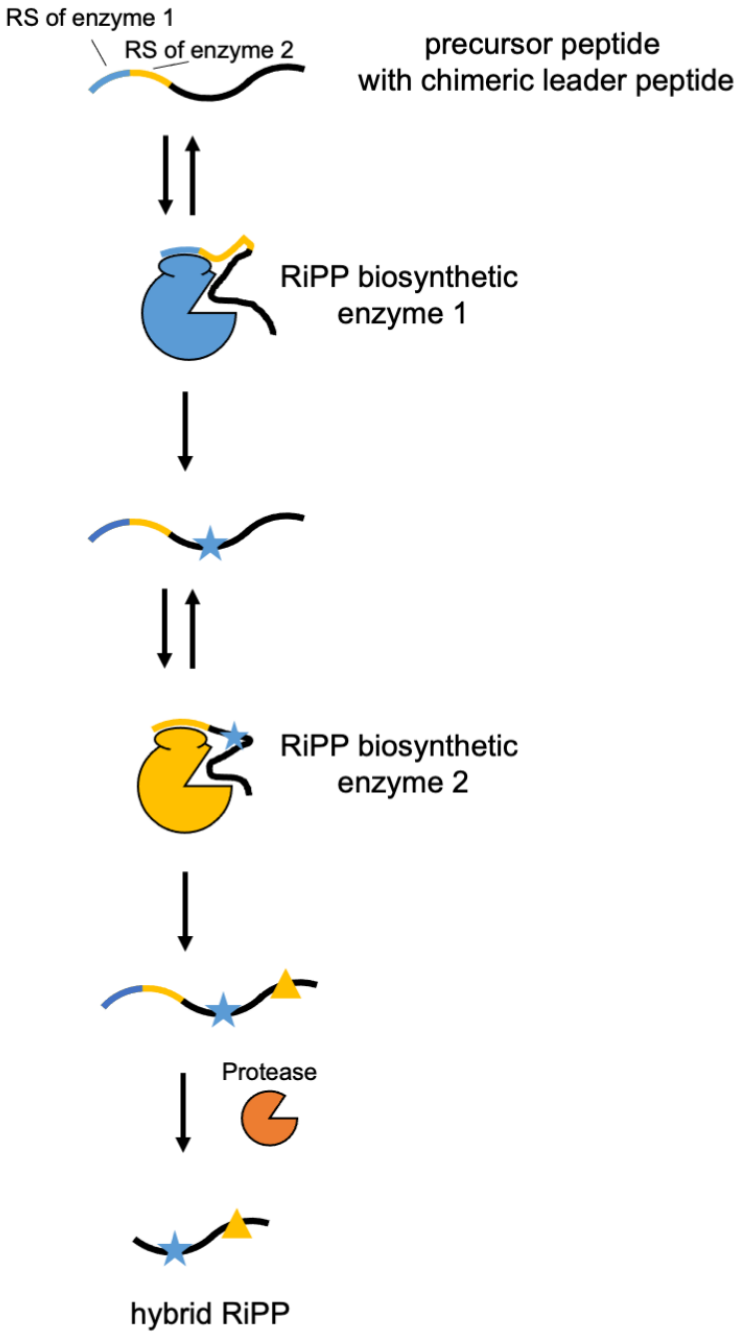
Supporting Information

Leader peptide exchange to produce hybrid, new-to-nature ribosomal natural products

Laura Franz[a] and Jesko Koehnke*[a,b]

[a] *Workgroup Structural Biology of Biosynthetic Enzymes, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), Saarland University, Campus Geb. E8.1, 66123 Saarbrücken, Germany*
[b] *School of Chemistry, University of Glasgow, Glasgow, UK. Email: Jesko.koehnke@glasgow.ac.uk*

Table of Contents



Scheme S1 Chimeric leader peptide strategy as introduced by Mitchell and co-workers.[1] The combination of recognition sequences (RS) of primary enzymes from unrelated RiPP pathways allows the generation of hybrid RiPPs. Thus far this strategy has been limited to the combination of two RSs (as shown in the Scheme) and requires an understanding of the substrate recognition requirements of each enzyme. Figure adapted from Burkhart *et al.*, 2017.[1]

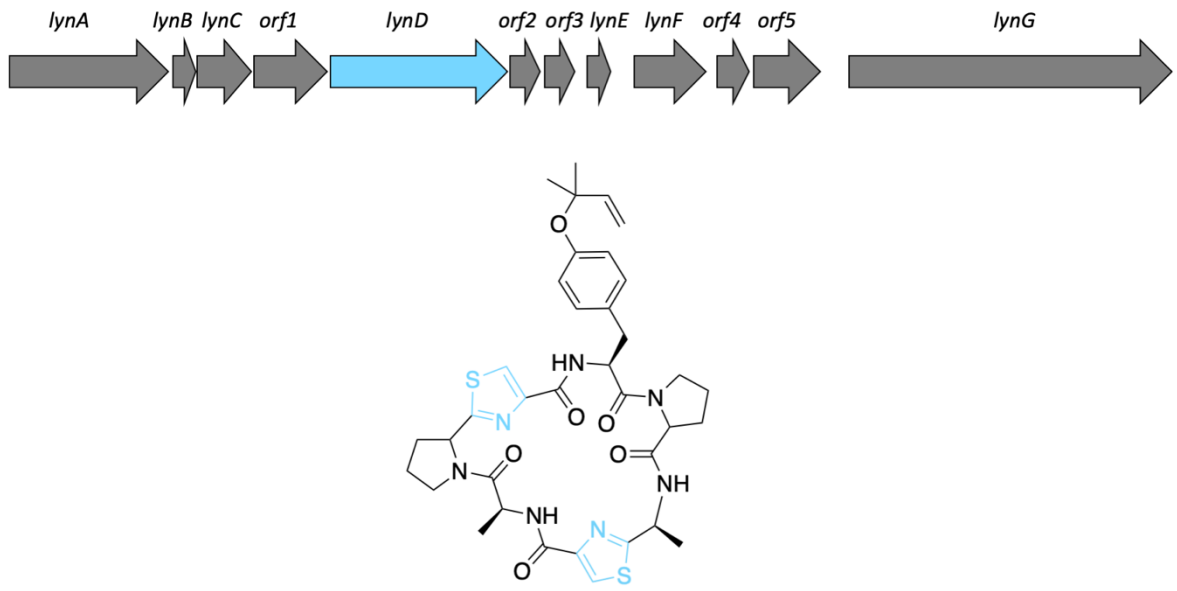


Fig. S1 Aesturamide structure (bottom) and the related cyanobactin biosynthetic gene cluster (BGC) (top) from *Lyngbya sp.*.[2] In the BGC the LynD-encoding gene is highlighted in light blue. LynD installs the cysteine-derived thiazolines, which are subsequently oxidised to thiazoles (light blue) by a different enzyme.

