ChemPubSoc Europe

DOI: 10.1002/cbic.201500179

CHEMBIOCHEM Communications

Ribosomal Synthesis of Natural-Product-Like Bicyclic Peptides in *Escherichia coli*

Nina Bionda and Rudi Fasan*[a]

Methods to access natural-product-like macrocyclic peptides can disclose new opportunities for the exploration of this important structural class for chemical biology and drug discovery applications. Here, the scope and mechanism of a novel strategy for directing the biosynthesis of thioether-bridged bicyclic peptides in bacterial cells was investigated. This method entails split intein-catalyzed head-to-tail cyclization of a ribosomally produced precursor peptide, combined with inter-side-chain crosslinking through a genetically encoded cysteine-reactive amino acid. This strategy could be successfully applied to achieve formation of structurally diverse bicyclic peptides with high efficiency and selectivity in *Escherichia coli*. Insights into the sequence of reactions underlying the peptide bicyclization process were gained from time-course experiments. Finally, the potential utility of this methodology toward the discovery of macrocyclic peptides with enhanced functional properties was demonstrated through the isolation of a bicyclic peptide with sub-micromolar affinity for streptavidin.

Macrocyclic peptides of both natural[1] and synthetic[2] origin have attracted significant and increasing interest as a potential source of biologically active molecules.[3] Among the bioactive peptides found in nature, several exhibit a bicyclic topology wherein a head-to-tail cyclic structure is further rigidified by an intramolecular linkage connecting two side chains of the peptide.[4] For example, α-amanitin, a fungal toxin that potently inhibits the activity of eukaryotic RNA polymerases II,[5] features a head-to-tail cyclic backbone constrained by an intramolecular bond between a modified trypthophan and cysteine residues (Scheme 1 A).[6] Other relevant examples include θ-defensins,[7] which act as antimicrobial agents as part of innate immune system in non-human primates, and plant-derived bicyclic peptides, such as members of the bouvardin and celogentin families, which were shown to possess anticancer activity.[8] In these molecules, the conformational constraints imposed by the bicyclic backbone are critical for their biological activity and beneficial toward increasing their cell penetration properties and stability against proteolysis. Owing to the attractive features of bicyclic peptides, a number of synthetic approaches have been investigated to access this type of compounds.[9] As an alternative approach, our group[10] and others[11] have made available strategies to obtain bicyclic peptides through the in vitro cyclization of ribosomally derived polypeptides. Despite this progress, viable approaches to direct the biosynthesis of bicyclic peptides of arbitrary sequence in living cells have been missing. Such approaches would be highly desirable toward acquiring the capability of generating genetically encoded libraries of natural-product-like bicyclic peptides, which could be then functionally interrogated by means of selection[12] or by phenotypic screens.[13]

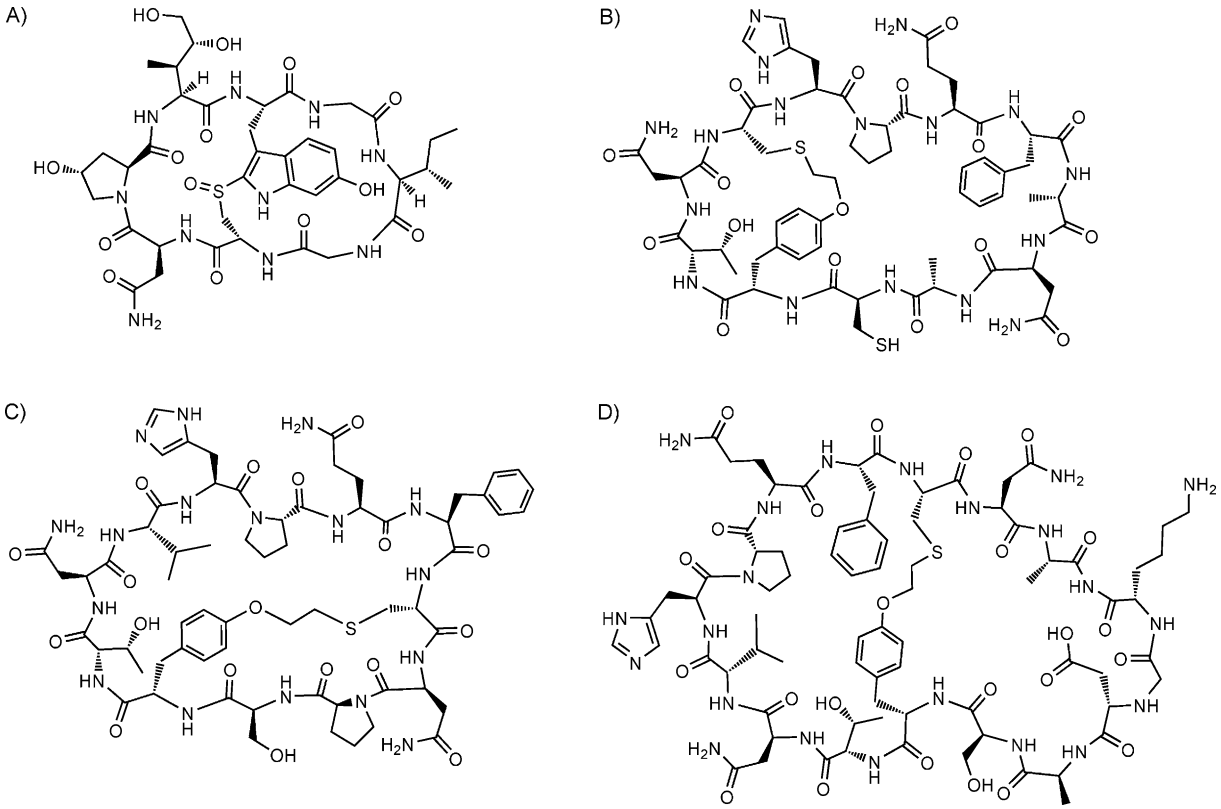
To bridge this gap, our group recently introduced a "biomimetic" strategy to enable the ribosomal synthesis of bicyclic peptides featuring a head-to-tail cyclic backbone as well as an inter-side-chain thioether linkage,[14] thereby resembling the overall topology of naturally occurring bicyclic peptides such as α-amanitin (Scheme 1 A). The scope of this methodology beyond the model dodecamer peptide sequence considered in those initial studies (Scheme 1 B) was not explored though. Here, we report the investigation and successful application of this approach to create bicyclic peptides featuring variable ring sizes and intramolecular connectivities. In addition, we elucidated the sequence of intracellular events that led to the cyclopeptide product and demonstrated the utility of this biosynthetic approach toward enabling the isolation of bicyclic peptides with improved functional properties.

Scheme 2 outlines the key steps of the aforementioned biomimetic method for orchestrating the biosynthesis of thioether-bridged bicyclic peptides in *Escherichia coli*. In this system, peptide N-to-C circularization is attained by framing a target peptide sequence between the C-terminal (Int_C) and N-terminal (Int_N) domains of the naturally occurring split intein DnaE (Figure S1 in the Supporting Information).[15] Installation of the inter-side-chain thioether linkage is achieved through a spontaneous reaction between a cysteine residue and the cysteine-reactive unnatural amino acid, *O*-(2-bromoethyl)-tyrosine (O2beY), which is incorporated into the precursor polypeptide through amber stop codon suppression.[14] O2beY is able to efficiently react with a cysteine positioned 2–8 residues apart through a nucleophilic substitution reaction, thus resulting in a stable, non-reducible thioether bridge.[14]

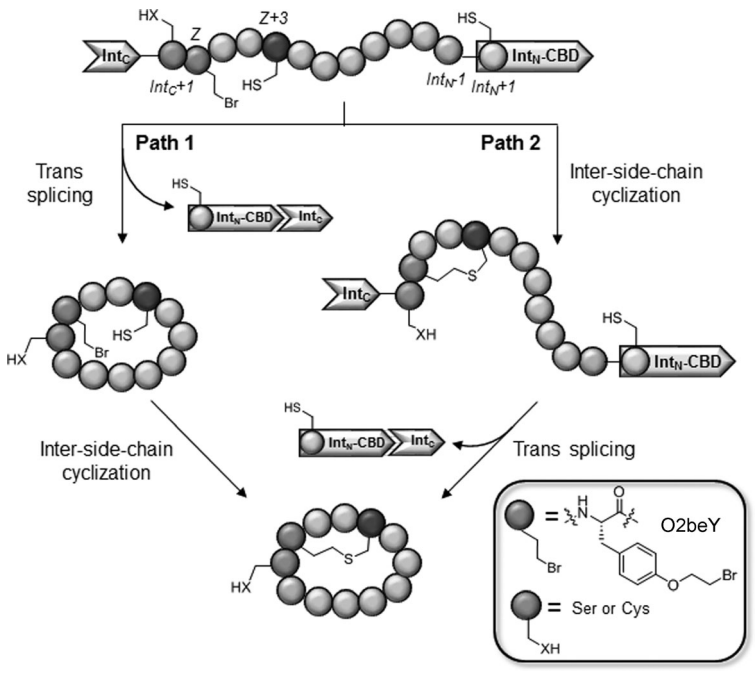
With regard to the mechanism, two possible pathways can be envisioned for the intracellular formation of the bicyclic peptide. As described in Scheme 2, a first pathway could involve initial head-to-tail cyclization of the target peptide sequence, followed by reaction of the cysteine residue with O2beY to create the thioether bridge (path 1). Alternatively, the crosslinking reaction between O2beY and the cysteine residue could occur before the peptide circularization step mediated by the split intein (path 2, Scheme 2). In our previous studies, expression of the protein constructs in *E. coli* corresponding to entries 1 and 2 of Table 1 led to accumulation of

[a] Dr. N. Bionda, Prof. Dr. R. Fasan
Department of Chemistry, University of Rochester
RC Box 270216, Rochester, NY 14627 (USA)
E-mail: fasan@chem.rochester.edu

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201500179.



Scheme 1. Natural and artificial bicyclic peptides. A) Structure of the natural product peptide α-amanitin. B)–D) Structures of representative thioether-linked bicyclic peptides obtained by using the biomimetic method investigated here and outlined in Scheme 2.



Scheme 2. Strategy for the ribosomal synthesis of thioether-bridged bicyclic peptides in *E. coli*. From N to C, the linear precursor polypeptide comprises the C-terminal domain of split intein DnaE (Int_C), a Ser or Cys residue at Int_C+1 site, the unnatural amino acid *O*-(2-bromoethyl)-tyrosine (O2beY or "Z"), a variable target sequence containing the reactive cysteine, and a DnaE N-terminal domain fused to a chitin binding domain (Int_N-CBD). The two envisioned pathways leading to the bicyclic product are indicated.

the corresponding bicyclic peptides in the cells as the nearly exclusive products.[14]

In addition, quantitative splicing of the precursor protein was observed in each case. Although promising, these results provided no information regarding the sequence of reactions underlying the bicyclization process. In order to shed light on this, the expression conditions were altered with the goal of trapping reaction intermediates resulting from partial cyclization. First, using the control construct Z3C_OpgY (entry 3, Table 1), we established that the extent of DnaE-catalyzed *trans* splicing decreases with temperature (Figure S3). Next, the precursor proteins Z3C_O2beY and Z3C(S1C)_O2beY were expressed at lower temperatures (20 vs. 27 °C) and for a shorter period of time (6 vs. 12 h). Under these conditions, both constructs underwent only partial splicing (64% for Z3C_O2beY; 59% for Z3C(S1C)_O2beY), as determined by SDS-PAGE analysis and LC-MS analysis of the corresponding full-length precursor protein and spliced protein after isolation from the cell lysate by using the C-terminal chitin binding domain (CBD; Figure S15).

Importantly, a streptavidin-binding His-Pro-Gln (HPQ) motif[16] was included in the target sequence in these constructs, enabling isolation of the bicyclic and monocyclic peptides, as well as any potential acyclic by-product(s), by affinity chromatography

Table 1. Name and target sequences of the protein constructs used in this study. Reported values refer to the percentage of intein splicing undergone by the precursor proteins after expression in *E. coli* and the relative amount of the corresponding monocyclic and bicyclic peptides, as determined by streptavidin-affinity isolation followed by LC-MS analysis.

