Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information

Semisynthetic head-to-tail cyclized peptides obtained by combining protein *trans*-splicing and intramolecular expressed protein ligation

Shubhendu Palei, Henning D. Mootz*

Institute of Biochemistry

University of Muenster

Corrensstraße 36

48149 Münster

Germany

*Henning.Mootz@uni-muenster.de

Table of contents

1. Supporting Tables

Table S1. Amino acid sequences of recombinant proteins used in this study

Protein name (number)	Parent vector (construct name)	Protein Sequence
SBP-linker-M86^{C}-SIEGFGSGG-*cis Mxe* GyrA (N198A)-linker-CBD	pET16b (pSP35)	MDEKTTGWRGGHVVEGLAGELEQLRARLEH HPQGQREPGASGGGGSSSNNNNNNNNNNLG IEGRISEFSTGKRVPIKDLLGEKDFEIWAINEQ TMKLESAKVSRVFCTGKKLVYTLKTRLGRTI KATANHRFLTIDGWKRLDELSLKEHIALPRKL ESSSLQLAPEIEKLPQSDIYWDPIVSITETGVEE VFDLTVPGLRNFVANDIIVHNSIEGFGSGGCIT GDALVALPEGESVRIADIVPGARPNSDNAIDL KVLDRHGNPVLADRLFHSGEHPVYTVRTVEG LRVTGTANHPLLCLVDVAGVPTLLWKLIDEI KPGDYAVIQRSAFSVDCAGFARGKPEFAPTT YTVGVPGLVRFLEAHHRDPDAQAIADELTDG RFYYAKVASVTDAGVQPVYSLRVDTADHAFI TNGFVSHATGLTGLNSGLTTNPGVSAWQVN TAYTAGQLVTYNGKTYKCLQPHTSLAGWEP SNVPALWQLQ
SBP-linker-M86^{C} -SIEGFGSGG- *Ssp* GyrB$^{1-150}$-linker-CBD	pET16b (pSP31)	SBP-linker-M86^{C}-SIEGFGSGGCFSGDTLVALTD GRSVSFEQLVEEEKQGKQNFCYTIRHDGSIGV EKIINARKTKTNAKVIKVTLDNGESIICTPDHK FMLRDGSYKCAMDLTLDDSLMPLHRKISTTE DSGHMEAVLNYNHRIVNIEAVSETIDVYDIEV PHTHNFALASTGMKIEEGKL-CBD
SBP- linker-M86^{C} -SIEGFSGGSGGSGGGSGG- *Ssp* GyrB$^{1-150}$- linker-CBD	pET16b (pSP32)	SBP-linker-M86^{C}-SIEGFSGGSGGSGGGSGG-*Ssp* GyrB$^{1-150}$-TGMKIEEGKL-CBD
M86^{C} -SIEGFGSGG- *Ssp* GyrB$^{1-150}$- linker-CBD	pET16b (pSP33)	MD-M86^{C}-SIEGFGSGG-*Ssp* GyrB$^{1-150}$-TGMKIEE GKL-CBD
SBP- linker-M86^{C}-SIEGFGSGG-*cis* M86 (N154A, S+1A)- linker-CBD	pET16b (pSP40)	SBP-linker-M86^{C}-SIEGFGSSGCISGDSLISLAST GKRVPIKDLLGEKDFEIWAINEQTMKLESAK VSRVFCTGKKLVYTLKTRLGRTIKATANHRF LTIDGWKRLDELSLKEHIALPRKLESSSLQLAP EIEKLPQSDIYWDPIVSITETGVEEVFDLTVPG LRNFVANDIIVHAAIETGMKIEEGKL-CBD
SBP- linker-M86^{C}-SIEGFGSEFE- *cis* AceL TerL (N131A, S+1A)- linker -CBD	pET16b (pSP45)	SBP-linker-M86^{C}-SIEGFGSEFECVYGDTMVET EDGKIKIEDLYKRLAMGMFRTNTNNIKILSPN GFSNFNGIQKVERNLYQHIIFDDDTEIKTSINH PFGKDKILARDVKVGDYLNSKKVLYNELVNE

		NIFLYDPINVEKESLYITNGVVSHAAEFLTGMKIEEGKL-CBD
SBP- linker-M86^{C}-SIEGFGSGY-*cis* gp41-1 (N125A, S+1A)- linker-CBD	pET16b (pSP48)	SBP-linker-M86^{C}-SIEGFGSGYCLDLKTQVQTPQGMKEISNIQVGDLVLSNTGYNEVLNVFPKSKKKSYKITLEDGKEIICSEEHLFPTQTGEMNISGGLKEGMCLYVKEMMLKKILKIEELDERELIDIEVSGNHLFYANDILTHAASSDVDIETGMKIEEGKL-CBD
SBP- linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)- linker-CBD (protein **2**)	pET16b (pSP49)	SBP- linker-M86^{C}- SIEGFGSCLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPNASMIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASAAFNGTTGMKIEEGKL-CBD
MBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE-linker- eGFP-H$_6$ (protein **4**)	pMAL-c2x (pSP130)	MKTEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNLGIEGRISEF-M86^{C}-SIEGFGS-*cis Npu* DnaE-AFNGTVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYKGSVDRSHHHHHH
MBP-linker-M86^{C}-SIEG-*cis Npu* DnaE- linker- eGFP-H$_6$ (**SP146**)	pMAL-c2x (pShP146)	MBP-linker-M86^{C}-SIEG-*cis Npu* DnaE- linker-eGFP-H$_6$
MBP-linker-M86^{C}-SIKGS-*cis Npu* DnaE-linker-eGFP-H$_6$ (**SP147**)	pMAL-c2x (pShP147)	MBP-linker-M86^{C}-SIKGS-*cis Npu* DnaE-linker eGFP-H$_6$
MBP-linker-M86^{C}-SIERGS-*cis Npu* DnaE- linker-eGFP-H$_6$ (**SP149**)	pMAL-c2x (pShP149)	MBP-linker-M86^{C}-SIERGS-*cis Npu* DnaE-linker eGFP-H$_6$

MBP-linker-M86^{C}-SIERGS-*cis Npu* DnaE-linker-eGFP-H_6 (protein **11**)	pMAL-c2x (pShP152)	MBP-linker-M86^{C}- SIPPICFPDGR -*cis Npu* DnaE-linker-eGFP-H_6

M86^{C} = Int^{C} of M86 mutant of *Ssp* DnaB intein[1].

Table S2. Synthetic peptides used in this study

Peptide	Sequence (Int^{N} sequence underlined)*
Ao-SG-Int^{N} (**1**)	Ao-SG-CISGDSLISLA-SWKA
CESG-Int^{N} (**3**)	CESG-CISGDSLISLA-SKKWA
CTK-Int^{N} (**9**)	CTK-CISGDSLISLA-SKKWA
CTO-Int^{N} (**10**)	CTO-CISGDSLISLA-SKKWA

* abbreviations: Ao = aminooxyacetic acid, O = L-Ornithine

HPLC-MS profiles of the purified peptides used in the study are provided in Figure S12.

Table S3: Intein pairs tested for splicing and thioester cleavage

Intein pair	Premature cleavage during expression and purification	protein *trans*-splicing activity of M86 intein	Cleavage of thioester by DTT
SBP-linker-M86^{C}-SIEGFGSGG-*cis Mxe* GyrA (N198A)-linker-CBD (pSP35)	high	not active	not tested
SBP-linker-M86^{C} - SIEGFGSGG- *Ssp* GyrB$^{1-150}$ - linker-CBD (pSP31)	negligible	not active	not tested
SBP- linker-M86^{C} - SIEGFSGGSGGSGGGSGG-*Ssp* GyrB$^{1-150}$- linker-CBD (pSP32)	negligible	not active	not tested
M86^{C} -SIEGFGSGG- *Ssp* GyrB$^{1-150}$- linker-CBD (pSP33)	high	active	very slow

SBP- linker-M86^{C}-SIEGFGSGG-*cis* M86 (N154A, S+1A)- linker-CBD (pSP40)	very high	not tested	not tested
SBP- linker-M86^{C}-SIEGFGSEFE-*cis* AceL TerL (N131A, S+1A)- linker - CBD (pSP45)	none	active (75%-24h)	slow (35%-24h)
SBP- linker-M86^{C}-SIEGFGSGY-*cis* gp41-1 (N125A, S+1A)- linker-CBD (pSP48)	negligible	active (60%-24h)	slow (<30%-72h)
SBP- linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)- linker-CBD (protein **2**)	negligible	active (85%-24h)	very good (70%-24h, 95%-72h)