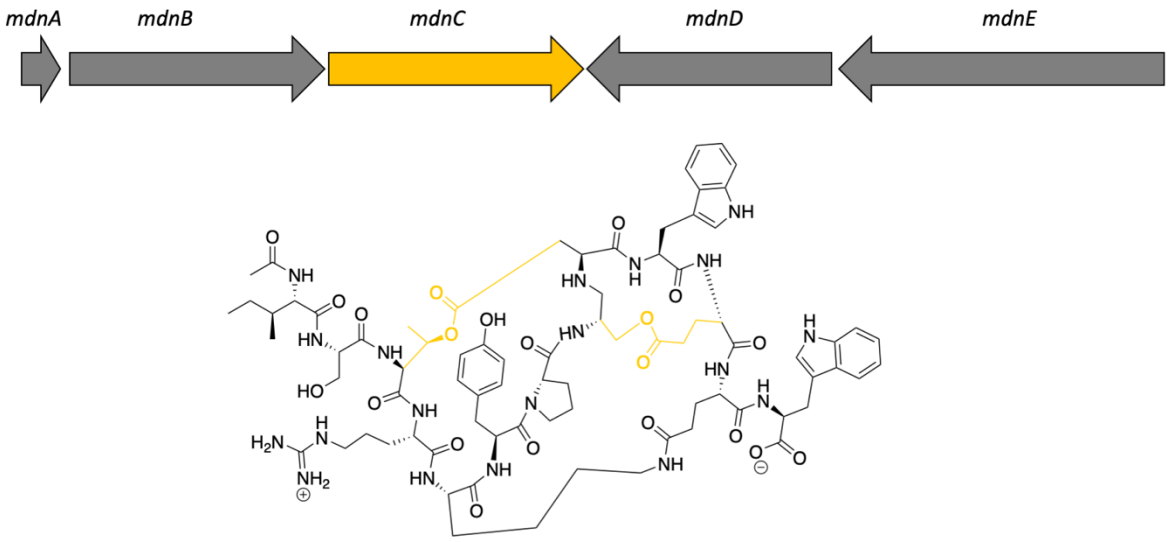


Fig. S2 The ATP grasp ligase MdnC (yellow) from the microviridin BGC (top) catalyses the formation of ω-ester bonds (yellow) from Ser/Thr and Asp/Glu side chains during the biosynthesis of microviridin J (bottom).[3, 4]

Fig. S3 Peptide sequences of full-length precursor peptides (PPs) PatE1 and MdnA and the peptides **1**-**3**. In the PP sequences, the leader peptides are shown in light blue resp. yellow, and core peptide (CP) regions are highlighted in bold. The respective recognition sequences (RS) of LynD and MdnC are underlined. Previous studies demonstrated that the PP PatE1 from related patellamide BGC is suitable for LynD.[5] We designed peptide **1** by combining a truncated PatE1 leader peptide (LP), the sortase A recognition motif (LPATG), a di-glycine and the MdnA CP sequence harbouring two point mutations (Ser2 and Pro7 to Cys).[3, 4] Peptide **2** was designed to contain a truncated MdnA LP and a C-terminal SrtA recognition motif. Peptide **3** can be obtained by sortase A-mediated transpeptidation from peptides **1** and **2**.