	Construct name	Target sequence[a]	Cys position[b]	Intein splicing [%]	Monocyclic peptide[c] [%]	Bicyclic peptide [%]
1	Z3C_O2beY	S(**O2beY**)TN**C**HPQFANA	*Z*+3	100/64[d]	4/35[d]	96/65[d]
2	Z3C(S1C)_O2beY	C(**O2beY**)TN**C**HPQFANA	*Z*+3	100/59[d]	3/33[d]	97/67[d]
3	Z3C_OpgY	S(**OpgY**)TN**C**HPQFANA	*Z*+3	100	96	–
4	Z8C_O2beY	S(**O2beY**)TNVHPQF**C**NA	*Z*+8	98	24	76
5	Z8C(S1C)_O2beY	C(**O2beY**)TNVHPQF**C**NA	*Z*+8	98	25	75
6	Z8C(A12P)_O2beY	S(**O2beY**)TNVHPQF**C**N**P**	*Z*+8	60	25	75
7	Z8C(A12P)_OpgY	S(**OpgY**)TNVHPQF**C**N**P**	*Z*+8	45	97	–
8	Z8C(A12N)_O2beY	S(**O2beY**)TNVHPQF**C**N**N**	*Z*+8	92	25	75
9	Z8C(A12N)_OpgY	S(**OpgY**)TNVHPQF**C**N**N**	*Z*+8	89	94	–
10	Z8C(16-mer)_O2beY	S(**O2beY**)TNVHPQF**C**NAKGDA	*Z*+8	96	28	72
11	Z8C(18-mer)_O2beY	S(**O2beY**)TNVHPQF**C**NAKGDTQA	*Z*+8	98	31	69

[a] Full-length precursor protein corresponds to Int_C-(target sequence)-Int_N-(chitin binding domain). [b] Position of reactive cysteine residue (underlined) in relation to O2beY or OpgY (= Z position). [c] For the OpgY-containing constructs, the remainder to 100% consists of the linear peptide. [d] From precursor proteins expressed for 6 h at 20 °C in *E. coli*. The standard errors for the percentage values reported in the table are ≤5%.

with streptavidin-coated beads. The low-molecular-weight peptide products were isolated from the corresponding cell lysates by using this procedure, and analyzed by LC-MS. Interestingly, under the altered expression conditions, a significantly larger fraction of the head-to-tail monocyclic peptides (65–67%, Table 1) was isolated for both the Z3C_O2beY and Z3C(S1C)_O2beY constructs compared to the thioether-bridged bicyclic peptide (33–35%), as estimated by the peak areas of the respective extracted-ion chromatograms (Figures S4–S5).[17] Together, these results and the splicing data clearly indicated that backbone cyclization precedes crosslinking of the Cys/O2beY side chains, that is, the overall reaction follows the course described by path 1 of Scheme 2.

Also noteworthy was the observation that nearly identical results were obtained for the Z3C_O2beY and Z3C(S1C)_O2beY constructs, which differ from each other by the residue at the Int_C+1 position (Ser or Cys, respectively). This residue is responsible for attacking the thioester linkage at the C-terminal end of the target sequence during the split intein-catalyzed *trans* splicing process (Figure S1). The similar behavior of these proteins thus showed that the overall bicyclization mechanism (that is, path 1 vs. path 2) is not affected by the inherently different nucleophilicity of the residue (Cys vs. Ser) at the Int_C+1 site. Finally, no detectable amounts of the acyclic peptide or thioether-bridged monocyclic peptide byproducts were observed in either case, indicating that premature splicing of the split intein fragments does not effectively compete with the desired bicyclization process.

To this point, all bicyclic peptides produced by the strategy outlined in Scheme 2 contained a thioether bridge between the O2beY and Cys residues in an *i*, *i*+*3* arrangement. To explore the possibility of generating alternative bicyclic topologies, two protein constructs in which the cysteine occupies the *i*+*8* position with respect to O2beY, namely Z8C_O2beY and Z8C(S1C)_O2beY (entries 4–5, Table 1), were prepared and tested. The corresponding proteins were produced in BL21(DE3) cells coexpressing an orthogonal aminoacyl-tRNA synthetase (AARS)/tRNA pair capable of mediating the ribosomal incorporation of O2beY in response to an amber stop codon.[14]

To our delight, both constructs were found to lead to the intracellular accumulation of the desired bicyclic product in high yields (~75%), with the remainder of the isolated peptides being in the head-to-tail monocyclic form (Figures 1, S7 and S8). In a parallel experiment, benzyl mercaptan (BnSH) was added immediately after lysis of cells expressing the Z8C_O2beY construct to quench any unreacted O2beY residue. This experiment yielded a distribution of monocyclic and bicyclic peptide products identical to that obtained in the absence of BnSH treatment (Figure S18), thus further confirming the formation of the bicyclic peptides directly inside the bacterial cells. The slight reduction in the bicyclic-to-monocyclic ratio observed with these *i*, *i*+*8* constructs, as compared to the *i*, *i*+*3* counterparts (>95%), is consistent with the observed dependence of the crosslinking efficiency on the spacing distance between the O2beY/Cys residues.[14] Importantly, the successful bicyclization of the target sequence in Z8C(S1C)_O2beY also highlights the high chemoselectivity of the O2beY-mediated crosslinking reaction toward the desired *i*+*8* cysteine, as opposed to the *i*−*1* and *i*+*11* cysteines occupying the Int_C+1 and Int_N+1 positions, respectively. This conclusion is further supported by the high levels of intein splicing observed for both Z8C(S1C)_O2beY and Z8C_O2beY constructs (98%, Figure S15A), a process that would be prevented by alkylation of the Int_N+1 or Int_C+1 cysteine by O2beY. Interestingly, further incubation of the isolated monocyclic peptide in phosphate buffer (pH 7.5) did not lead to any appreciable increase in the amount of bicyclic product after up to 24 h, suggesting a beneficial effect of macromolecular crowding in the intracellular milieu[18] toward facilitating O2beY-mediated crosslinking of the peptide.

Whereas our time-course experiments indicated that backbone cyclization is faster than O2beY-induced crosslinking, we wondered whether the latter could actually facilitate the former reaction in the presence of target sequences less prone to circularization. In order to investigate this aspect, two addi-

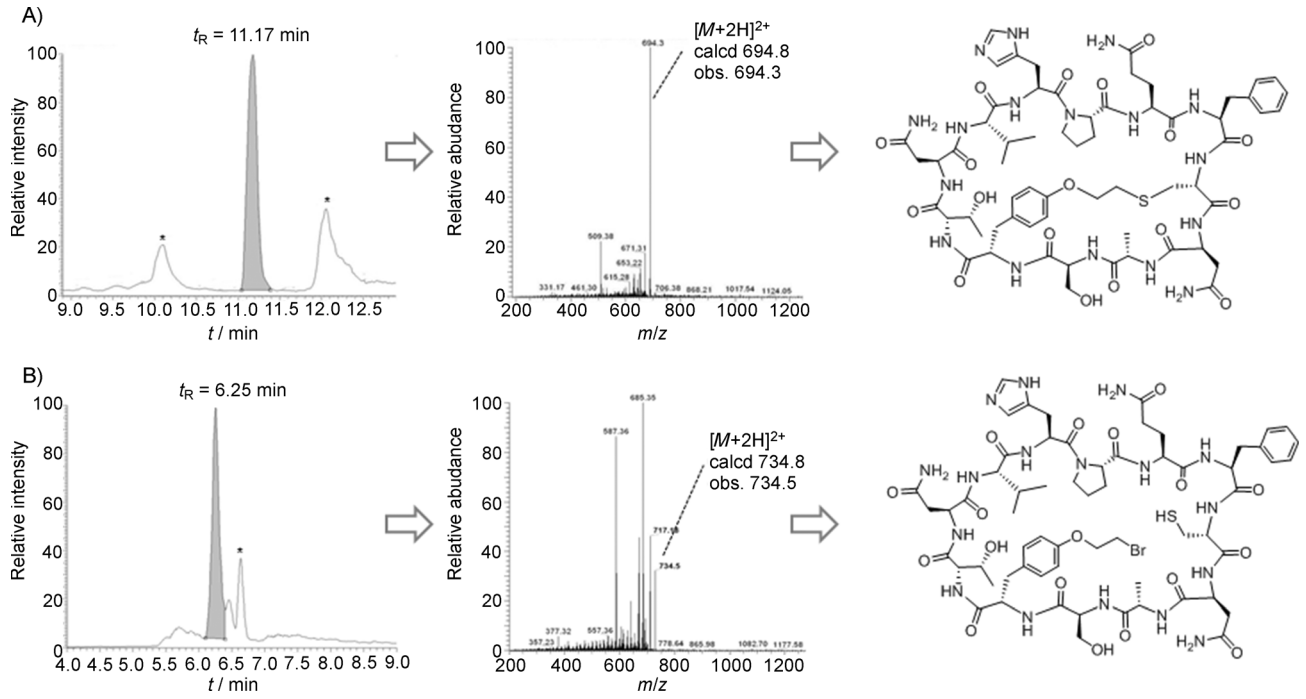


Figure 1. LC-MS extracted-ion chromatogram (left), MS/MS fragmentation spectrum (center), and chemical structure (right) of the A) bicyclic and B) monocyclic peptide products isolated by streptavidin affinity from cells expressing the construct Z8C_O2beY. Peaks labeled with * correspond to unrelated multicharged ions from adventitious proteins. Data for the other constructs can be found in Figures S7–S13.

tional *i*, *i*+*8* constructs were prepared in which the terminal alanine at position Int_N-1 was substituted for Asn and Pro (entries 6 and 8 in Table 1, respectively). These types of Int_N-1 substitutions were previously reported to disfavor DnaE split intein-mediated peptide circularization.[19] As controls, two precursor proteins were also prepared in which O2beY was replaced with a structurally similar unnatural amino acid that is unable to react with cysteine, namely *O*-propargyl-tyrosine, or OpgY (entries 7 and 9, Table 1). Surprisingly, the O2beY- and OpgY-containing A12N variants showed equally high levels of protein splicing (~90%, Figure S15B), suggesting that the previously reported deleterious effect of Asn at the Int_N-1 site is most likely sequence-dependent. More insightful results were obtained for the O2beY- and OpgY-containing A12P variants, which showed a reduced and differential degree of split intein *trans* splicing, as illustrated by the MS spectra in Figure S15C–D. In this case, the O2beY-containing variant showed a higher degree of protein splicing as compared to the OpgY-containing counterpart (60 vs. 45%), suggesting that O2beY/Cys-mediated crosslinking can indeed facilitate DnaE-mediated head-to-tail peptide cyclization when the latter becomes less favorable. This effect may be attributed to the thioether bridge causing a rigidification of the target peptide sequence and/or a closer proximity of the Int_N+1 cysteine to the ester group at the level of the Int_C-1 residue, thereby facilitating the *trans* splicing reaction (Figure S1). Importantly, for both of the O2beY-containing precursor proteins, the bicyclic peptide constituted the largely predominant product (>75%) isolated from the cell lysates by streptavidin affinity (Figures S9–S12).

In the interest of further exploring the reaction scope of this biosynthetic strategy, the precursor proteins corresponding to entries 10 and 11 of Table 1 were investigated. In the corresponding bicyclic peptide products, the size of the second ring was expanded to include seven or nine residues, respectively, as compared to the three-residue rings of the previous (*i*, *i*+8)-bridged constructs. To our delight, both constructs were found to undergo nearly quantitative splicing (96–98%), resulting in the intracellular formation of the desired 16-mer and 18-mer bicycles as the major product (72 and 69%, respectively), as determined by LC-MS analysis (Figures S13–S14). Together, the results from these and the previous constructs support the versatility of this method toward enabling the biosynthesis of structurally diverse bicyclic peptide structures (Scheme 1).

