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 therapeutics.[1] 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 spectrometry (MS/MS), which is rapid and efficient for sequencing linear peptides,[8] becomes considerably more laborious and error-prone for analyzing macrocyclic peptides.[9] Under collision-induced dissociation (CID), indiscriminate fragmentation may occur at any peptide bond within the macrocycle to generate a family of complex degenerate ions.[9] 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 and Simpson, among others, have demonstrated that site-specific linearization can be done chemoselectively at Met residues;[10] 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 conditions,[12] 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 the challenges of existing analytical techniques for highly stable macrocyclic and lasso peptides. This site-selective cleavage methodology is applicable to peptides containing Ser, Thr, Cys, or Glu residues, due to their ability to form a cyclic urethane-derived moiety that is susceptible to cleavage (Figure 1). These residues are present in approximately 60% of naturally occurring macrocyclic peptides and 66% of lasso peptides as determined by a survey of NORINE,[13] a cyclic peptide database, and a comprehensive list of all lasso peptides discovered to date.[3] This method is also applicable to several less common amino acids found in nonribosomal peptides containing reactive β- or γ-hydroxyl side chains, such as β-hydroxylphenylalanine and homoserine (Hse).

Due to difficulties in sequencing macrocycles by MS/MS, our method first converts them to their linear counterparts.

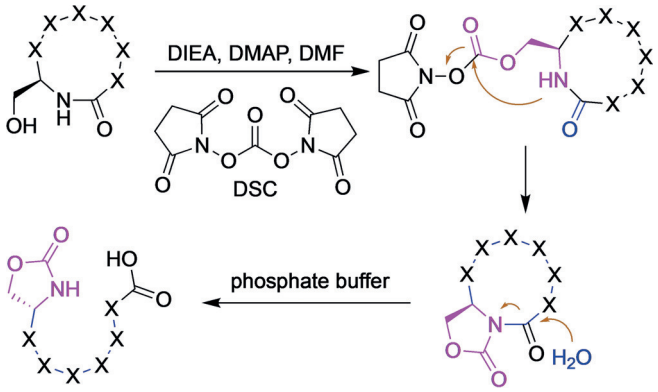

Figure 1. Strategy for sequencing macrocyclic peptides (X = amino acid).

[*] Prof. Dr. M. Raj
Department of Chemistry and Biochemistry
Auburn University
Auburn, AL 36830 (USA)
E-mail: mzr0068@auburn.edu
H. E. Elashal, H. E. Elashal
Department of Chemistry and Biochemistry
Seton Hall University
South Orange, NJ 07079 (USA)
R. D. Cohen
Department of Process Research and Development
Merck & Co., Inc.
Rahway, NJ 07065 (USA)
Dr. C. Zong, Prof. Dr. A. J. Link
Department of Chemistry and
Department of Chemical and Biological Engineering
Princeton University
Princeton, NJ 08544 (USA)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201801299.

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

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[14a] and for a variety of cyclic peptides on solid support.[14b] To test the methodology on cyclic peptides in solution, reaction conditions were first evaluated on a macrocyclic, model peptide, *cyc*-Ser-Phe-Arg-Tyr-Ala-Glu, and complete conversion to its linear analogue was observed in 48 h (Figures S1 and S2 in the Supporting Information).

The method was then tested on polycarponin C, a mono-N-methylated cyclic peptide with anticancer activity.[15] 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 (Figure S3). Notably, scission proceeded smoothly even in the presence of a neighboring Pro residue, unlike enzymatic digestion.[16] 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 analogue (Figure 3).[17] Again, after trypsin or chymotrypsin incubation for 48 h, the peptide remained intact. As expected, direct MS/MS analysis was troublesome due to the complications previously noted. Conversely, our method completely cleaved the peptide to its linear analogue following Thr activation and hydrolysis (Figure S4). Thus, both examples demonstrate the method's efficiency for sequencing macrocyclic peptides containing multiple N-methyl groups.

To demonstrate applicability with less common amino acids that are prevalent in nonribosomally 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).