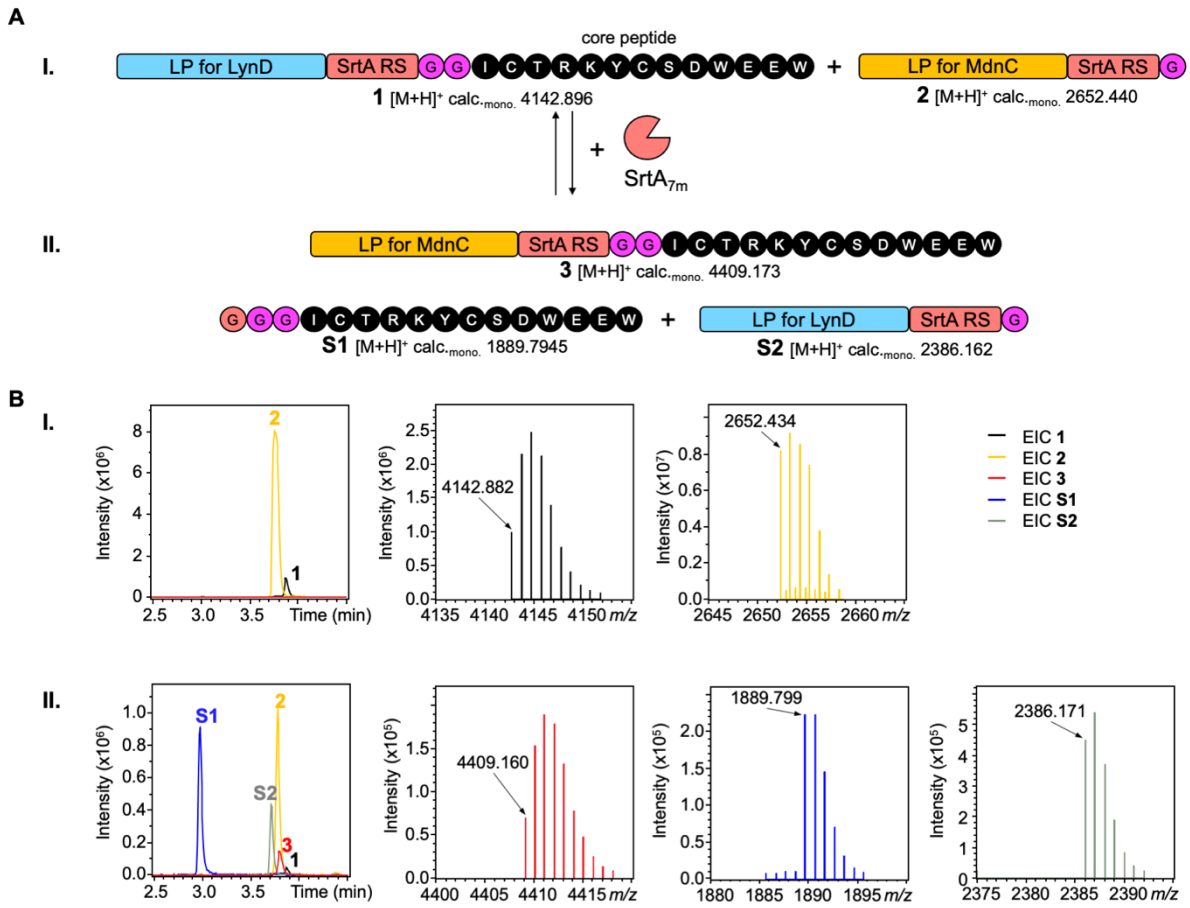


Fig. S4 SrtA$_{7m}$-catalysed leader peptide (LP) exchange. **A** SrtA$_{7m}$ cleaves the threonine-glycine bond in the SrtA recognition motif (LPAT|G) in peptides **1** and **2**, releasing peptide **S1** (from peptide **1**) or di-glycine (from peptide **2**), and forming acyl-enzyme intermediates with the threonine of the peptides ([LP]-LPAT-SrtA). The intermediates are then attacked by **S1** or di-glycine, generating peptide **3** and **S2** (or the peptides **1** and **2** are again produced). In the experiment, peptides **1** and **2** were incubated with (II.) and without (I.) the addition of SrtA$_{7m}$ (2 h, 37 °C). The reactions were analysed by LC-ESI-MS (**B**). Deconvoluted mass spectra for the peptides at the EIC maxima are shown. The masses of the singly-charged monoisotopic ions $[M+H]^+$ are stated in the spectra.

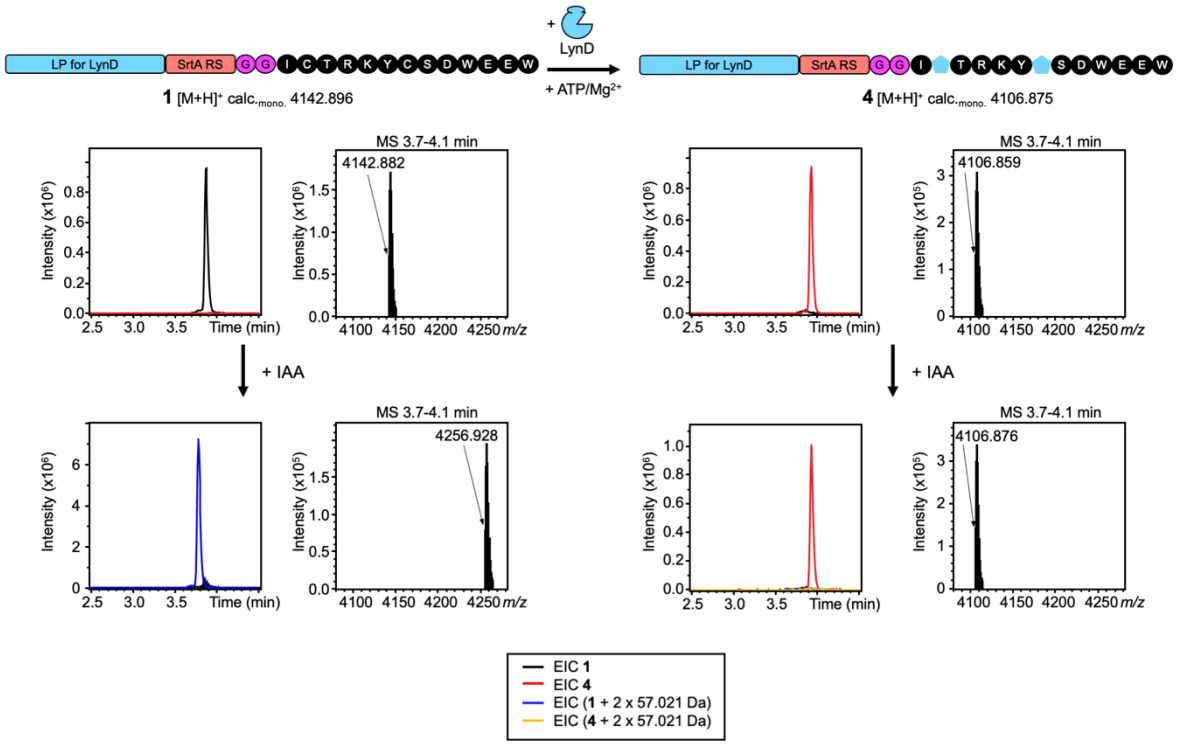


Fig. S5 LynD converts the Cys residues in peptide **1** into thiazolines (light blue pentagons). Peptide **1** was incubated in presence of ATP/Mg^{2+} with and without (negative control) LynD. Analysis of the reactions by LC-ESI-MS showed a loss of 36 Da in the LynD-treated sample, which is in agreement with the formation of two thiazolines. Treatment of the reaction products with iodoacetamide, which alkylates free Cys residues, led to a mass shift in the negative control sample, but not in the LynD-treated sample. This strongly implies that the Cys residues are heterocyclized by LynD as expected. EICs were generated for the possible peptide masses. In the deconvoluted mass spectra the masses of the singly-charged monoisotopic ions $[M+H]^+$ are stated.

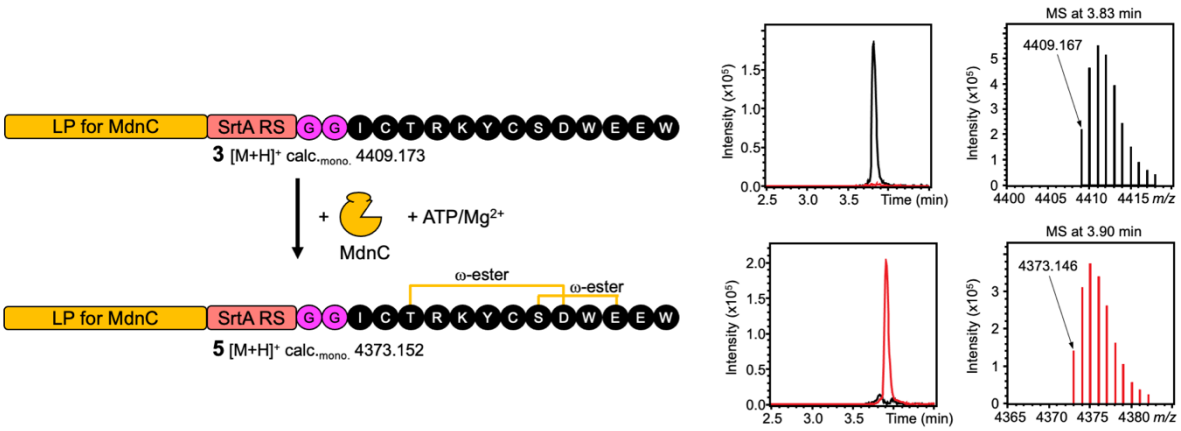


Fig. S6 MdnC introduces two ω-ester bonds into peptide **3.** The incubation of MdnC with peptide **3** in presence of ATP/Mg^{2+} results in the loss of 36 Da, as analysed by LC-ESI-MS. EICs for the expected masses of **1** (black) and **3** (red) were generated. Deconvoluted mass spectra at the EIC maxima state the masses of the singly-charged monoisotopic ions $[M+H]^+$.