In these compounds, the conformational rigidification induced by the presence of the thioether bridge significantly alters the properties of the peptide, as suggested by the large shift in polarity observed for the bicyclic product as compared to the monocyclic counterpart (Δt_R~5 min in reversed-phase C_{18} column, Figure 1). Another beneficial effect potentially arising from an increased conformational rigidity of the molecule could be an enhanced protein binding affinity as a result of the reduced entropic costs associated with complex formation. To examine this aspect, milligram amounts of the (*i, i*+*3*)- and (*i, i*+*8*)-bridged bicyclic peptides corresponding to entries 1 and 4 in Table 1, called bicyclo-Z3C and bicyclo-Z8C, respectively, were isolated from large-scale cultures (1 L) of *E. coli* cells expressing the precursor proteins Z3C_O2beY and Z8C_O2beY, respectively. The isolated yield for these peptides under non-optimized conditions was about 0.6–0.7 mg L^{-1} culture, which compares favorably with that reported for the production of cyclopeptide natural products from *E. coli* cells expressing heterologous biosynthetic pathways (e.g., 0.1–1 mg L^{-1} culture for

patellamides).[1c] As a control, a head-to-tail monocyclic peptide encompassing the same target sequence as bicyclo-Z3C was obtained from cells expressing the precursor protein Z3C_OpgY (entry 3, Table 1), called cyclo-Z3C. All of these macrocyclic peptides encompass the streptavidin-binding HPQ sequence but differ from each other based on the relative location (bicyclo-Z3C vs. bicyclo-Z8C) or presence (cyclo-Z3C) of the O2beY/Cys crosslink. The relatively affinity of these compounds toward streptavidin was then evaluated by using an in-solution competition assay.[16] As illustrated in Figure 2A, a streptavidin-binding surface was first generated by immobilizing bicyclo-Z8C(S1C) (obtained from cells expressing the construct corresponding to entry 5, Table 1) onto maleimide-coated microtiter plates. Inhibition curves were then obtained by titrating each of the peptides in the presence of a streptavidin-horseradish peroxidase (HRP) conjugate and measuring the concentration-dependent decrease in fluorescence after incubation with the HRP substrate. After validation of the assay with D-desthiobiotin (IC_{50}: 12 nM), the ability of the monocyclic peptide cyclo-Z3C to inhibit binding of the streptavidin to the functionalized surface was determined, yielding an IC_{50} value of 1.9 μM (Figure 2B). Interestingly, bicyclo-Z3C was found to display a comparatively weaker affinity for streptavidin, as determined by the twofold higher IC_{50} value (3.7 μM). In stark contrast, bicyclo-Z8C exhibited a significantly higher inhibitory activity ($IC_{50} = 0.77$ μM) as compared to both the monocyclic peptide (2.5-fold) and the (*i*, *i*+*3*)-bridged bicyclic peptide (4.8-fold, Figure 2B). Based on the available crystal structure of streptavidin in complex with an HPQ-containing peptide (Figure S2),[20] neither of the sites occupied by the unnatural amino acid or by the *i*+*3* (or *i*+*8*) cysteine are expected to establish direct contacts with the protein surface. Accordingly, we derived that the observed effect of the inter-side-chain thioether bridge on the streptavidin binding affinity of these compounds mainly stems from its ability to pre-organize the HPQ pharmacophore. More importantly, these results illustrate the potential utility of this method in the context of the discovery of bicyclic peptides with improved functional properties.

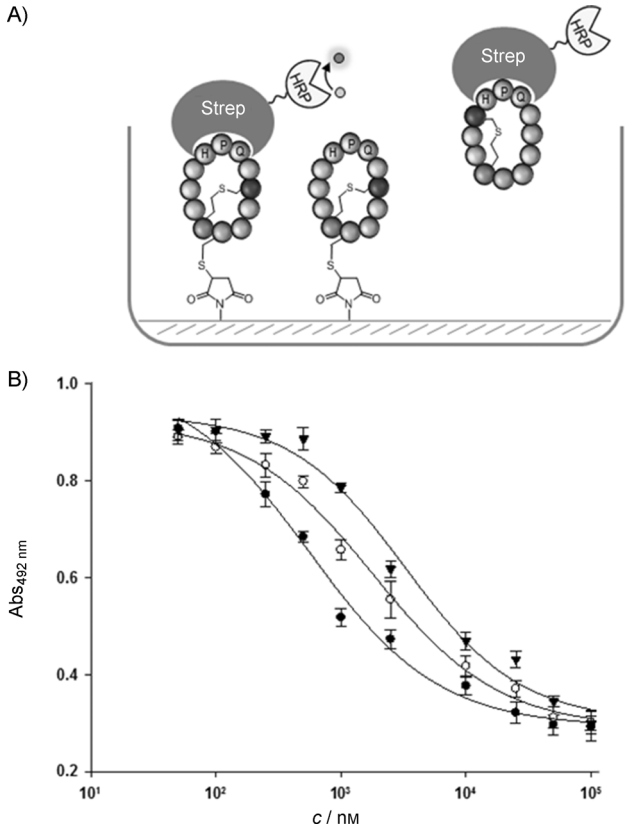


Figure 2. Streptavidin binding assay. A) Schematic representation of the inhibition assay. B) Inhibition curves for monocyclic peptide cyclo-Z3C (○; $IC_{50} = (1.9 \pm 0.2)$ μM) and bicyclic peptides bicyclo-Z3C (▼; $IC_{50} = (3.7 \pm 0.2)$ μM) and bicyclo-Z8C (●; $IC_{50} = (0.77 \pm 0.08)$ μM).

In conclusion, the functionality and versatility of a new strategy for directing the biosynthesis of natural-product-like bicyclic peptides were demonstrated through the generation of structurally diverse bicyclic peptides featuring varying ring size and thioether bridge connectivities with high efficiency and selectivity in *E. coli*. From a mechanistic standpoint, time-course experiments revealed that these macrocyclic peptides are generated at the post-translational level by head-to-tail cyclization followed by inter-side-chain crosslinking, with the latter increasing the efficiency of the former reaction in the presence of target sequences less prone to split intein-mediated circularization. Finally, the potential utility of this methodology toward the discovery of functional, conformationally constrained peptides was illustrated through the production and isolation of a bicyclic peptide with enhanced affinity toward the model target protein streptavidin.

Acknowledgements

This work was supported by the U.S. National Institutes of Health (grant R21 CA187502). MS instrumentation was supported by the U.S. National Science Foundation (grants CHE-0840410 and CHE-0946653).

Keywords: bicyclic peptides · ribosomal synthesis · split inteins · thioether bridges · unnatural amino acids

[1] a) L. Xie, L. M. Miller, C. Chatterjee, O. Averin, N. L. Kelleher, W. A. van der Donk, *Science* **2004**, *303*, 679–681; b) R. J. Clark, H. Fischer, L. Dempster, N. L. Daly, K. J. Rosengren, S. T. Nevin, F. A. Meunier, D. J. Adams, D. J. Craik, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13767–13772; c) M. S. Donia, B. J. Hathaway, S. Sudek, M. G. Haygood, M. J. Rosovitz, J. Ravel, E. W. Schmidt, *Nat. Chem. Biol.* **2006**, *2*, 729–735.

[2] a) F. Al-Obeidi, A. M. D. Castrucci, M. E. Hadley, V. J. Hruby, *J. Med. Chem.* **1989**, *32*, 2555–2561; b) R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, D. Obrecht, P. R. E. Mittl, M. G. Grutter, J. A. Robinson, *ChemBioChem* **2006**, *7*, 515–526; c) R. L. A. Dias, R. Fasan, K. Moehle, A. Renard, D. Obrecht, J. A. Robinson, *J. Am. Chem. Soc.* **2006**, *128*, 2726–2732; d) O. Demmer, A. O. Frank, F. Hagn, M. Schottelius, L. Marinelli, S. Cosconati, R. Brack-Werner, S. Kremb, H. J. Wester, H. Kessler, *Angew. Chem. Int. Ed.* **2012**, *51*, 8110–8113; *Angew. Chem.* **2012**, *124*, 8234–8237; e) V. Dewan, T. Liu, K. M. Chen, Z. Qian, Y. Xiao, L. Kleiman, K. V. Mahasenan, C. Li, H. Matsuo, D. Pei, K. Musier-Forsyth, *ACS Chem. Biol.* **2012**, *7*, 761–769; f) N. Bionda, M. Stawikowski, R. Stawikowska, M. Cudic, F. Lopez-Vallejo, D. Treitl, J. Medina-Franco, P. Cudic, *ChemMedChem* **2012**, *7*, 871–882; g) D. S. Nielsen, H. N. Hoang, R. J. Lohman, T. A. Hill, A. J. Lucke, D. J.

Craik, D. J. Edmonds, D. A. Griffith, C. J. Rotter, R. B. Ruggeri, D. A. Price, S. Liras, D. P. Fairlie, *Angew. Chem. Int. Ed.* **2014**, *53*, 12059 – 12063; *Angew. Chem.* **2014**, *126*, 12255 – 12259; h) A. Muppidi, Z. Wang, X. Li, J. Chen, Q. Lin, *Chem. Commun.* **2011**, *47*, 9396 – 9398; i) A. J. Kamens, R. J. Eisert, T. Corlin, J. D. Baleja, J. A. Kritzer, *Biochemistry* **2014**, *53*, 4758 – 4760.

[3] a) E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discovery* **2008**, *7*, 608 – 624; b) E. Marsault, M. L. Peterson, *J. Med. Chem.* **2011**, *54*, 1961 – 2004; c) E. A. Villar, D. Beglov, S. Chennamadhavuni, J. A. Porco, Jr., D. Kozakov, S. Vajda, A. Whitty, *Nat. Chem. Biol.* **2014**, *10*, 723 – 731.

[4] N. H. Tan, J. Zhou, *Chem. Rev.* **2006**, *106*, 840 – 895.

[5] D. A. Bushnell, P. Cramer, R. D. Kornberg, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1218 – 1222.

[6] H. E. Hallen, H. Luo, J. S. Scott-Craig, J. D. Walton, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19097 – 19101.

[7] Y. Q. Tang, J. Yuan, G. Osapay, K. Osapay, D. Tran, C. J. Miller, A. J. Ouellette, M. E. Selsted, *Science* **1999**, *286*, 498 – 502.

[8] a) S. D. Jolad, J. J. Hoffmann, S. J. Torrance, R. M. Wiedhopf, J. R. Cole, S. K. Arora, R. B. Bates, R. L. Gargiulo, G. R. Kriek, *J. Am. Chem. Soc.* **1977**, *99*, 8040 – 8045; b) J. Kobayashi, H. Suzuki, K. Shimbo, K. Takeya, H. Morita, *J. Org. Chem.* **2001**, *66*, 6626 – 6633.

[9] a) Y. Sun, G. S. Lu, J. P. Tam, *Org. Lett.* **2001**, *3*, 1681 – 1684; b) W. D. Kohn, L. Zhang, J. A. Weigel, *Org. Lett.* **2001**, *3*, 971 – 974; c) J. Ruiz-Rodríguez, J. Spengler, F. Albericio, *Angew. Chem. Int. Ed.* **2009**, *48*, 8564 – 8567; *Angew. Chem.* **2009**, *121*, 8716 – 8719; d) M. Bartoloni, R. U. Kadam, J. Schwartz, J. Furrer, T. Darbre, J. L. Reymond, *Chem. Commun.* **2011**, *47*, 12634 – 12636; e) B. K. W. Chung, J. L. Hickey, C. C. G. Scully, S. Zaretsky, A. K. Yudin, *MedChemComm* **2013**, *4*, 1124 – 1128; f) T. Karskela, P. Virta, H. Lonnberg, *Curr. Org. Synth.* **2006**, *3*, 283 – 311.

[10] J. M. Smith, N. C. Hill, P. J. Krasniak, R. Fasan, *Org. Biomol. Chem.* **2014**, *12*, 1135 – 1142.

[11] a) Y. Sako, J. Morimoto, H. Murakami, H. Suga, *J. Am. Chem. Soc.* **2008**, *130*, 7232 – 7234; b) K. Iwasaki, Y. Goto, T. Katoh, H. Suga, *Org. Biomol. Chem.* **2012**, *10*, 5783 – 5786; c) C. Heinis, T. Rutherford, S. Freund, G. Winter, *Nat. Chem. Biol.* **2009**, *5*, 502 – 507; d) A. Gould, Y. Li, S. Majumder, A. E. Garcia, P. Carlsson, A. Shekhtman, J. A. Camarero, *Mol. Biosyst.* **2012**, *8*, 1359 – 1365; e) I. R. Rebollo, A. Angelini, C. Heinis, *Medchemcomm* **2013**, *4*, 145 – 150.

[12] a) A. R. Horswill, S. N. Savinov, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15591 – 15596; b) A. Tavassoli, S. J. Benkovic, *Angew. Chem. Int. Ed.* **2005**, *44*, 2760 – 2763; *Angew. Chem.* **2005**, *117*, 2820 – 2823; c) A. Tavassoli, Q. Lu, J. Gam, H. Pan, S. J. Benkovic, S. N. Cohen, *ACS Chem. Biol.* **2008**, *3*, 757 – 764; d) T. S. Young, D. D. Young, I. Ahmad, J. M. Louis, S. J. Benkovic, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11052 – 11056; e) J. H. Appleby-Tagoe, I. V. Thiel, Y. Wang, H. D. Mootz, X. Q. Liu, *J. Biol. Chem.* **2011**, *286*, 34440 – 34447.

