**Supporting Information**

# Roads to Rome: Role of Multiple Cassettes in Cyanobactin RiPP Biosynthesis

Wenjia Gu, Debosmita Sardar, Elizabeth Pierce, Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

## Table of contents

## Supporting Materials and Methods

### Genes and cloning

***General method***

All primers are listed in **Table S1** and were obtained from DNA synthesis core, University of Utah. gBlocks (**Table S1**) for Gibson cloning are synthesized by IDT. Vector synthetic schemes are shown in **Scheme S1** and **S2**.

***Plasmids for heterologous expression of* tru *pathway derivatives***

*tru* pathway was cloned from pTru-SD1[1] and placed into a pUC19 vector under *lac* promoter control. The AatII digestion site in the pUC19 backbone was removed by PCR (primer pUC-AatII-fwd and pUC-AatII-rev) and a NsiI site was added at the lacZ-alpha start codon (primer pUC-NsiI-fwd and pUC-NsiI-rev) to create pTru-p3p2. It contains the *tru* pathway genes required to make patellins 2 (p2, **2**) and 3 (p3, **3**) in *E. coli* DH10β. The precursor gene has a SalI restriction site before cassette-1, a MluI site between the two cassettes and an AatII site after cassette-2. Alternative cassettes were substituted for the native *truE-p3p2* gene as described in **Scheme S1**. Single cassette substitutions follow route **a** (pTru-p3p3, pTru-p3tk). Double cassette substitutions follow route **b** (pTru-p2p3, pTru-p2tk, pTru-p2p2, pTru-tkp3, pTru-tkp2, pTru-tktk; tk, trunkamide, **1**). Cassette-2 p2 mutant substitutions follow route **c** (pTru-p2-p2L2A, pTru-p2-p2P4A, pTru-p2P4A-p2P4A). P2 mutant double cassette substitutions follow route **d** (pTru-p2L2A-p2L2A, pTru-p2P4A-p2P4A). Single cassette precursor plasmids encoding either patellin 3, 2 or trunkamide sequence follow route **e** (pTru-p3, pTru-p2, pTru-tk). Gibson cloning was done by mixing Gibson master mix (15 µL),[2] insert (1 µL, 200 ng) and backbone (4 µL, 100 ng) and incubating at 50 ˚C for 2 h. PCR and ligation conditions followed the vendor's protocols. A total of 13 plasmids (**Table S2**) with different combinations of the precursor genes were synthesized and verified by Sanger sequencing.

***Plasmids for production of proteins and peptides***

Plasmids for expression of His-tagged proteins TruD, RSI-TruD, PatA, PatG, and TruG were previously described.[3–6] His-tagged TruE substrate derivatives were synthesized as follows in pET28b(+) (**Scheme S2**). Vector for pTruE-p3p2 was published previously.[1] The published vector pTruE-p3Ly1[7] was used to make pTruE-p2p3, pTruE-p3p3 and pTruE-p23 by the around horn method shown in **Scheme S2** route **b**. Deletion of the second cassette was done following route **a** to make pTruE-p3. pTruE-p2p2 was made by route **c**.

### Substrate synthesis

Synthetic peptides were purchased from Genscript and are listed in **Table S3**.

## Protein expression and purification

RSI-TruD, TruD, PatA, PatG and TruG were purified following the method described previously.[3–6] The plasmids were transformed into electrocompetent *E. coli* BL21 Rosetta™ 2 (DE3) (Millipore Sigma). The cells were plated on LB agar with kanamycin (50 mg/mL), incubated overnight at 30 ˚C, then inoculated into 2xYT (20 mL) media and grown overnight at 30 ˚C, 150 rpm. 10 mL of seed culture was used per liter of 2xYT media, and a total of 2 L culture were used for each protein. The cultures were incubated at 30 ˚C, 150 rpm until $OD_{600}$ reached 0.8. After cooling to 18˚C, isopropyl β-D-thiogalactopyranoside (IPTG, 0.1 mM) was added, and the culture was incubated overnight. The cells were harvested by centrifugation, and the pellets were collected and frozen at -80 ˚C. After thawing, cell pellets were resuspended in the lysis buffer (5 g/mL, 50 mL in total; Tris-Cl (50 mM), NaCl (200 mM), imidazole (10 mM), 5% glycerol, lysozyme (600 μg/mL), pH 7.5) and centrifuged at 21728 g. The supernatant was collected and filtered, then loaded onto the Ni-NTA affinity resin (Qiagen). The resin was rinsed twice with wash buffer (25 mL; NaCl (1 M), imidazole (30 mM), pH 8.0), and elution buffer (12 mL; NaCl (1 M), imidazole (200 mM), pH 8.0) were used in four batches to elute the protein out. The eluents were analyzed by tris-glycine SDS-PAGE and dialyzed at 4 ˚C in the following buffer: HEPES (25 mM), NaCl (500 mM), 10% glycerol, pH 8.0. The dialysis buffer was changed four times and the proteins were flash-frozen and stored at -80 ˚C.

TruE substrates were expressed and purified by the same method as the enzymes except 1 mM of IPTG was added during induction and was analyzed by tris-tricine SDS-PAGE (**Figure S15**).[8]

## *tru* pathway heterologous expression in *E. coli*

*tru* pathway plasmids were transformed into chemically competent *E. coli* DH10β, and the resulting transformants were streaked onto LB agar containing ampicillin (50 mg/mL). Nine different plasmids were used with the following combinations: p3p2 (p3 sequence in cassette-1 and p2 sequence in cassette-2), p2p3, p2p2, p3p3, tkp2, p2tk, tktk, p3tk, tkp3. Single cassette plasmids (p3, p2 and tk) were expressed as a comparison. Plasmids encoding patellin 2 mutants (pTru-p2-p2L2A, pTru-p2L2A-p2L2A, pTru-p2-p2P4A, pTru-p2P4A-p2P4A) were also expressed. For inoculation, five colonies from the agar plate were picked and incubated to 2xYT media (5 mL) overnight, in a shaker that was set to 30 ˚C, 150 rpm. Next, the inoculations were combined and cells (50 μL) were added per 5 mL of 2xYT media in a 24-well plate, with addition of sterilized L-cysteine hydrochloride (Amresco, 10 mM). All experiments were performed at least twice, in triplicate technical replicates, with representative results shown. Cultures were incubated at 30 ˚C with shaking at 150 rpm for 5 days. On day 2, 3, 4, and 5, three independent wells representing each condition were harvested by centrifugation. Cultures without cysteine additives were also harvested as a negative control. For p2 mutants, cells were only harvested on day 5. For single-cassette expression shown in Figure S4, cultures were harvested on days 3 and 4 and performed in triplicate technical replicates in a single experiment. After washing with phosphate-buffered saline (NaCl (137 mM), KCl (2.7 mM), $Na_2HPO_4$ (10 mM), $KH_2PO_4$ (1.8 mM)), the cells were extracted with acetone (2.5 mL), centrifuged for 1 h, then the acetone fraction was removed by pipet, dried, and redissolved in methanol (200 μL).

## Enzymatic reactions

### *TruD enzymatic reactions*

TruE-p3p2, TruE-p2p3, TruE-p3p3 and TruE-p2p2 were heterocyclized by TruD under the following conditions: substrate (50 μM), TruD (1 μM), Tris-Cl (50 mM, pH 7.5), $MgCl_2$ (5 mM), ATP (1 mM) and dithiothreitol (DTT, 6 mM), 37 ˚C in a water bath. The reactions were done at least twice, in triplicate technical replicates. During the following time points 30 μL of sample were taken: 10 s, 1, 3, 5, 10, 20, 40, 60 min. The reactions were quenched for 10 min in an oil bath heated to 110 ˚C. A portion of the samples (5 μL) was analyzed by HPLC-MS, while the rest (~25 uL) were passed through P-6 gel resin slurry (100 μL; Bio-rad) on a 96-well filter plate. The ion-exchange resin was pre-equilibrated with Tris-Cl (50 mM, pH 7.5) and $MgCl_2$ (5 mM) to remove DTT that inhibits PatA. After desalting, the reaction mixtures were fully digested by PatA at 37 ˚C in a thermocycler (Bio-rad, C1000 Touch) for 2 h under the following conditions: TruD reaction mixture (25 μM of the original precursor peptide), PatA (1 μM), Tris-Cl (50 mM, pH 7.5) and $MgCl_2$ (5 mM).

### *PatA enzymatic reactions*

For PatA digestion reactions on fully heterocyclized precursor peptides, TruE-p3p2, TruE-p2p3, TruE-p3p3 and TruE-p2p2 were heterocyclized by TruD under the following conditions: substrate (90 μM), TruD (2 μM), Tris-Cl (50 mM, pH 7.5), $MgCl_2$ (5 mM), ATP (1 mM) and DTT (6 mM), 37 ˚C in a water bath for 6 h, and quenched in a 100 ˚C heat block for 10 min. Products were verified by HPLC-MS to ensure full heterocyclization on both cassettes. To remove DTT that inhibits PatA reactions, the reaction mixtures were passed through glass pasteur pipettes filled with P-6 gel resin slurry (2 mL; Bio-rad) pre-equilibrated with Tris-Cl (50 mM, pH 7.5) and $MgCl_2$ (5 mM). PatA reactions were performed in triplicates in a water bath at 37 ˚C, with substrate (50 μM), PatA (0.5 μM), Tris-Cl (pH 7.5, 50 mM) and $MgCl_2$ (5 mM). The following time points were used: 10 s, 1, 3, 5, 10, 15, 20, 30, 50, 60 min. At each time point, 10 μL of the reaction mixture was taken and quenched by methanol (10 μL).

For PatA digestion reactions on unmodified precursor peptides, TruE-p3p2, TruE-p2p3, TruE-p3p3 or TruE-p2p2 were directly added to the reaction mixture: substrate (50 μM), PatA (0.5 μM), Tris-Cl (50 mM, pH 7.5) and $MgCl_2$ (5 mM). When exploring the disulfide bridge effect, reducing agent TCEP was added to the mixture with different concentrations: 0, 50 μM (1:1 molar ratio to substrate), 100 μM (2:1) and 200 μM (4:1).

### *TruD enzymatic reactions after PatA proteolysis*

Unmodified precursor peptides (TruE-p3p2, TruE-p2p3, TruE-p3p3 or TruE-p2p2) were fully proteolyzed in the reaction mixture: substrate (90 μM), PatA (1 μM), Tris-Cl (50 mM, pH 7.5) and $MgCl_2$ (5 mM) for 6 h in a 37 ˚C water bath and confirmed by HPLC-MS. Then, mixture is added to a TruD reaction mixture: substrate (50 μM), TruD (2 μM), Tris-Cl (50 mM, pH 7.5), $MgCl_2$ (5 mM), ATP (1 mM) and DTT (6 mM), and incubated in a water bath at 37 ˚C. The following time points were used: 10 s, 20 min, 40 min.

***RSI-TruD enzymatic reactions***

Synthesized short peptides including p2-RSIII-RSII, p2-RSIII, p3-RSIII and p3-RSIII-RSII (150 μM) were each fully heterocyclized in a reaction mixture (500 μL) containing RSI-TruD (2 μM), Tris-Cl (50 mM, pH 7.5), $MgCl_2$ (5 mM), ATP (1 mM) and DTT (6 mM) in a water bath for at least 6 h at 37 ˚C. The reactions were then quenched for 10 min heat inactivation in a 100 ˚C heat block. The fully heterocyclized peptides were used as standards or substrates for PatG/TruG reactions.

***TruD enzymatic reactions with or without RSI***

Synthesized short peptides including p2-RSIII-RSII, p2-RSIII, p3-RSIII and p3-RSIII-RSII (50 μM) were heterocyclized in a reaction mixture containing TruD or RSI-TruD (2 μM), Tris-Cl (50 mM, pH 7.5), $MgCl_2$ (5 mM), ATP (1 mM) and DTT (6 mM) in a water bath at 37 ˚C. Time points were taken at 5, 10, 20, 30 and 40 min.