Table S4: Stabilization of *Npu* DnaE intein constructs

***cis-Npu* DnaE constructs**	**Split *Npu* DnaE constructs**
SBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-H_6	SBP-linker-M86^{C}-SIEGFGS-*Npu* DnaE Int^{N}
SBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-PML-SBP	SBP-linker-*Npu* DnaE Int^{C} (N138A, S+1A)-CBD
M86^{C}-linker-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-PML-SBP	DnaE Int^{C} (N138A, S+1A)-Trx-H_6
M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-Trx-H_6	MBP-linker-M86^{C}- SIRKGS- *Npu* DnaE Int^{N}
SBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-eGFP-H_6	*Npu* DnaE Int^{C} (N138A, S+1A)-eGFP-H_6
MBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A) -MBP-H_6	
MBP-linker-M86^{C}-SIEGFGS-*cis Npu* DnaE (N138A, S+1A)-eGFP-H_6	

Table S5: Head-to-tail cyclic peptides with respective yields

Cyclic peptide	Synthetic part (peptide)	Recombinant part (protein)	% Yield (HPLC)*
5 (11-mer)	**3**	**4**	50
6 (8-mer)	**3**	**SP146**	85
7 (9-mer)	**3**	**SP147**	86
8 (10-mer)	**3**	**SP149**	80

* The yield was calculated based on all the products originating only from the precursor dual intein as the Int^N peptide was used in excess (1.5 fold) (Figure 3 and S8). Based on the HPLC chromatogram (210 nm) the yield of corresponding cyclic peptide (%) = ((sum of area under the peak of both head-to-tail and unarranged thioester form of cyclic peptide) / (sum of area under the peak of both head-to-tail and unarranged thioester form of cyclic peptide + area under C-cleavage Ex^C (SIEGFGS) peak + area under Ex^C-MESH thioester + area under cyclic Ex^C (SIEGFGS) peak)) × fraction of dual intein precursor conversion in PTS × 100. The unrelated products, which are originating from Int^N peptide (e.g. Int^N, by products SKKWA, SLA-SKKWA) were excluded from the calculation of the final yield.

2. Supporting Figures

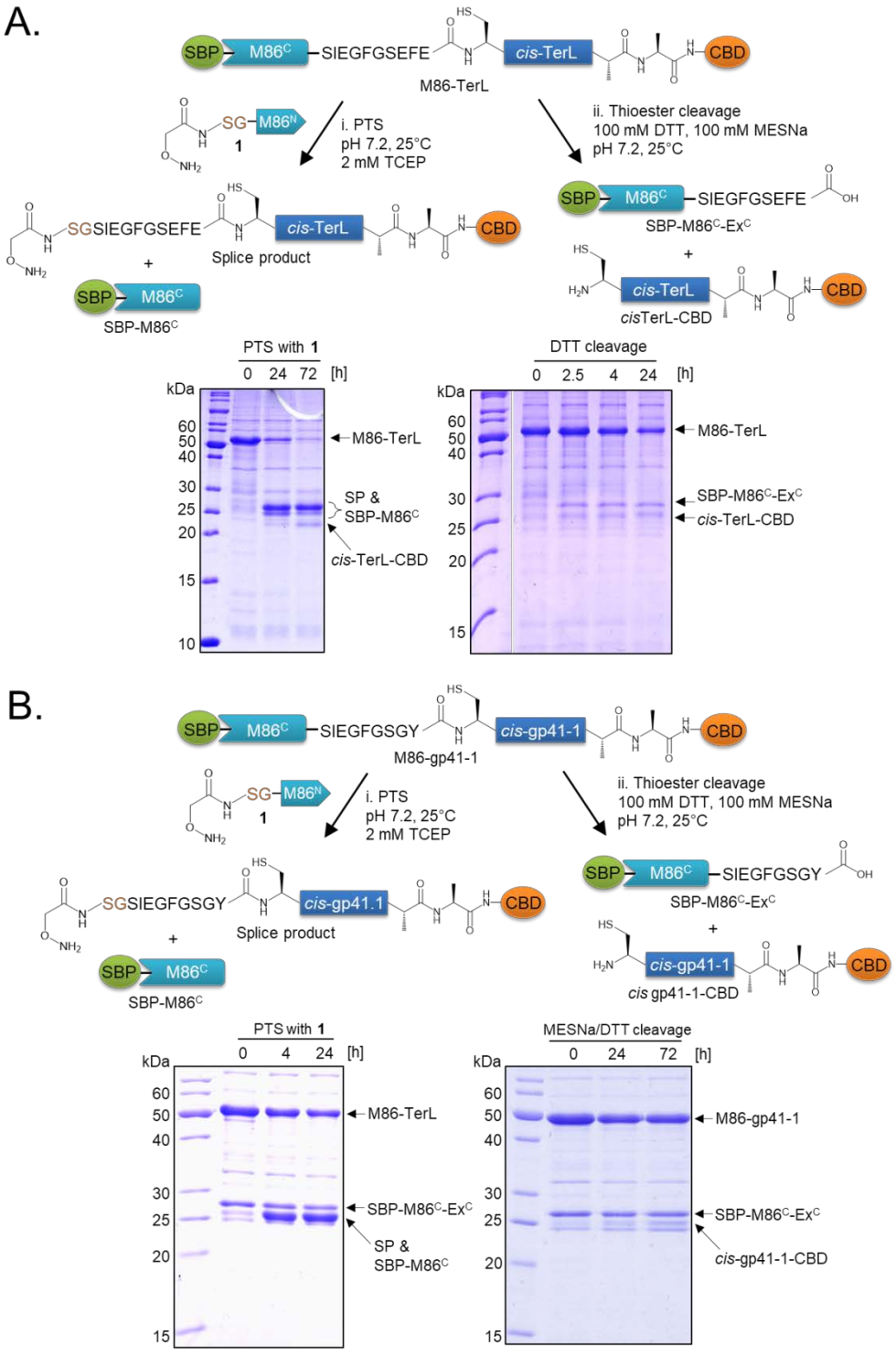

Figure S1. PTS and thioester cleavage activity of M86-AceL TerL and M86-gp41-1 dual intein. PTS reaction with peptide **1** and thioester cleavage activity in presence of 100 mM DTT and 100 mM MESNa at RT of the dual intein constructs A) SBP-linker-M86^{C}-SIEGFGSEFE-*cis* AceL TerL- linker-CBD, B) SBP-linker-M86^{C}-SIEGFGSGY-*cis* gp41-1-linker-CBD.

