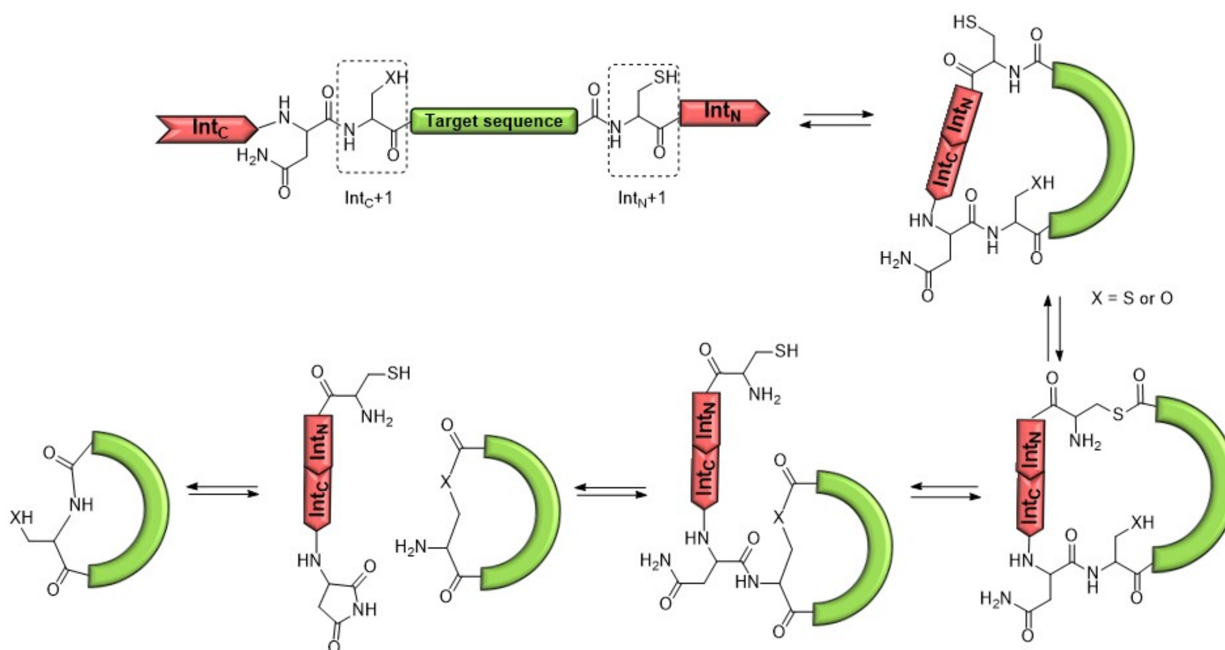
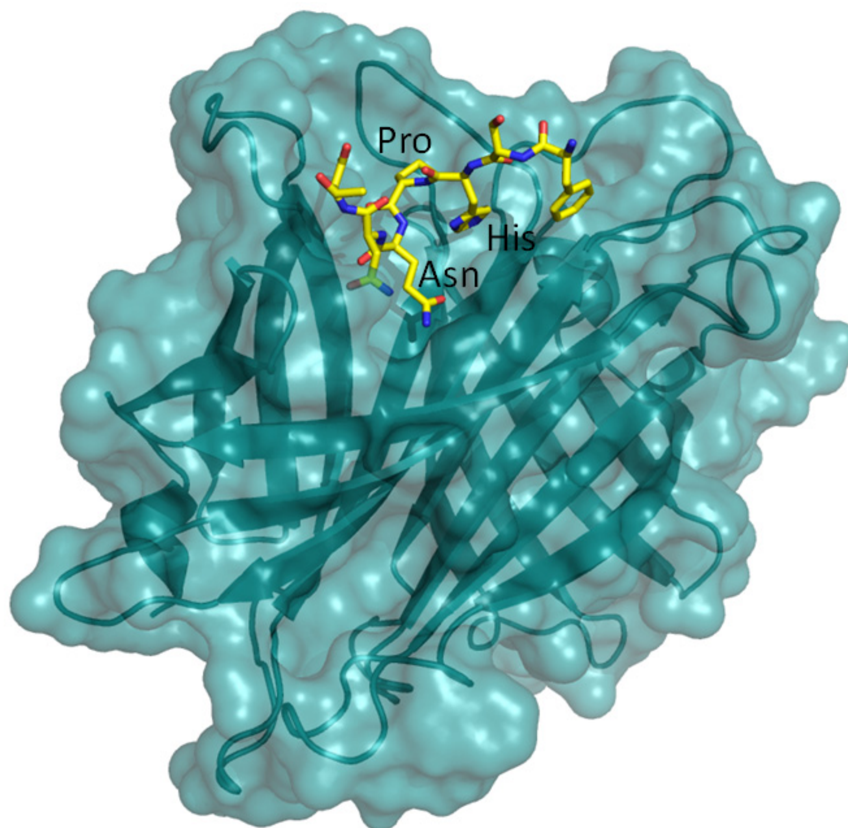
Supplementary Table S1. Oligonucleotide primers