***Standard curves for short peptides***

A mixture of the same cassette, heterocyclized or unheterocyclized short peptides (e.g., p3-RSIII and p3*-RSIII mixed together; * represents heterocyclized cassette) were prepared with concentrations of 0.5, 1, 5, 10, and 20 μM to determine the linear range response on the mass spectrometer. For p3-RSIII-RSII and p3*-RSIII-RSII set, an extra set of concentrations were used: 0.5, 1, 2.5, 5, 7.5 μM. Different concentrations were prepared and measured in triplicates.

***PatG and TruG enzymatic reactions***

The heterocyclized short peptides (final concentration 50 μM of each substrate) were added to PatG reaction mixture (Tris-Cl (50 mM) pH 7.5, $MgCl_2$ (5 mM), and PatG (10 μM))[9] or TruG reaction mixture (NaCl (500 mM), bicine (20 mM) pH 8.0, LiCl (1 mM), 5% DMSO and TruG (10 μM))[10] with different combinations: p2*-RSIII-RSII with p2*-RSIII (representing p2*p2*), p3*-RSIII-RSII with p2*-RSIII (p3*p2*), p2*-RSIII-RSII with p3*-RSIII (p2*p3*), p3*-RSIII-RSII with p3*-RSIII (p3*p3*). Same cassette sets were also used in the following combination for PatG reactions: p2*-RSIII with p3*-RSIII (cassette-2 sets), and p2*-RSIII-RSII with p3*-RSIII-RSII (cassette-1 sets). Reactions were performed in triplicate in a water bath at 37 ˚C. For PatG, six different time points were taken: 10 s, 3, 10, 21, 29, and 53 h. For TruG, the following time points were used: 10 s, 3, 13, 15, 28.25 and 37 h. In each time point, aliquot of the mixtures (10 μL or 20 μL) were removed and quenched in equal volume of methanol.

## Analysis of reaction products

High-performance liquid chromatography-electrospray ionization MS (HPLC-ESI-MS) was performed using a Micromass Q-TOF Micro Mass Spectrometer (Waters) in positive mode. Solvent A was 0.1% formic acid in Optima LC-MS grade water and solvent B was Optima LC-MS grade acetonitrile. A flow rate of 0.3 mL/min was used in all cases. An internal standard e8 (0.25 μg/mL) was added to all samples to minimize autosampler injection variations.

***For* tru *pathway heterologous expression***

From the 200 µL of sample dissolved in methanol, 50 microliters were added to the LCMS vial and with the e8 internal standard cyclic peptide (5 µL, 0.25 µg/mL). a C18 column (Kinetex® 2.6 µm C18 100 Å, 50 x 2.1 mm) was used with the following gradient: **1)** 0-1 min, 90% A and 10% B; **2)** 1-11 min, a linear gradient from 10% B to 90% B; **3)** 11-20 min, 90% B; **4)** 20-21 min, a linear gradient from 90% B to 10% B; **5)** 21-25 min, equilibration of the column. The observed *m/z* is shown in **Table S4**.

***For enzyme reactions***

**Standards.** A C18 column (Kinetex® 2.6 µm C18 100 Å, 50 x 2.1 mm) was used. Samples were run in random orders in triplicate: **1)** 0-1 min, 10% B; **2)** 1-15 min, a linear gradient from 10% B to 95% B; **3)** 15-16 min, 95% B; **4)** 16-18 min, a linear gradient from 95% B to 10% B; **5)** 18-20 min, equilibration of the column. The observed *m/z* is shown in **Table S10** and the result is shown in **Figure S16**.

**TruD reactions.** Samples were diluted to a 1-5 µM range, which is in the linear response range for the cassette fragments. For TruD reactions, a C4 column (Aries® widepore 4.6 µm C4 200 Å, 100 x 2.1 mm) was used with the following gradient: **1)** 0-1 min, 10% B; **2)** 1-11 min, a linear gradient from 10% B to 90% B; **3)** 11-11.5 min, 90% B; **4)** 11.5 - 12 min, a linear gradient from 90% B to 10% B; **5)** 18-20 min, equilibration of the column. The observed *m/z* is shown in **Table S5**.

**PatA and PatG reactions.** A C18 column (Kinetex® 2.6 µm C18 100 Å, 50 x 2.1 mm) was used. For PatA reactions, the gradient was **1)** 0-1 min, 10% B; **2)** 1-15 min, a linear gradient from 10% B to 95% B; **3)** 15-16 min, 95% B; **4)** 16-18 min, a linear gradient from 95% B to 10% B; **5)** 18-20 min, equilibration of the column. The expected *m/z* for PatA reaction products is shown in **Table S6 – S8**. For PatG reactions, the gradient is as follows: **1)** 0-2 min, 0.5% B; 2-7 min, a linear gradient from 0.5% B to 70% B; **2)** 7-8.5 min, a linear gradient from 70% B to 100% B; **3)** 8.5 - 9 min, a linear gradient from 100% B to 0.5% B; and **4)** 18-20 min, equilibration of the column. The observed *m/z* is shown in **Table S9**.

The integrated areas of the substrate/intermediate/product peaks were quantified with MassLynx 4.0 Quantification function. The areas were normalized to the internal standard e8 to remove injection variations from the autosampler. For multiple charge ions, the area of the highest response was used. To normalize the ionization differences between each peptide, a standard, a 2-fold and a 4-fold dilution of the standards were run in each batch, and the ratio of the standard peptides was used to eliminate the ionization differences between short peptides. The standard includes 4 µM of p2-RSIII-RSII, p2-RSIII, p3-RSIII, p3-RSIII-RSII, p2*-RSIII-RSII, p2*-RSIII, p3*-RSIII, p3*-RSIII-RSII, RSIII, and RSIII-RSII. The data were exported and processed in Microsoft Excel and the figures were plotted with Prism GraphPad 7.0.

For TruD reactions time course by PatA digestions, the amount of core1*-RSIII-RSII (heterocyclized cassette-1) and core2*-RSIII-RSII was compared. Core1*-RSIII-RSII is digestion product of TruE-cassette1*-cassette2 and TruE-cassette1*-cassette2*, while core2*-RSIII-RSII comes from TruE-cassette1-

cassette2* and TruE-cassette1*-cassette2*. As a result, the comparison between these two fragments represents the comparison of TruE-cassette1*-cassette2 and TruE-cassette1-cassette2*.

For PatA time course on heterocyclized precursors, masses corresponding to core1*-RSIII-RSII, core2*-RSIII-RSII, core1*-RSIII-RSII-core2*-RSIII-RSII, leader-core1*-RSIII-RSII were extracted. The amount of core1*-RSIII-RSII and core2*-RSIII-RSII was also compared. If PatA follows strict N-to-C directionality, then only core1*-RSIII-RSII-core2*-RSIII-RSII would be detected and the amount of core1*-RSIII-RSII should equal to core2*-RSIII-RSII. If PatA follows strict C-to-N directionality, then only leader-core1*-RSIII-RSII would be detected and before full digestion, the amount of core1*-RSIII-RSII is lower than core2*-RSIII-RSII.

For PatG time course, the production of RSIII-RSII and RSIII was monitored. Since there were no linear proteolysis products and only the cyclic products were observed, the detection of RSIII-RSII and RSIII is sufficient to show the enzyme reaction process. The cyclic products were not quantified due to inconsistent ionization. For analysis, the ratio of RSIII-RSII over RSIII was used. A ratio over 1 indicates more consumption of cassette-1 and a ratio lower than 1 indicates less consumption of cassette-1. For sets with only cassette-1 or cassette-2 pieces, since both cassettes produce the same RSIII-RSII or RSIII piece, the substrate was monitored instead. A ratio over 1 indicates more leftover of cassette-1 thus less favorable of cassette-1 and ratio lower than 1 indicates more use of cassette-1. For TruG time course, amount of the substrate was monitored throughout the sets since the reaction was diluted and the RSIII-RSII and RSIII were hard to detect.

For PatA time points on unheterocyclized precursor peptides set, the samples were run by the MS core in University of Utah, with a C8 column (SunFire$^{TM}$ C18 5 µm 3.0 x 50 mm) over the UPLC-MS (Agilent 1290 Infinity II LC; Agilent 6540 Accurate-Mass Q-TOF) with a 0.1 mL/min flow rate. The following gradient was used: **1)** 0-0.5 min, 20% B; **2)** 0.5-3 min, a linear gradient from 20% B to 75% B; **3)** 3-3.2 min, 75% to 99% B; **4)** 3.2-4 min, a linear gradient from 99% to 95% B; **5)** 4-4.5 min, 95% to 40% B. The peak areas of extracted ion chromatogram were quantified with MassHunter Quantitative Analysis Software (v B.07.00) and plotted with Prism GraphPad 7.0. To normalize, extracted peak areas were multiplied by normalization factor (e8 standard area from one sample over the average of e8 standard area with all the samples). The expected *m/z* is shown in **Table S8**.

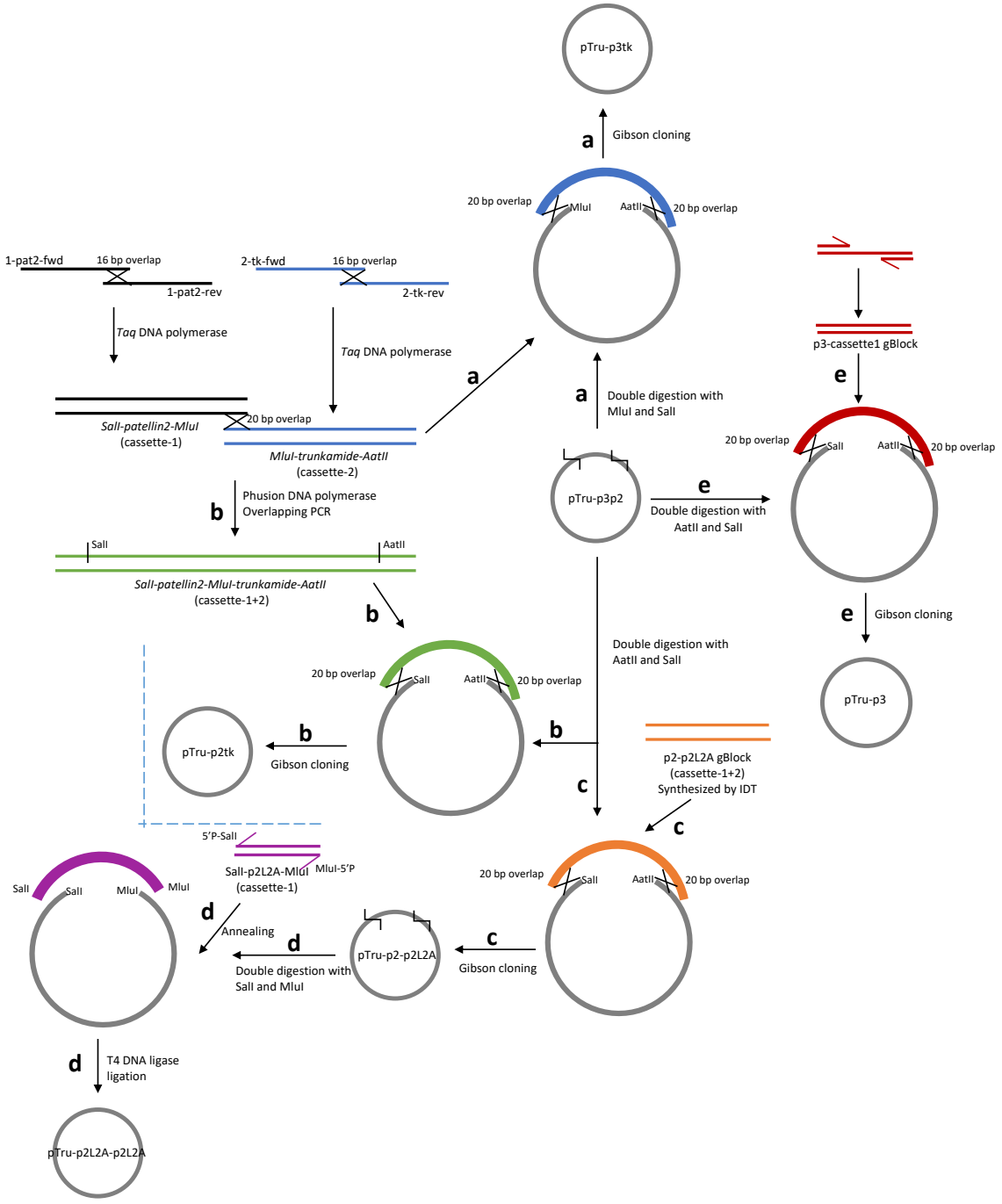
***Hypothesis testing of TruG and PatG reactions***

For TruG and PatG sets, the ratio of initial substrate area was used in each time course to show the consumption of the two cassettes. *P*-value for each time point (**Table S11** and **S12**) was calculated with the one sample *t*-test using the following equation in excel. The two-sided confidence interval was set to 95%, the hypothesized value was set to 1 and the degree of freedom was 2.