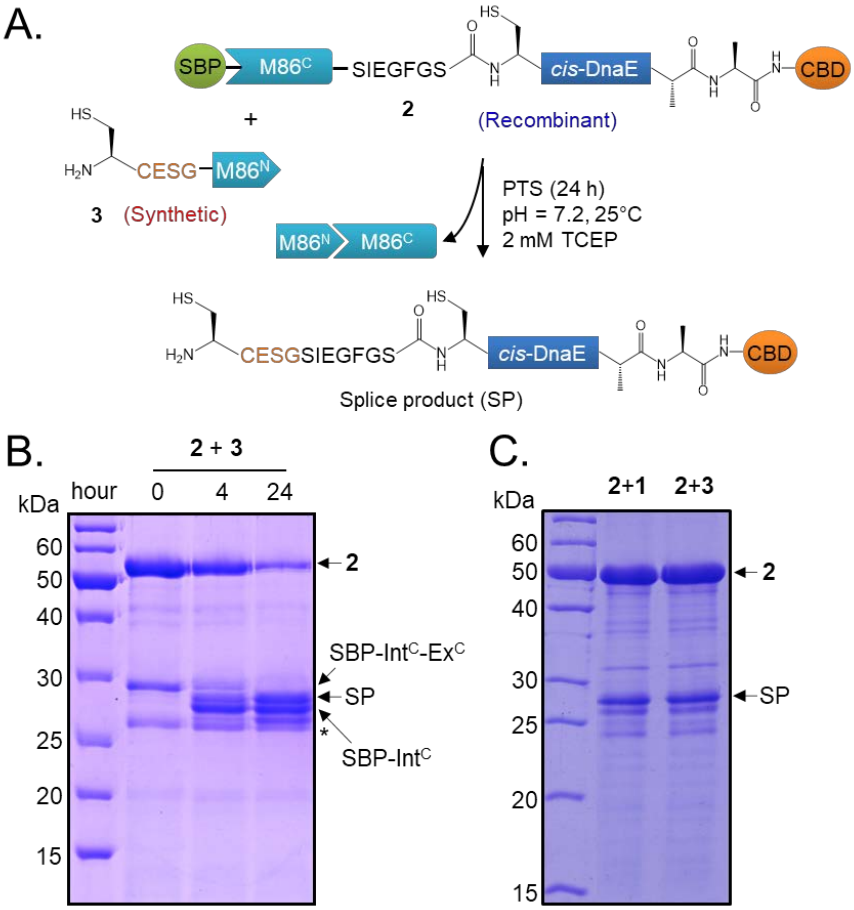

Figure S2. Protein *trans*-splicing reaction of protein 2 with peptide 3. A) Schematic representation of PTS of **2** with **3** resulting linear splice product (SP). B) Coomassie stained SDS-PAGE analysis of the splice reaction described in 'A'. * = unknown impurity. C) The precipitate that appeared after 24 h splice reaction of protein **2** with peptide **1** and **3**, was centrifuged and the pellet was run on SDS-PAGE to identify the precipitated proteins. Coomassie stained SDS-PAGE analysis revealed the precipitated proteins were precursor **2** and corresponding splice products (SP).

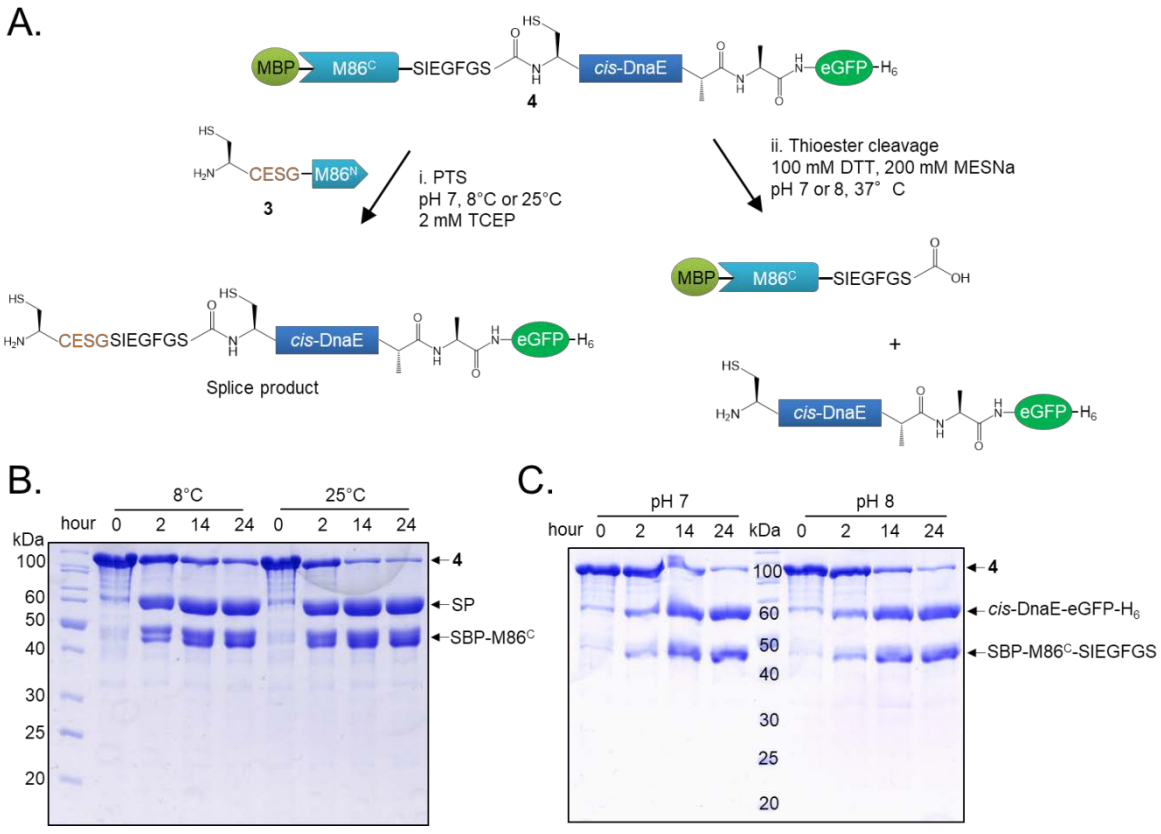

Figure S3. Protein *trans*-splicing and thioester cleavage of protein 4. A) Schematic representation of PTS of protein **4** with peptide **3** at pH 7 and 8°C or 25°C; and thioester cleavage of protein **4** with 100 mM DTT in 200 mM MESNa at 37°C. B) Coomassie stained SDS-PAGE analysis of PTS reaction described in 'A'. Yield after 24h at 8°C = 73%, at 25°C = 90%. C) Coomasssie stained SDS-PAGE analysis of thioester cleavage reaction described in 'A'. Yield after 24h at pH 7 = 86%, pH 8 = 86.3%.

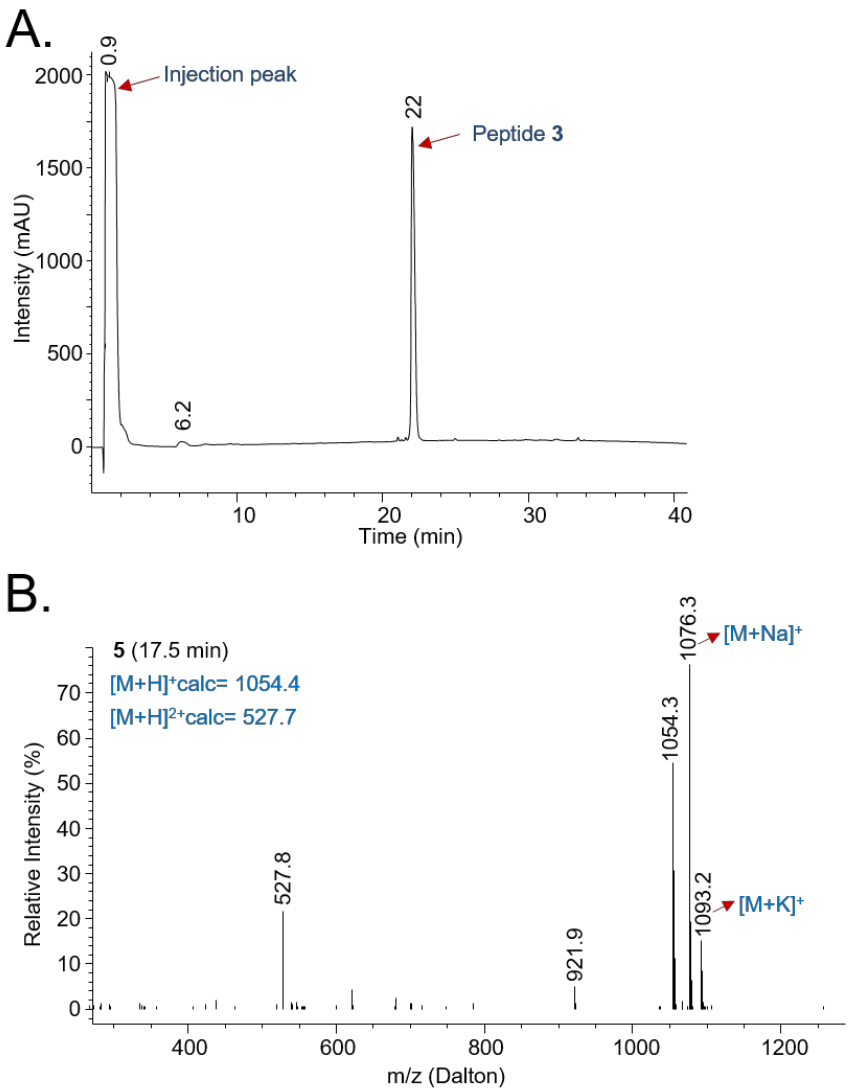

Figure S4. Characterization of cyclization reaction of cyclic peptide 5. A) HPLC chromatogram (210 nm) of a control splice reaction (after 0 hour) of protein **4** and peptide **3**. B) ESI-MS chromatogram of the minor thioester isoform of cyclic peptide **5** at 17.5 min in Figure 3B. $[M+H]^+$calc = 1054.4 Da, obs = 1054.3 Da; $[M+H]^{2+}$calc = 527.7 Da, obs = 527.8 Da. Da = Dalton. HPLC chromatogram (210 nm) of the cyclization reaction described in Figure 3A with reaction conditions