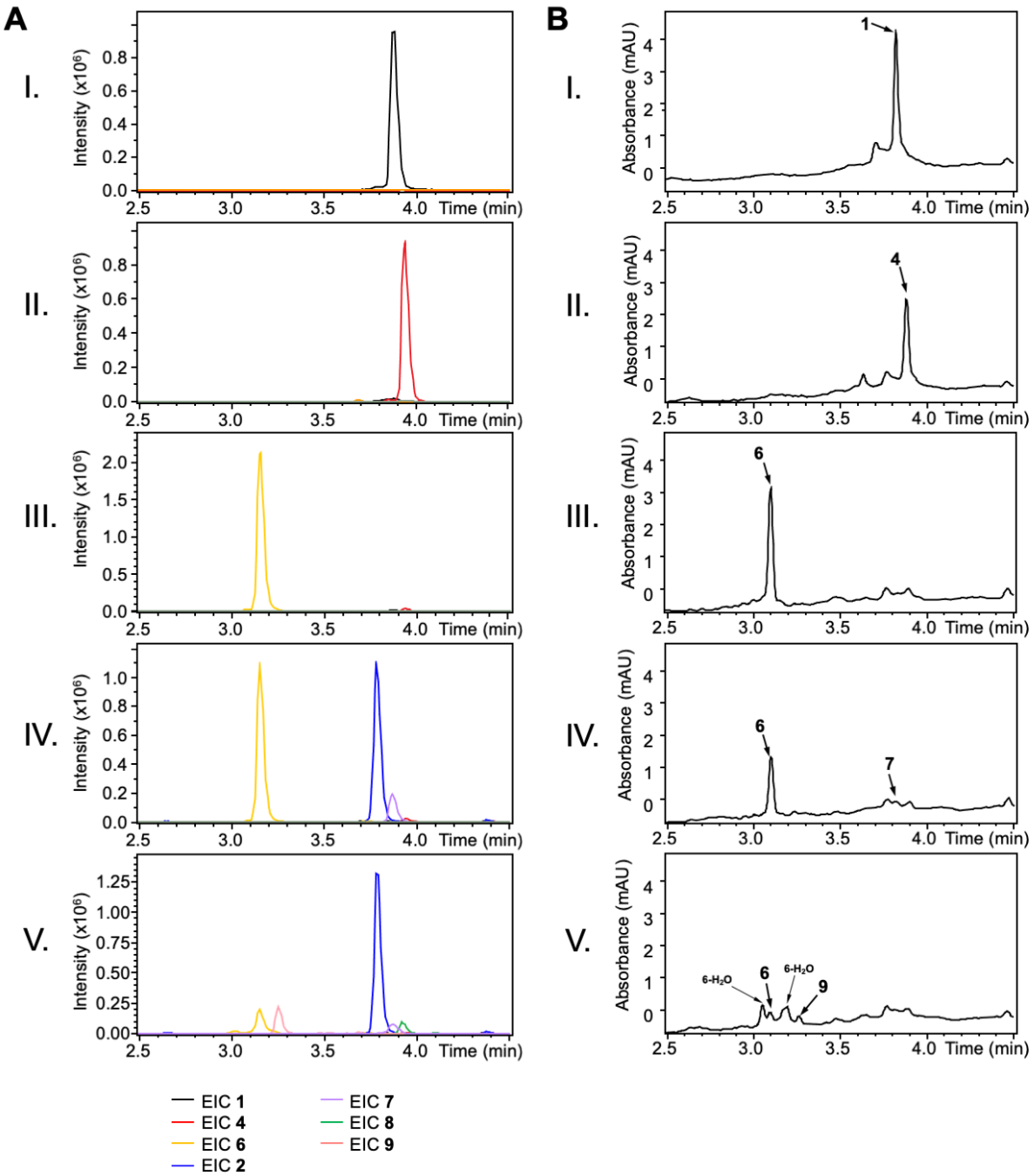
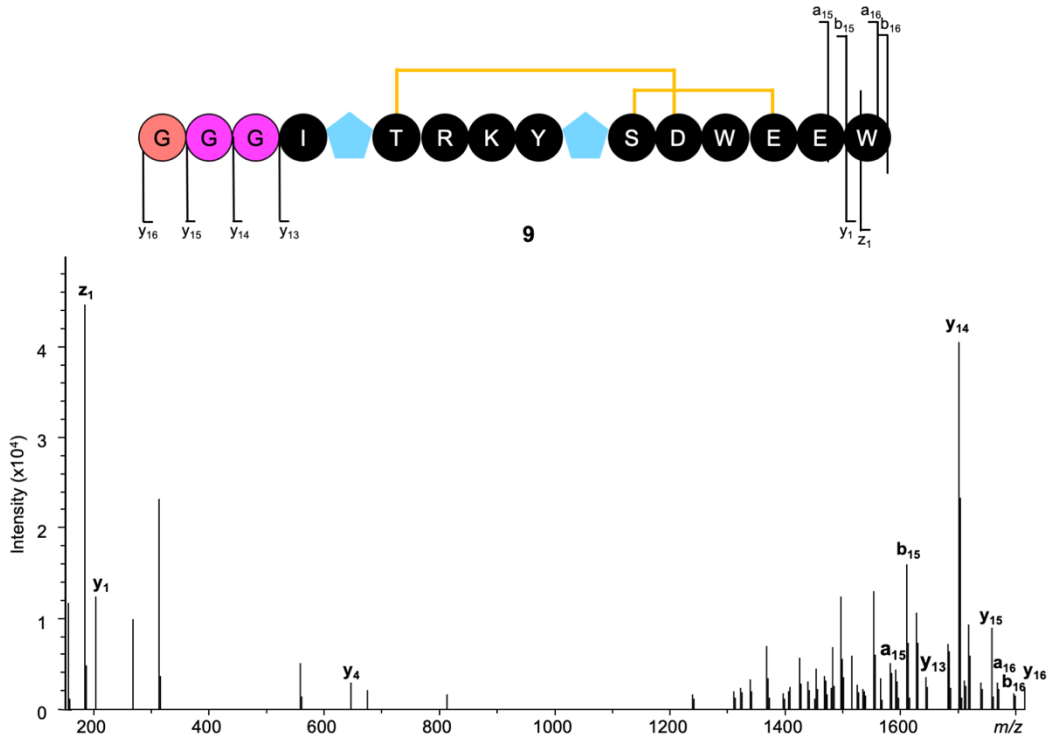