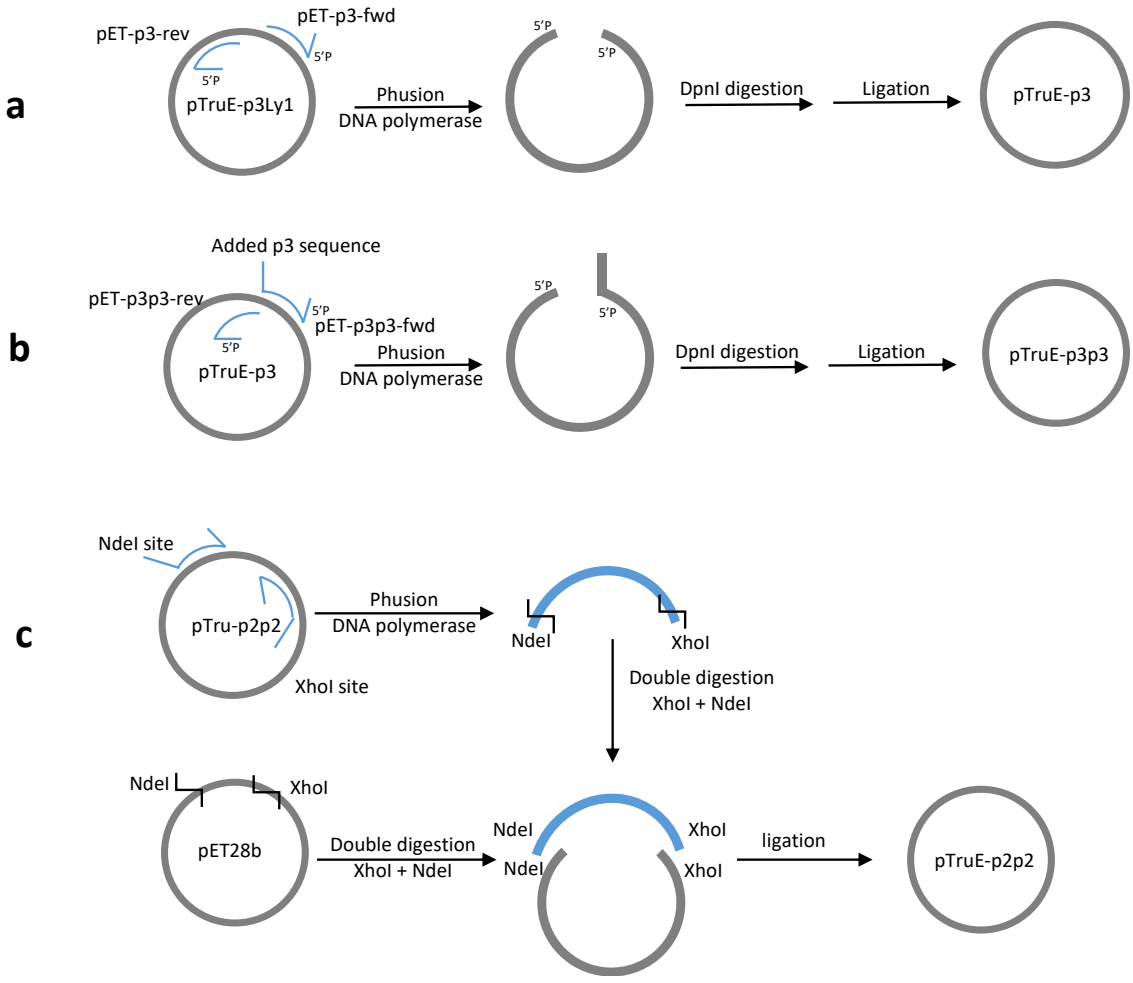
$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \qquad (1)$$

## Supporting Schemes

**Scheme S1.** Cloning strategies for pTru-p3p2 derived plasmids. **(a)** strategy to change the cassette-2 sequence; **(b)** strategy to change the cassette-1 and -2 sequence; **(c)** method to clone patellin2 mutant plasmids in cassette-2; (**d**) method to clone patellin2 mutant plasmids in both cassettes; **(e)** method to clone single cassette plasmids.



**Scheme S2.** Cloning strategies for pTruE-p3p2 derived plasmids. **(a)** Cloning of pTruE-p3. **(b)** Cloning of pTruE-p3p3 and pTruE-p2p3, pTruE-p23. **(c)** Cloning of pTruE-p2p2.



## Supporting Tables

**Table S1.** Primers and gBlocks used in this study

| Primer name | Sequence | Description |
|---|---|---|
| pUC-NsiI-fwd | CTATGACCATGcatACGCCAAGC | For amplifying the backbone of pUC19 and adding an NsiI restriction site (ATGCAT) at the lacZ-alpha start codon. |
| pUC-NsiI-rev | GCTTGGCGTATGCATGGTCATAG | For amplifying the backbone of pUC19 and adding an NsiI restriction site (ATGCAT) at the lacZ-alpha start codon. |
| pUC-AatII-fwd | GTTTCTTAGACGACAGGTGGCAC | For amplifying pUC19 and removing an AatII site after lacZ-alpha |
| pUC-AatII-rev | GTGCCACCTGTCGTCTAAGAAAC | For amplifying pUC19 and removing an AatII site after lacZ-alpha |
| 1-pat2-fwd | CGAGGAAGCATTGGGTGGTGTCG ACGCCTCGACCGTGCCGACCT | To make *SalI-patellin2-MluI* to put into *tru* pathway; encoding first part of the cassette-1 with p2 sequence |
| 1-pat2-rev | AGACGCGTCAACGCCATCATAGCT GCACAAGGTCGGCACGGTCGA | To make *SalI-patellin2-MluI* to put into *tru* pathway; encoding second part of the cassette-1 with p2 sequence |
| 1-tk-fwd | CGAGGAAGCATTGGGTGGTGTCG ACGCCTCGACCAGCATTGCTCC | To make *SalI-trunkamide-MluI* to put into *tru* pathway; encoding first part of the cassette-1 with tk sequence |
| 1-tk-rev | AGACGCGTCAACGCCATCATAGCT GCAAAACGGAGCAATGCTGGTC | To make *SalI-trunkamide-MluI* to put into *tru* pathway; encoding second part of the cassette-1 with tk sequence |
| 1-pat3-fwd | CGAGGAAGCATTGGGTGGTGTCG ACGCCTCGACCCTGCCGGTGCCG AC | To make *SalI-patellin3-MluI* to put into *tru* pathway; encoding first part of the cassette-1 with p3 sequence |
| 1-pat3-rev | AGACGCGTCAACGCCATCATAGCT GCAGAGGGTCGGCACCGGCAG | To make *SalI-patellin3-MluI* to put into *tru* pathway; encoding second part of the cassette-1 with p3 sequence |
| 2-pat2-fwd | AGCTATGATGGCGTTGACGCGTCT ACCGTGCCGACCTTGTGCTCCTAT GACGAC | To make *MluI-patellin2-AatII* to put into *tru* pathway; encoding first part of the cassette-2 with p2 sequence |
| 2-pat2-rev | ATGGTCCCGGCGTGGACTTGACGT CGGTCGCTAGATTAGTCGTCATAG GAGCACA | To make *MluI-patellin2-AatII* to put into *tru* pathway; encoding second part of the cassette-2 with p2 sequence |
| 2-tk-twd | AGCTATGATGGCGTTGACGCGTCT ACCAGCATTGCTCCGTTTTGCTCC TATGACG | To make *MluI-trunkamide-AatII* to put into *tru* pathway; encoding first part of the cassette-2 with tk sequence |
| 2-tk-rev | ATGGTCCCGGCGTGGACTTGACGT CGGTCGCTAGATTAGTCGTCATAG GAGCAAAACG | To make *MluI-trunkamide-AatII* to put into *tru* pathway; encoding second part of the cassette-2 with tk sequence |
| 2-pat3-fwd | AGCTATGATGGCGTTGACGCGTCT ACCCTGCCGGTGCCGACCCTCTGC | To make *MluI-patellin3-AatII* to put into *tru* pathway; encoding first part of the cassette-2 with p3 sequence |

| | | |
|---|---|---|
| | TCCTATG | |
| 2-pat3-rev | ATGGTCCCGGCGTGGACTTGACGT CGGTCGCTAGATTAGTCGTCATAG GAGCAGAGGGT | To make *MluI-patellin3-AatII* to put into *tru* pathway; encoding second part of the cassette-2 with p3 sequence |
| p2-p2L2A gBlock | CCTGCCGGTGCCGACCCTCTGCAG CTATGATGGCGTTGACGCGTCTAC CGTGCCGACCGCCTGCTCCTATGA CGACTAATCTAGCGACCGACGTCA AGTCCACGCCGGGACCATTTCAAT TTTGCACGGAGCCG | To make pTru-p2-p2L2A plasmid |
| p2-p2P4A gBlock | CCTGCCGGTGCCGACCCTCTGCAG CTATGATGGCGTTGACGCGTCTAC CGTGgccACCTTGTGCTCCTATGAC GACTAATCTAGCGACCGACGTCAA GTCCACGCCGGGACCATTTCAATT TTGCACGGAGCCG | To make pTru-p2-p2P4A plasmid |
| p2L2A-cassette1-fwd | 5'P-tcgacgcctcgACCGTGCCGACCGCCTG Cagctatgatggcgttga | To make *SalI- p2L2A-MluI* for pTru-p2L2A-p2L2A plasmid; forward primer |
| p2L2A-cassette1-rev | 5'P-CGCGTCAACGCCATCATAGCTGCA GGCGGTCGGCACGGTCGAGGCG | To make *SalI-p2L2A-MluI* for pTru-p2L2A-p2L2A plasmid; reverse primer |
| p2P4A-cassette1-fwd | 5'P-tcgacgcctcgACCGTGGCCACCTTGTG Cagctatgatggcgttga | To make *SalI-p2P4A-MluI* fragment for pTru-p2P4A-p2P4A plasmid; forward primer |
| p2P4A-cassette1-rev | 5'P-CGCGTCAACGCCATCATAGCTGCA CAAGGTGGCCACGGTCGAGGCG | To make *SalI-p2P4A-MluI* fragment for pTru-p2P4A-p2P4A plasmid; reverse primer |
| pET-p3-fwd | 5'-P-AGCTACGATGATTAAGGATCCG | To make pTruE-p3 from pTruE-TruLy1; forward primer |
| pET-p3-rev | 5'-P-GCAGAGGGTAGGAACAGGTAAG | To make pTruE-p3 from pTruE-TruLy1; reverse primer |
| pET_p23-fwd | 5'-P-ACCGTGCCGACCTTGTGCACCTTA CCTGTTCCTACCCTCTGC | To make pTruE-p23 from pTruE-p3; forward primer |
| pET_p2p3-fwd | 5'-P-TCTTACGACGGCGTGGATGCCTCG ACCTTACCTGTTCCTACCCTCTGC | To make pTruE-p2p3 from pTruE-p23; forward primer |
| pET_p2p3-rev | 5'-P-GCACAAGGTCGGCACGGTCG | To make pTruE-p2p3 from pTruE-p23; reverse primer |

| | | |
|---|---|---|
| pET_p3p3-fwd | 5'-P-ACCTTACCTGTTCCTACCCTCTGCAGCTACGATGATTAAGGATCCGAATTC | To make pTruE-p3p3 from pTruE-p3p2; forward primer |
| pET_p3p3-rev | 5'-P-CGAGGCATCCACGCCGTCG | To make pTruE-p3p3 from pTruE-p3p2; reverse primer |
| truEp2p2-NdeI-fwd | gcggcagccatATGAATAAAAAGAATATTCTCCC | To make *truE-p2p2* from pTruE-p2p2 with NdeI restriction site; forward primer |
| truEp2p2-XhoI-rev | cagtcagtctcgagTTA GTC GTC ATA GGA GCA CA | To make *truE-p2p2* from pTruE-p2p2 with XhoI restriction site; reverse primer |
| cassette p3_fwd | gtccgaggaagcattgggtggtgtcgacgcctcgaccctgccggtgccgaccctctgctcctatgacgactaatctagcgaccgacgtcaagtccacgccgggaccatt | Template to make the single cassette p3 gBlock for pTru-p3 |
| cassette p2_fwd | gtccgaggaagcattgggtggtgtcgacgcctcgaccgtgccgaccttgtgctcctatgacgactaatctagcgaccgacgtcaagtccacgccgggaccatt | Template to make the single cassette p2 gBlock for pTru-p2 |
| cassette tk_fwd | gtccgaggaagcattgggtggtgtcgacgcctcgACCTCTATTGCTCCCTTCTGCtcctatgacgactaatctagcgaccgacgtcaagtccacgccgggaccatt | Template to make the single cassette tk gBlock for pTru-tk |
| cassette1_fwd | gtccgaggaagcattggg | To make the single cassette gBlock; forward primer |
| cassette1_rev | AAT GGT CCC GGC GTG G | To make the single cassette gBlock; reverse primer |