[13] a) L. Cheng, T. A. Naumann, A. R. Horswill, S. J. Hong, B. J. Venters, J. W. Tomsho, S. J. Benkovic, K. C. Keiler, *Protein Sci.* **2007**, *16*, 1535 – 1542; b) I. K. Nordgren, A. Tavassoli, *Mol. Biosyst.* **2014**, *10*, 485 – 490; c) J. A. Kritzer, S. Hamamichi, J. M. McCaffery, S. Santagata, T. A. Naumann, K. A. Caldwell, G. A. Caldwell, S. Lindquist, *Nat. Chem. Biol.* **2009**, *5*, 655 – 663.

[14] N. Bionda, A. L. Cryan, R. Fasan, *ACS Chem. Biol.* **2014**, *9*, 2008 – 2013.

[15] C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638 – 13643.

[16] T. A. Naumann, S. N. Savinov, S. J. Benkovic, *Biotechnol. Bioeng.* **2005**, *92*, 820 – 830.

[17] This procedure assumes that the monocyclic and bicyclic peptide products have similar ionization properties, an assumption which was corroborated by comparison of the product distribution as calculated by HPLC/UV–visible analysis and LC-MS ion-exchange chromatography (IEC) analysis (Figure S15).

[18] H. X. Zhou, G. Rivas, A. P. Minton, *Annu. Rev. Biophys. Bioeng.* **2008**, *37*, 375 – 397.

[19] C. P. Scott, E. Abel-Santos, A. D. Jones, S. J. Benkovic, *Chem. Biol.* **2001**, *8*, 801 – 815.

[20] B. A. Katz, R. T. Cass, *J. Biol. Chem.* **1997**, *272*, 13220 – 13228.

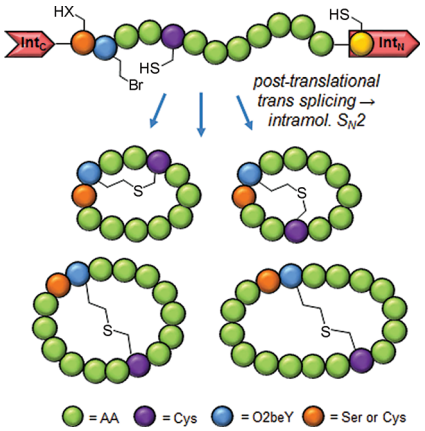
Manuscript received: April 5, 2015

Accepted article published: July 14, 2015

Final article published: ■■ ▮▮, 0000

COMMUNICATIONS

Ringing up peptides: The combination of split intein-mediated peptide circularization with intramolecular crosslinking through a genetically encoded, cysteine-reactive amino acid (O2beY) provides a versatile strategy to produce structurally diverse bicyclic peptides in bacterial cells. By using this approach, a bicyclic peptide with enhanced binding affinity for streptavidin was isolated.



*N. Bionda, R. Fasan**

■■ – ■■

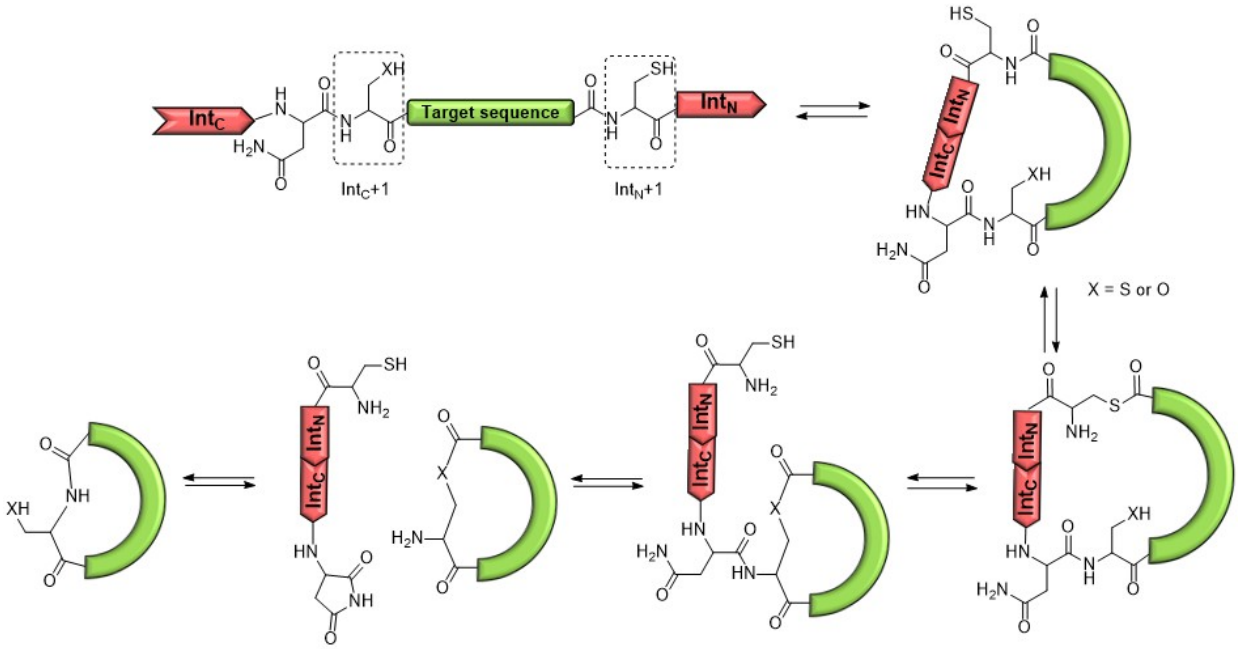
Ribosomal Synthesis of Natural-Product-Like Bicyclic Peptides in *Escherichia coli*

These are not the final page numbers! ↗↗

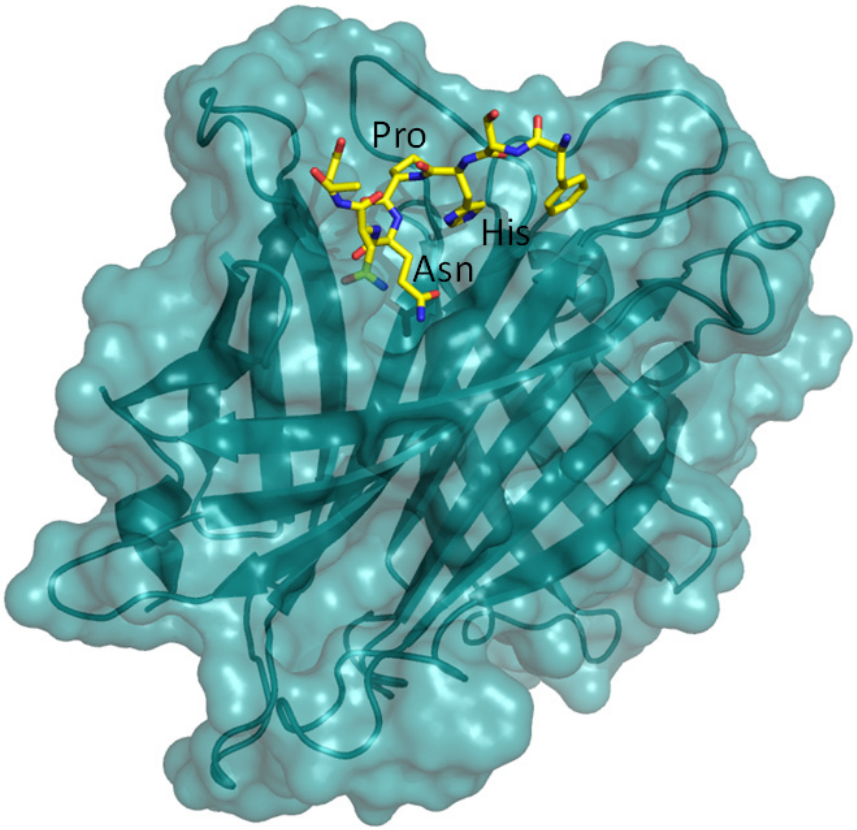
Supplementary Table S1. Oligonucleotide primers

Primer	Sequence
SICLOPPS_for	5'-CAGGTCATATGGTTAAAGTTATCGGTCGTCGATCC-3'
SICLOPPS_rev	5'-CAACAGGTACCTTTAATTGTACCTGCGTCAAGTAATGGAAAG-3'
Z3C_for	5'-CGCAGTTCGCGAACGCGTGCTTAAGTTTTGGCACCGAAATT-3'
Z3C_1/2_rev	5'-GGATGGCAGTTGGTCTAGCTATTGTGGGCGATAGCACCATTAGC-3'
Z3C_2/2_rev	5'-CGCGTTCGCGAACTGCGGATGGCAGTTGGTCTAGCTATTGTG-3'
Z3C(S1C)_1/2_rev	5'-GATGGCAGTTGGTCTAGCAATTGTGGGCGATAGCACCATTAGC-3'
Z3C(S1C)_2/2_rev	5'-CGCGTTCGCGAACTGCGGATGGCAGTTGGTCTAGCAATTG-3'
Z8C(A12P)_for	5'-TTCGCGAACCCGTGCTTAAG-3'
Z8C(A12P)_rev	5'-CTTAAGCACGGGTTCGCGAA-3'
Z8C(A12N)_for	5'-GTTCTGCAACAACTGCTTAAGTTTGGC-3'
Z8C(A12N)_rev	5'-GCCAAACTTAAGCAGTTGTTGCAGAAC-3'
Z8C(16mer)_rev	5'-GAACTTCTTAAGCACGCATCGCCTTTCGCGTTGCAGAACTGCGGA-3'
Z8C(18mer)_rev	5'-GAACTTCTTAAGCACGCCTGGGTATCGCCTTTCGCGTTGCAGAACTGCGGA-3'
SICLOPPS_S(CatU+8)_for	GCCCACAATAGCTAGACCAACGTG
SICLOPPS_S(CatU+8)_rev	CACGTTGGTCTAGCTATTGTGGGC

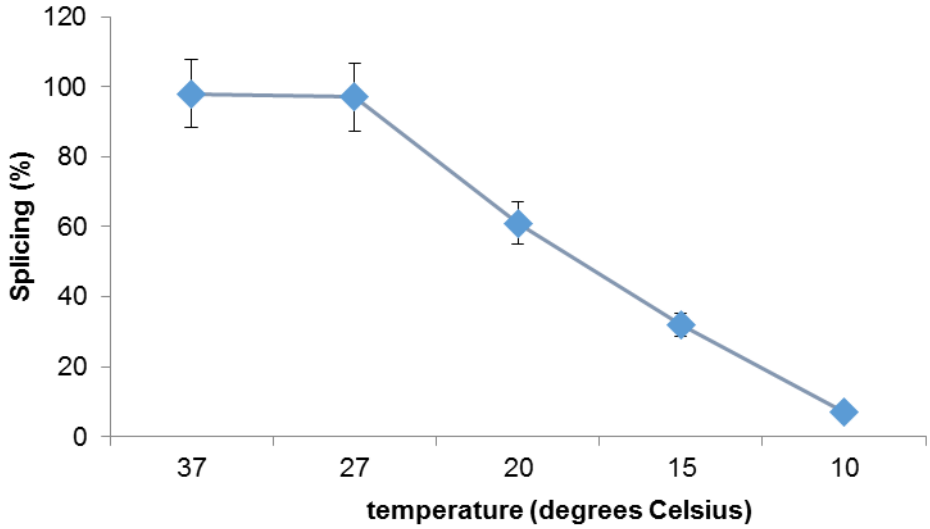
Supplementary Figure S1. Overview of the mechanism of split intein-mediated peptide circularization method (SICLOPPS).[1] Int_N and Int_C correspond to the N-domain and C-domain respectively, of *Synechocystis* sp. PCC6803 DnaE split intein.[2] The green segment corresponds to a variable peptide sequence. The Int_N+1 cysteine and Int_C+1 cysteine (or serine) are indicated.