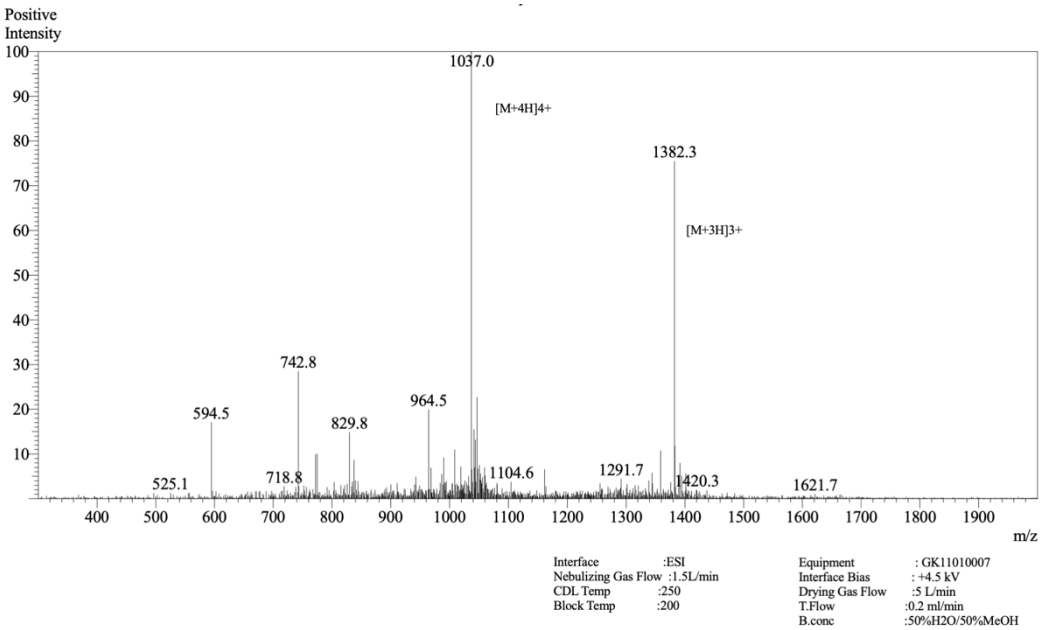
Fig. S7 Combination of the cyanobactin heterocyclase LynD and the microviridn J ATP-grasp ligase MdnC using the LPX strategy to produce a heterocycle-containing graspetide (**9**). **A** EICs for peptides **1**-**2**, **4**, **6**-**9** for the steps I.-V. (see Fig. 1). **B** UV chromatograms (λ=280 nm) for the steps I.-V.



fragment	sum formula	theo. mass	obs. mass	error (ppm)
z_1	$C_{11}H_{10}NO_2^+$	188.0712	188.0710	-1.06
y_1	$C_{11}H_{13}N_2O_2^+$	205.0977	205.0973	-1.95
y_4	$C_{32}H_{37}N_6O_9^+$	649.2622	649.2617	-0.77
a_{15}	$C_{70}H_{97}N_{20}O_{19}S_2^+$	1585.6680	1585.6643	-2.33
b_{15}	$C_{71}H_{97}N_{20}O_{20}S_2^+$	1613.6629	1613.6582	-2.91
y_{13}	$C_{76}H_{100}N_{19}O_{19}S_2^+$	1646.6884	1646.6829	-3.34
y_{14}	$C_{78}H_{103}N_{20}O_{20}S_2^+$	1703.7099	1703.7056	-2.52
y_{15}	$C_{80}H_{106}N_{21}O_{21}S_2^+$	1760.7314	1760.7284	-1.70
a_{16}	$C_{81}H_{107}N_{22}O_{20}S_2^+$	1771.7473	1771.7383	-5.07
b_{16}	$C_{82}H_{107}N_{22}O_{21}S_2^+$	1799.7423	1799.7367	-3.11
y_{16}	$C_{82}H_{109}N_{22}O_{22}S_2^+$	1817.7528	1817.7461	-3.69

Fig. S8 MS fragmentation of **9**. The detected fragments of the a, b and y series are shown in the schematic of **9**, in the MS spectrum and in the table.

A



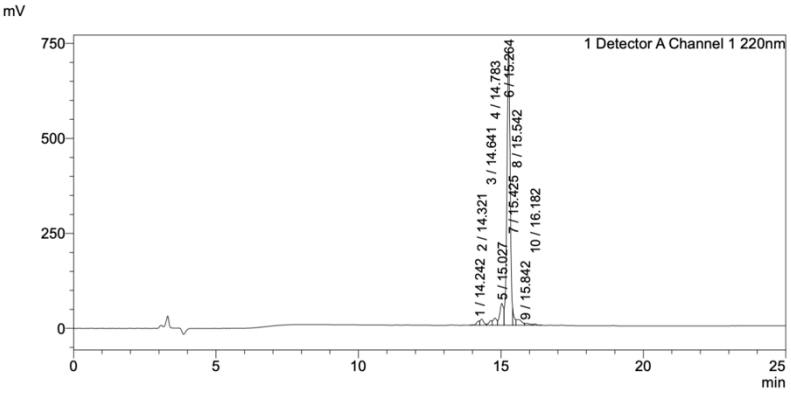
B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm
<<LC Time Program>>

Time	Module	Command	Value
0.01	Pumps	Pump A B.Conc	15
25.00	Pumps	Pump A B.Conc	75
25.01	Pumps	Pump A B.Conc	95
31.00	Pumps	Pump A B.Conc	95
31.01	Pumps	Pump A B.Conc	15
40.00	Pumps	Pump A B.Conc	15
40.01	Controller	Stop	