**Table S2.** Plasmids used in this study

| **Plasmid name** | **backbone** | **pathway/precursor/enzyme** | **cassette-1 encoding** | **cassette-2 encoding** |
|---|---|---|---|---|
| pTru-p3p2 | pUC19 | *tru* pathway | p3 (TLPVPTLC) | p2 (TVPTLC) |
| pTru-p2p3 | pUC19 | *tru* pathway | p2 (TVPTLC) | p3 (TLPVPTLC) |
| pTru-p3p3 | pUC19 | *tru* pathway | p3 (TLPVPTLC) | p3 (TLPVPTLC) |
| pTru-p2p2 | pUC19 | *tru* pathway | p2 (TVPTLC) | p2 (TVPTLC) |
| pTru-tkp2 | pUC19 | *tru* pathway | tk (TSIAPFC) | p2 (TVPTLC) |
| pTru-p2tk | pUC19 | *tru* pathway | p2 (TVPTLC) | tk (TSIAPFC) |
| pTru-tktk | pUC19 | *tru* pathway | tk (TSIAPFC) | tk (TSIAPFC) |
| pTru-p3tk | pUC19 | *tru* pathway | p3 (TLPVPTLC) | tk (TSIAPFC) |
| pTru-tkp3 | pUC19 | *tru* pathway | tk (TSIAPFC) | p3 (TLPVPTLC) |
| pTru-p3 | pUC19 | *tru* pathway | p3 (TLPVPTLC) | N/A |
| pTru-p2 | pUC19 | *tru* pathway | p2 (TVPTLC) | N/A |
| pTru-tk | pUC19 | *tru* pathway | tk (TSIAPFC) | N/A |
| pTru-p2-p2L2A | pUC19 | *tru* pathway | p2 (TVPTLC) | p2L2A (TVPTAC) |
| pTru-p2L2A- p2L2A | pUC19 | *tru* pathway | p2L2A (TVPTAC) | p2L2A (TVPTAC) |

| | | | | |
|---|---|---|---|---|
| pTru-p2-p2P4A | pUC19 | *tru* pathway | p2 (TVPTLC) | p2P4A (TVATLC) |
| pTru-p2P4A- p2P4A | pUC19 | *tru* pathway | p2P4A (TVATLC) | p2P4A (TVATLC) |
| pTruE-p3p2 | pET28b(+) | precursor | p3 (TLPVPTLC) | p2 (TVPTLC) |
| pTruE-p3 | pET28b(+) | precursor | p3 (TLPVPTLC) | N/A |
| pTruE-p2p3 | pET28b(+) | precursor | p2 (TVPTLC) | p3 (TLPVPTLC) |
| pTruE-p3p3 | pET28b(+) | precursor | p3 (TLPVPTLC) | p3 (TLPVPTLC) |
| pTruE-p2p2 | pET28b(+) | precursor | p2 (TVPTLC) | p2 (TVPTLC) |
| pTruE-p23 | pET28b(+) | precursor | p23 (TVPTLCTLPVPTLC) | N/A |
| pET28b-his-TruD | pET28b(+) | TruD enzyme | N/A | N/A |
| pET28b-his-RSI-TruD | pET28b(+) | RSI-TruD enzyme | N/A | N/A |
| pET28b-his-PatA | pET28b(+) | PatA enzyme | N/A | N/A |
| pET28b-his-PatG | pET28b(+) | PatG enzyme | N/A | N/A |
| pET28b-his-TruG | pET28b(+) | TruG enzyme | N/A | N/A |

**Table S3.** Peptides used in this study

| **Peptide name** | **Peptide sequence** | **source** |
|---|---|---|
| TruE-p3p2 | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSS QLAELSEEALGGVDAS**TLPVPTLC**SYDGVDAS**TVPTLC**SYDD | Protein expression and purification |
| TruE-p2p3 | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSS QLAELSEEALGGVDAS**TVPTLC**SYDGVDAS**TLPVPTLC**SYDD | Protein expression and purification |
| TruE-p3p3 | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSS QLAELSEEALGGVDAS**TLPVPTLC**SYDGVDAS**TLPVPTLC**SY DD | Protein expression and purification |
| TruE-p2p2 | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSS QLAELSEEALGGVDAS**TVPTLC**SYDGVDAS**TVPTLC**SYDD | Protein expression and purification |
| p2-RSIII-RSII | **TVPTLC**SYDGVDAS | From Genscript |
| p3-RSIII-RSII | **TLPVPTLC**SYDGVDAS | From Genscript |
| p2-RSIII | **TVPTLC**SYDD | From Genscript |
| p3-RSIII | **TLPVPTLC**SYDD | From Genscript |
| RSI | acetyl-LAELSEEALG-amide | From Genscript |
| RSIII-RSII | SYDGVDAS | From Genscript |
| RSIII | SYDD | From Genscript |
| e8 standard[11] | A synthetic cyclic peptide | From peptide core, University of Utah |
| p2*-RSIII-RSII | **TVPTLC***SYDGVDAS | Fully heterocyclized by RSI-TruD |
| p3*-RSIII-RSII | **TLPVPTLC***SYDGVDAS | Fully heterocyclized by RSI-TruD |
| p2*-RSIII | **TVPTLC***SYDD | Fully heterocyclized by RSI-TruD |

| | | |
|---|---|---|
| p3*-RSIII | **TLPVPTLC***SYDD | Fully heterocyclized by RSI-TruD |

Core sequences in bold. * represents heterocyclization.

**Table S4.** Cyanobactin *m/z* from heterologous expression in *E. coli*

| **cyanobactin name** | **Peptide sequence** | **Observed *m/z*** |
|---|---|---|
| trunkamide (**1**) | c-[TSIAPFC*] | $[M+H]^+$ = 838.1 |
| **1** – (single prenylation) | c-[TSIAPFC*] or c-[TSIAPFC*] | $[M+H]^+$ = 770.3 |
| **1** – – (non-prenylated product) | c-[TSIAPFC*] | $[M+H]^+$ = 702.5 |
| patellin 2 (**2**) | c-[TVPTLC*] | $[M+H]^+$ = 733.5 |
| **2** – | c-[TVPTLC*] or c-[TVPTLC*] | $[M+H]^+$ = 665.4 |
| **2** – – | c-[TVPTLC*] | $[M+H]^+$ = 597.3 |
| patellin 3 (**3**) | c-[TLPVPTLC*] | $[M+H]^+$ = 943.5 |
| **3** – | c-[TLPVPTLC*] or c-[TLPVPTLC*] | $[M+H]^+$ = 875.4 |
| **3** – – | c-[TLPVPTLC*] | $[M+H]^+$ = 807.2 |
| p2L2A (**4**) | c-[TVPTAC*] | $[M+H]^+$ = 691.8 |
| **4** – | c-[TVPTAC*] or c-[TVPTAC*] | $[M+H]^+$ = 623.5 |
| **4** – – | c-[TVPTAC*] | $[M+H]^+$ = 555.6 |
| p2P4A (**5**) | c-[TVATLC*] | $[M+H]^+$ = 706.5 |
| **5** – | c-[TVATLC*] or c-[TVATLC*] | $[M+H]^+$ = 638.4 |
| **5** – – | c-[TVATLC*] | $[M+H]^+$ = 570.3 |

*represents heterocyclization; c-[X] represents N-to-C terminal cyclization. Underscore (X) represents prenylation.

**Table S5.** TruD heterocyclization of peptides and the observed *m/z*

| **Peptide name** | **Peptide sequence** | **Observed *m/z*** |
|---|---|---|
| TruE-p3p2 (substrate) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC**SYDGVDAS**TVPTLC**SYDD | $[M+8H]^{8+}$ = 1116.5 |
| TruE-p3*p2 and TruE-p3p2* (intermediates) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC***SYDGVDAS**TVPTLC**SYDD and GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC**SYDGVDAS**TVPTLC***SYDD | $[M+8H]^{8+}$ = 1114.5 |
| TruE-p3*p2* (product) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC***SYDGVDAS**TVPTLC***SYDD | $[M+8H]^{8+}$ = 1112.0 |
| TruE-p2p3 (substrate) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC**SYDGVDAS**TLPVPTLC**SYDD | $[M+8H]^{8+}$ = 1116.5 |
| TruE-p2*p3 and | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC***SYDGVDAS**TLPVPTLC**SYDD | $[M+8H]^{8+}$ = 1114.5 |

| | | |
|---|---|---|
| TruE-p2p3* (intermediates) | and GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC**SYDGVDAS**TLPVPTLC***SYDD | |
| TruE-p2*p3* (product) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC***SYDGVDAS**TLPVPTLC***SYDD | $[M+8H]^{8+}$ = 1112.0 |
| TruE-p3p3 (substrate) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC**SYDGVDAS**TLPVPTLC**SYDD | $[M+8H]^{8+}$ = 1142.5 |
| TruE-p3*p3 and TruE-p3p3* (intermediates) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC***SYDGVDAS**TLPVPTLC**SYDD and GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC**SYDGVDAS**TLPVPTLC***SYDD | $[M+8H]^{8+}$ = 1140.5 |
| TruE-p3*p3* (product) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TLPVPTLC***SYDGVDAS**TLPVPTLC***SYDD | $[M+8H]^{8+}$ = 1138.0 |
| TruE-p2p2 (substrate) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC**SYDGVDAS**TVPTLC**SYDD | $[M+8H]^{8+}$ = 1090.5 |
| TruE-p2*p2 and TruE-p2p2* (intermediates) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC***SYDGVDAS**TVPTLC**SYDD and GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC**SYDGVDAS**TVPTLC***SYDD | $[M+8H]^{8+}$ = 1088.5 |
| TruE-p2*p2* (product) | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELSEE ALGGVDAS**TVPTLC***SYDGVDAS**TVPTLC***SYDD | $[M+8H]^{8+}$ = 1085.9 |
| e8 (standard) | A synthetic cyclic peptide | $[M+H]^{+}$ = 883.5 |