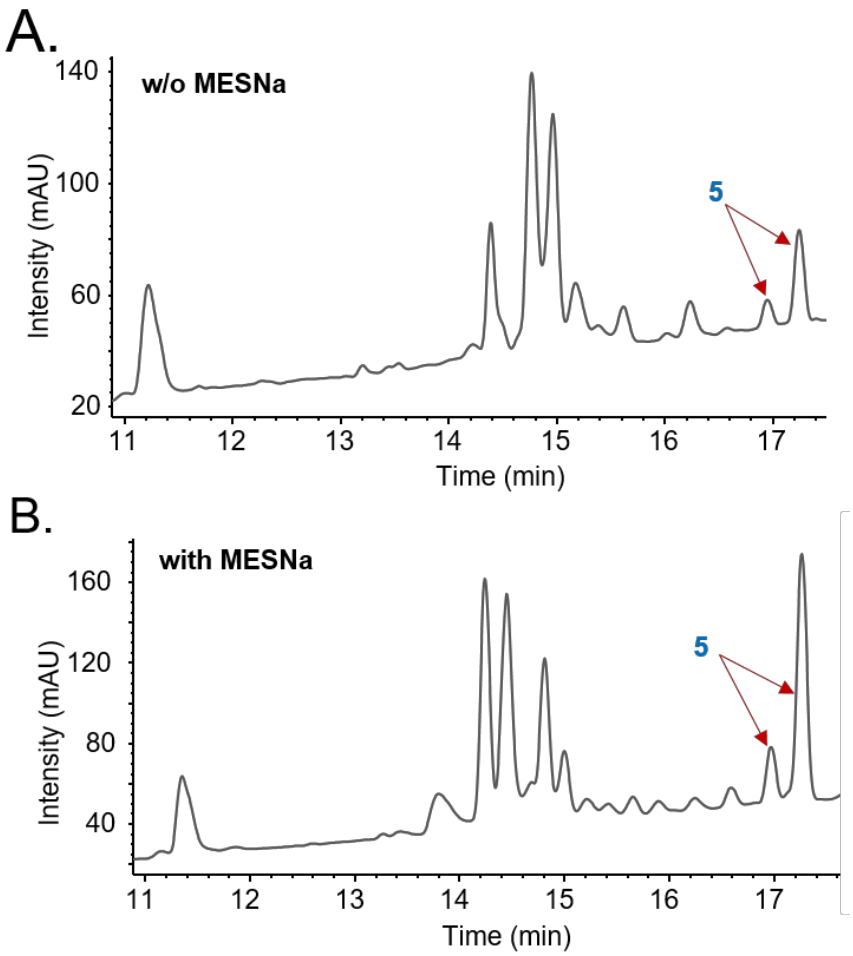

Figure S5. Role of MESNa in cyclization. HPLC chromatogram (210 nm) of splice reaction of protein **4** and peptide **3** (8°C, 24 hours, pH 7.4 Tris buffer) followed by cyclization (37°C, 24 hours, pH 7.4 Tris buffer) A) without MESNa, yield of cyclic peptide **5** = 16%. B) with 200 mM MESNa, yield of cyclic peptide **5** = 40%. All other peaks in the HPLC chromatogram are characterized in Fig 3B.

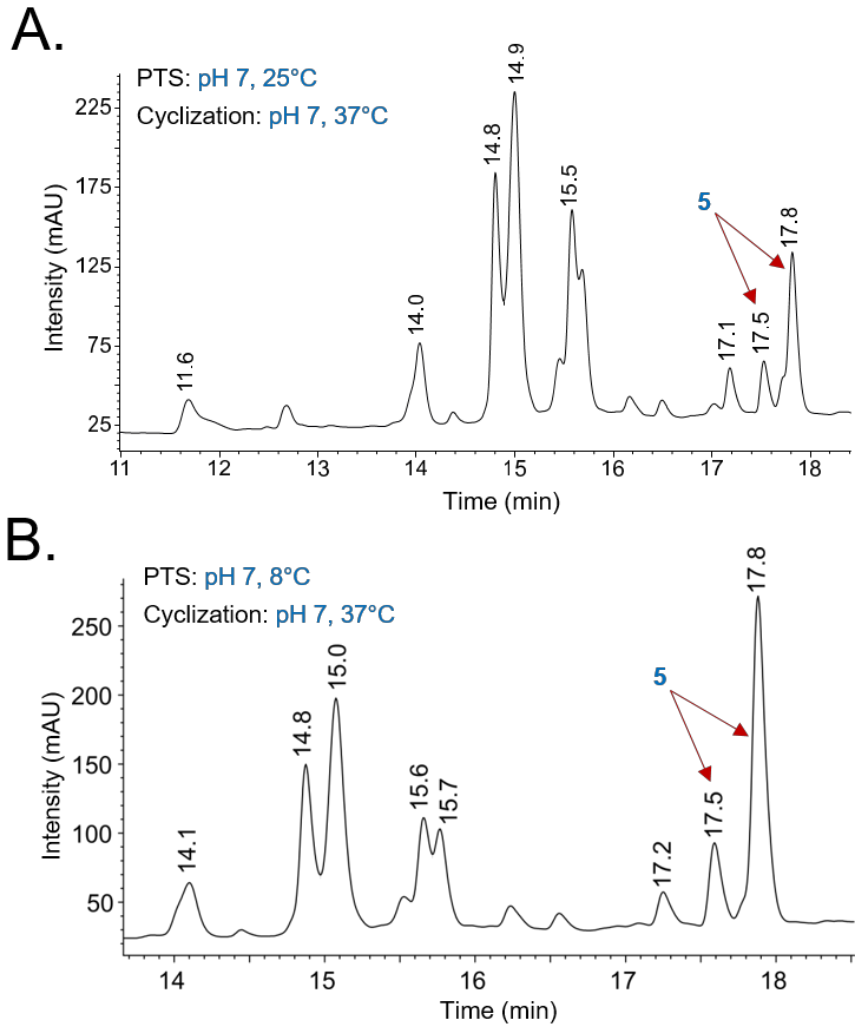

Figure S6. Optimization of pH and temperature of PTS and cyclization. A) PTS: pH 7, 25°C, Cyclization: pH 7, 37°C. Yield of cyclic peptide **5** = 21%. B) PTS: pH 7, 8°C, Cyclization: pH 7, 37°C. Yield of cyclic peptide **5** = 45%