<<Column Performance>>
<Detector A>
Column : Inertsil ODS-3 4.6 x 250 mm
Equipment:ZJ19010324

<Chromatogram>



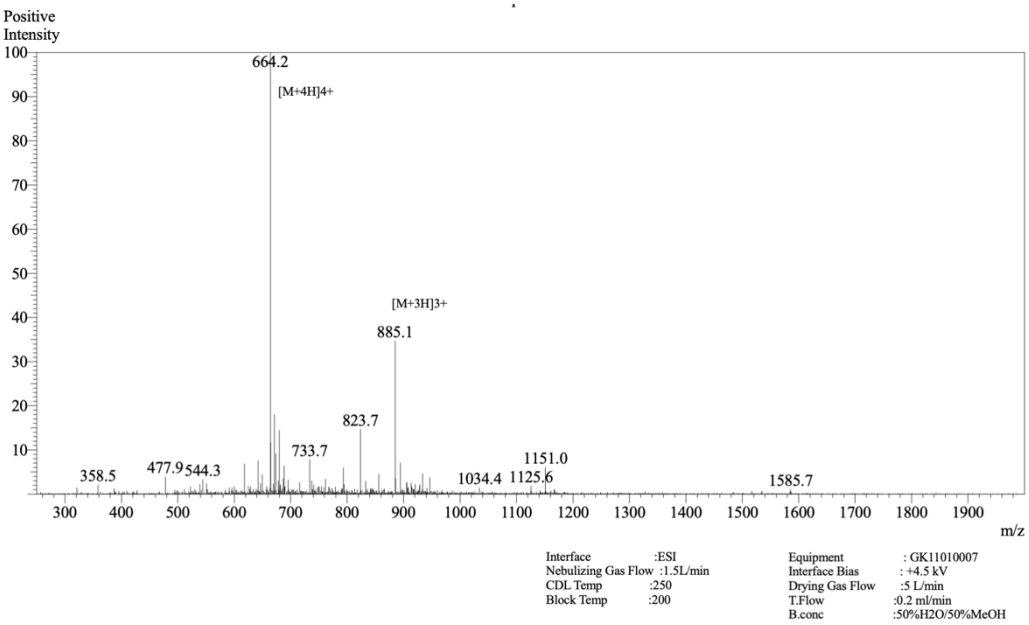
<Peak Table>

Detector A Channel 1 220nm

Peak#	Ret. Time	Area	Height	Area%
1	14.242	91565	11057	1.271
2	14.321	131846	16085	1.829
3	14.641	95453	12011	1.325
4	14.783	167047	18643	2.318
5	15.027	535779	56791	7.434
6	15.264	5746482	718456	79.738
7	15.425	155882	36828	2.163
8	15.542	198961	14451	2.761
9	15.842	59251	4816	0.822
10	16.182	24465	2948	0.339
Total		7206732	892085	100.000

Fig. S9 Analytic certificate for peptide **1**. The analytic certificate was created and provided by the vendor GenScript.

A

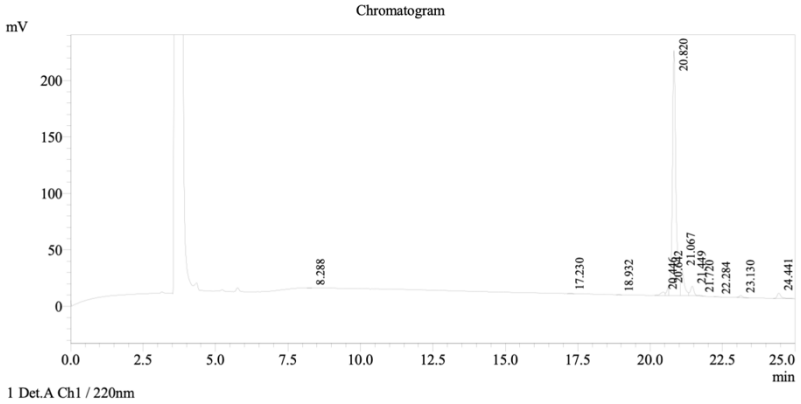


B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm

Time	Unit	Command	Value	Comment
0.01	Pumps	Pump A B.Conc	5	
25.00	Pumps	Pump A B.Conc	65	
25.01	Pumps	Pump A B.Conc	95	
27.00	Pumps	Pump A B.Conc	95	
27.01	Pumps	Pump A B.Conc	5	
35.00	Pumps	Pump A B.Conc	5	
35.01	Controller	Stop		

<<Column Performance>>
<Detector A>
Column : Inertsil ODS-3 4.6 x 250 mm
Equipment: GK1101009