*represents heterocyclization

**Table S6.** PatA digestion peptides of TruD reactions and the observed *m/z*

| **Peptide name** | **Peptide sequence** | **Observed *m/z*** |
|---|---|---|
| leader | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLAELS EEALGGVDAS | $[M+7H]^{7+}$ = 887.9 |
| p2-RSIII-RSII | **TVPTLC**SYDGVDAS | $[M+2H]^{2+}$ =714.3 |
| p3-RSIII-RSII | **TLPVPTLC**SYDGVDAS | $[M+2H]^{2+}$ = 819.4 |
| p2-RSIII | **TVPTLC**SYDD | $[M+H]^{+}$ =1113.5 |
| p3-RSIII | **TLPVPTLC**SYDD | $[M+2H]^{2+}$ =662.3 |
| p2*-RSIII-RSII | **TVPTLC***SYDGVDAS | $[M+2H]^{2+}$ =705.3 |
| p3*-RSIII-RSII | **TLPVPTLC***SYDGVDAS | $[M+2H]^{2+}$ = 810.4 |
| p2*-RSIII | **TVPTLC***SYDD | $[M+2H]^{2+}$ =548.4 |
| p3*-RSIII | **TLPVPTLC***SYDD | $[M+2H]^{2+}$ =653.3 |
| e8 standard | A synthetic cyclic peptide | $[M+H]^{+}$ =883.5 |

*represents heterocyclization

**Table S7.** PatA digestion peptides of heterocyclized peptides and the observed *m/z*

| **Peptide name** | **Peptide sequence** | **Observed *m/z*** |
|---|---|---|
| leader | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLA ELSEEALGGVDAS | $[M+7H]^{7+}$ = 887.9 |
| p2*-RSIII-RSII | **TVPTLC***SYDGVDAS | $[M+2H]^{2+}$ =705.3 |
| p3*-RSIII-RSII | **TLPVPTLC***SYDGVDAS | $[M+2H]^{2+}$ = 810.4 |
| p2*-RSIII | **TVPTLC***SYDD | $[M+2H]^{2+}$ = 548.4 |
| p3*-RSIII | **TLPVPTLC***SYDD | $[M+2H]^{2+}$ = 653.3 |
| p3*-RSIII-RSII-p2*-RSIII | **TLPVPTLC***SYDGVDAS**TVPTLC***SYDD | $[M+3H]^{3+}$ = 899.4 |
| p2*-RSIII-RSII-p3*-RSIII | **TVPTLC***SYDGVDAS**TLPVPTLC***SYDD | $[M+3H]^{3+}$ = 899.4 |
| p3*-RSIII-RSII-p3*-RSIII | **TLPVPTLC***SYDGVDAS**TLPVPTLC***SYDD | $[M+3H]^{3+}$ = 969.4 |
| p2*-RSIII-RSII-p2*-RSIII | **TVPTLC***SYDGVDAS**TVPTLC***SYDD | $[M+3H]^{3+}$ = 830.0 |
| leader-RSII-p3*-RSIII-RSII | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLA ELSEEALGGVDAS**TLPVPTLC***SYDGVDAS | $[M+8H]^{7+}$ = 976.8 |
| leader-RSII-p2*-RSIII-RSII | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQLA ELSEEALGGVDAS**TVPTLC***SYDGVDAS | $[M+8H]^{8+}$ = 951.8 |
| e8 standard | A synthetic cyclic peptide | $[M+H]^{+}$ = 883.5 |

*represents heterocyclization

**Table S8.** PatA digestion peptides of intact precursor peptide and the observed *m/z*

| **Peptide name** | **Peptide sequence** | **Observed *m/z*** |
|---|---|---|
| p2-RSIII-RSII | **TVPTLC**SYDGVDAS | $[M+2H]^{2+}$ = 714.3 |
| p3-RSIII-RSII | **TLPVPTLC**SYDGVDAS | $[M+2H]^{2+}$ = 819.4 |
| p2-RSIII | **TVPTLC**SYDD | $[M+H]^{+}$ = 1113.5 |
| p3-RSIII | **TLPVPTLC**SYDD | $[M+2H]^{2+}$ = 662.3 |
| p3-RSIII-RSII-p2-RSIII | **TLPVPTLC**SYDGVDAS**TVPTLC**SYDD | $[M+3H]^{3+}$ =911.4 |
| p2-RSIII-RSII-p3-RSIII | **TVPTLC**SYDGVDAS**TLPVPTLC**SYDD | $[M+3H]^{3+}$ =911.4 |
| p3-RSIII-RSII-p3-RSIII | **TLPVPTLC**SYDGVDAS**TLPVPTLC**SYDD | $[M+3H]^{3+}$ =981.5 |
| p2-RSIII-RSII-p2-RSIII | **TVPTLC**SYDGVDAS**TVPTLC**SYDD | $[M+3H]^{3+}$ =841.4 |
| leader | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQL AELSEEALGGVDAS | $[M+7H]^{7+}$ = 887.9 |
| leader-RSII-p3-RSIII-RSII | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQL AELSEEALGGVDAS**TLPVPTLC**SYDGVDAS | $[M+8H]^{7+}$ = 979.1 |
| leader-RSII-p2-RSIII-RSII | GSSHHHHHHSSGLVPRGSHMNKKNILPQLGQPVIRLTAGQLSSQL AELSEEALGGVDAS**TVPTLC**SYDGVDAS | $[M+8H]^{8+}$ = 952.8 |
| e8 standard | A synthetic cyclic peptide | $[M+H]^{+}$ = 883.5 |

**Table S9.** PatG or TruG reaction peptides and the observed *m/z*

| Peptide name | Peptide sequence | Observed *m/z* |
|---|---|---|
| p2*-RSIII-RSII | **TVPTLC***SYDGVDAS | $[M+2H]^{2+}$ =705.3 |
| p3*-RSIII-RSII | **TLPVPTLC***SYDGVDAS | $[M+2H]^{2+}$ = 810.4 |
| p2*-RSIII | **TVPTLC***SYDD | $[M+2H]^{2+}$ = 548.4 |
| p3*-RSIII | **TLPVPTLC***SYDD | $[M+2H]^{2+}$ = 653.3 |
| RSIII-RSII | SYDGVDAS | $[M+H]^{+}$ = 813.5 |
| RSIII | SYDD | $[M+H]^{+}$ = 499.2 |
| **3** – – | c-[**TLPVPTLC***] | $[M+H]^{+}$ =807.5 |
| **2** – – | c-[**TVPTLC***] | $[M+H]^{+}$ =597.5 |
| e8 standard | A synthetic cyclic peptide | $[M+H]^{+}$ = 883.5 |

– – represents non-prenylated product; *represents heterocyclization; c-[X] represents head-to-tail cyclization.

**Table S10.** Standard observed *m/z*

| Peptide name | Peptide sequence | Observed *m/z* |
|---|---|---|
| p2*-RSIII-RSII | **TVPTLC***SYDGVDAS | $[M+2H]^{2+}$ =705.3 |
| p3*-RSIII-RSII | **TLPVPTLC***SYDGVDAS | $[M+2H]^{2+}$ = 810.4 |
| p2*-RSIII | **TVPTLC***SYDD | $[M+2H]^{2+}$ = 548.4 |
| p3*-RSIII | **TLPVPTLC***SYDD | $[M+2H]^{2+}$ = 653.3 |
| p2-RSIII-RSII | **TVPTLC**SYDGVDAS | $[M+2H]^{2+}$ = 714.3 |
| p3-RSIII-RSII | **TLPVPTLC**SYDGVDAS | $[M+2H]^{2+}$ = 819.4 |
| p2-RSIII | **TVPTLC**SYDD | $[M+H]^{+}$ = 1113.5 |
| p3-RSIII | **TLPVPTLC**SYDD | $[M+2H]^{2+}$ = 662.3 |
| e8 standard | A synthetic cyclic peptide | $[M+H]^{+}$ = 883.5 |

**Table S11.** One sample *t*-test *P*-value for TruG reactions

| time (h) | 0 | 3 | 13 | 15 | 28.25 | 37 |
|---|---|---|---|---|---|---|
| **p3p2** | 0.3471 | 0.4650 | 0.2787 | 0.0620 | 0.0333 | 0.0051 |
| **p2p3** | 0.0031 | 0.0066 | 0.0005 | 0.0001 | 0.0025 | 0.0006 |
| **p2p2** | 0.0089 | 0.0101 | 0.0058 | 0.0200 | 0.0024 | 0.0003 |
| **p3p3** | 0.3225 | 0.0154 | 0.0248 | 0.0129 | 0.0094 | 0.1823 |

Significant *P*-values (<0.05) are marked as red.

**Table S12.** One sample *t*-test *P*-value for PatG reactions

| time (h) | 0 | 3 | 10 | 21 | 29 | 53 |
|---|---|---|---|---|---|---|
| **p3p2** | N/A | 0.4767 | 0.1922 | 0.5765 | 0.0582 | 0.2654 |
| **p2p3** | N/A | 0.0003 | 0.0067 | 0.0277 | 0.0016 | 0.1112 |
| **p2p2** | N/A | 0.2442 | 0.4613 | 0.4815 | 0.1678 | 0.1837 |
| **p3p3** | N/A | 0.0076 | 0.0407 | 0.0088 | 0.0207 | 0.1463 |
| **cassette-1** | 0.2036 | 0.0739 | 0.4288 | 0.6504 | N/A | 0.5755 |
| **cassette-2** | 0.0666 | 0.4235 | 0.2627 | 0.2846 | 0.4786 | N/A |

Significant *P*-values (<0.05) are marked as red.

## Supporting Figures

**Figure S1.** Collagen biosynthesis and cyanobactin biosynthesis comparison.

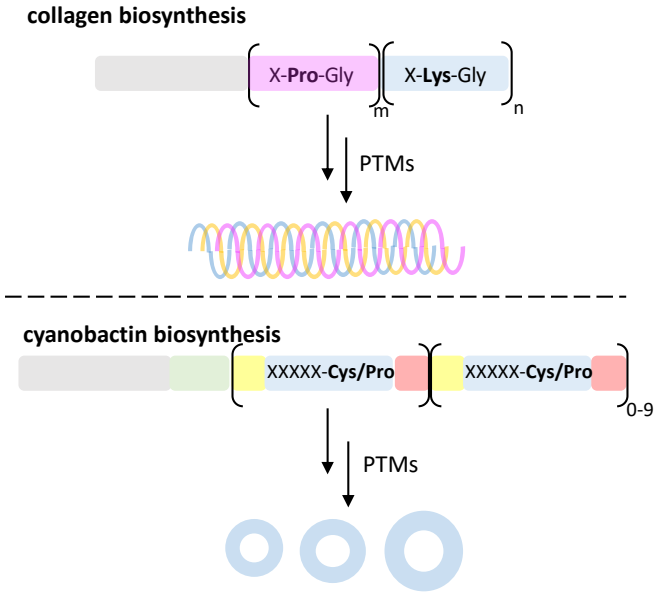
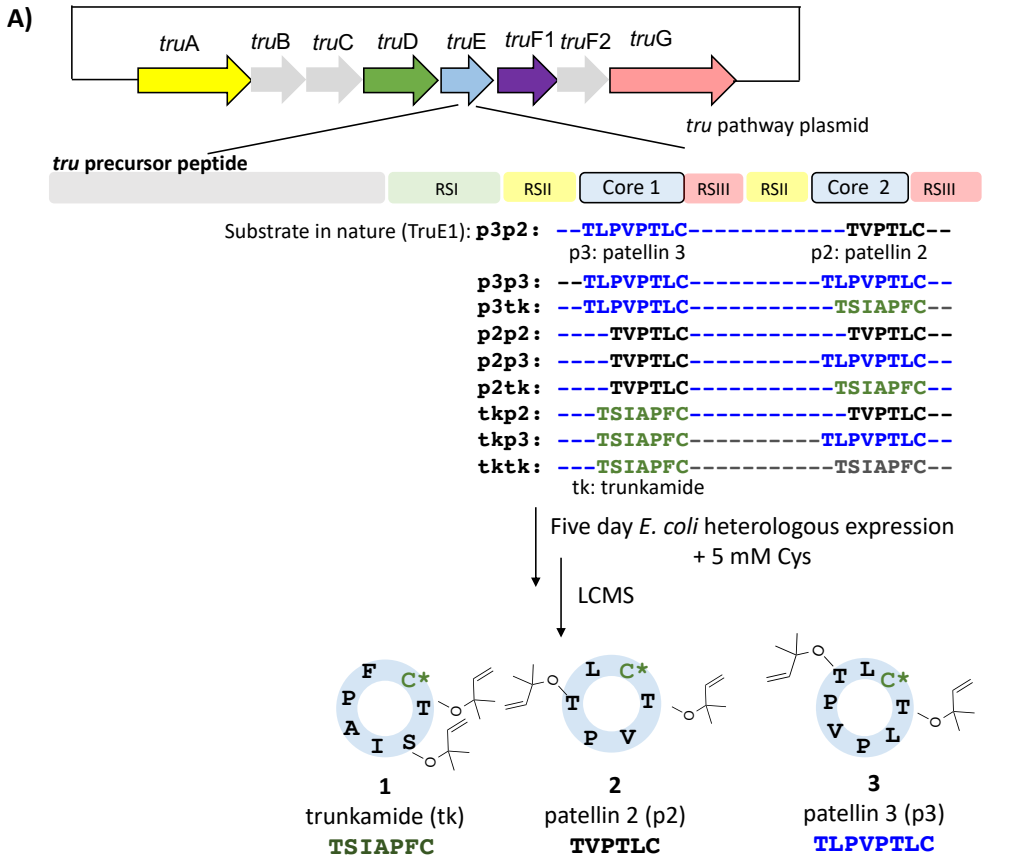


**Figure S2.** *tru* pathway precursor designs. **A)** Heterologous expression design of *tru* pathway in *E. coli*. **B)** Precursor peptides for *in vitro* enzyme assays.