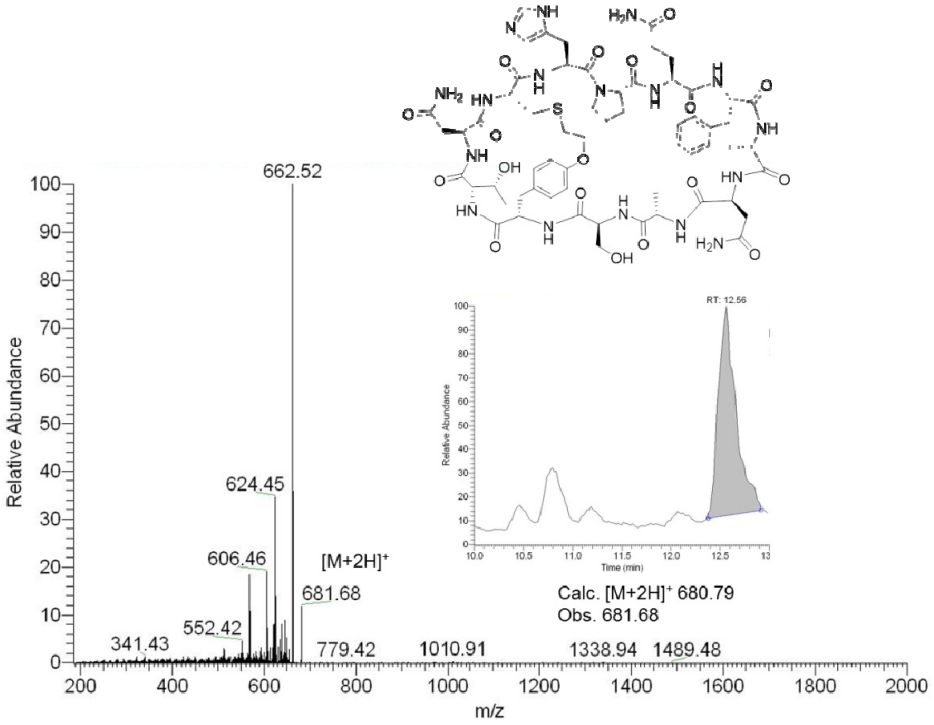
Supplementary Figure S2. Crystal structure of streptavidin (green, surface model) in complex with the linear peptide Ac-FSHPQNT-NH_2 (yellow, stick model) (pdb 1VWA). The residues corresponding to the HPQ binding motif are labeled.



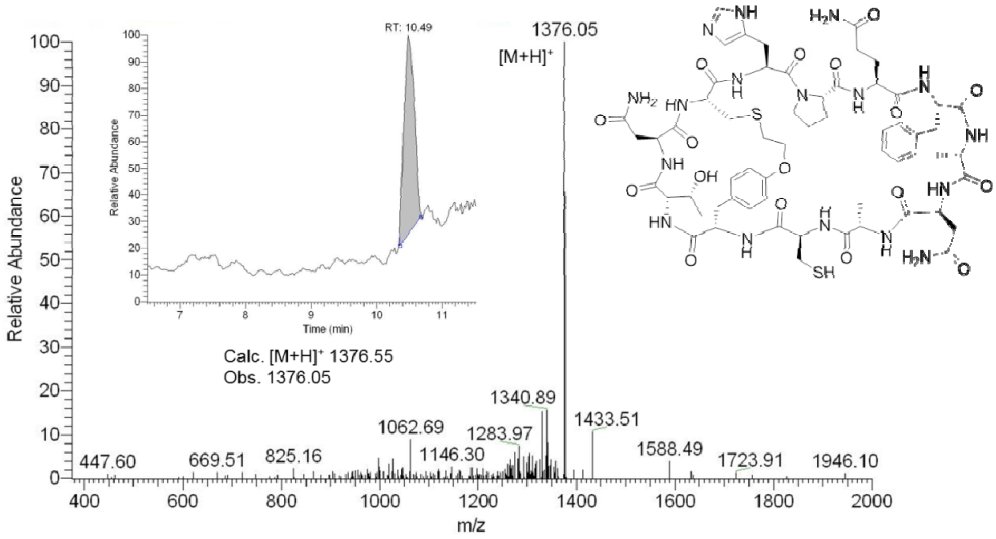
Supplementary Figure S3. Temperature dependence of the extent of DnaE-catalyzed trans splicing. *E. coli* cells expressing the construct Z3C_OpgY (Entry 3, Table 1) were grown at the indicated temperature for 6 hours after IPTG induction. The extent of protein splicing was determined by LC-MS analysis of the full-length precursor protein and spliced protein after isolation from the cell lysate using the C-terminal chitin binding domain.



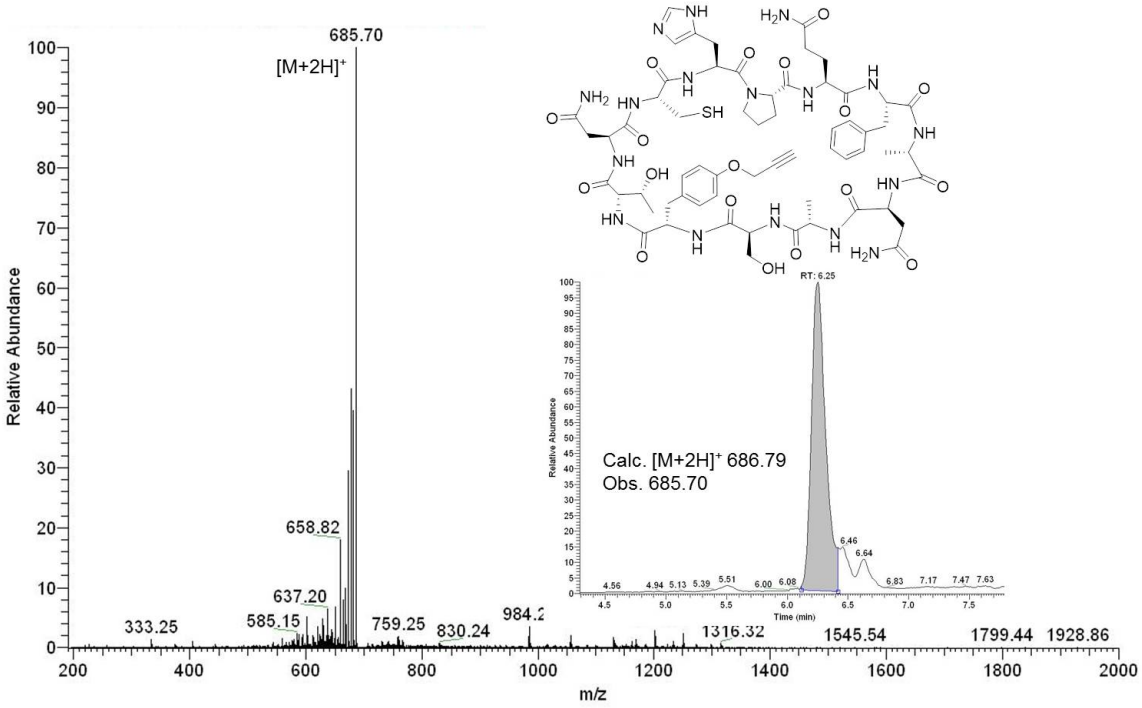
Supplementary Figure S4. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C_O2beY.



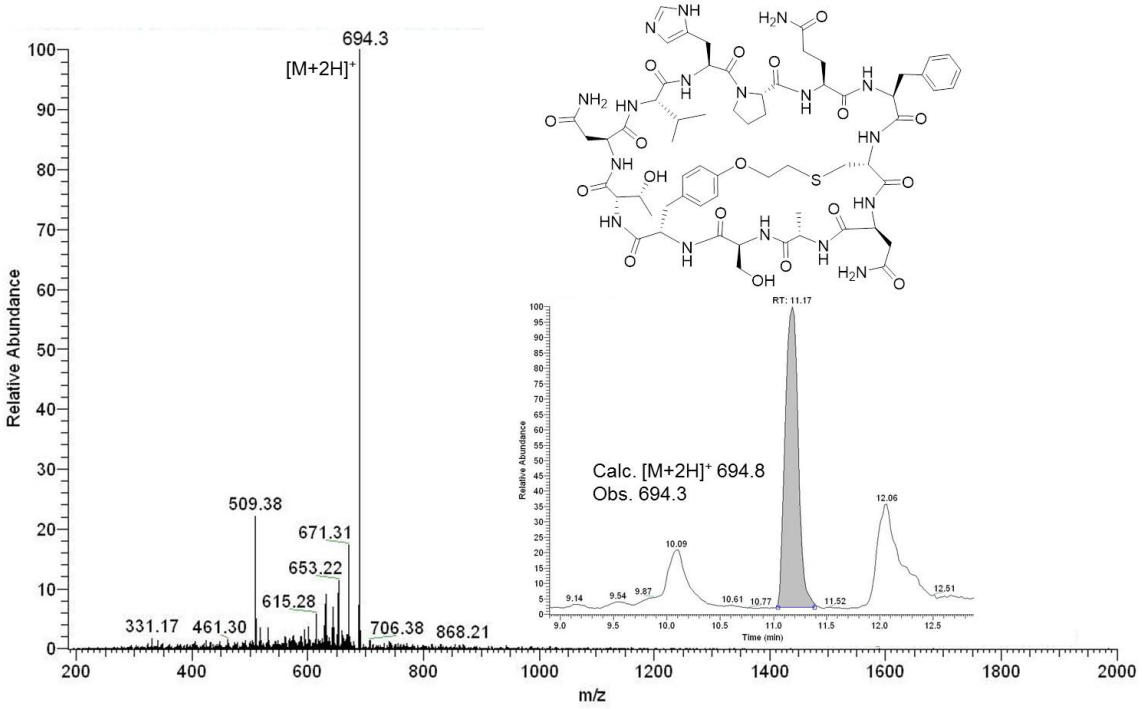
Supplementary Figure S5. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C(S1C)_O2beY.



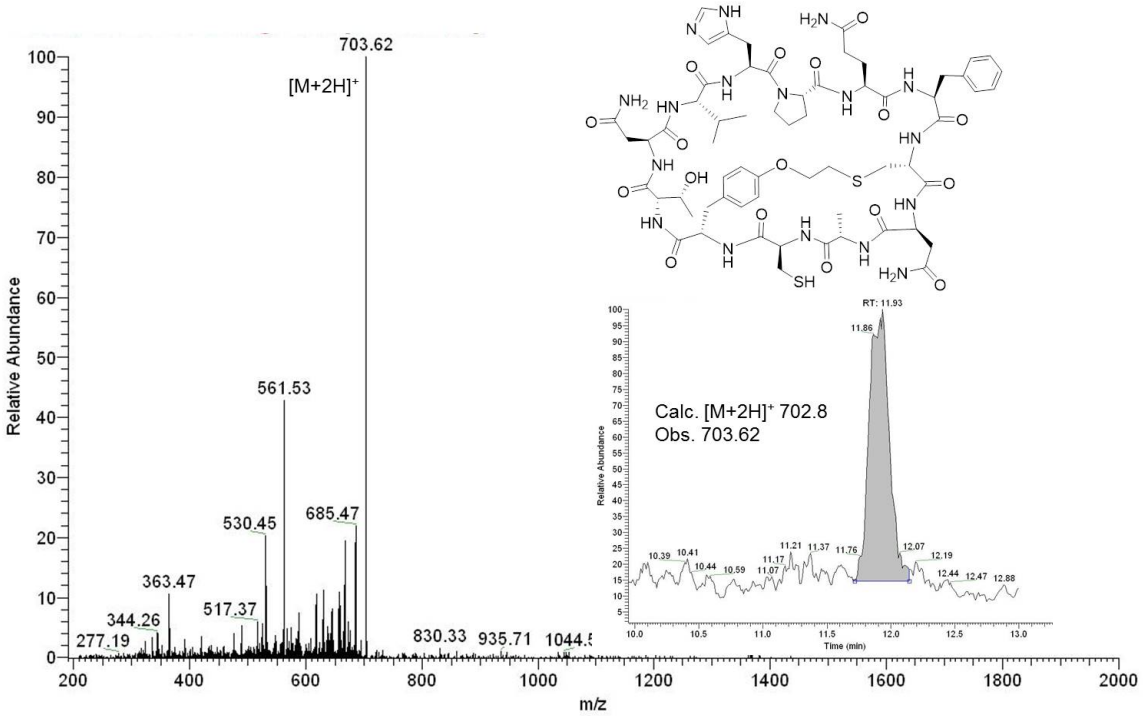
Supplementary Figure S6. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C_OpgY.