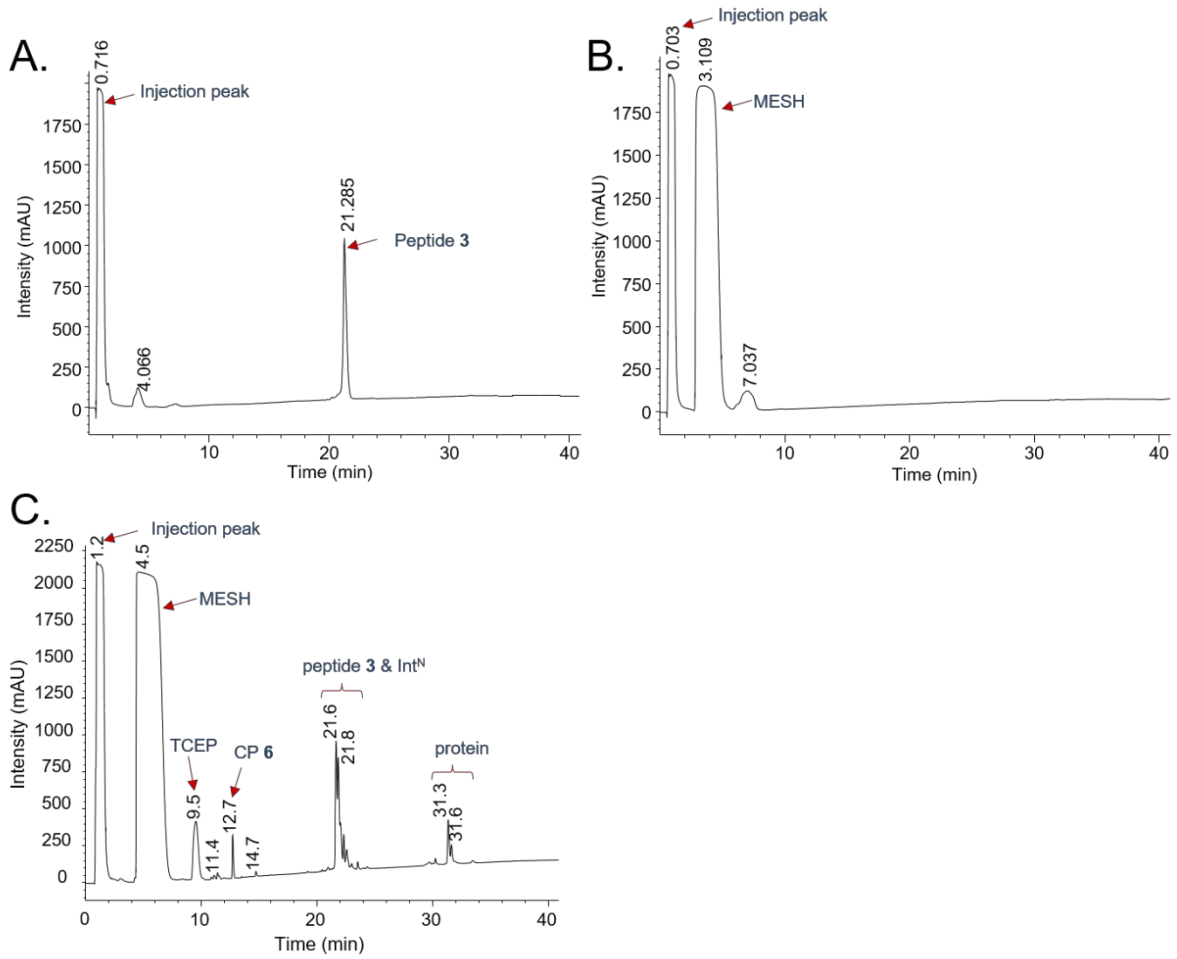

Figure S7. HPLC analysis of cyclization reaction of cyclic peptide 6. A) HPLC chromatogram (210 nm) of control splice reaction (after 0 hour) of protein **SP146** and peptide **3**. B) HPLC chromatogram (210 nm) of control cleavage reaction of protein **SP146** incubated at 37°C for 24 hours at pH 8 in presence of 100 mM TCEP and 100 mM MESNa. C) Full HPLC chromatogram of one pot PTS and cyclization of **SP146** and peptide **3** for the formation of cyclic peptide **6**. MESH = 2-Mercaptoethansulfonic acid

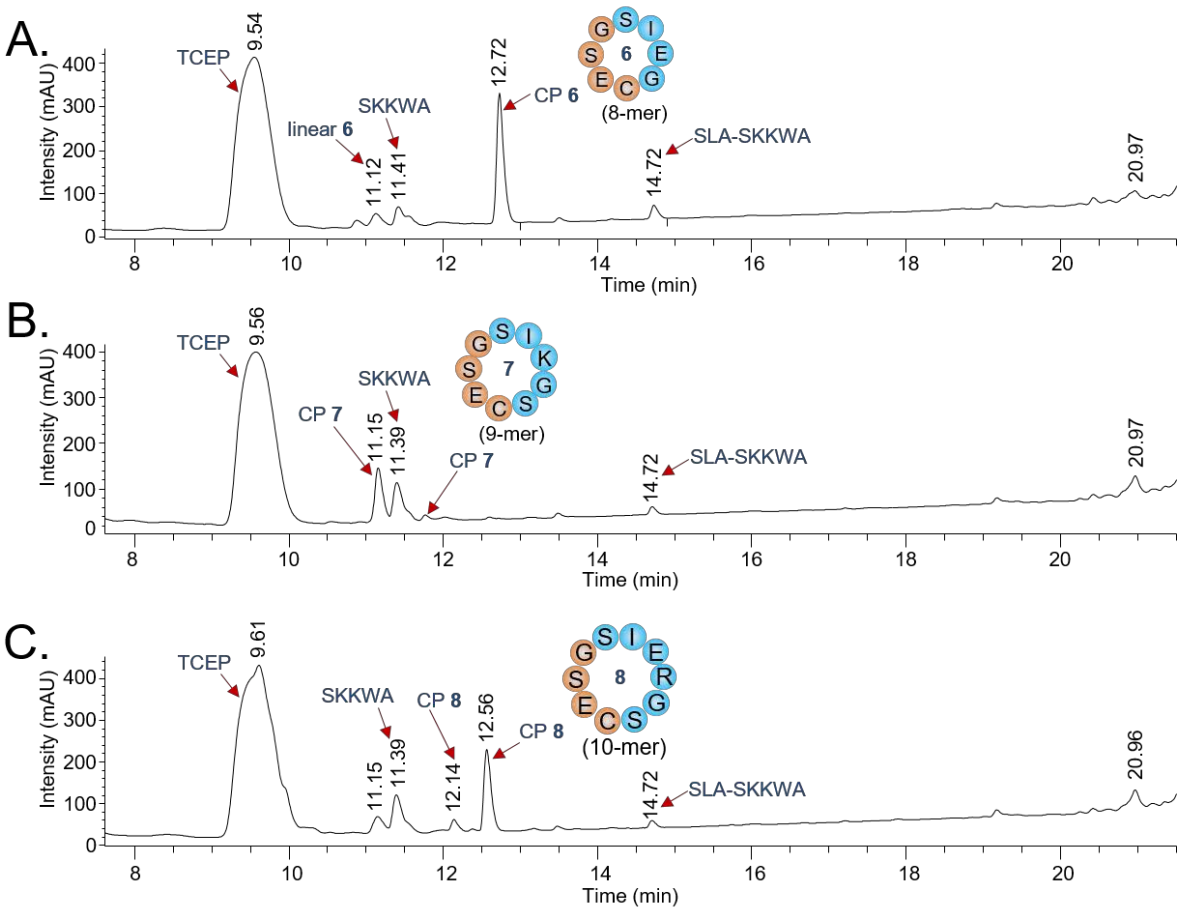

Figure S8. Annotation of peaks in the HPLC chromatogram (210 nm) for characterization of cyclic peptide (CP) A) **6**, B) **7** and C) **8**.

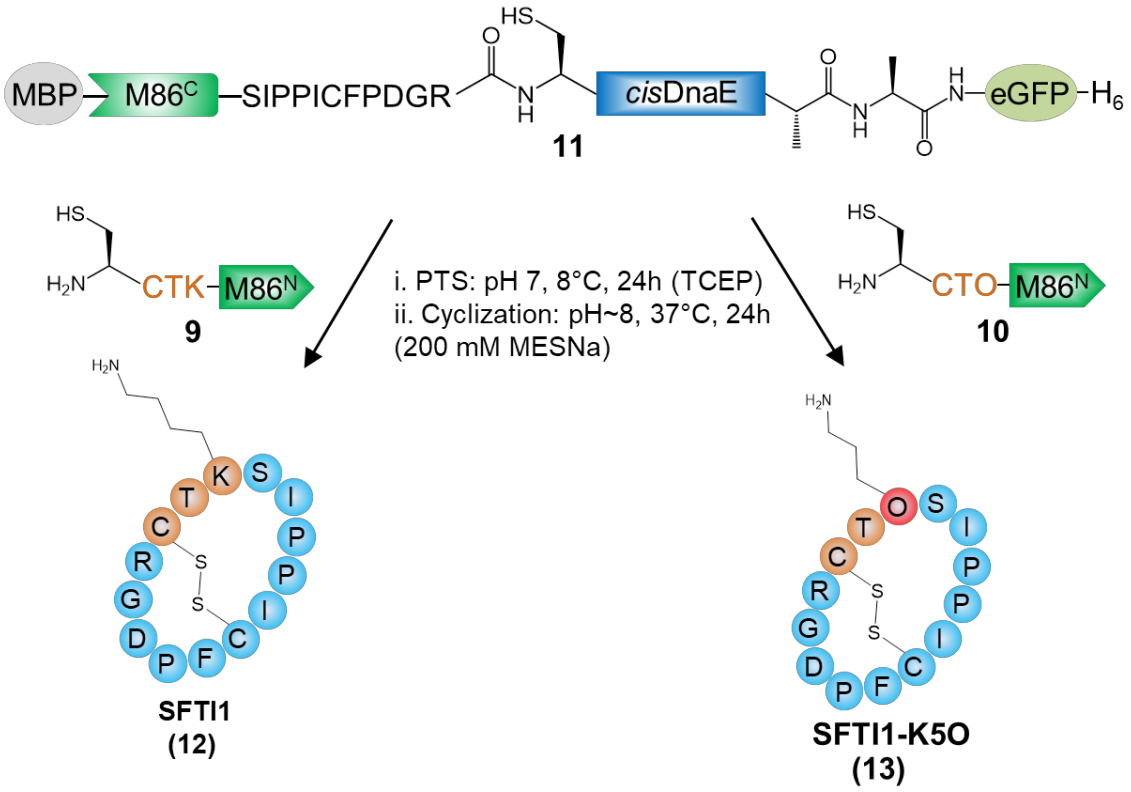

Figure S9. Schematic representation of one pot cyclization of semisynthetic SFTI1 and SFTI-K5O analogues: from respective synthetic peptides (**9, 10**) and recombinant protein precursors (**11**).