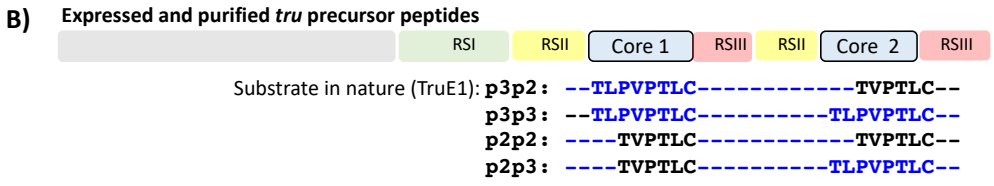



**Figure S3.** Compound **1**, **2**, **3** production in *E. coli*. **(A)** patellin 2 (**2**), single prenylated patellin 2 (**2–**) and non-prenylated patellin 2 (**2– –**) . Production on **B)** day two, **C)** day three, **D)** day four and **E)** day five. Compound yields for different plasmids were quantified by extracted MS area and normalized with internal standard. p2 represents cassette encoding patellin 2, p3 for cassette encoding patellin 3, tk for cassette encoding trunkamide. – – represents no prenylation and – represents single prenylation.

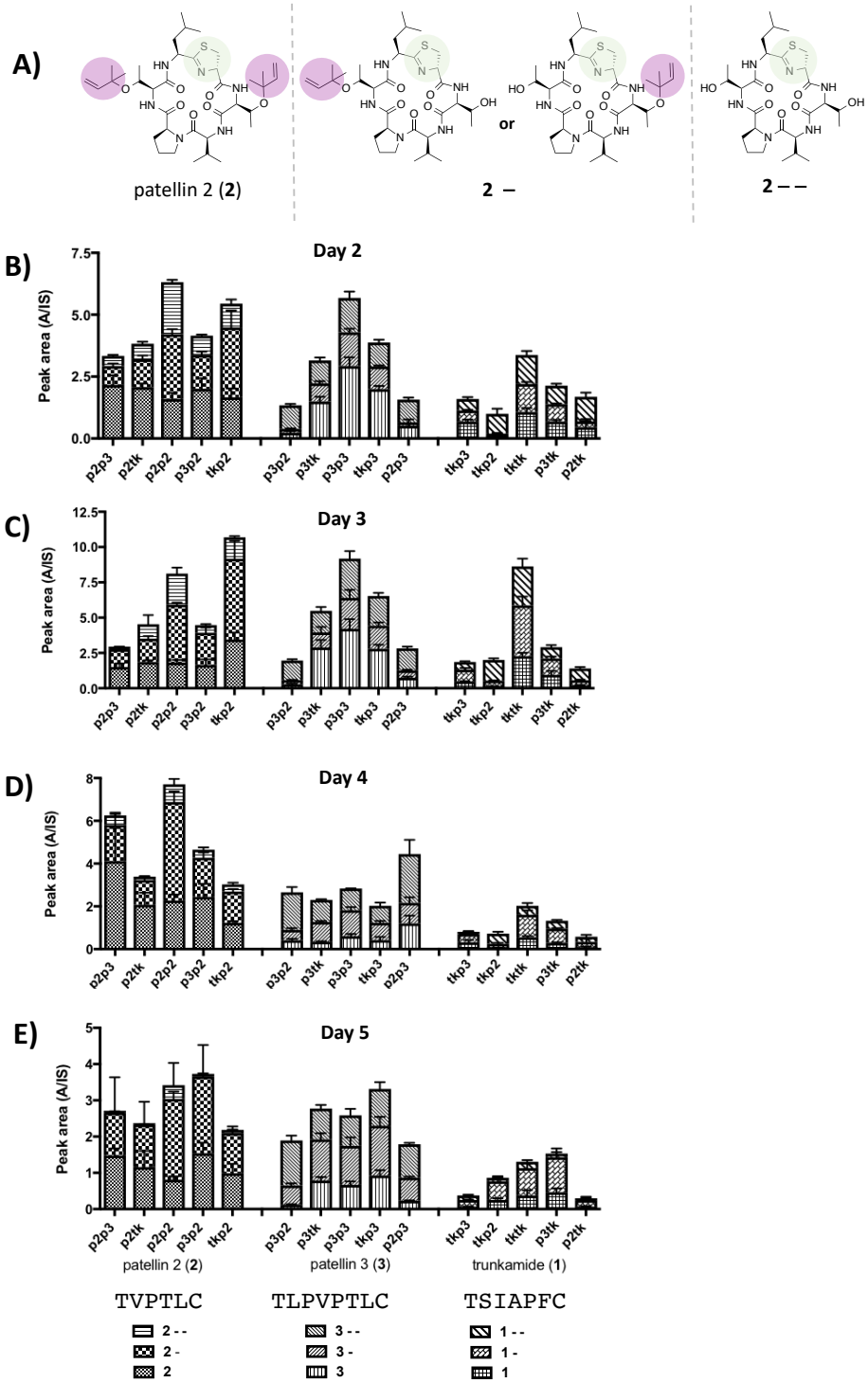


**Figure S4.** Compound production in *E. coli* with single cassette *tru* pathway variants on day three (top) or day 4 (p3 and p3p3 only; bottom).

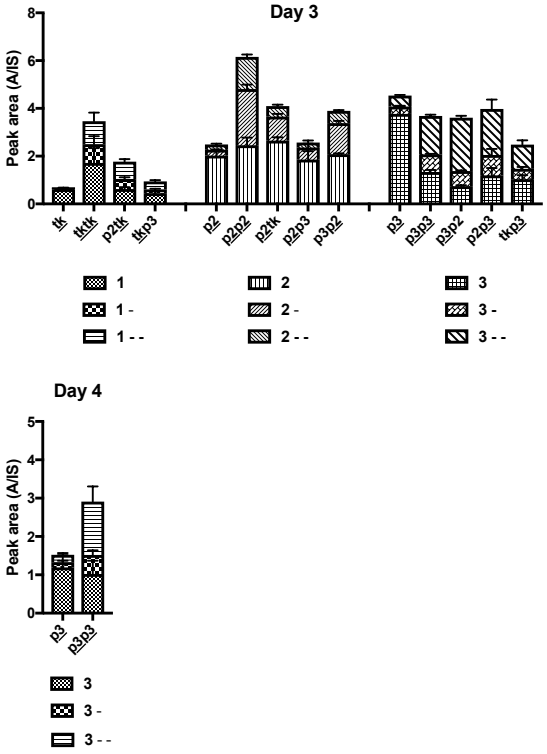


**Figure S5.** Patellin 2 mutant production in *E. coli* on day five. Compound yields for different plasmids were quantified by extracted MS area and normalized with internal standard. p2 represents cassette encoding patellin 2 (TVPTLC), p2L2A for mutant TVPTAC and p2P4A for mutant TVATLC. – – represents no prenylation and – represents single prenylation.

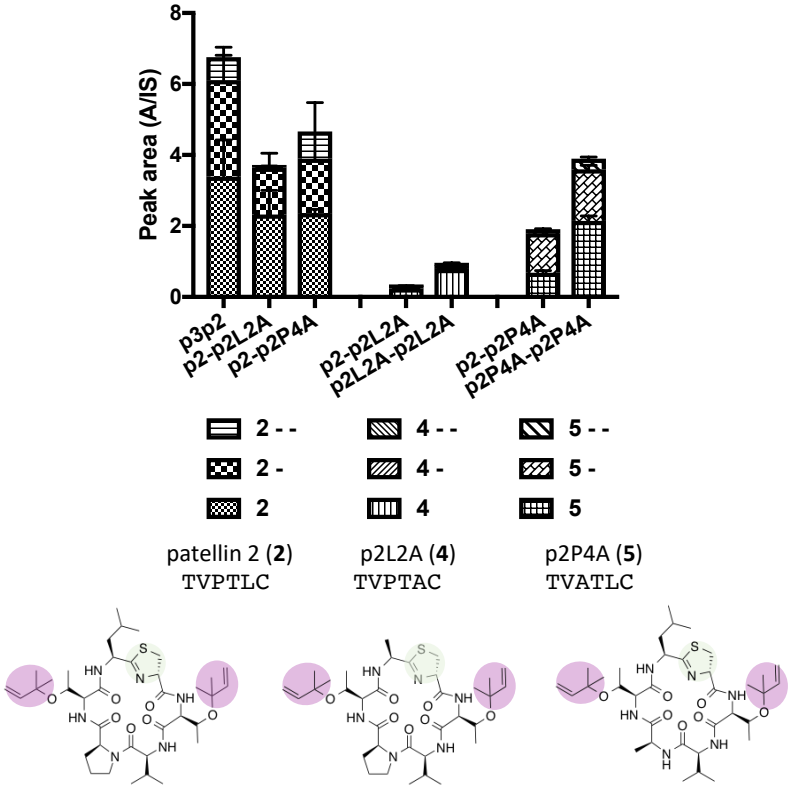


**Figure S6.** MS traces of TruD time course. Time points were taken at 10 s, 1, 5, 10, 20, 30, 40 and 60 min. Substrate (heterocycle # 0), single heterocyclized product (heterocycle # 1) and double heterocyclized product (heterocycle # 2) are marked and aligned: **A)** TruE-p2p2; **B)** TruE-p3p2; **C)** TruE-p2p3; **D)** TruE-p3p3.