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

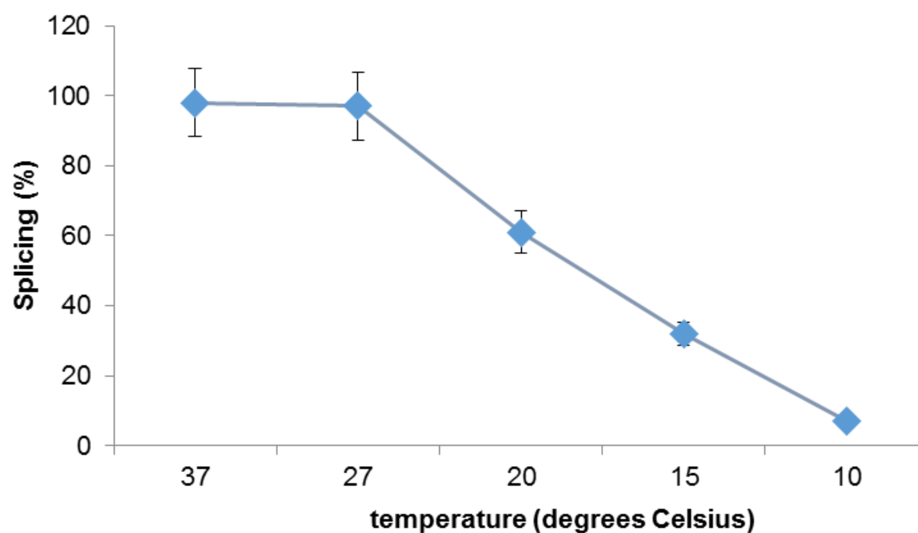
Supplementary Figure S1. Overview of the mechanism of split intein-mediated peptide circularization method (SICLOPPS).[1] Int_N and Int_C correspond to the N-domain and C-domain respectively, of *Synechocystis* sp. PCC6803 DnaE split intein.[2] The green segment corresponds to a variable peptide sequence.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



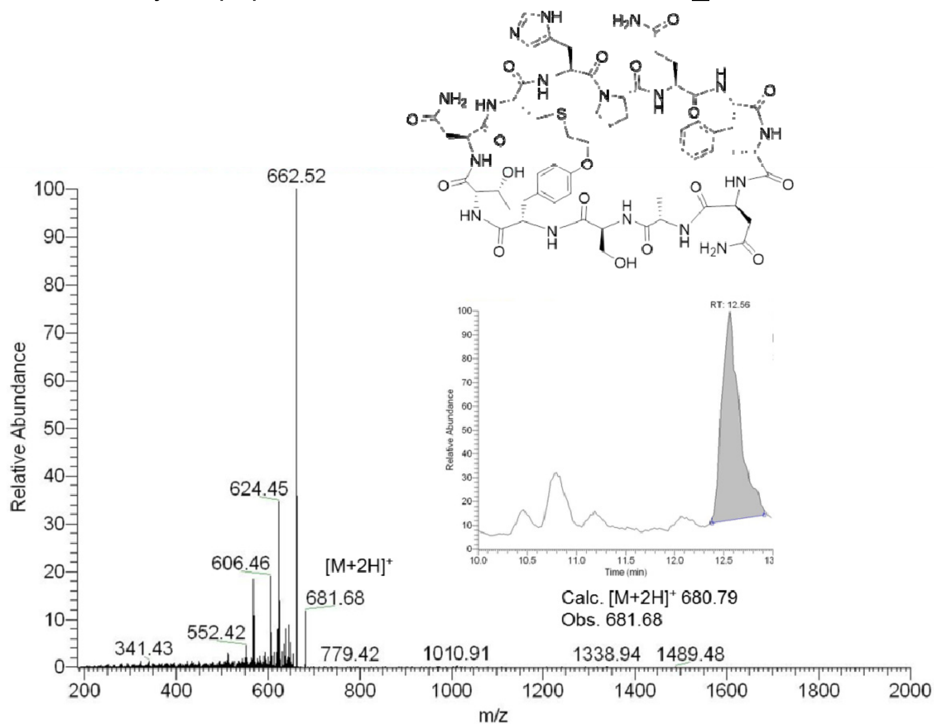
Supplementary Figure S2. Crystal structure of streptavidin (green, surface model) in complex with the linear peptide Ac-FSHPQNT-NH_2 (yellow, stick model) (pdb 1VWA). The residues corresponding to the HPQ binding motif are labeled.



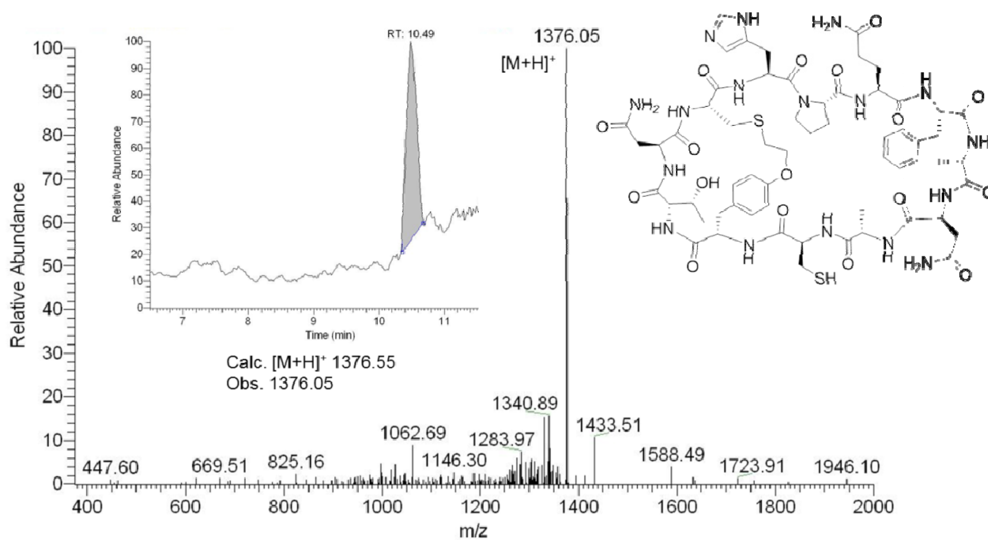
Supplementary Figure S3. Temperature dependence of the extent of DnaE-catalyzed trans splicing. *E. coli* cells expressing the construct Z3C_OpgY (Entry 3, Table 1) were grown at the indicated temperature for 6 hours after IPTG induction. The extent of protein splicing was determined by LC-MS analysis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