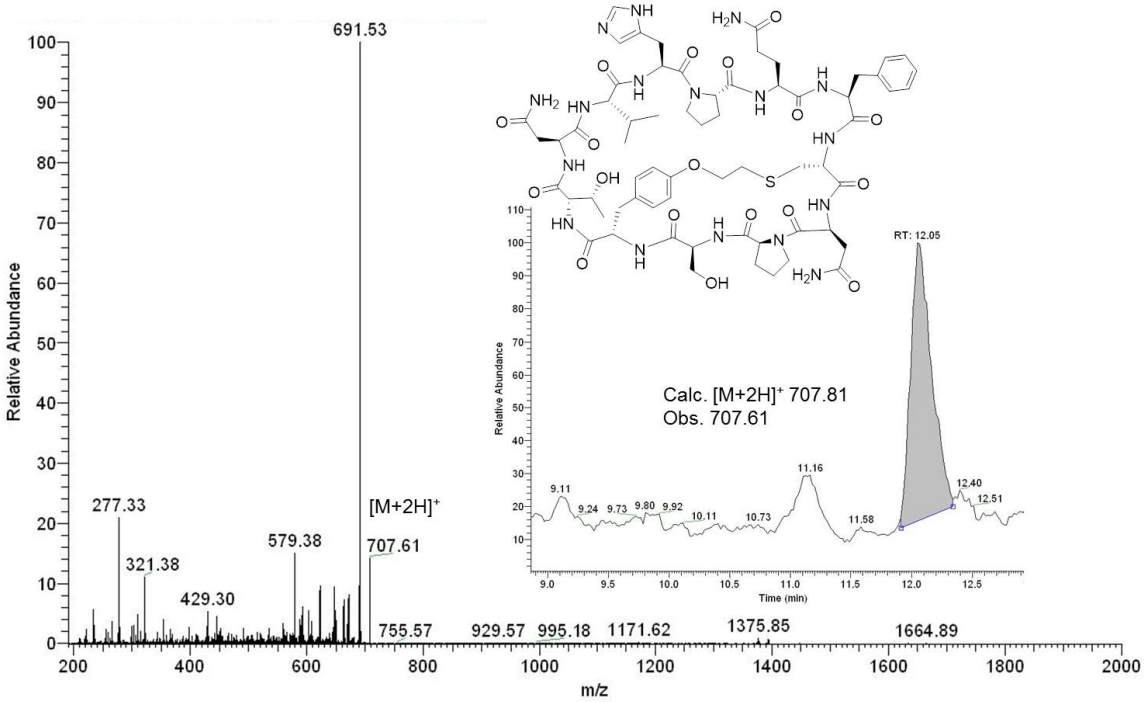
Supplementary Figure S7. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C_O2beY.



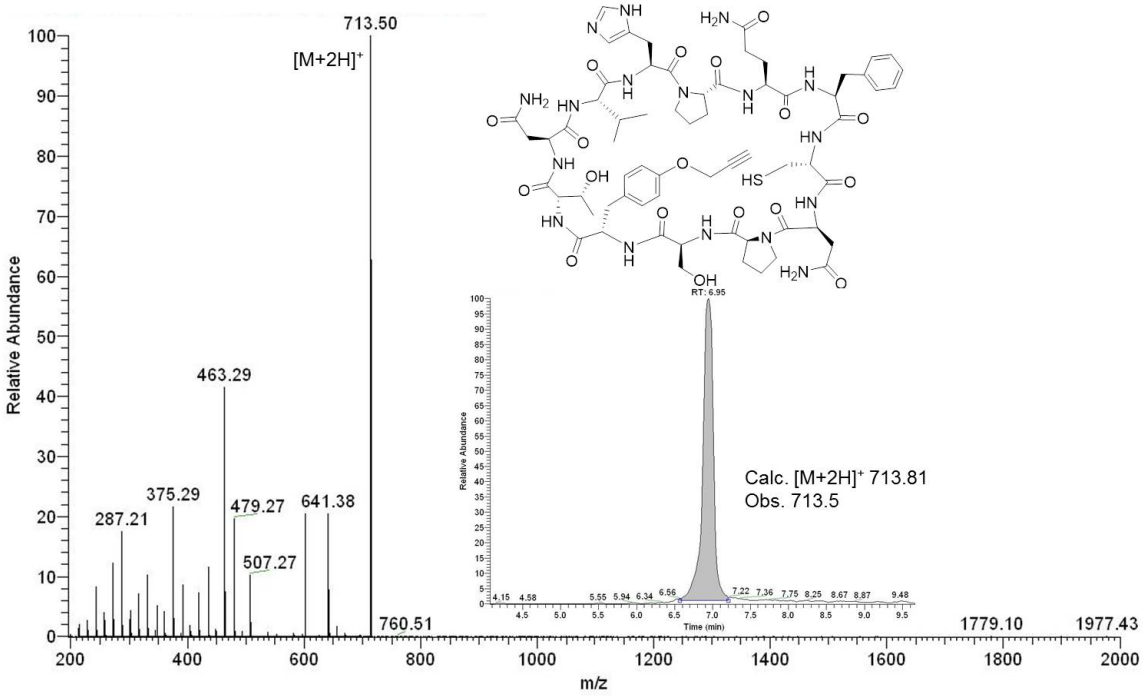
Supplementary Figure S8. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(S1C)_O2beY.



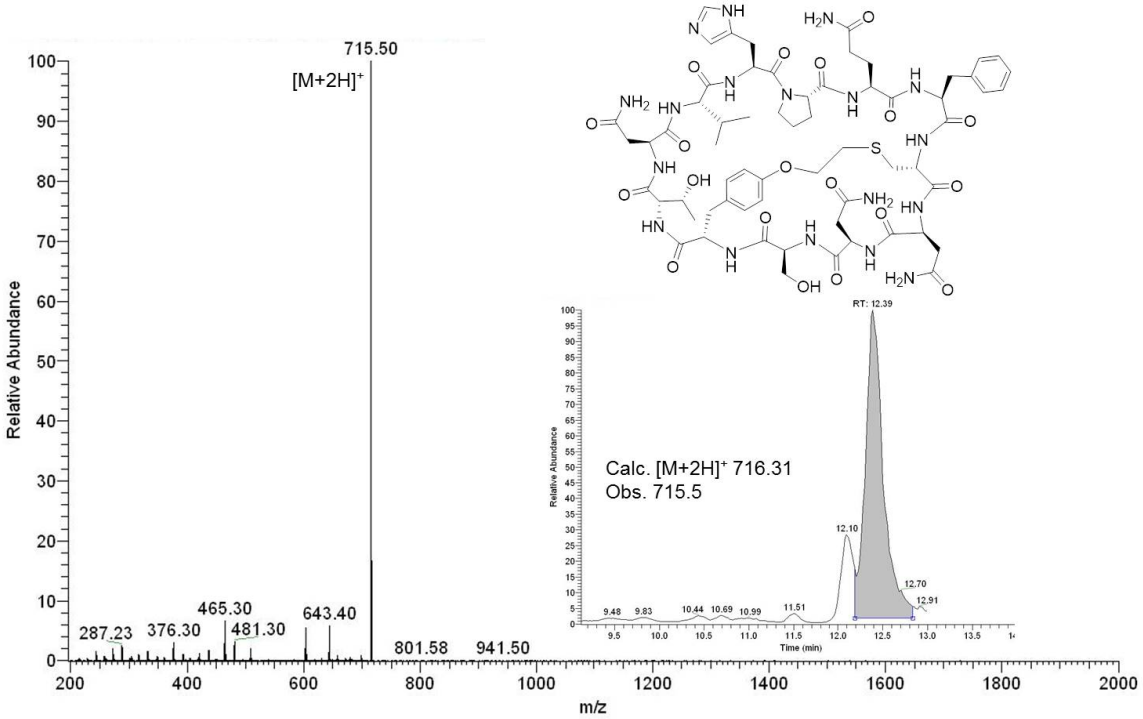
Supplementary Figure S9. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12P)_O2beY.



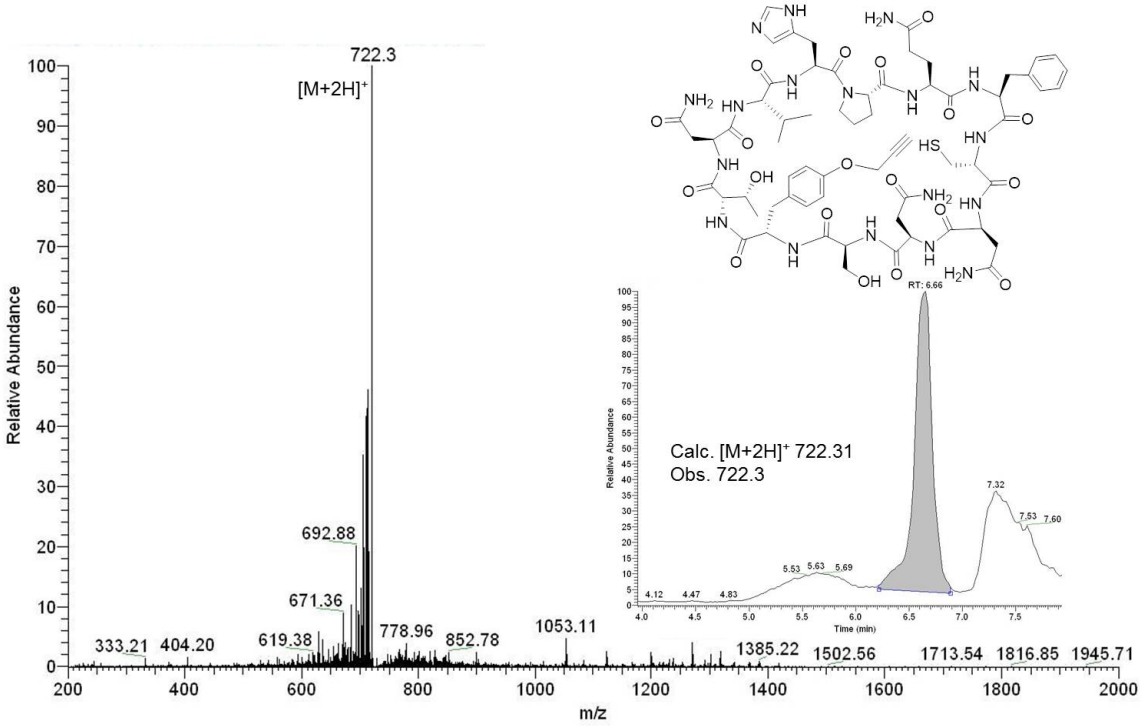
Supplementary Figure S10. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12P)_OpgY.



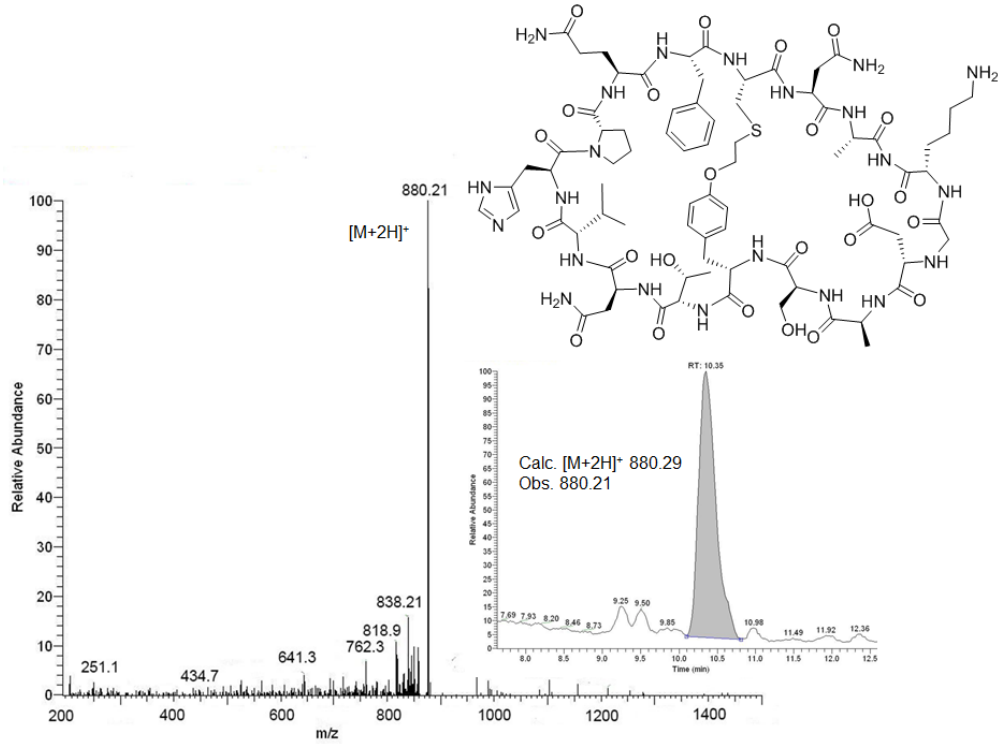
Supplementary Figure S11. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12N)_O2beY.