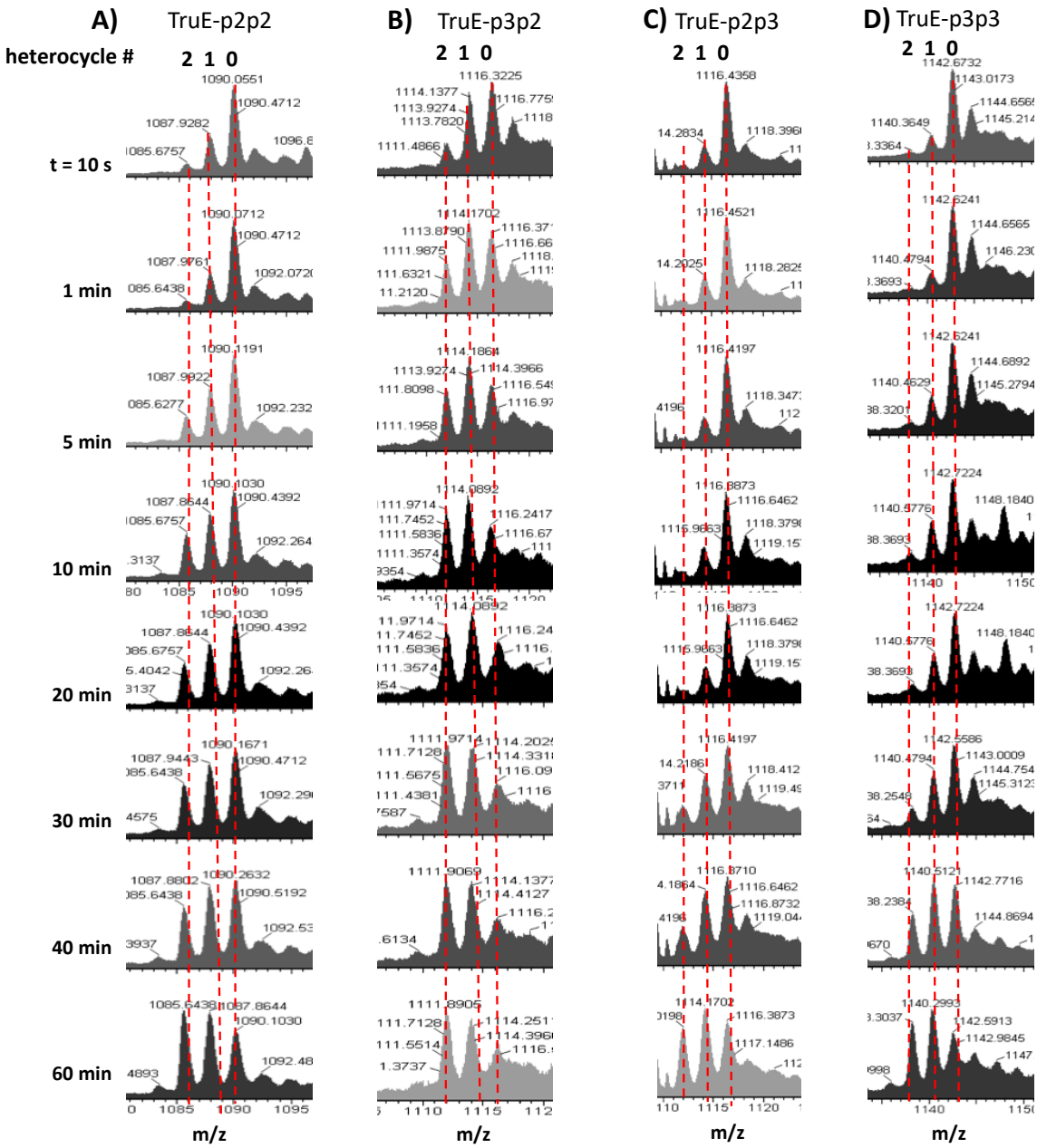


**Figure S7.** MS traces of TruD time course of TruE-p23. **A)** Schematic steps of TruD over TruE-p23 precursor peptide. Two possible routes **a** and **b** are presented. **B)** Time points were taken at 10 s, 5, 10 and 60 min. Substrate, single heterocyclized product and double heterocyclized product are marked and aligned.

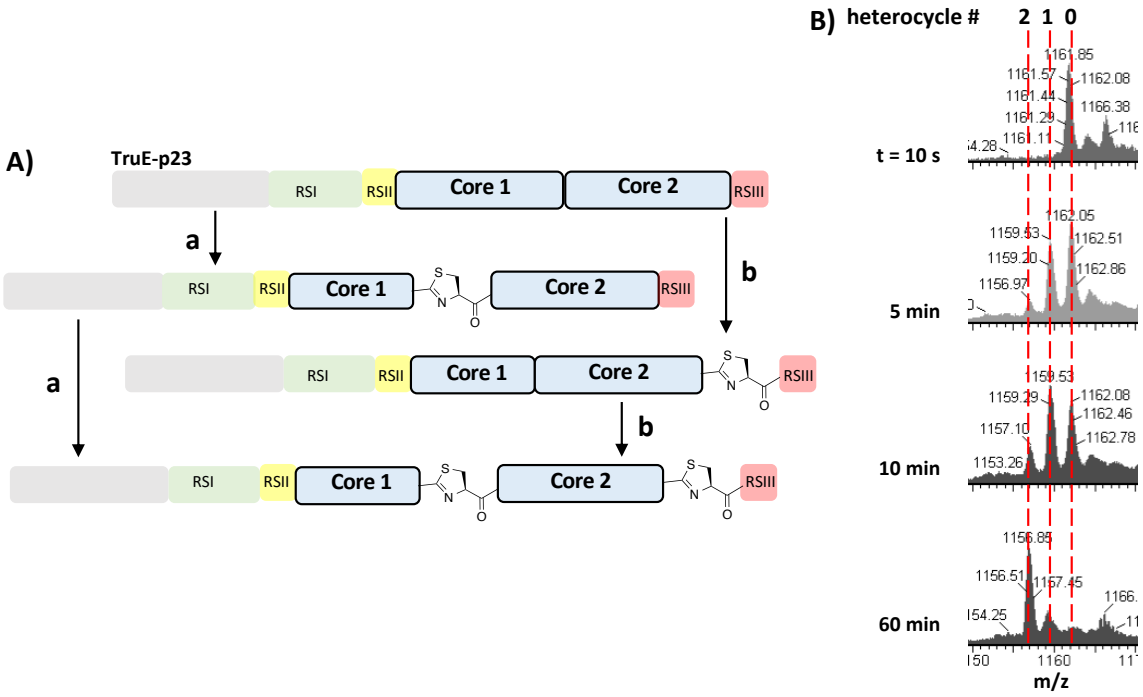
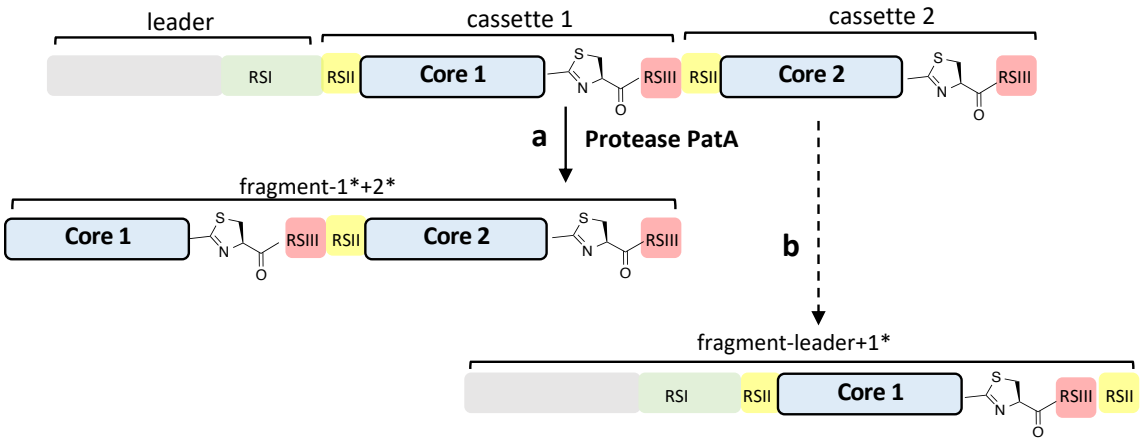


**Figure S8.** PatA proteolysis of TruE-p2*p3*. The proteolysis steps above show the major route **a** and minor route **b**. Lower graph shows MS trace of the proteolysis fragments. Fragment-leader-1* was also observed besides fragment-1*+2*.



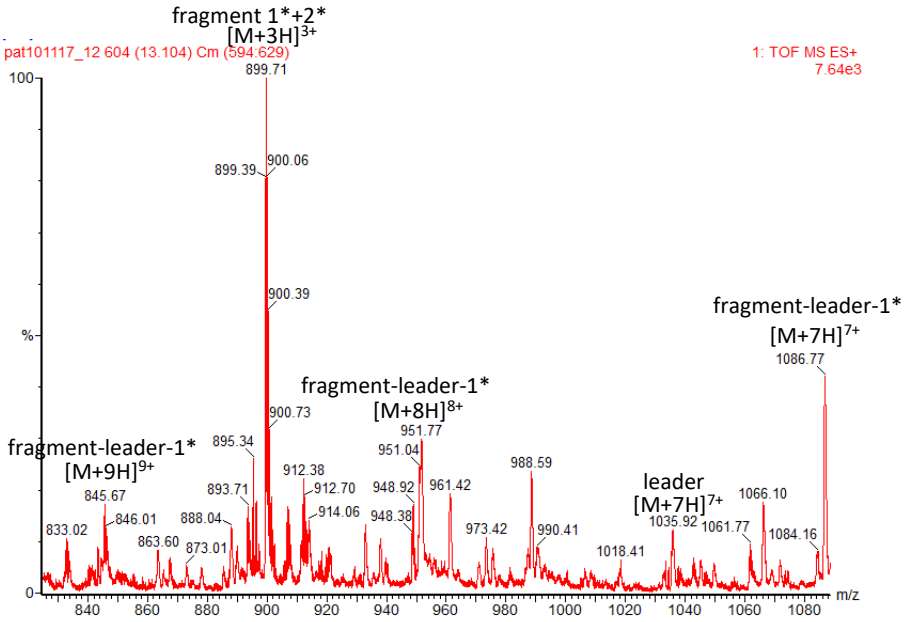



**Figure S9.** PatA proteolysis of the intact precursor peptide. **A)** Proteolysis steps. **B)** TruE-p3p2; **C)** TruE-p2p3; **D)** TruE-p3p3. The production of core1-RSIII-RSII and core2-RSIII-RSII was inhibited by intra disulfide bridge between cassette-1 and cassette-2. Area was normalized with normalization factor (internal standard area from one sample over the average of e8 standard area with all the samples).

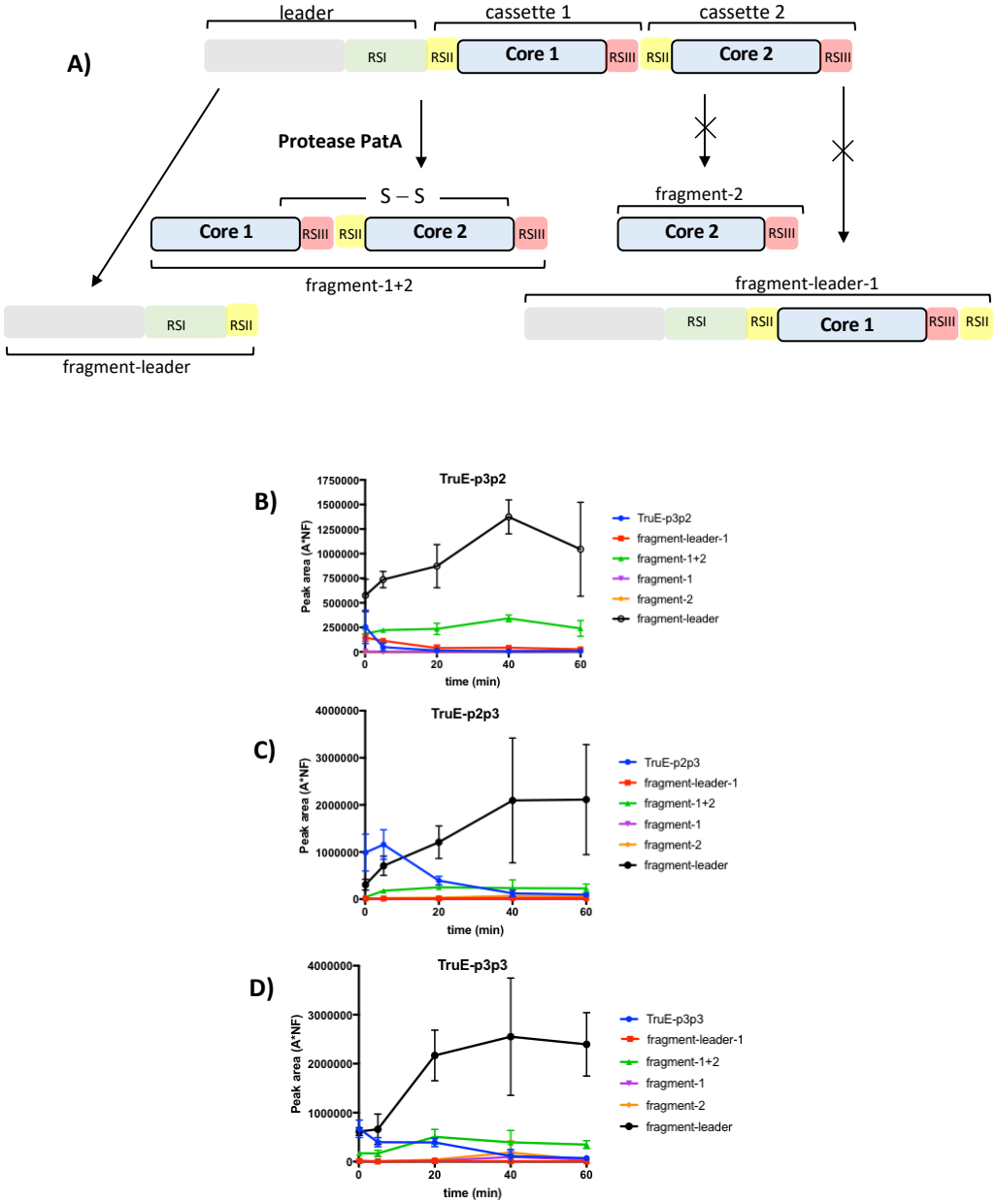


**Figure S10.** TruD reaction time course after PatA treatment. Digested TruE-p2p2 was heterocyclized during time course and almost reach completion after 40 min.