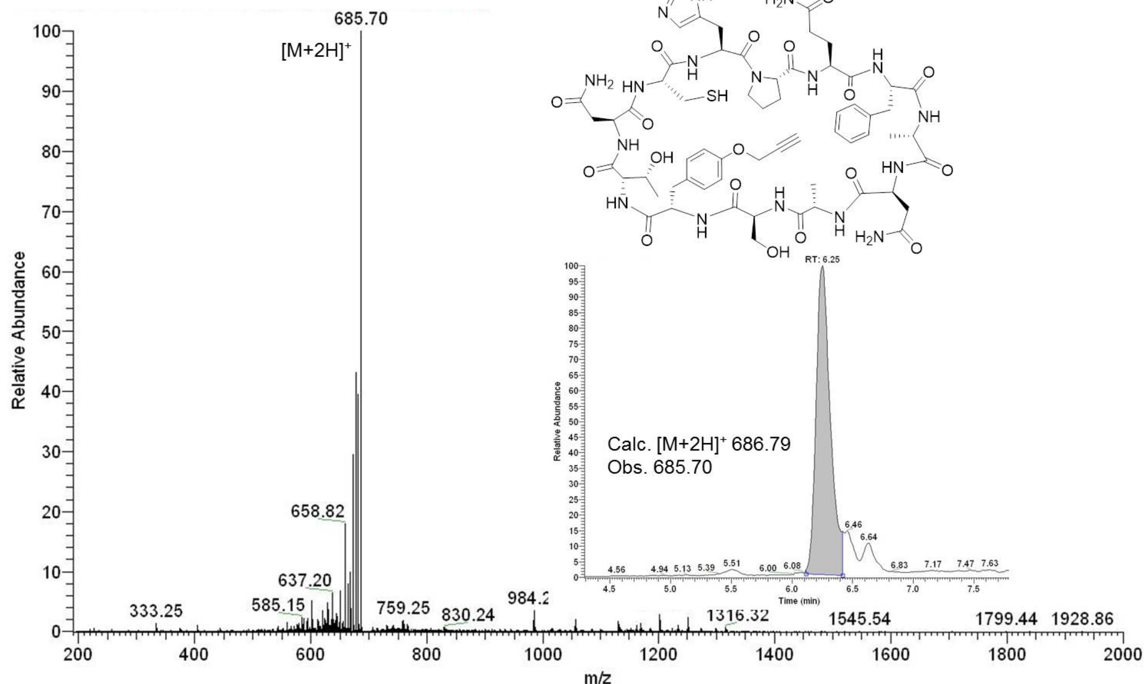
Supplementary Figure S4. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C_O2beY.



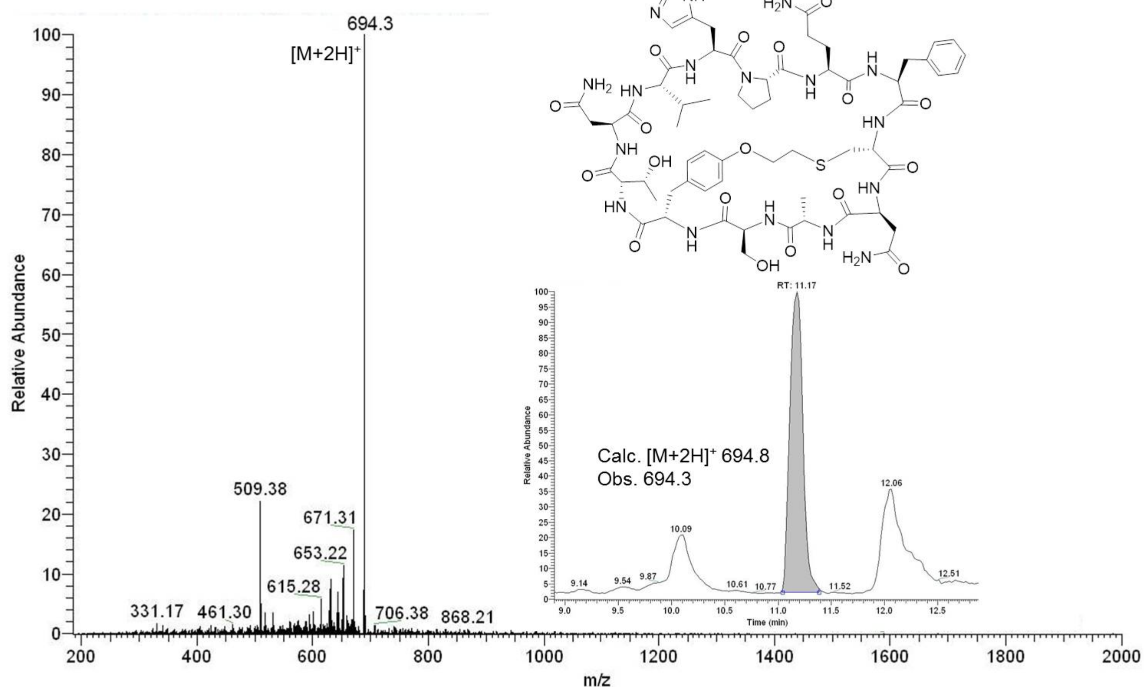
Supplementary Figure S5. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C(S1C)_O2beY.