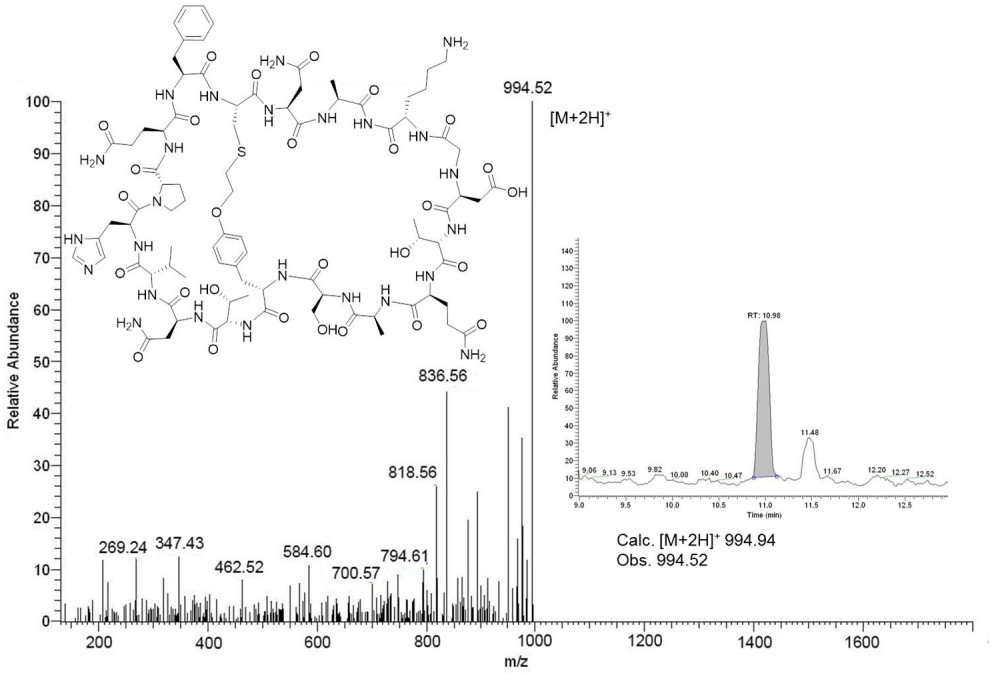
Supplementary Figure S12. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12N)_OpgY.



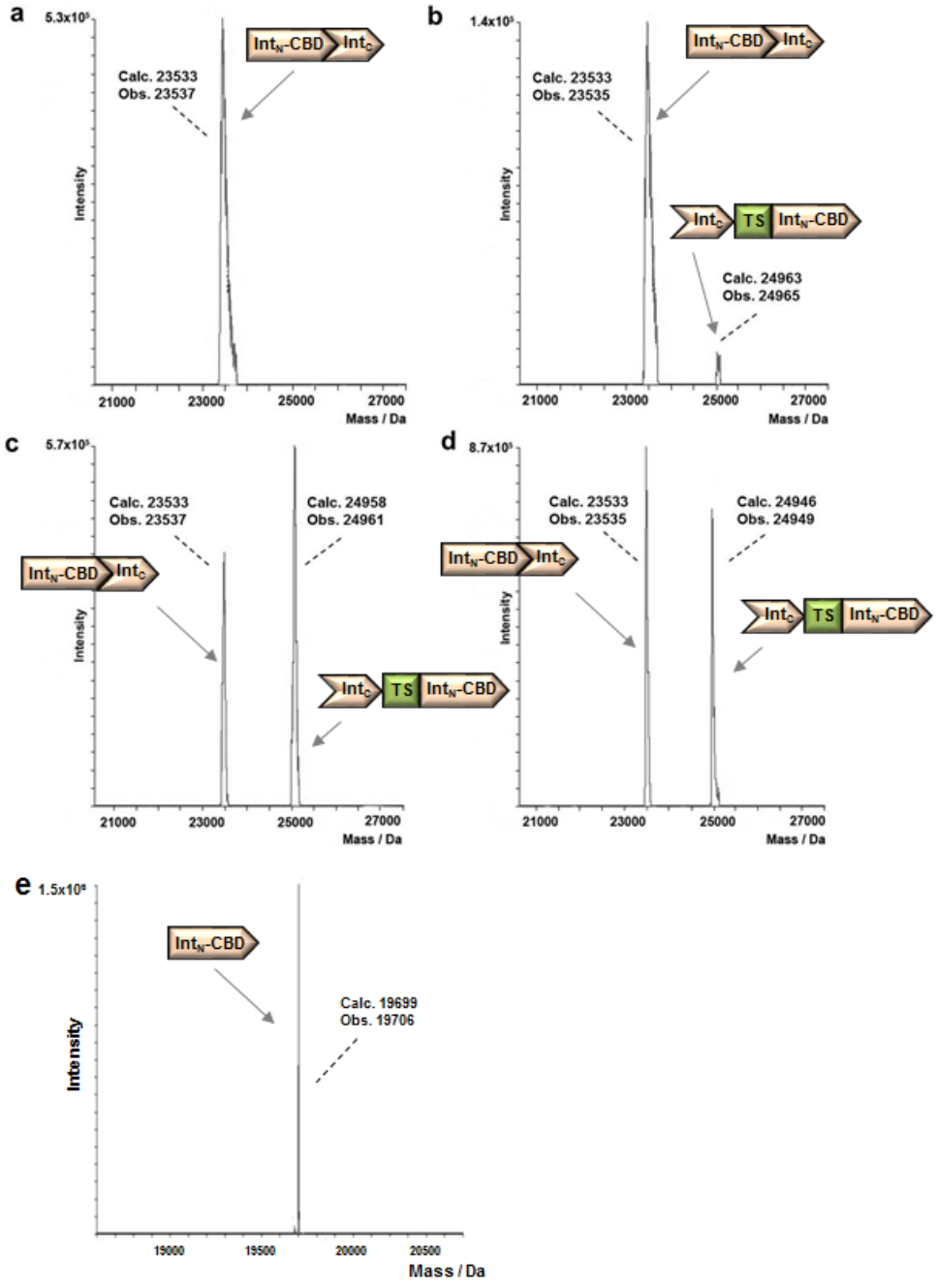
Supplementary Figure S13. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(16mer)_O2beY.



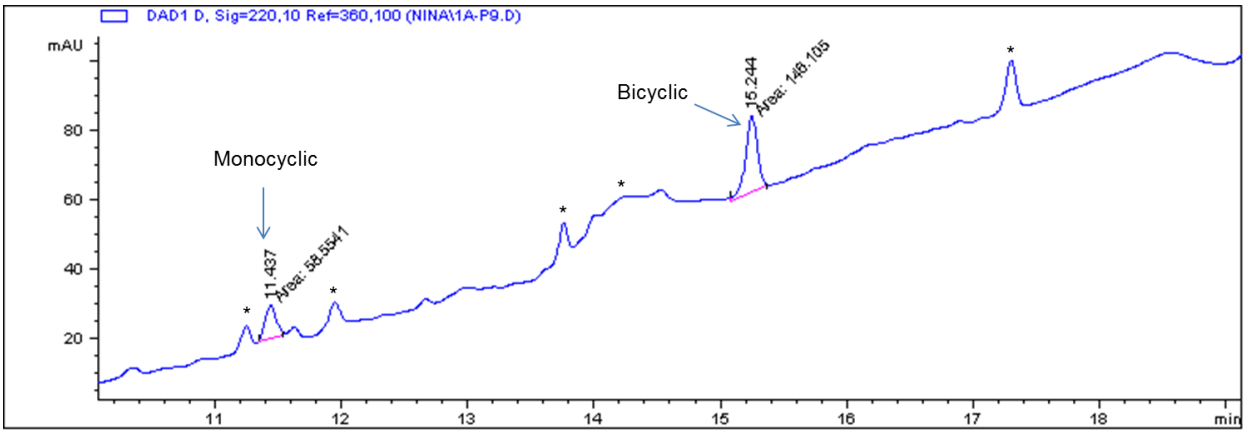
Supplementary Figure S14. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(18mer)_O2beY.



Supplementary Figure S15. Deconvoluted MS spectra of CBD-containing proteins isolated from cells expressing: (a) Z8C_O2beY construct, (b) Z8C(A12N)_O2beY construct, (c) Z8C(A12P)_OpgY construct, (d) Z8C(A12P)_O2beY construct. Under standard LC-MS conditions, the DnaE Int_C/Int_N-CBD complex elutes in associated form (calc. $[M+H]^+$ *m/z* 25533), thus enabling direct comparison of the signal corresponding to the full-length protein (~24950 Da) with that of the spliced DnaE intein. Under denaturing conditions, the complex dissociates and the N- and C-terminal domains of DnaE elute separately as illustrated by panel (e), which provides the LC-MS spectrum of the same protein as in (a) but after incubation in 6 M guanidine hydrochloride buffer.

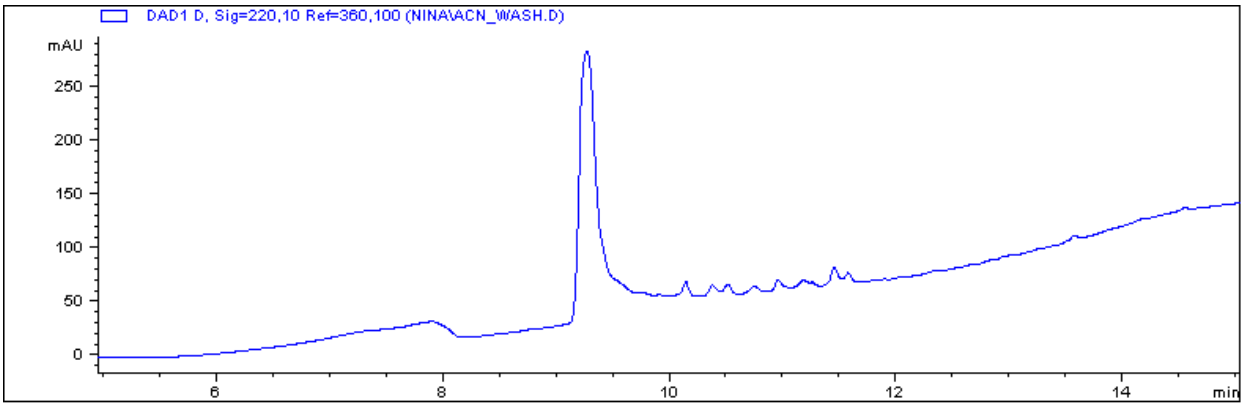


Supplementary Figure S16. HPLC analysis of the butanol extract from lysate of cells expressing the Z8C_O2beY construct. The identity of the monocyclic and bicyclic peptide (labeled) was confirmed by LCMS. Comparison of the peak areas yielded a product distribution consisting of 72% monocyclic peptide and 28% bicyclic peptide which is in excellent agreement with the relative amounts of these product as determined by LC-MS ion-extract chromatogram analysis (76% and 24%, respectively, Table 1). Peaks labeled with * correspond to unrelated components of the cell lysate.

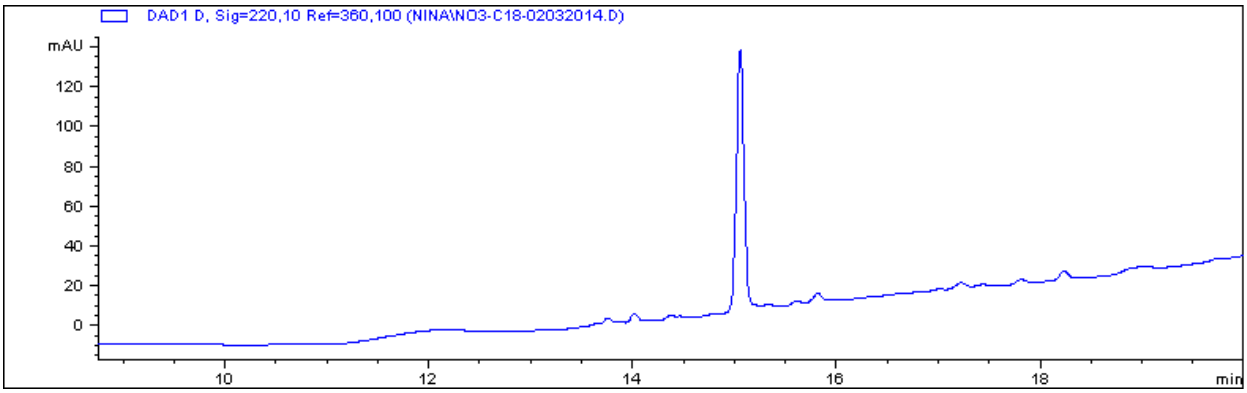


Supplementary Figure S17. RP HPLC traces corresponding to purified *cyclo*-Z3C (A) and *bicyclo*-Z8C (B).