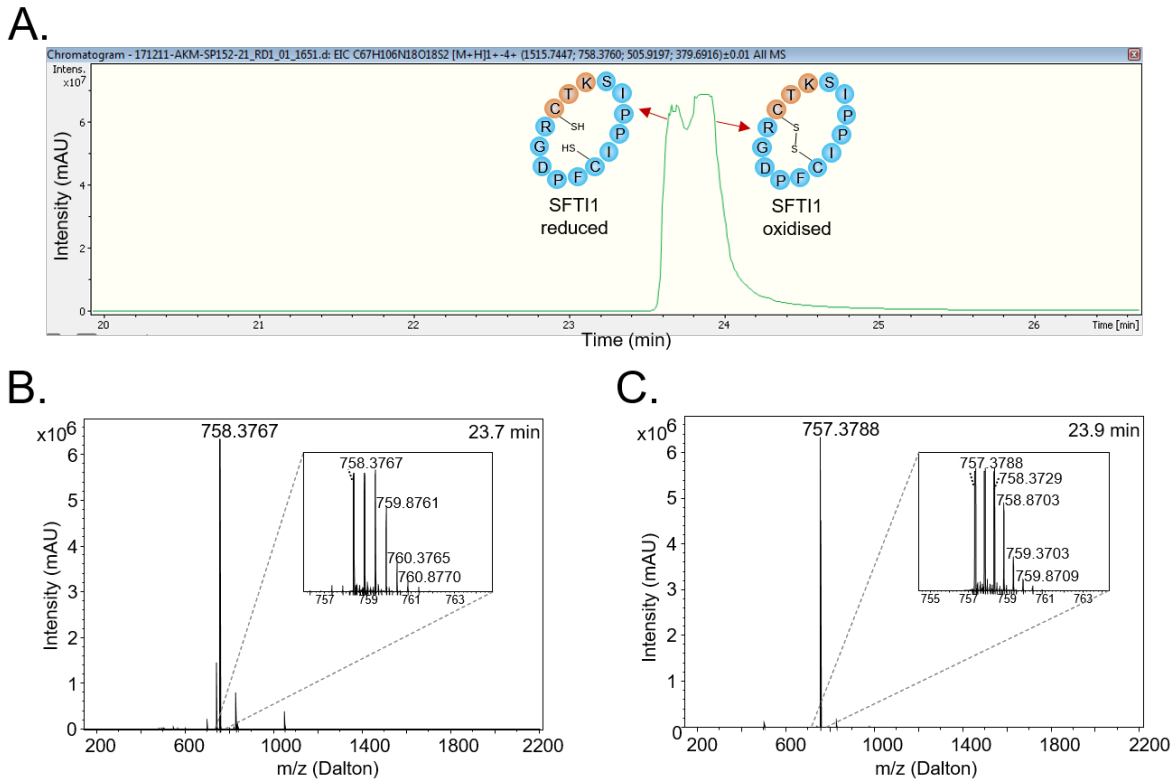

Figure S10. nanoLC-HRMS (ESI-TOF) analysis of semisynthetic SFTI1 peptide. A) Extracted-Ion-Chromatogram (EIC) of SFTI1 revealed formation of both reduced and oxidised forms of SFTI1 with retention time 23.7 min and 23.9 min respectively. B) HRMS spectrum of reduced SFTI1. $[M+H]^{2+}$ calc = 758.3760 Da, obs = 758.3767. C) HRMS spectrum of oxidised SFTI1. $[M+H]^{2+}$ calc = 757.3681 Da, obs = 757.3788 Da. Da = Dalton

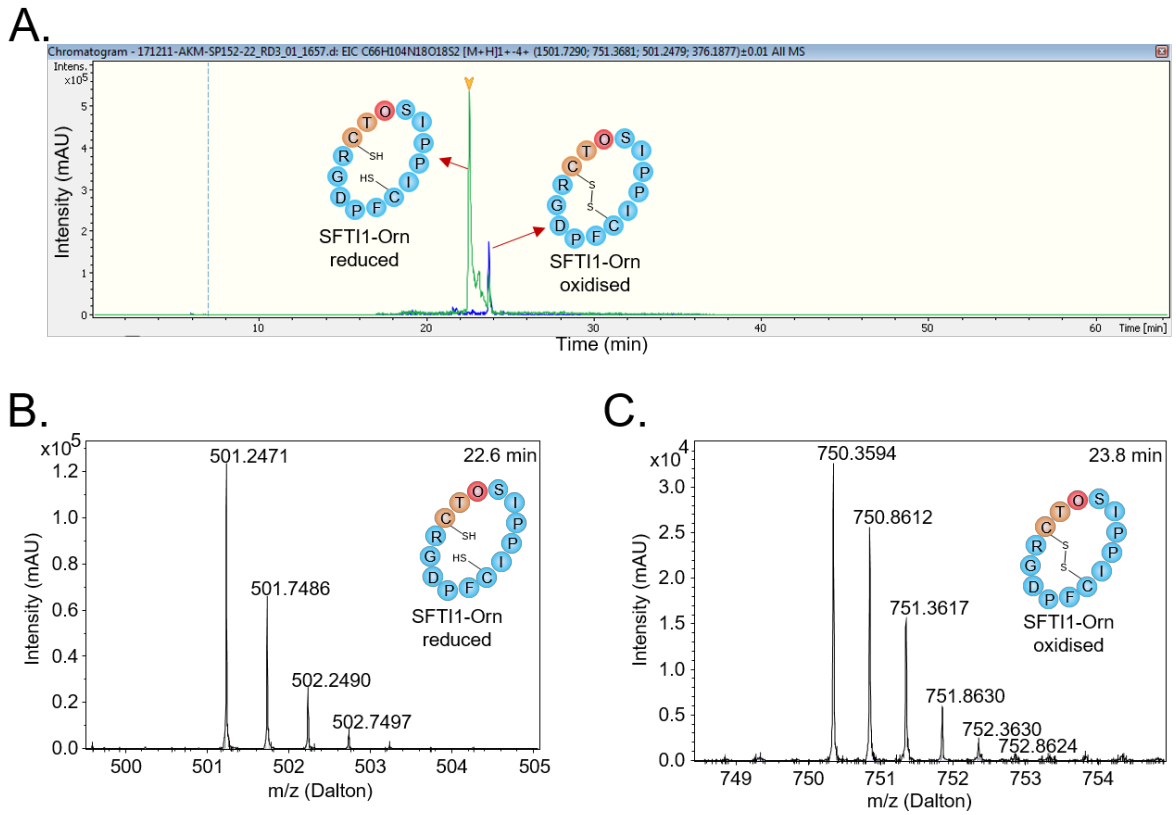

Figure S11. nanoLC-HRMS (ESI-TOF) analysis of semisynthetic SFTI1-K5O analog. A) Extracted-Ion-Chromatogram (EIC) of SFTI1-K5O analog revealed formation of both reduced and oxidised forms of SFTI1 with retention time 22.6 min and 23.8 min respectively. B) HRMS spectrum of reduced SFTI1-K5O analog. $[M+H]^{3+}$ calc = 501.2479 Da, obs = 501.2471. C) HRMS spectrum of oxidised SFTI1-K5O analogue. $[M+H]^{2+}$ calc = 750.3603 Da, obs = 750.3594 Da. Da = Dalton