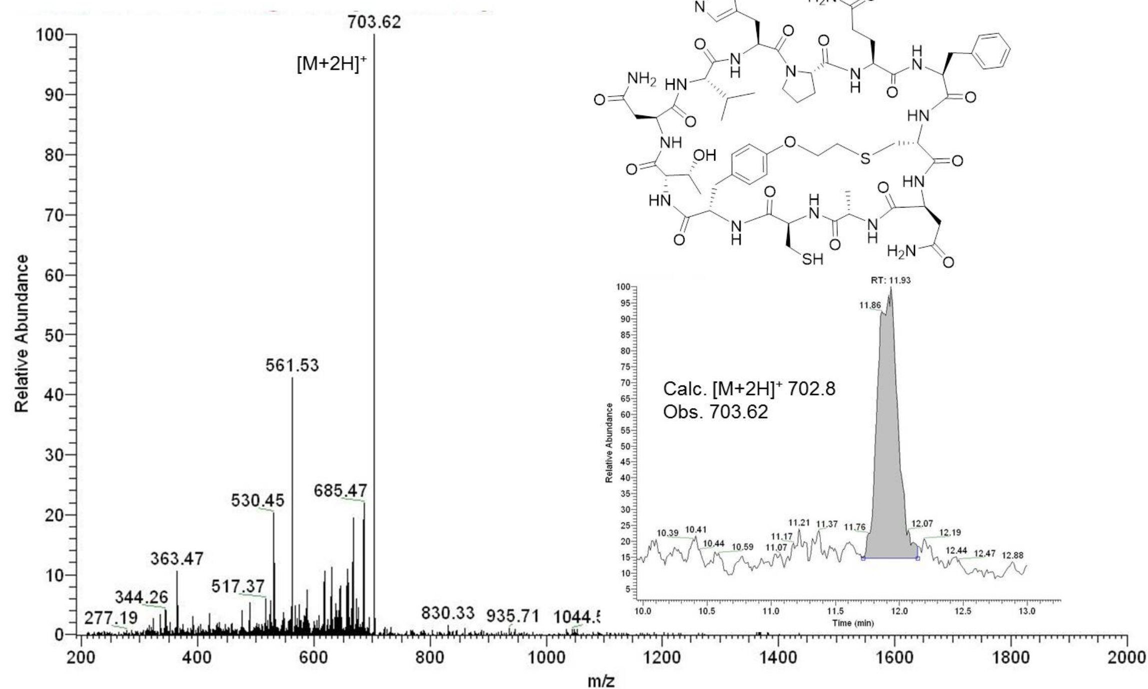
Supplementary Figure S6. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z3C_OpgY.



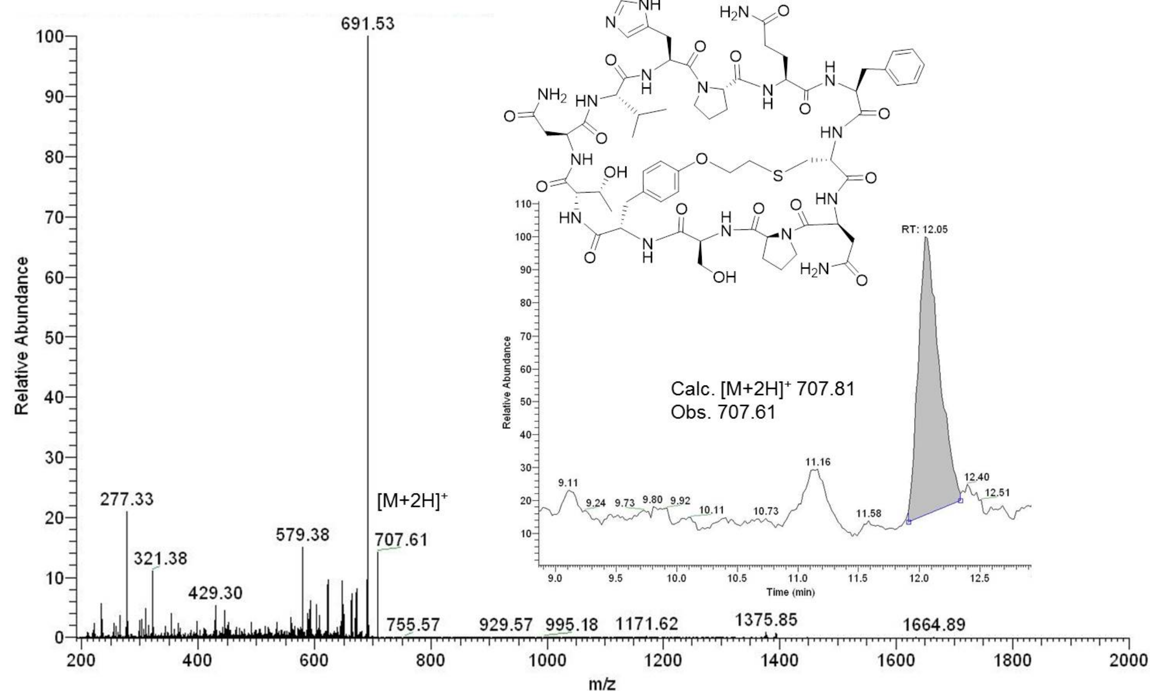
Supplementary Figure S7. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C_O2beY.



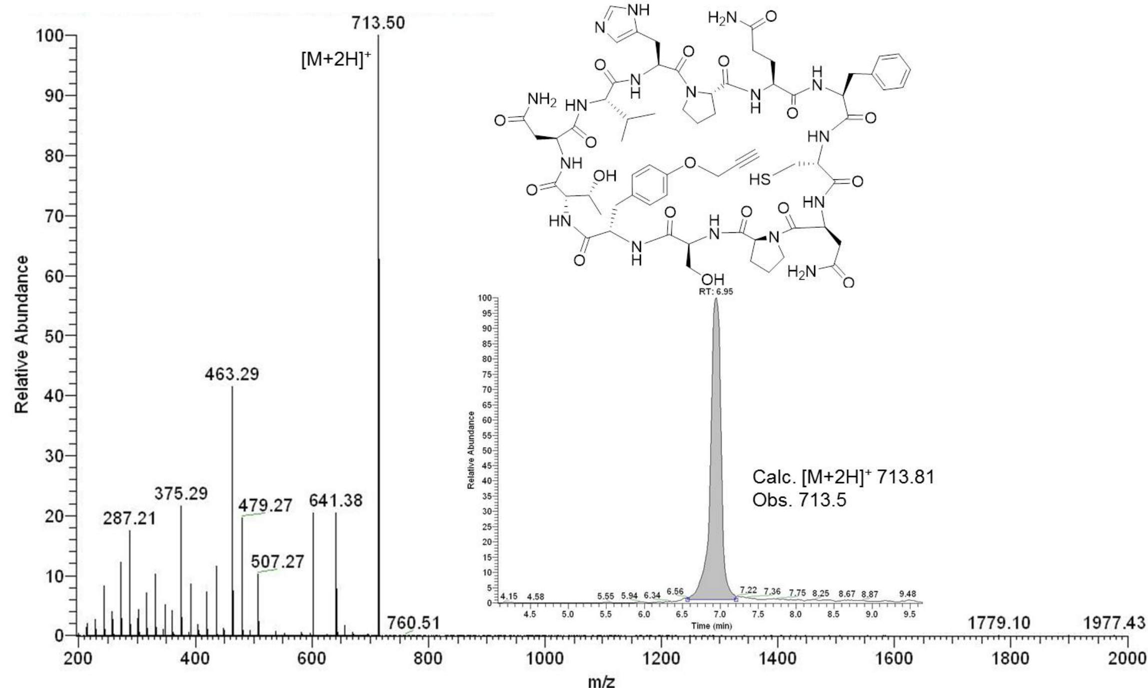
Supplementary Figure S8. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(S1C)_O2beY.