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

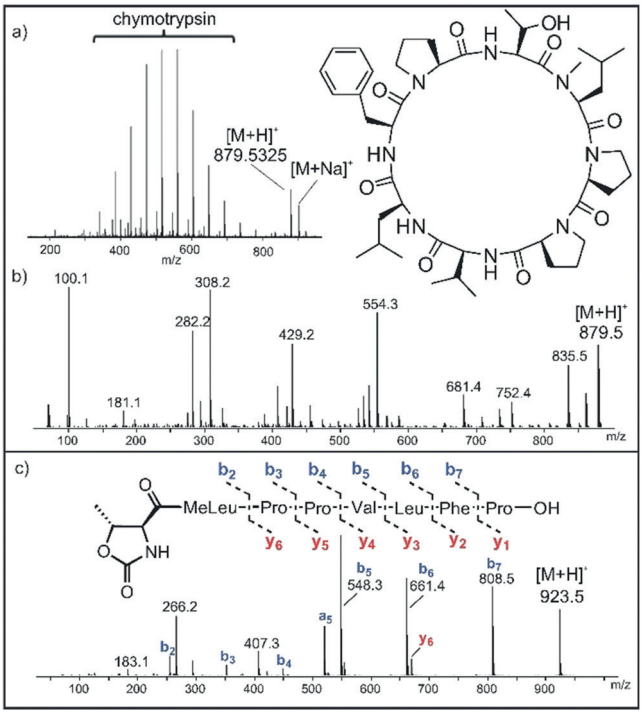

Figure 2. Sequencing polycarponin C. a) HRMS spectrum after 48 h incubation with chymotrypsin showing proteolytic resistance, b) complex MS/MS spectrum of polycarponin C, and c) simplified MS/MS spectrum of its ring-opened analogue.

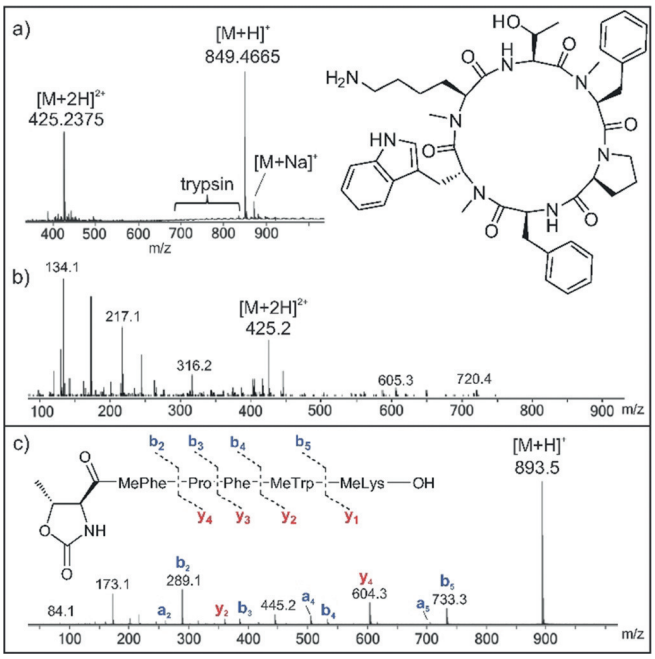

Figure 3. Sequencing data for a somatostatin analogue. a) HRMS spectrum after protease incubation showing no detectable cleavage, b) MS/MS spectrum of cyclized peptide, and c) MS/MS spectrum of ring-opened peptide.

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

15213757, 2018, 21, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ange.201801299 by Emory-Eresources, Wiley Online Library on [05/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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 clusters.[19–21] Therefore, there is a considerable need for a chemical cleavage method since no general procedure currently exists.

The proteolytic stability of the 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 ring.[22] 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).

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 Ser in the loop should be less accessible. To our surprise, the major product arose from 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 Lys-17 was converted into a cycloocta-urea moiety. Previously on short model peptides treated with DSC, conversion of the Lys side chain amino group to the corresponding carbamic acid was observed which subsequently would regenerate Lys under the hydrolytic conditions.[14a] 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 chain.[23] Reaction with DSC and subsequent hydrolysis generated the desired linear, ring-opened peptide (Figure S7). 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 forms on the Asp adjacent 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 Glu.[23] 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 ΔC5, a lasso peptide containing three possible cleavage sites (Ser, Thr, and Glu) in the loop region,[19] was investigated. The major product arose from 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 ring.[24] In the case of benenodin-1 ΔC5, Glu and Gln serve as steric locks that prevent unthreading. We have developed the chemical approach for the synthesis of peptide-based rotaxane (both cyclic and threaded) for the first time. All the current methods for the synthesis of rotaxanes and catenanes require nonpeptidic 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 and thus has potential for