1 Det.A Ch1 / 220nm

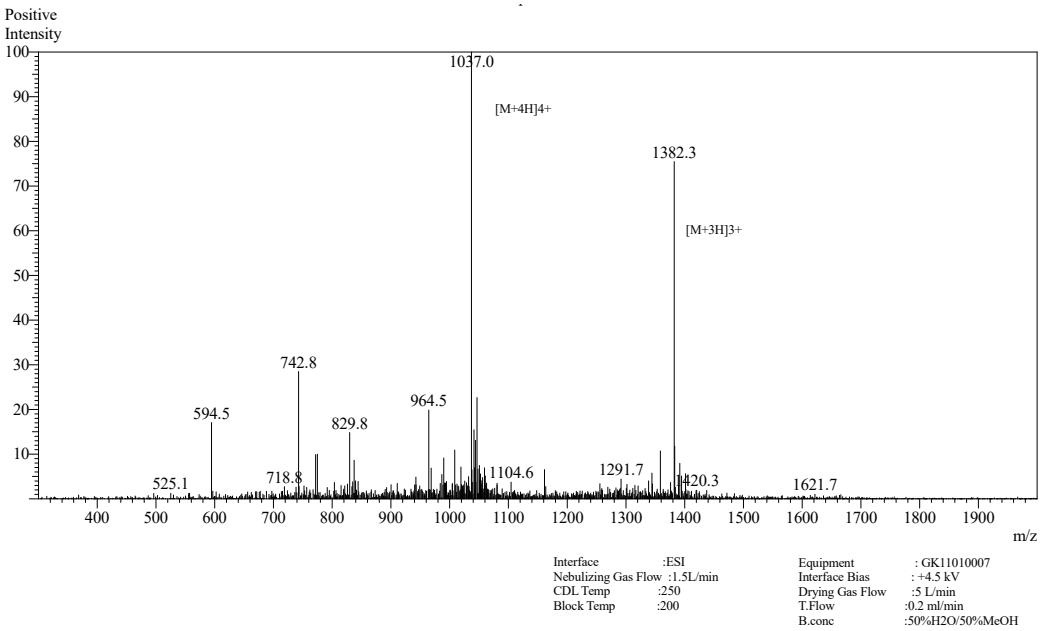
Peak Table

Detector A Ch1 220nm

Peak#	Ret. Time	Area	Height	Area %
1	8.288	1123	140	0.049
2	17.230	3518	546	0.155
3	18.932	1442	249	0.063
4	20.446	32872	2864	1.445
5	20.642	39320	8608	1.729
6	20.820	1859997	217078	81.773
7	21.067	206576	24332	9.082
8	21.449	66638	8364	2.930
9	21.720	3058	477	0.134
10	22.284	1054	159	0.046
11	23.130	14257	1884	0.627
12	24.441	44743	4754	1.967
Total		2274600	269455	100.000

Fig. S10 Analytic certificate for peptide **2**. The analytic certificate was created and provided by the vendor GenScript.

A



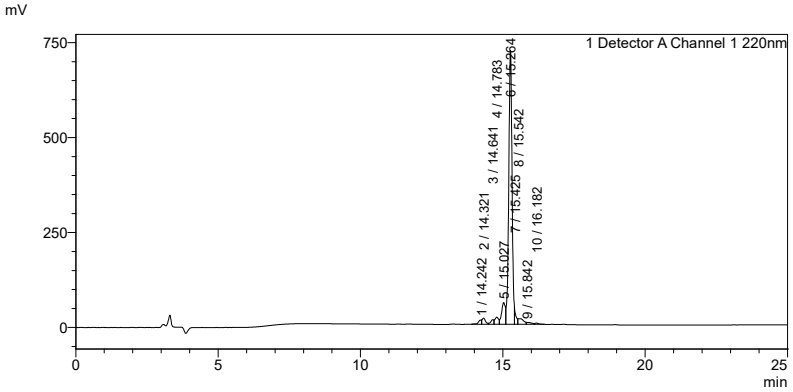
B

Pump A : 0.065% trifluoroacetic in 100% water (v/v)
Pump B : 0.05% trifluoroacetic in 100% acetonitrile (v/v)
Total Flow:1 ml/min
Wavelength:220 nm
<<LC Time Program>>

Time	Module	Command	Value
0.01	Pumps	Pump A B.Conc	15
25.00	Pumps	Pump A B.Conc	75
25.01	Pumps	Pump A B.Conc	95
31.00	Pumps	Pump A B.Conc	95
31.01	Pumps	Pump A B.Conc	15
40.00	Pumps	Pump A B.Conc	15
40.01	Controller	Stop	

<<Column Performance>>
<Detector A>
Column : Inertsil ODS-3 4.6 x 250 mm
Equipment:ZJ19010324

<Chromatogram>



<Peak Table>

Detector A Channel 1 220nm

Peak#	Ret. Time	Area	Height	Area%
1	14.242	91565	11057	1.271
2	14.321	131846	16085	1.829
3	14.641	95453	12011	1.325
4	14.783	167047	18643	2.318
5	15.027	535779	56791	7.434
6	15.264	5746482	718456	79.738
7	15.425	155882	36828	2.163
8	15.542	198961	14451	2.761
9	15.842	59251	4816	0.822
10	16.182	24465	2948	0.339
Total		7206732	892085	100.000

Fig. S11 Analytic certificate for peptide **3**. The analytic certificate was created and provided by the vendor GenScript.

Table S1 Calculated and observed peptide masses in this study. The masses for the singly-charged monoisotopic ions $[M+H]^+$ are stated. The observed singly-charged masses were obtained by deconvolution using the Bruker Compass DataAnalysis software (Version 4.4).

peptide	corresponding figure	$[M+H]^+calc._{mono.}$ (Da)	$[M+H]^+obs._{mono.}$ (Da)	error (ppm)
1	Fig. S4	4142.896	4142.882	-3.38
2	Fig. S4	2652.440	2652.434	-2.26
3	Fig. S4	4409.173	4409.160	-2.94
4	Fig. S5	4106.875	4106.859	-3.89
5	Fig. S6	4373.152	4373.146	-1.37
1	Fig.1+S7	4142.896	4142.882	-3.38
4	Fig.1+S7	4106.875	4106.866	-2.19
6	Fig.1+S7	1853.773	1853.778	+2.70
7	Fig.1+S7	4373.152	4373.141	-2.52
8	Fig.1+S7	4337.131	4337.131	-0.05
9	Fig.1+S7	1817.752	1817.745	-3.85

Materials & Methods

Protein expression and purification

Synthetically produced, codon-optimized genes coding for $SrtA_{7m}$[6] and MdnC (PDB 5IG9) were obtained from Eurofins Genomics. *$srtA_{7m}$* was cloned into the pHisTEV plasmid[7], *mdnC* was cloned into the pBMS plasmid. Both plasmids were a gift from Dr. Huanting Liu (University of St Andrews, UK). The synthetic gene for LynD was a gift from Prof. James Naismith (University of Oxford, UK). The resulting constructs were verified by DNA sequencing using T7 promotor and terminator (LGC genomics) and transformed into chemically competent *E. coli* Lemo21(DE3) cells (New England Biolabs) using a standard heat shock procedure. A single colony was inoculated into LB medium supplemented with the appropriate antibiotics and incubated overnight at 37 °C, 200 rpm for an overnight culture.

For expression of $SrtA_{7m}$, the overnight culture was diluted 1:100 into fresh LB medium supplemented with the appropriate antibiotics. The resulting cultures were grown at 37 °C, 180 rpm until an OD_{600} of 0.8, at which point the cultures were transferred to a pre-cooled 20 °C shaker (180 rpm) and protein expression was induced by the addition of 0.5 mM IPTG. Cultures were harvested after 16 h incubation by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 20 mM Imidazole, 3 mM β-mercaptoethanol, 10% glycerol) supplemented with cOmplete EDTA-free protease inhibitor tablets (Roche) and DNase (0.4 mg/g wet cells, Sigma). The cell suspension was lysed via passage through a cell disrupter (30 kpsi, Constant Systems), and cell debris was removed by centrifugation (40,000 x g, 4 °C, 15 min). The supernatant was loaded onto a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated in lysis buffer. After an extensive column wash (30 CV), the His_6-tagged protein was eluted using lysis buffer supplemented with 250 mM imidazole. The imidazole was removed by passing the eluate through a 16/10 desalting column (GE Healthcare), equilibrated in lysis buffer. To cleave the His_6-tag, the protein was incubated with TEV protease (ratio 1:10 TEV : protein, 4 °C, 16 h). After the incubation, the solution was applied to a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated in lysis buffer. The flow through was concentrated and loaded onto a gel filtration column (HiLoad 16/600 Superdex 200 pg, GE healthcare) pre-equilibrated in gel filtration buffer (20 mM HEPES, 200 mM NaCl, 10% glycerol, 1 mM TCEP, pH 7.4). Fractions of highest purity were determined by SDS-PAGE, pooled and concentrated to 500 µM.