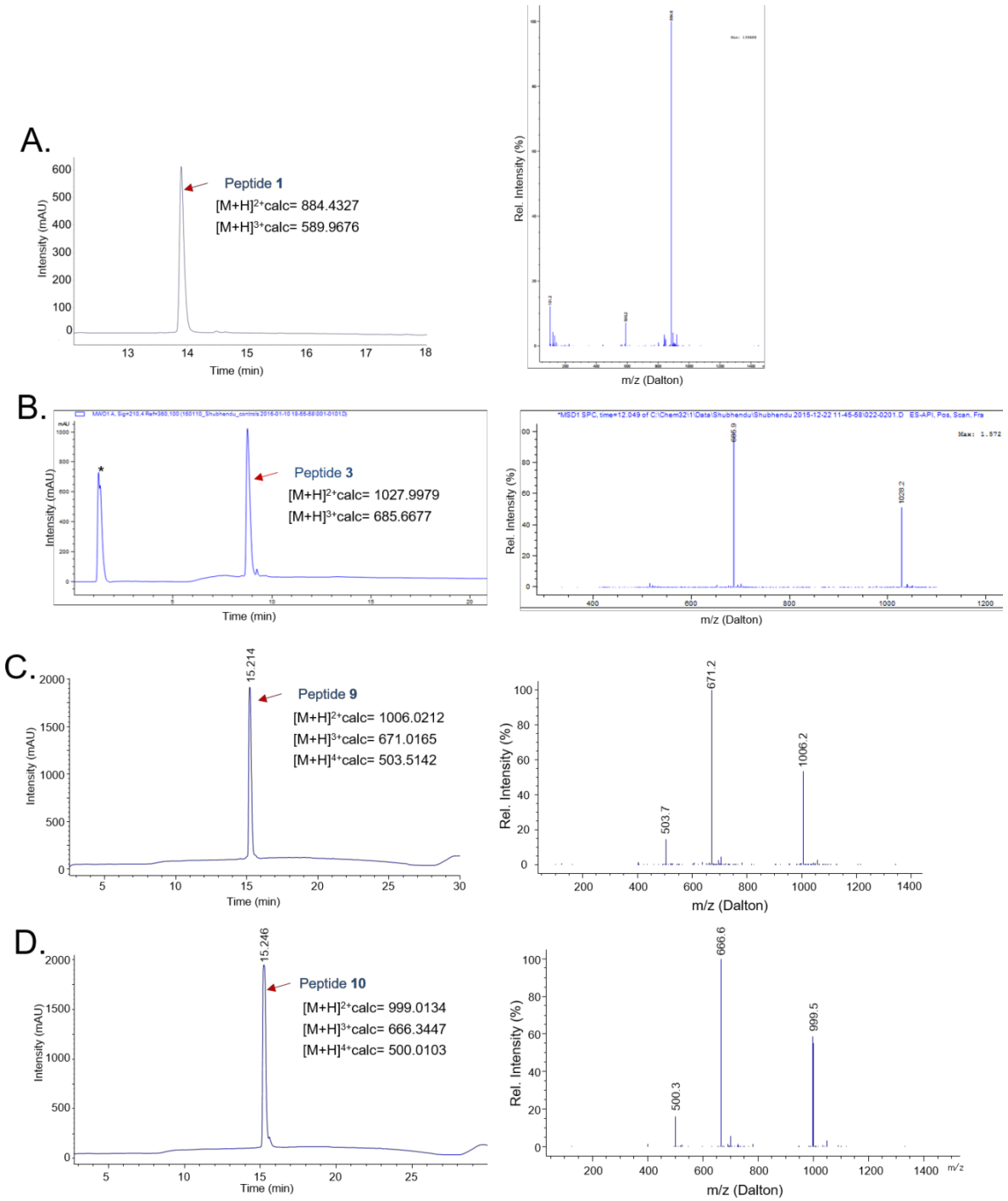

Figure S12. LC ESI-MS analysis of synthetic M86^{N} peptides used. A) peptide **01**, B) peptide **03**, C) peptide **09**, D) peptide **10**. * = non-peptidic injection peak.

3. Materials and Methods

3.1 General Remarks

Unless otherwise specified, standard protocols were used. Oligonucleotides were purchased from Biolegio (Netherlands) and all plasmids were verified by DNA sequencing by Seqlab (Göttingen). All chemically synthesized peptides were analyzed by ESI-MS and purity > 95 % was verified by analytical HPLC or LC-MS. DMF, amino acids and coupling reagents were purchased from Novabiochem (Hessen, Germany), Iris Biotech GmbH (Marktredwitz, Germany), Sigma-Aldrich (Germany), Acros Organics (Germany). DCM and HPLC grade acetonitrile were purchased from VWR (Darmstadt, Germany) or Carl Roth (Karlsruhe, Germany), Isopropyl-D-1-thiogalactopyranoside (IPTG) from Fermentas (Thermo Scientific). Agarose beads immobilized TPCK trypsin was purchased from Thermo Scientific (20230). Phenylmethyl sulfonylfluoride (PMSF) was purchased from Applichem (Germany). Strep-tactin sepharose beads and D-desthiobiotin were purchased from IBA life science (Goettingen, Germany). Ni-NTA beads were purchased from Cube Biotech (Manheim, Germany). Amylose and Chitin beads were purchased from NEB (Frankfurt, Germany). MESNa was purchased from Sigma Aldrich (Hamburg, Germany) and TCEP from Fluka (Taufkirchen, Germany). All restriction enzymes used for cloning were purchased from Thermo Scientific.

3.2 Peptide Synthesis and HPLC Purification

Standard Fmoc solid-phase peptide synthesis (SPPS) was performed at a scale of 0.1 mmol. Peptides **1**, **3**, **9** and **10** were synthesized on a Liberty microwave-assisted peptide synthesizer (CEM, Kamp-Lintford, Germany) using standard coupling conditions protocols (23W, 5 min, 70°C; and 50°C for the cysteine building blocks). Amino acids were coupled to Ala-TG-Wang resin in 5:1 molar excess. All peptides were purified by preparative HPLC (Varian ProStar Series or Knauer Smartline-Series) under acidic conditions (buffer A: 0.05 % TFA in H_2O; buffer B: 0.05 % TFA in 100 % acetonitrile) with an Agilent C18 (10 µm) preparative column and confirmed by MALDI-TOF or LC-MS analysis. All peptides were of >95 % purity as monitored by analytical HPLC (Agilent 1100 series) using an Eclipse XDB-C18 3.5 µm column (Agilent) with detection at 210 nm. Purified peptides were dissolved in splice buffer (pH 7) and concentration was determined photometrically at 280 nm (ε_{280}= 5500 $M^{-1}cm^{-1)}$.

3.3 Plasmids and Protein Expression

For cloning of the plasmid backbones *E. coli* DH5α or Top10 cells were used. *E. coli* BL21 Gold (DE3) cells were used for protein expression. Plasmids encoded the protein sequences listed in Table S1, including the C-terminal fragment of the M86 intein ($M86^{C}$),[1] the *Mxe* GyrA intein,[2] the *Ssp* GyrB intein,[3] the artificially fused *cis*-AceL TerL intein,[4] the artificially fused *cis*-gp41-1 intein,[5] and the artificially fused *cis-Npu* DnaE intein.[6] In general, for the expression of proteins overnight cultures were inoculated in fresh LB medium with 100 µg/mL ampicillin or 50 µg/mL kanamycin, in the ratio 1:100. Bacterial cells were grown at 37°C with shaking (180 rpm) until OD_{600} of 0.5 was reached. Gene expression was induced with 400 µM IPTG for 4 hours at 30°C or overnight (16 hours) at 18°C.

3.4 Protein purification (Ni-NTA, CBD, MBP, SBP)

Ni-NTA: Cells were harvested by centrifugation. The cell pellet from 600 ml expression culture was resuspended in 10-15 ml Ni-NTA buffer (300 mM NaCl, 50mM Tris/HCl, pH 7.2) or phosphate buffer (150 mM NaCl, 50 mM KH_2PO_4, 1 mM EDTA, pH 7). Cells were ruptured mechanically using an Avestin EmulsiFlex C-5 emulsifier (Avestin Europe GmbH, Mannheim, Germany). After centrifugation the supernatant was equilibrated with 20 mM imidazole and was passed over Ni-NTA column pre-equilibrated with Ni-NTA buffer (20 mM imidazole). The column was washed two times with the same buffer and finally with Ni-NTA buffer with 40 mM imidazole. The purified protein was eluted with Ni-NTA buffer containing 250 mM imidazole.