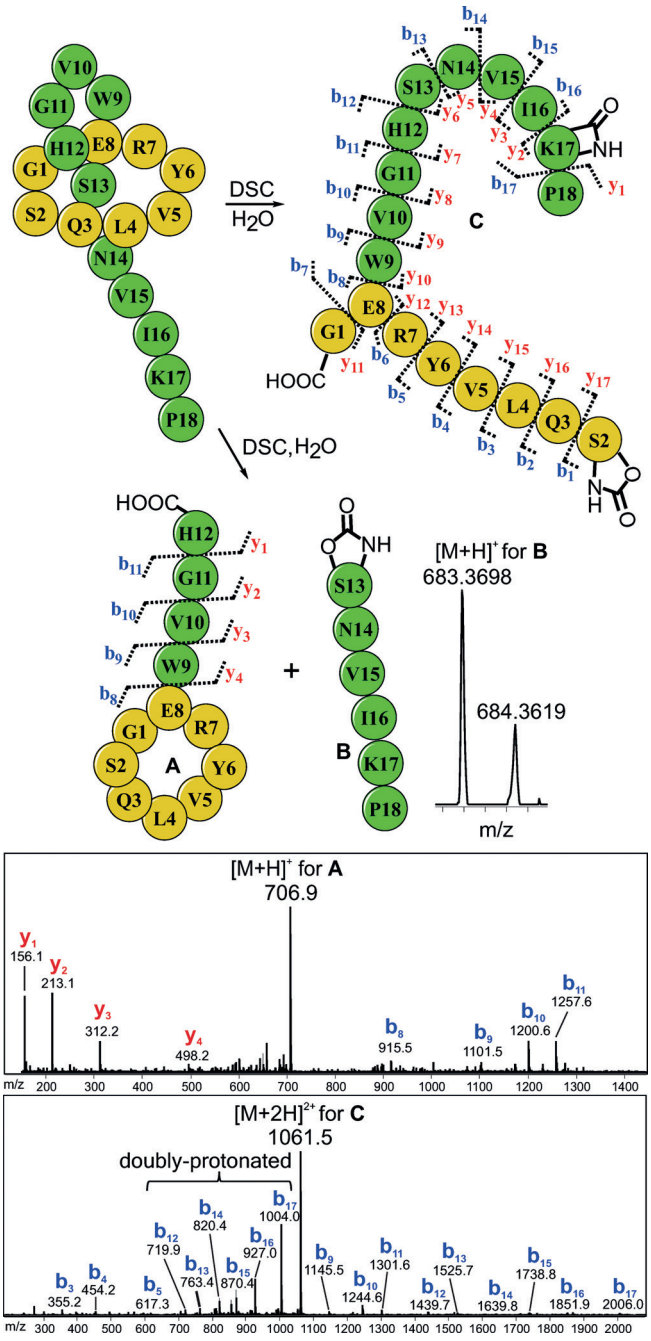

Figure 4. Sequencing data for the lasso peptide lariatin A. MS/MS spectra of the fully ring-opened peptide, **C**, and cleavage product, **A**, from reaction at Ser-13, which is present at the steric lock region.

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

15213757, 2018, 21, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ange.201801299 by Emory-Eresources, Wiley Online Library on [05/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

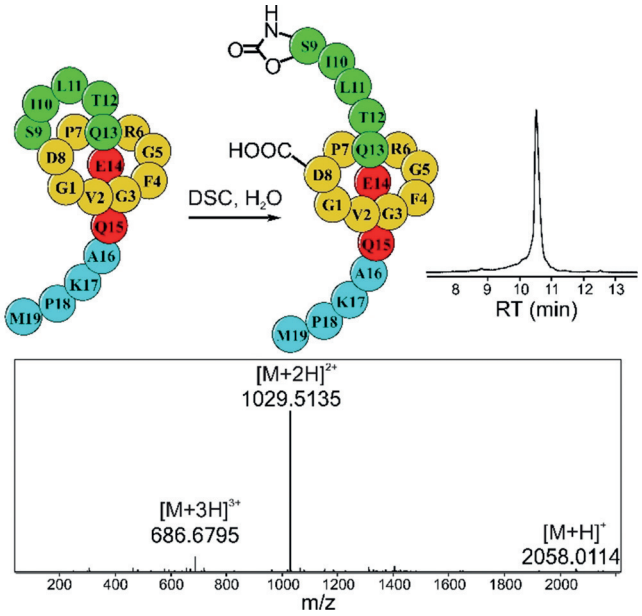

Figure 5. LC-MS analysis of a rotaxane obtained from the modification/cleavage at Ser-9 of benendonin-1 ΔC5.

the synthesis of peptide-based rotaxanes and various other interlocked peptide-based molecular machines. Small amounts of other benenodin-1 ΔC5 cleavage products were also observed from Ser, Thr, and Glu residues (Figure S8). 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. Such synthetic peptide rotaxanes may provide insight into natural interlocked systems.

Notably, reaction of DSC with benenodin-1 ΔC5 and subsequent cleavage generated some unexpected results. While our method has shown compatibility with both Ser and Thr residues, it was surprising 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-NH_2 **1** and OAc-FTYFSYF-NH_2 **2**. The relative reactivity of Ser, Thr, and Glu residues towards modification by DSC and hydrolysis was 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 (see the Supporting Information). These studies showed that serine underwent DSC modification at a 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 than at modified Glu. On the other hand, modified Ser and modified Thr did not exhibit a difference in the rate of hydrolysis ().

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 displays 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 rotaxanes (both cyclic and threaded) for the first time from lasso peptide benenodin-1 ΔC5. 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. Work in this direction is currently underway in our laboratories.