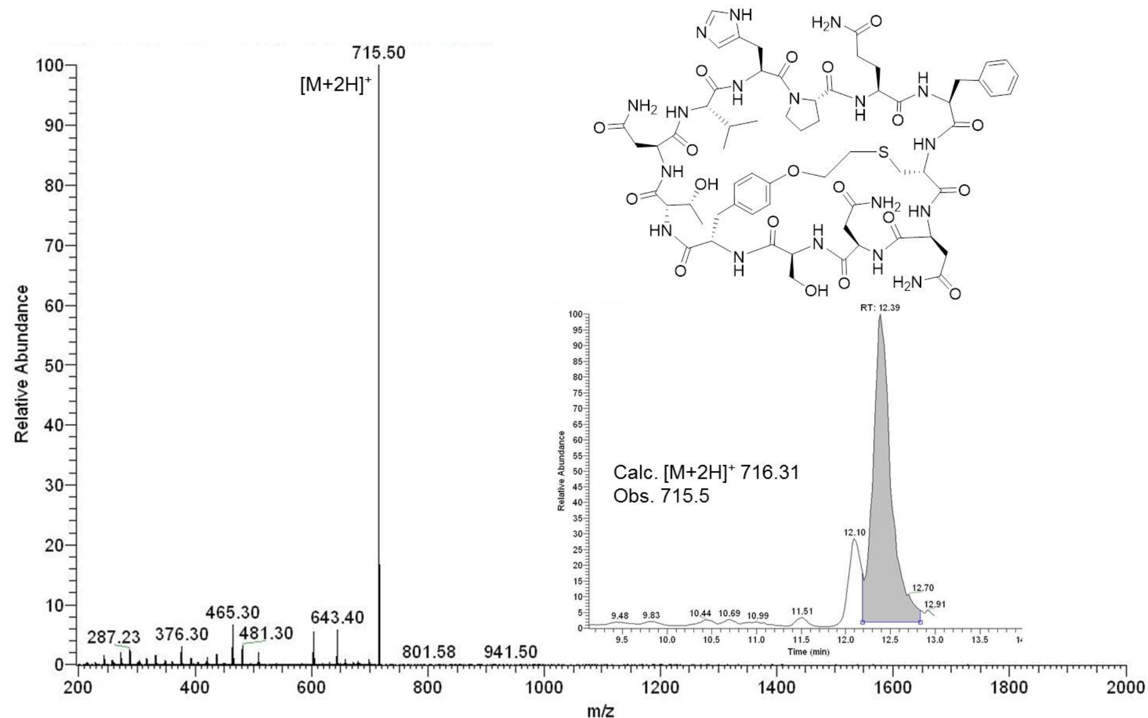
Supplementary Figure S9. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12P)_O2beY.



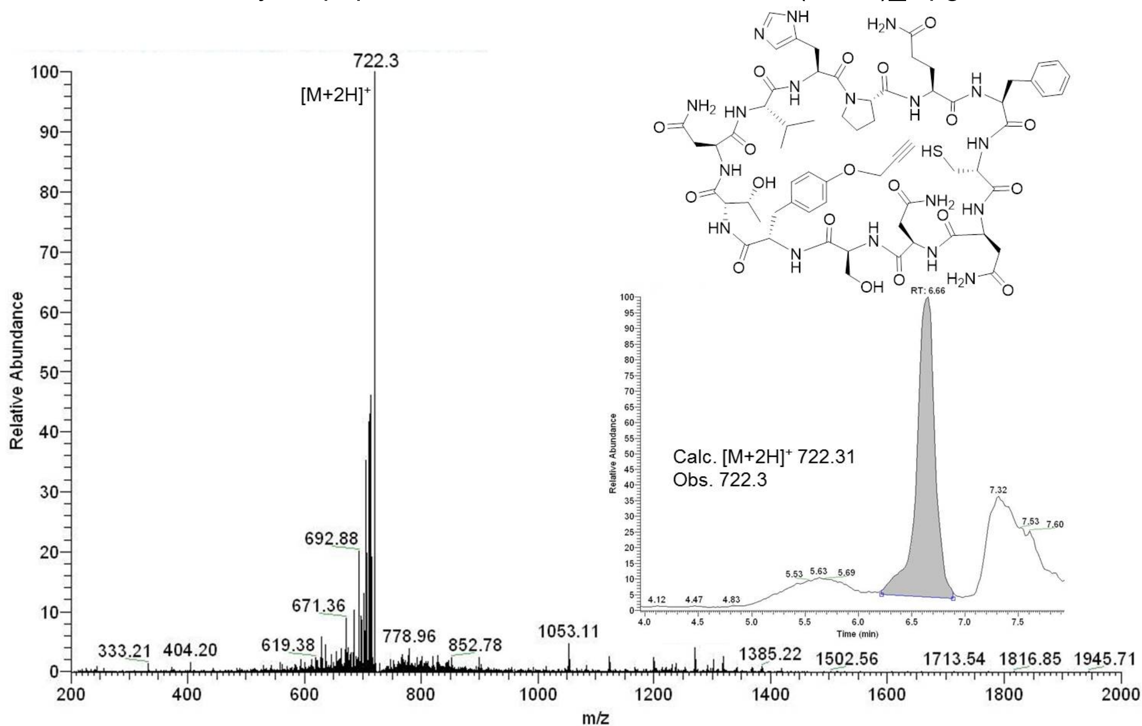
Supplementary Figure S10. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12P)_OpgY.