MdnC and LynD were expressed and purified similarly using the expression conditions and buffers as previously described.[4, 5]

In vitro experiments

Peptides used in this study were purchased from GenScript. The peptide purity was: peptide **1** 79.7%, peptide **2** 81.7%, peptide **3** 94.1%. The vendor's analytic certificates are shown in figures S9-S11.

To probe the $SrtA_{7m}$ catalysed exchange of the LP for LynD with the LP for MdnC, 100 µM peptide **1** and 100 µM peptide **2** were incubated with and without the addition of 5 µM $SrtA_{7m}$ in reaction buffer 1 (20 mM HEPES, 200 mM NaCl, 2.5 mM TCEP, pH 7.4) for 2 h at 37 °C.

To test the ability of LynD to convert the two Cys residues in the peptide **1** into thiazolines, 100 µM peptide **1** was incubated with and without the addition of LynD in reaction buffer 2 (5 mM ATP, 5 mM $MgCl_2$, 20 mM HEPES, 200 mM NaCl, 2.5 mM TCEP, pH 7.4) for 2 h at 37 °C.

The ability of MdnC to introduce two ω-ester bonds into peptide **3** was tested by incubating 50 µM peptide **3** with 5 µM MdnC in reaction buffer 2 for 8 h at 37°C.

For the combination of all components, 100 µM peptide **1** was incubated with 5 µM LynD in reaction buffer 2 for 2 h at 37 °C. Next, 100 µM peptide **2**, 5 µM $SrtA_{7m}$ and 5 µM MdnC were added and incubated for 16 h at 37 °C. Negative controls were set up analogously without the addition of $SrtA_{7m}$.

For iodoacetamide alkylation, 20 mM iodoacetamide was added to the reaction solutions and incubated for 1 h at 20 °C in the dark.

All reactions were stopped by the addition of 2 volumes ACN and frozen at -80 °C until analysis by LC-ESI-MS. All reactions were run at least three times in independent experiments with similar results.

LC-MS and MS^2 analysis

All measurements were performed on a Dionex Ultimate 3000 RSLC system (Thermo Fisher Scientific) using a flow rate of 600 µL min^{-1} and a column oven temperature of 45 °C. Samples were separated by a gradient from (A) H_2O + 0.1% formic acid to (B) ACN + 0.1% formic acid using a BEH C18, 50 x 2.1 mm, 1.7 µm dp column equipped with a C18 precolumn (Waters). The linear gradient was initiated by a 1 min isocratic step at 5% B, followed by an increase to 95% B in 9 min to end up with a 1.5 min plateau step at 95% B before re-equilibration to the initial conditions.

For MS measurements on a maXis-4G hr-qToF mass spectrometer (Bruker Daltonics), the LC flow was split 1:8 before entering the mass spectrometer using the Apollo II ESI source. In the source region, the temperature was set to 200 °C, the capillary voltage was 4000 V, the dry-gas flow was 5.0 L / min and the nebulizer was set to 1 bar. Ion transfer settings were set to Funnel 1

RF 350 Vpp and Multipole RF 400 Vpp, quadrupole settings were set to an ion energy of 5.0 eV and a low mass cut of 120 m/z. The collision cell was set to an energy of 5.0 eV and the pulse storage time was 5 µs. Data were recorded in centroid mode ranging from 150 to 2500 m/z at a 2 Hz scan rate. Calibration of the maXis4G qTOF spectrometer was achieved with sodium formate clusters before every injection to avoid mass drifts. All MS analyses were acquired in the presence of the lock masses $C_{12}H_{19}F_{12}N_3O_6P_3$, $C_{18}H_{19}O_6N_3P_3F_2$ and $C_{24}H_{19}F_{36}N_3O_6P_3$ which generate the $[M+H]^+$ ions of 622.028960, 922.009798 and 1221.990638.

LC-MS^2 fragmentation spectra were recorded using a scheduled precursor list (SPL). The SPL entries and parameters were set to fragment the $[M+6H]^{6+}$ ion of **9**. SPL tolerance parameters for precursor ion selection were 0.15 min and 0.05 m/z. The CID energy was ramped from 35 eV for 500 m/z to 45 eV for 1000 m/z.

Data were displayed and analyzed using the Bruker Compass DataAnalysis software (Version 4.4). Shown MS spectra are deconvoluted and are single spectra (taken at the EIC maximum), except for Figure S5, where spectra are averaged. The signals in the MS-spectra are labelled with the observed monoisotopic mass. Extracted-ion chromatograms were generated using the calculated monoisotopic masses of the ions with the charge states 1-6 and a mass range of 0.01 Da.

References

1. B. J. Burkhart, N. Kakkar, G. A. Hudson, W. A. van der Donk and D. A. Mitchell, *ACS Cent Sci*, 2017, **3**, 629-638.
2. J. Martins and V. Vasconcelos, *Journal*, 2015, **13**, 6910-6946.
3. N. Ziemert, K. Ishida, A. Liaimer, C. Hertweck and E. Dittmann, *Angewandte Chemie International Edition*, 2008, **47**, 7756-7759.
4. K. Li, H. L. Condurso, G. Li, Y. Ding and S. D. Bruner, *Nature chemical biology*, 2016, **12**, 973-979.
5. J. Koehnke, G. Mann, A. F. Bent, H. Ludewig, S. Shirran, C. Botting, T. Lebl, W. Houssen, M. Jaspars and J. H. Naismith, *Nat Chem Biol*, 2015, **11**, 558-563.
6. H. Hirakawa, S. Ishikawa and T. Nagamune, *Biotechnology Journal*, 2015, **10**, 1487-1492.
7. H. Liu and J. H. Naismith, *Protein Expr Purif*, 2009, **63**, 102-111.