Acknowledgements

This work was supported by NIH grant GM107036 to A.J.L. M.R. thanks Gary E. Martin for helpful suggestions.

Conflict of interest

The authors declare no conflict of interest.

Keywords: peptides · rotaxanes · sequencing

How to cite: *Angew. Chem. Int. Ed.* **2018**, *57*, 6150–6154
Angew. Chem. **2018**, *130*, 6258–6262

[1] A. Henninot, J. C. Collins, J. M. Nuss, *J. Med. Chem.* **2017**, DOI: https://doi.org/10.1021/acs.jmedchem.7b00318.
[2] F. Giordanetto, J. Kihlberg, *J. Med. Chem.* **2014**, *57*, 278.
[3] J. D. Hegemann, M. Zimmermann, X. Xie, M. A. Marahiel, *Acc. Chem. Res.* **2015**, *48*, 1909.
[4] M. S. Butler, *J. Nat. Prod.* **2004**, *67*, 2141.
[5] X. Liang, A. Girard, E. Biron, *ACS Comb. Sci.* **2013**, *15*, 535.
[6] S. W. Millward, S. Fiacco, R. J. Austin, R. W. Roberts, *ACS Chem. Biol.* **2007**, *2*, 625.
[7] P. Edman, G. Begg, *Eur. J. Biochem.* **1967**, *1*, 80.
[8] K. F. Medzihradszky, R. J. Chalkley, *Mass Spectrom. Rev.* **2015**, *34*, 43.
[9] L. C. M. Ngoka, M. L. Gross, *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 732.
[10] L. S. Simpson, T. Kodadek, *Tetrahedron Lett.* **2012**, *53*, 2341.
[11] J. L. King, T. H. Jukes, *Science* **1967**, *164*, 788.
[12] R. Kaiser, L. Metzka, *Anal. Biochem.* **1999**, *266*, 1.
[13] S. Caboche, M. Pupin, V. Leclère, A. Fontaine, P. Jacques, G. Kucherov, *Nucleic Acids Res.* **2008**, *36*, 326.
[14] a) H. E. Elashal, M. Raj, *Chem. Commun.* **2016**, *52*, 6304; b) H. E. Elashal, R. D. Cohen, H. E. Elashal, M. Raj, *Org. Lett.* **2018**, DOI: https://doi.org/10.1021/acs.orglett.8b00717.
[15] N. V. Shinde, A. S. Dhake, K. P. Havalavalaval, *Orient. J. Chem.* **2016**, *32*, 515.

 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

15213757, 2018, 21, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ange.201801299 by Emory-Eresources, Wiley Online Library on [05/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[16] N. M. Milović, N. M. Kostić, *J. Am. Chem. Soc.* **2003**, *125*, 781.
[17] J. Chatterjee, B. Laufer, H. Kessler, *Nat. Protoc.* **2012**, *7*, 432.
[18] M. O. Maksimov, S. J. Pan, A. J. Link, *Nat. Prod. Rep.* **2012**, *29*, 996.
[19] C. Zong, M. J. Wu, J. Z. Qin, A. J. Link, *J. Am. Chem. Soc.* **2017**, *139*, 10403.
[20] J. R. Chekan, J. D. Koos, C. Zong, M. O. Maksimov, A. J. Link, S. K. Nair, *J. Am. Chem. Soc.* **2016**, *138*, 16452.
[21] M. O. Maksimov, A. J. Link, *J. Am. Chem. Soc.* **2013**, *135*, 12038.
[22] M. Iwatsuki, H. Tomoda, R. Uchida, H. Gouda, S. Hirono, S. Omura, *J. Am. Chem. Soc.* **2006**, *128*, 7486.
[23] C. Zong, W. Cheung-Lee, H. E. Elashal, M. Raj, A. J. Link, *Chem. Commun.* **2018**, *54*, 1339.
[24] M. Xue, Y. Yang, X. Chi, X. Yan, F. Huang, *Chem. Rev.* **2015**, *115*, 7398.
[25] V. Aucagne, D. A. Leigh, J. S. Lock, A. R. Thomson, *J. Am. Chem. Soc.* **2006**, *128*, 1784.
[26] C. J. Bruns, H. Liu, M. B. Francis, *J. Am. Chem. Soc.* **2016**, *138*, 15307.
[27] I. R. Fernando, M. Frasconi, Y. Wu, W.-G. Liu, M. R. Wasielewski, W. A. Goddard, J. F. Stoddart, *J. Am. Chem. Soc.* **2016**, *138*, 10214.

Manuscript received: January 31, 2018
Accepted manuscript online: April 12, 2018
Version of record online: April 2018, 27

© 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

15213757, 2018, 21, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ange.201801299 by Emory-Eresources, Wiley Online Library on [05/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License