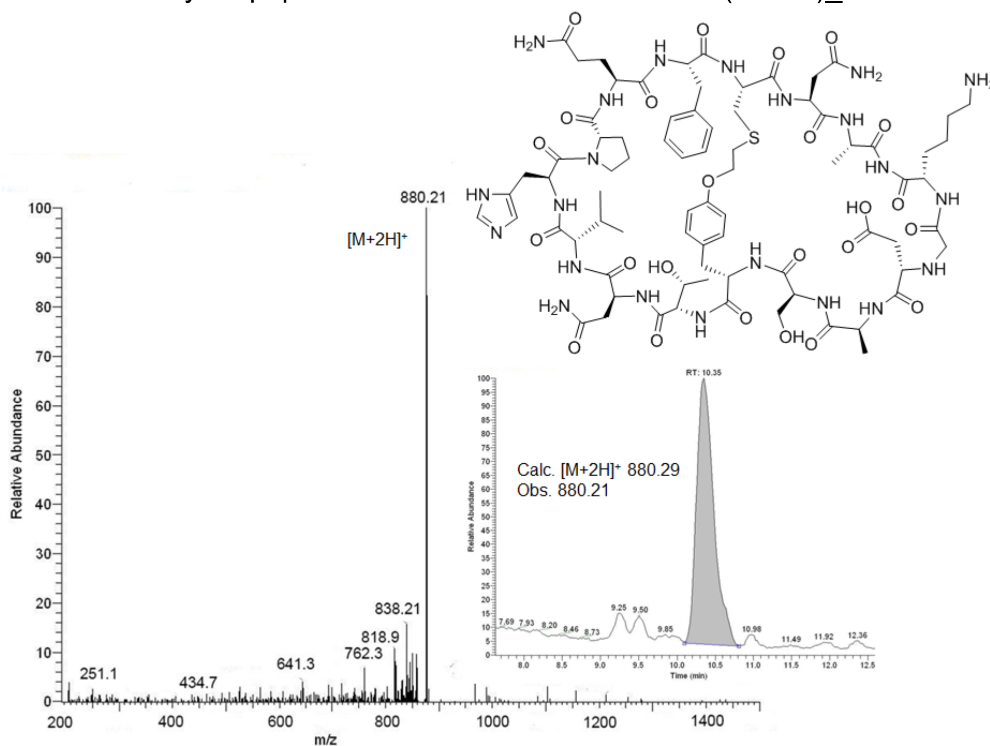
Supplementary Figure S11. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12N)_O2beY.



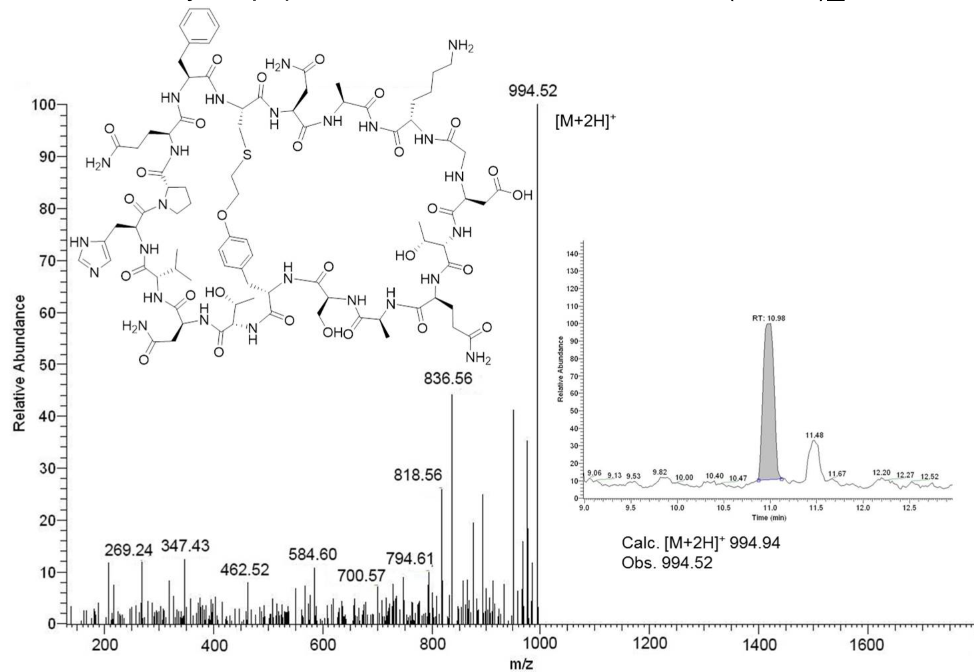
Supplementary Figure S12. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(A12N)_OpgY.