CBD tag: The cell lysate was resuspended in splice buffer (50 mM Tris-HCl, 300 mM NaCl, 1 mM EDTA, pH 7.2) and passed over chitin beads. Unspecific binding to chitin beads were washed with excessed volume of splice buffer.

MBP tag: Ni-NTA elution fractions were passed over pre-equilibrated amylose resins (with Ni-NTA elution buffer (250 mM imidazole in Ni-NTA buffer or phosphate buffer) followed by amylose column buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4)) 3-4 times. After washing 2-3 times with amylose column buffer, protein was eluted with 10 mM Maltose in amylose column buffer.

SBP tag: For SBP tag purification, the Ni-NTA elution fractions were passed 3 times over streptavidin column pre-equilibrated with buffer W (100 mM Tris-HCl, 150 mM NaCl, 1 mM

EDTA, pH 8). After washing the column with buffer W, protein was eluted 2.5 mM Desthiobiotin in buffer W.

After affinity purifications, eluted fractions were checked by SDS-PAGE, pooled and dialyzed (2 times) either in splice buffer or in phosphate buffer without reducing agent (2 mM DTT) followed by dialysis (1 time) in 10% glycerol. The concentration of purified proteins was determined photometrically by absorbance at 280 nm using Beer's-Lambert law. Dialyzed fractions were further concentrated by using spin columns (vivaspin 4) with a MWCO of 5000 Da from Satorius. Purified aliquots were stored at -80°C after being flash frozen in liquid nitrogen.

3.5 Protein *trans*-splicing (PTS) reaction

In general, protein (recombinant $M86^{C}$) and peptides (synthetic $M86^{N}$) were incubated in 1:3 ratios (15 µM protein : 45 µM peptide) in splice buffer (pH 7 or 7.2) or phosphate buffer (pH 7) condition with 2 mM TCEP. In the scaled-up splice reactions (100 µL volume or more) for detection of cyclic peptides, protein (recombinant $M86^{C}$) and peptides (synthetic $M86^{N}$) were incubated in 1:1.5 ratios (30 µM protein : 45 µM peptide) with 2 mM TCEP. The splice reactions were incubated for the specified periods of time at 8°C or 25°C and monitored by SDS-PAGE. To quantify splice reactions by SDS-PAGE, Coomassie-stained gels were densitometrically analyzed using the software GelAnalyzer (www.gelanalyzer.com).

3.6 Thioester cleavage reaction

The accessibility of the N-terminal thioester formed by *cis*-intein of the dual-intein construct was checked by a cleavage assay. For the N-cleavage assay, 15 µM of the protein was incubated with 10 mM TCEP, 200 mM MESNa and 100 mM DTT in phosphate buffer (pH 8). TCEP helps to keep all the Cys including the catalytic residue in reduced states, which is essential for formation of thioester and subsequent nucleophilic attack of thiol. The cleavage reaction mixture was incubated at 25°C or 37°C depending on the type of *cis*-intein used. To quantify splice reactions by SDS-PAGE, Coomassie-stained gels were densitometrically analyzed using the software GelAnalyzer (www.gelanalyzer.com).

3.7 One pot protein *trans*-splicing (PTS) and expressed protein ligation (EPL) cyclization

For preparation of head-to-tail semisynthetic cyclic peptide, protein *trans*-splice reaction of the dual-intein constructs (recombinant $M86^{C}$) with peptide **3, 9** or **10** (synthetic $M86^{N}$) was performed followed by cyclization in presence of MESNa. For the *trans*-splice reaction 30 μM of the dual-intein precursor protein was incubated with 45 μM of peptide (protein : peptide = 1:1.5) in presence of 2 mM TCEP at pH 7. The splice reaction was incubated at 25°C or 8°C (reduced C-cleavage) for 24 h. Thereafter, 200 mM MESNa and 10 mM TCEP were added for cyclization and the reaction mixture (~100 μL) was incubated at 37°C for another 24 h. The cyclization reaction was quenched with 0.1% TFA for characterization by LC-MS.

3.8 Characterization of the cyclic peptides by LC-MS

5% acetonitrile was added to the above cyclization reaction mixture (quenched with 0.1% TFA) to achieve the starting condition of HPLC gradient. Samples were centrifuged (14000 rpm, 2 min) to remove any precipitate and the supernatant was injected into HPLC. A linear HPLC gradient (5-80% B in 50 min; solution A: 0.1% TFA in water, solution B: 0.1% TFA in acetonitrile)) with a flow rate of 0.4 mL/min was used for separation on a C18 column (ZORBAX SB-C18 RR HT, 3 x 50 mm, 1.8 μm, AgilentTechnologies, Waldbronn, Germany) coupled to MS (Agilent, 1200 infinite series; 6130 Quadrupole LC/MS, Agilent Technologies, Waldbronn, Germany). Quantification was done by integrating the chromatograms at 210 nm.

3.9 Characterization of semisynthetic SFTI1 and SFTI1-K5O by ESI-MS

For ESI-MS, an UltiMateTM 3000 RSLCnano system (Thermo Fisher Scientific Inc., MA, USA) was used that was operated as a capillary (cap) system, which was connected to a maXis II UHR-TOF LC-MS system (Bruker Corp., MA, USA) with a nano ESI-source (CaptiveSpray nanoBooster, Bruker Corp., MA, USA). 20 μL (cap-LC) of the TPCK trypsin eluted fraction was loaded on a C18 trapping column (cap-LC: Acclaim PepMap100, 5 μm, 100 Å, 300 μm i.d. x 5 mm, Thermo Scientific) at a flow rate of 20 μL/min (cap-LC) in 5% solution B (solution A: 0.1% formic acid in water; solution B: 0.1% formic acid in acetonitrile) and washed for 5 min. A linear gradient was applied (5-95% B in 85 min) at a flow rate of 4 μL/min (cap-LC) for the separation on a C18 capillary column (cap-LC: Acclaim PepMap, 2 μm, 300 μm x 150 mm, Thermo Scientific). MS settings: capillary voltage 1600 V, mass range: m/z 150-3000.

Data acquisition and manipulation were carried out with Bruker software of the Compass suite (OtofControl 4.0, DataAnalysis 4.3, Bruker Daltonik GmbH). Signals of low abundance were processed by means of the Find Molecular Features algorithm (S/N threshold 3, correlation coefficient 0.7, minimum compound length 10, smoothing width 5) of the Bruker DataAnalysis software.

3.9.1 Activity based binding of semisynthetic SFTI1 and SFTI1-K5O immobilised trypsin

Agarose beads immobilized TPCK trypsin were washed three times with 0.1M NH_4HCO_3 (pH 8) buffer. The cyclization reaction mixture of SFTI1 or SFTI1-K5O were added to the washed beads and incubated at 37°C with regular mix (1000 rpm) for 1 hours. Loaded beads were extensively washed with 0.1M NH_4HCO_3 (pH 8) buffer and SFTI1 or SFTI1-K5O peptide was eluted with 0.2M HCl (pH < 1). The eluted fraction was neutralized with NaOH and vacuum dried. For characterization of the cyclic peptides in LC-MS or ESI-MS the cyclic peptides were redissolved in HPLC starting condition (5% solution B) as described above.

4. Supporting References

1. J. H. Appleby-Tagoe, I. V. Thiel, Y. Wang, Y. Wang, H. D. Mootz and X. Q. Liu, *J Biol Chem*, 2011, **286**, 34440-34447.
2. A. Telenti, M. Southworth, F. Alcaide, S. Daugelat, W. R. Jacobs, Jr. and F. B. Perler, *J Bacteriol*, 1997, **179**, 6378-6382.
3. J. H. Appleby, K. Zhou, G. Volkmann and X. Q. Liu, *J Biol Chem*, 2009, **284**, 6194-6199.
4. I. V. Thiel, G. Volkmann, S. Pietrokovski and H. D. Mootz, *Angew Chem Int Ed Engl*, 2014, **53**, 1306-1310.
5. a) B. Dassa, N. London, B. L. Stoddard, O. Schueler-Furman and S. Pietrokovski, *Nucleic Acids Res*, 2009, **37**, 2560-2573; b) P. Carvajal-Vallejos, R. Pallisse, H. D. Mootz and S. R. Schmidt, *J Biol Chem*, 2012, **287**, 28686-28696.
6. J. Zettler, V. Schütz and H. D. Mootz, *FEBS Lett*, 2009, **583**, 909-914.