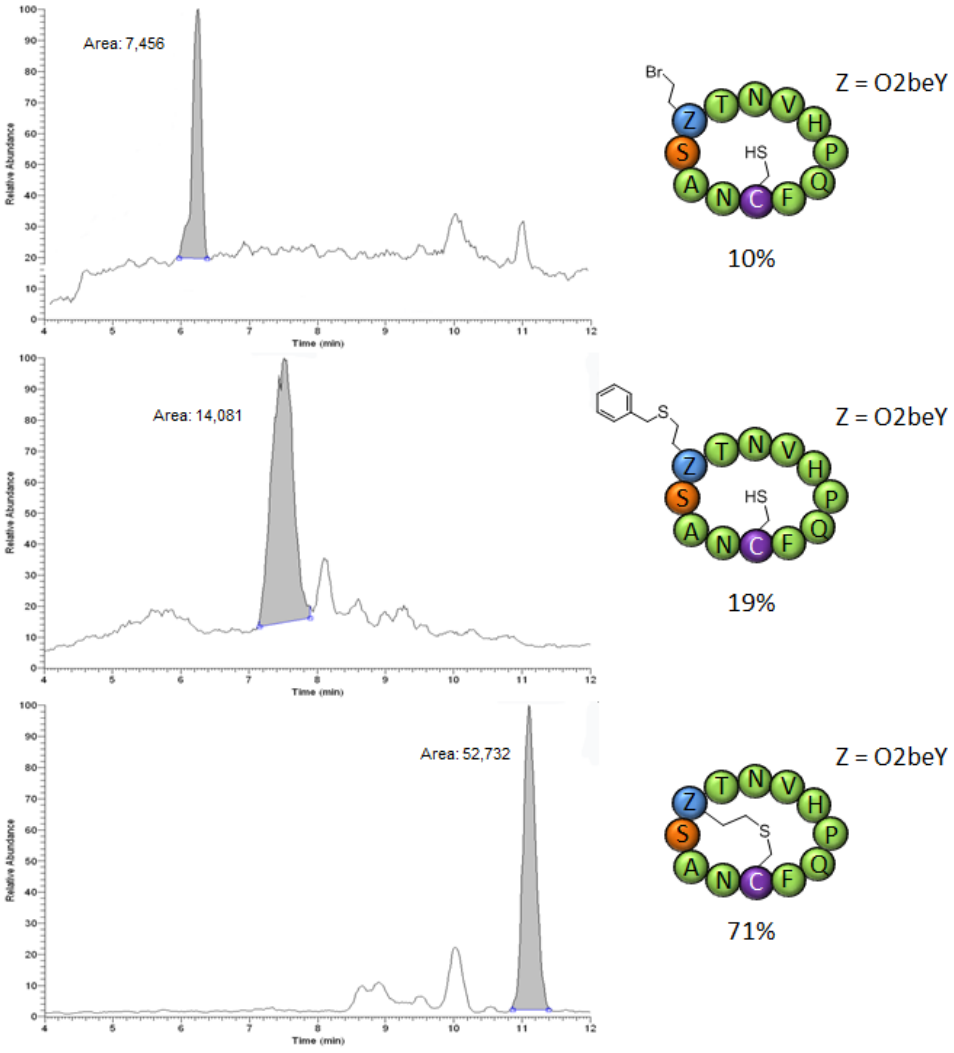
A



B



Supplementary Figure S18. Extracted-ion chromatograms corresponding to monocyclic and bicyclic peptide products isolated via streptavidin-affinity chromatography from cell lysate of cells expressing Z8C_O2beY pre-treated with benzyl mercaptan (1 mM, 1 hour, room temperature). The ratio of monocyclic : bicyclic peptide (29:71) is comparable to that obtained in the absence of benzyl mercaptan treatment (24:76, Table 1), confirming that the bicyclization reaction occurs intracellularly.



Experimental procedures

Reagents and analytical methods. Chemical reagents and solvents were purchased from Sigma-Aldrich, and Chem-Impex. Silica gel chromatography was carried out using AMD Silica Gel 60 (230-400 mesh). Streptavidin-binding agarose beads were purchased from Pierce and chitin beads were obtained from BioLabs. ^{1}H and ^{13}C NMR spectra were recorded on the Bruker Avance spectrometers (400 and 125 MHz, respectively) using solvent peak as reference. The NMR data are reported as chemical shifts (δ ppm). LC-MS analyses were performed on Thermo Scientific LTQ Velos ESI/ion-trap mass spectrometer coupled to an Accela U-HPLC. Peptides and proteins were analyzed using Thermo Scientific HyPurity C4 column (particle size 5 µm, 100 x 2.1 mm) or Thermo Scientific Hypersil Gold C4 (particle size 5 µm, 100 x 2.1 mm) and a linear gradient 5% to 95% ACN (with 0.1% formic acid) in water (with 0.1% formic acid) over 13 min. The unnatural amino acids *O*-(2-bromoethyl)-tyrosine (O2beY) and *O*-propargyltyrosine (OpgY) were synthesized as described previously.[3]

Cloning and plasmid construction. Oligonucleotides were purchased from Integrated DNA Technologies and their sequences are provided in Supplementary Table S1. The genes encoding the *N*-terminal and *C*-terminal DnaE inteins from *Synechocystis sp*. were extracted from pSFBAD09 and pJJDuet30 plasmids[4] (Addgene #11963 and #11962, respectively). DnaE-Int_N was amplified using primers SICLOPPS_for as forward primer and Z3C_1/2_rev and Z3C(S1C)_1/2_rev, respectively, as reverse primers. The PCR products (0.6 Kbp) were used as templates for a second PCR reaction using forward primer SICLOPPS_for and Z3C_2/2_rev and Z3C(S1C)_2/2_rev as reverse primers, respectively. The DnaE-Int_C was amplified using primers Z3C_for as forward primer and SICLOPPS_rev. The fragments were fused to generate a gene encoding the DnaE-Int_N fused to the DnaE-Int_C through the desired target sequence, as described previously.[3a] The resulting products (0.5 Kbp) were digested with *Nde* I and *Kpn* I and cloned into pSFBAD09 to provide the plasmids p_Z3C and p_Z3C(S1C). Ligation at the *Kpn* I site also introduces the chitin-binding domain (CBD). The p_Z8C construct was prepared the same way using Z8C_1/2_rev and Z8C_2/2_rev and SICLOPPS_for. p_Z8C(S1C) construct was prepared similarly using Z8C(S1C)_1/2_rev and Z8C(S1C)_2/2_rev and SICLOPPS_for. The gene encoding for Z8C(A12P) was constructed using p_Z8C as template and SICLOPPS_for and Z8C(A12P)_for as forward primers and SICLOPPS_rev and Z8C(A12P)_rev as reverse primers. Similarly, the gene encoding for Z8C(A12N) was constructed using p_Z8C as template and SICLOPPS_for and Z8C(A12N)_for as forward

primers and SICLOPPS_rev and Z8C(A12N)_rev as reverse primers. The resulting products (0.5 Kbp) were digested with *Nde* I and *Kpn* I and cloned into pSFBAD09 to provide the plasmids p_Z8C(A12P) and p_Z8C(A12N). The gene encoding for Z8C(16mer) was constructed using p_Z8C as a template and SICLOPPS_for and Z8C_16mer_rev primers. The PCR product (0.15 Kbp) was digested using *Nde* I and *Afl* II and cloned into pSFBAD09 to provide the p_ Z8C(16mer). Similarly, the gene encoding for Z8C(18mer) was constructed using p_Z8C as a template and SICLOPPS_for and Z8C_18mer_rev primers. In all vectors the genes encoding for the SICLOPPS construct are under the control of an arabinose-inducible AraC promoter.

Preparation and isolation of bicyclic macrocycles. Chemically competent BL21(DE3) *E. coli* cells were co-transformed with pEVOL_O2beY and the appropriate pBAD-based vector encoding for the desired precursor polypeptide. Overnight cultures were grown in Luria-Bertani (LB) medium supplemented with 50 mg/L of ampicillin and 26 mg/L of chloramphenicol and used to inoculate 0.2 L of M9 media containing the same concentration of antibiotics and supplemented with 1% glycerol. Bacterial cultures were grown at 37°C until OD_{600} reached 0.6, at which point O2beY (2 mM), L-arabinose (0.06% m/v) were added to induce protein expression. Cultures were grown for additional 12 hours at 27°C followed by additional 3 hours at 37°C after overnight growth. Cells were harvested by centrifugation at 4,000 rpm for 25 min, resuspended in 50 mM Tris, 300 mM NaCl, 20 mM imidazole buffer (pH 7.5) and lysed by sonication. Upon centrifugation at 13,000 rpm for 30 min, cell lysate was incubated with streptavidin-coated beads for 3 hours under gentle shaking on ice. Beads were washed two times with the same buffer followed by incubation with acetonitrile:H_2O (70:30 v/v) for one minute to release any streptavidin-bound peptides. Eluates were lyophilized and the identity of the peptides determined by LC-MS analysis. For measuring the relative amount of the mono- and bicyclic peptide products all of these peptides were searched and taken into consideration for determining the extent of cyclization, as well as potential adducts resulting from nucleophilic displacement of O2beY side-chain 2-bromoethyl group with cysteine, glutathione, or water. None of the latter adducts was observed under the applied experimental conditions. To analyze the amount of protein splicing occurred *in vivo*, the same cell lysate samples were incubated with chitin beads for one hour on ice. Beads were washed two times with buffer followed by incubation with acetonitrile:H_2O (70:30 v/v) for one minute to release any chitin-bound protein. Eluates were analyzed by LC-MS.

Solid-phase binding assay

The streptavidin-binding affinity of bicyclic peptides was evaluated using an adapted version of a previously reported assay[5]. Constructs Z3C_O2beY, Z8C_O2beY, Z8C(S1C)_O2beY and Z3C_OpgY were expressed in 1 L cultures as described above. After harvesting, cells were resuspended in 40 mL of 100 mM sodium phosphate buffer (150 mM NaCl, pH 7.5) and lysed by sonication. After clarification by centrifugation, the cell lysate was extracted twice with *n*-butanol (1:1 v/v) and the butanol extract was lyophilized. The cyclic peptide products used for competitive binding (*cyclo*-Z3C, *bicyclo*-Z8C, and *bicyclo*-Z3C) were subjected to a round of additional purification using streptavidin-coated agarose beads according to the procedure described above. The bicyclic peptides were eluted from the beads using an acetonitrile:water (7:3) solution, lyophilized and resuspended in the binding buffer (50 mM Tris buffer with 10% DMSO and 1 mM TCEP, pH 7.5). Prior to use, white 96-well maleimide activated plates (Pierce Biotechnology, Cat. No. 15152) were washed with 50 mM Tris buffer containing 0.05% Tween-20; the same buffer was used for all subsequent wash steps. The plates were then coated using the butanol extract of cell lysate from cells expressing the Z8C(S1C)_O2beY construct, after lyophilization and resuspension of the lyophilized peptide in binding buffer. Typically, the butanol extract from 1 L of culture was resuspended in 10 mL of binding buffer, 100 µL of which was added to each well of the maleimide-activated plate and left to react overnight at 4°C. After coating, the plates were washed once with 200 µL/well of wash buffer and two blocking steps were performed. First, the wells were blocked with 200 µL of 1 mM cysteine solution in 50 mM Tris buffer (pH 7.5) for 30 min at room temperature; then incubation with 200 µL/well of 5% bovine serum albumin in 100 mM sodium phosphate buffer (150 mM NaCl, pH 7.5) for 30 min at room temperature. Plates were washed three times prior to performing the binding experiments. The competition assay was performed by co-incubating streptavidin-conjugated horseradish peroxidase (streptavidin-HRP) and D-destiobiotin, *bicyclo*- Z3C or *bicyclo*-Z8C, or *cyclo*-Z3C peptide. The concentration range used for D-destiobiotin was 0.4–200 nM and for the monocyclic and bicyclic peptides was 0.05–100 µM. The assay readout was performed as per manufacturer's instructions. Briefly, horseradish peroxidase substrate *O*-phenylenediamine dihydrochloride (OPD) was dissolved in water, added to each well (100 µL) and incubated for 20 min followed by the addition of 50 µL/well of 2.5 M sulfuric acid. The absorbance was measured at 492 nm.

References

[1] C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13638-13643.

[2] H. Wu, Z. Hu, X. Q. Liu, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9226-9231.

[3] a) N. Bionda, A. L. Cryan, R. Fasan, *ACS chemical biology* **2014**, *9*, 2008-2013; b) J. M. Smith, F. Vitali, S. A. Archer, R. Fasan, *Angewandte Chemie* **2011**, *50*, 5075-5080.

[4] S. Zuger, H. Iwai, *Nat Biotechnol* **2005**, *23*, 736-740.

[5] T. A. Naumann, S. N. Savinov, S. J. Benkovic, *Biotechnol. Bioeng.* **2005**, *92*, 820-830.