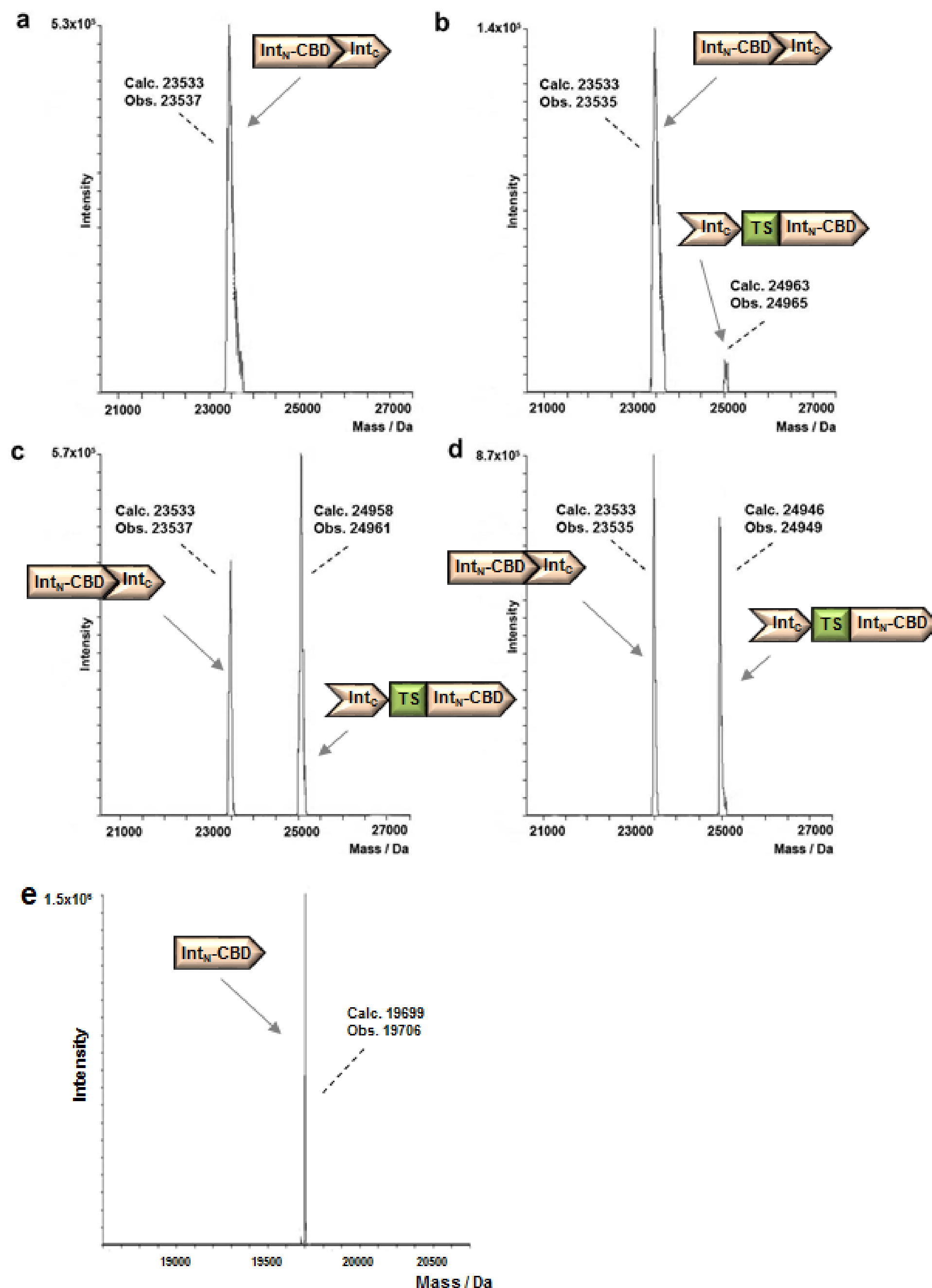
Supplementary Figure S13. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(16mer)_O2beY.



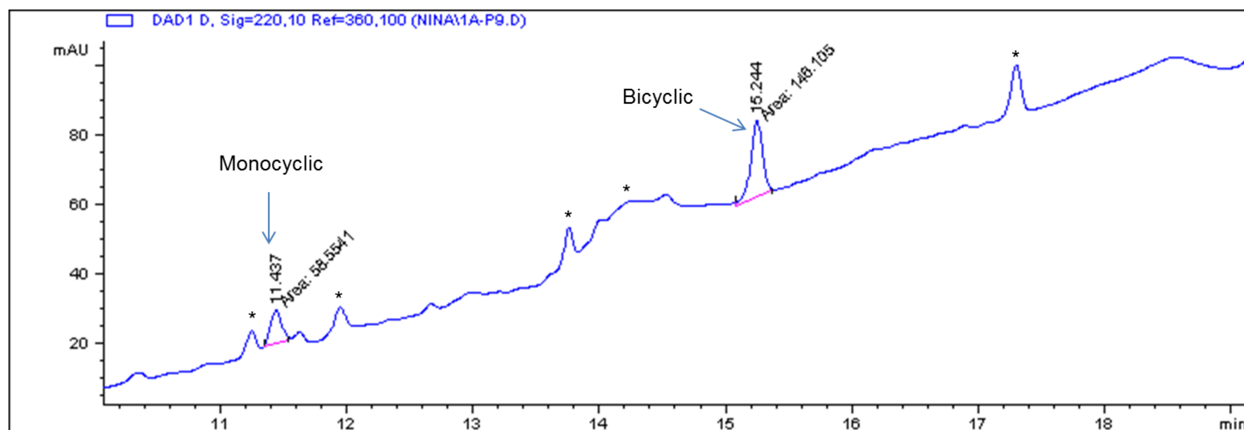
Supplementary Figure S14. Chemical structure, extracted-ion chromatogram, and MS/MS spectrum of the macrocyclic peptide obtained from construct Z8C(18mer)_O2beY.



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.

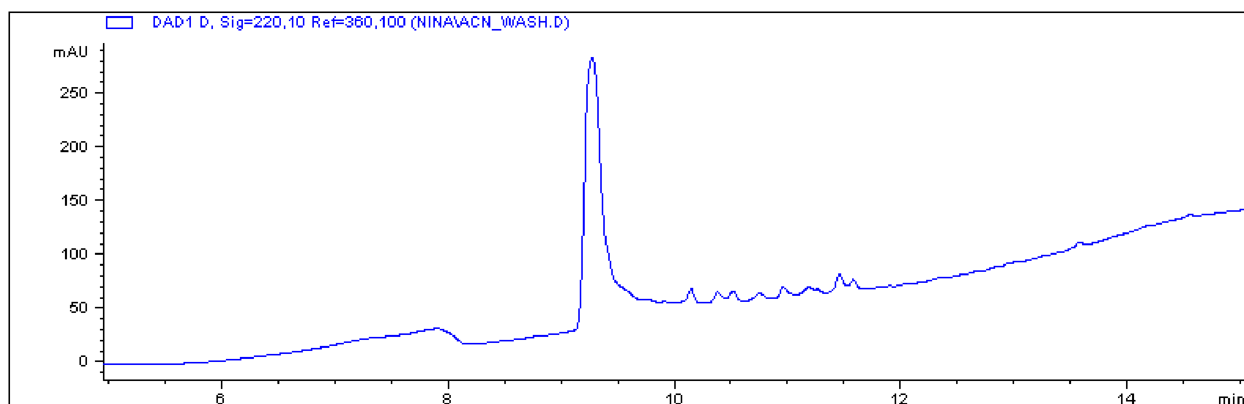


Supplementary Figure S16. HPLC analysis of the butanol extract from lysate of cells expressing the Z8C_O2beY construct. The identity of the monocyclic and bicyclic peptide (labeled) was confirmed by LCMS. Comparison of the peak areas yielded a product distribution consisting of 72% monocyclic peptide 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.

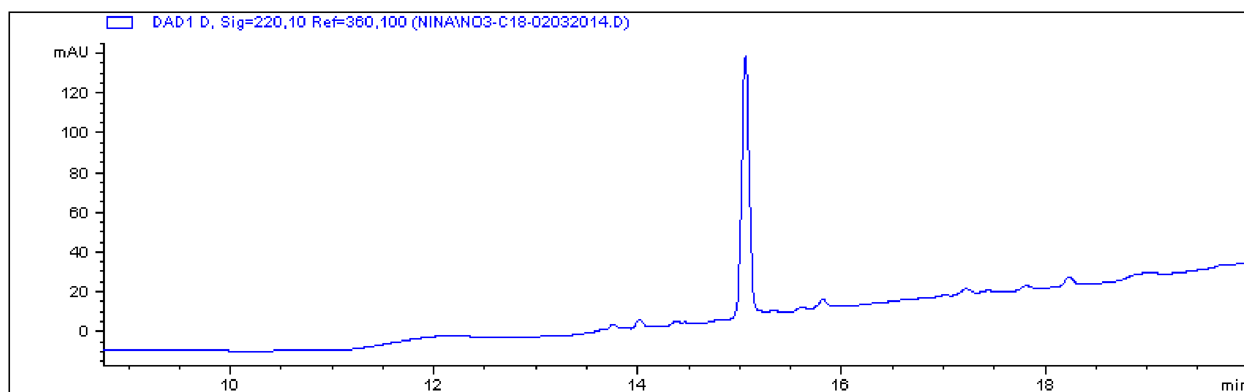


Supplementary Figure S17. RP HPLC traces corresponding to purified *cyclo*-Z3C (A) and *bicyclo*-Z8C (B).

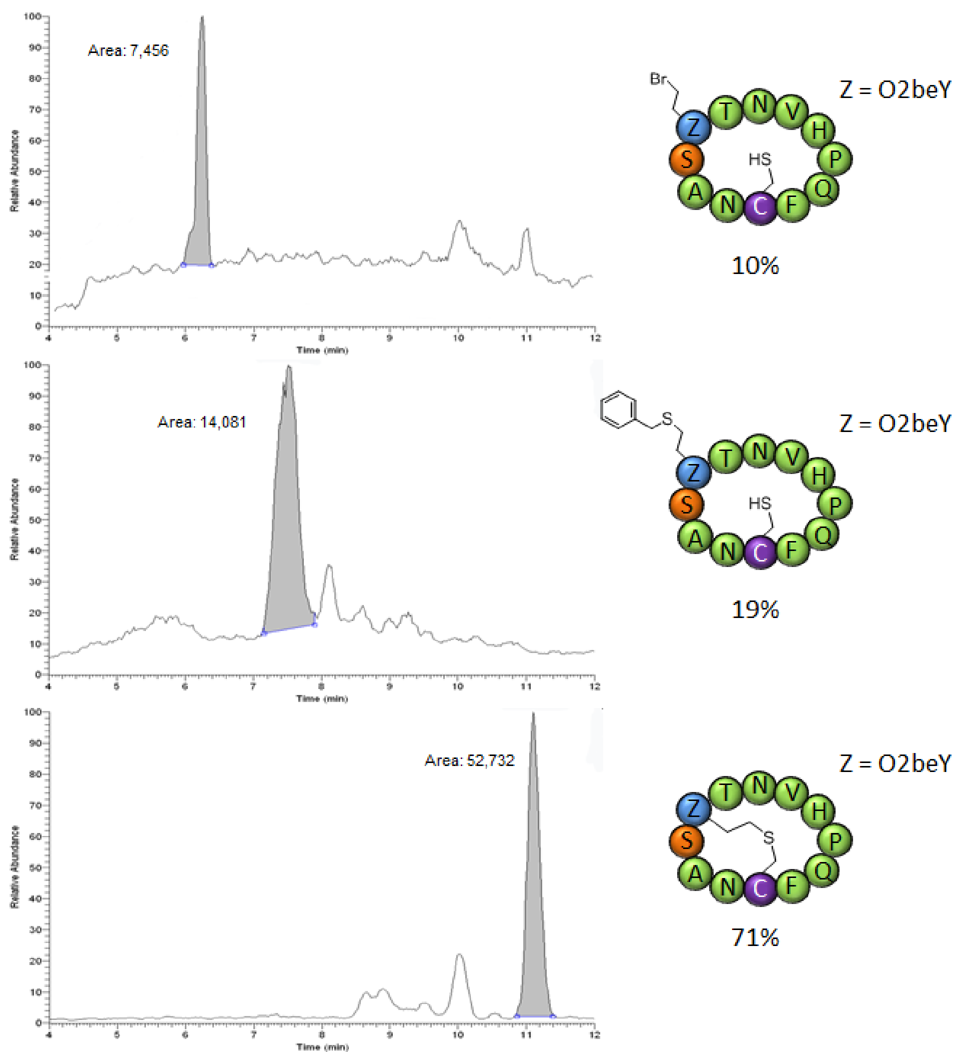
A



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.