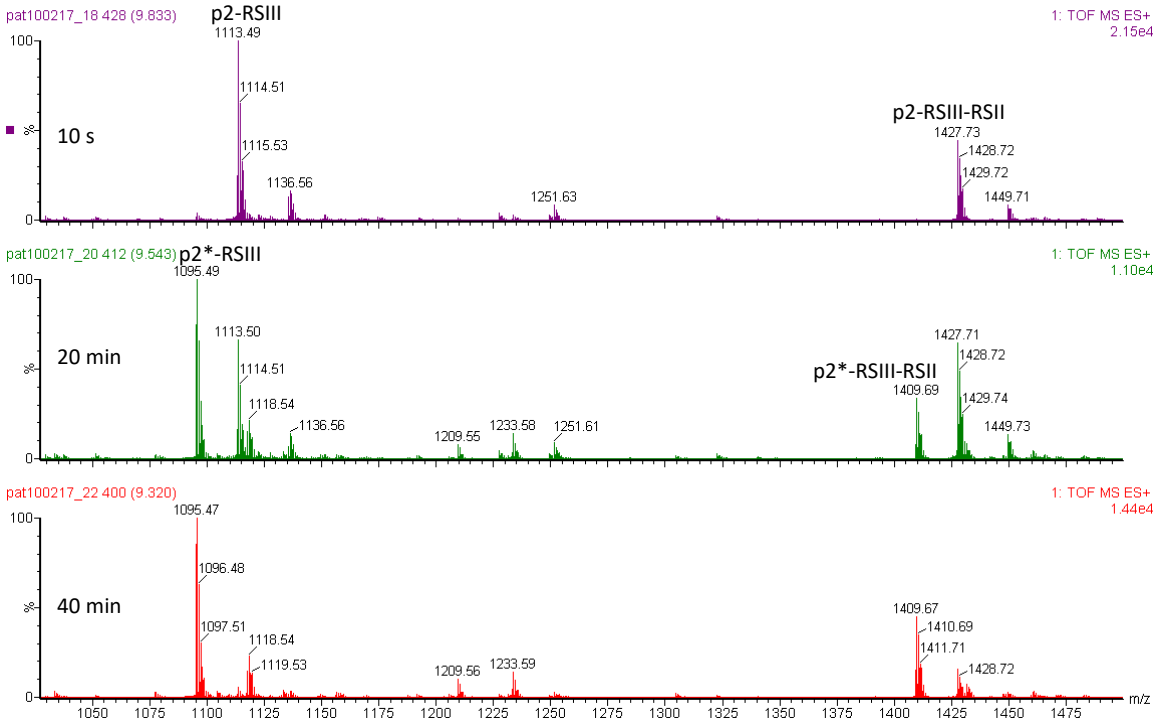


**Figure S11.** TruD reaction with and without RSI. Different combinations (p3p2, p2p3, p2p2, p3p3, cassette-1 and cassette-2) were used. The ratio of heterocyclized and unheterocyclized fragments is compared, which shows the conversion from substrate to product (ionization difference was not normalized). Without RSI, the heterocyclization event was usually slower.

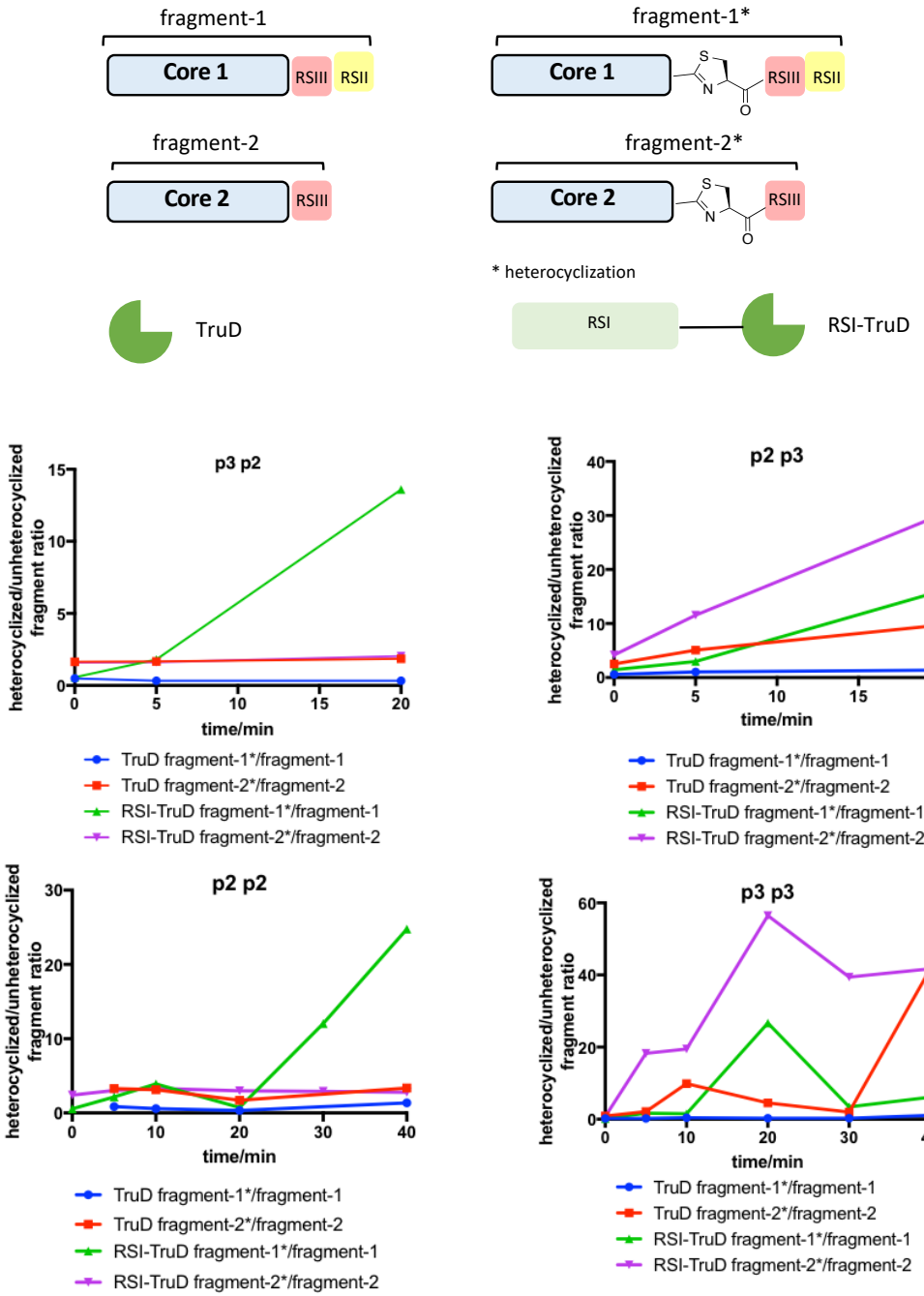


**Figure S12.** TruG reaction time course. Time points were taken at 0, 3, 13, 15, 28.25 and 37 h. Extracted peak area is normalized to internal standards and to four different synthetic peptides to correct for ionization differences.

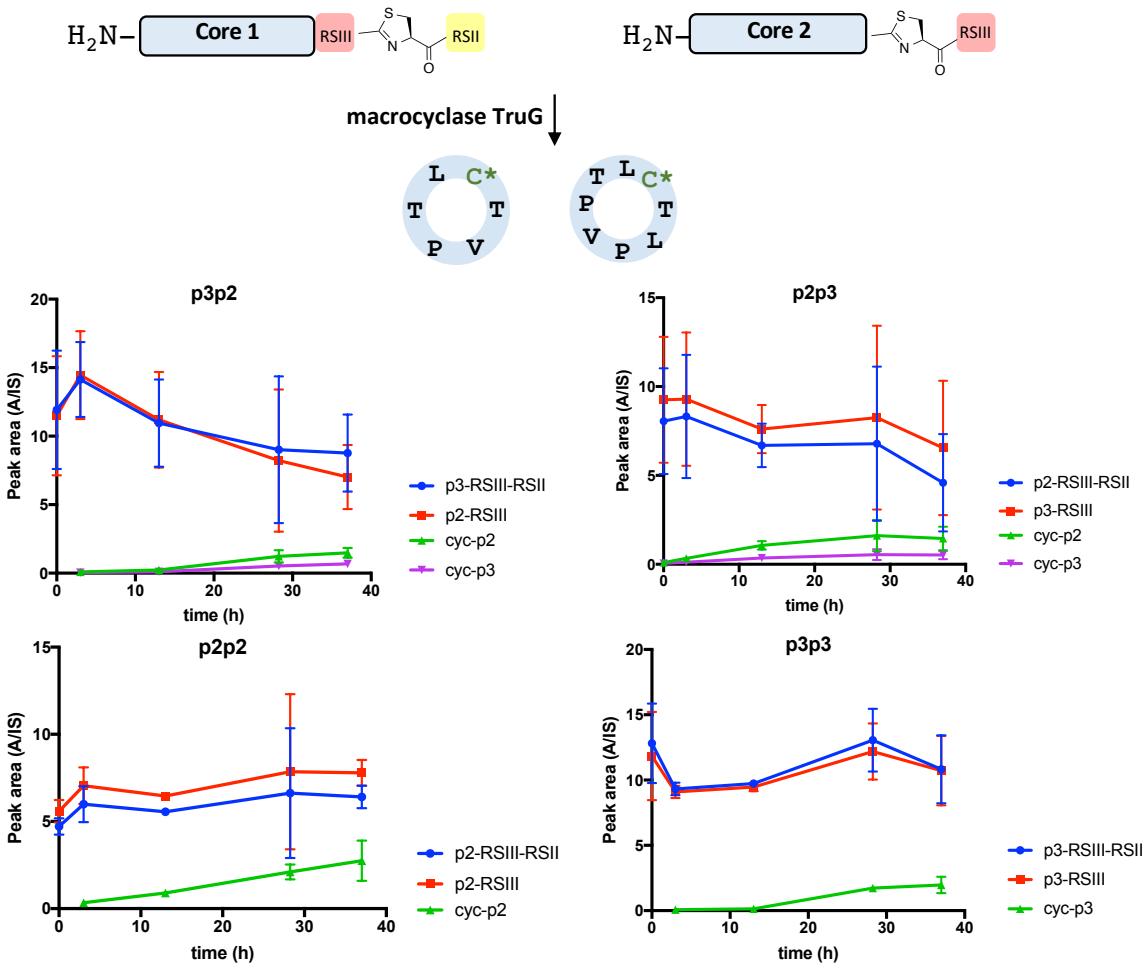


**Figure S13.** PatG reaction time course. Time points were taken at 3, 10, 21, 29 and 53 h. For **A)-D)**, a ratio of the RSIII-RSII over RSIII is used, and a ratio over 1 indicates more production of cassette-1 product; for **E)-F)**, a ratio of the substrate is used, and a ratio over 1 indicates less cassette-1 product. The ratio of 1 is marked with a red dashed line. Bars with a significance level <0.05 are marked with *. On the right shows the reaction fragment peak area, which were normalized to standards and corrected for ionization differences.

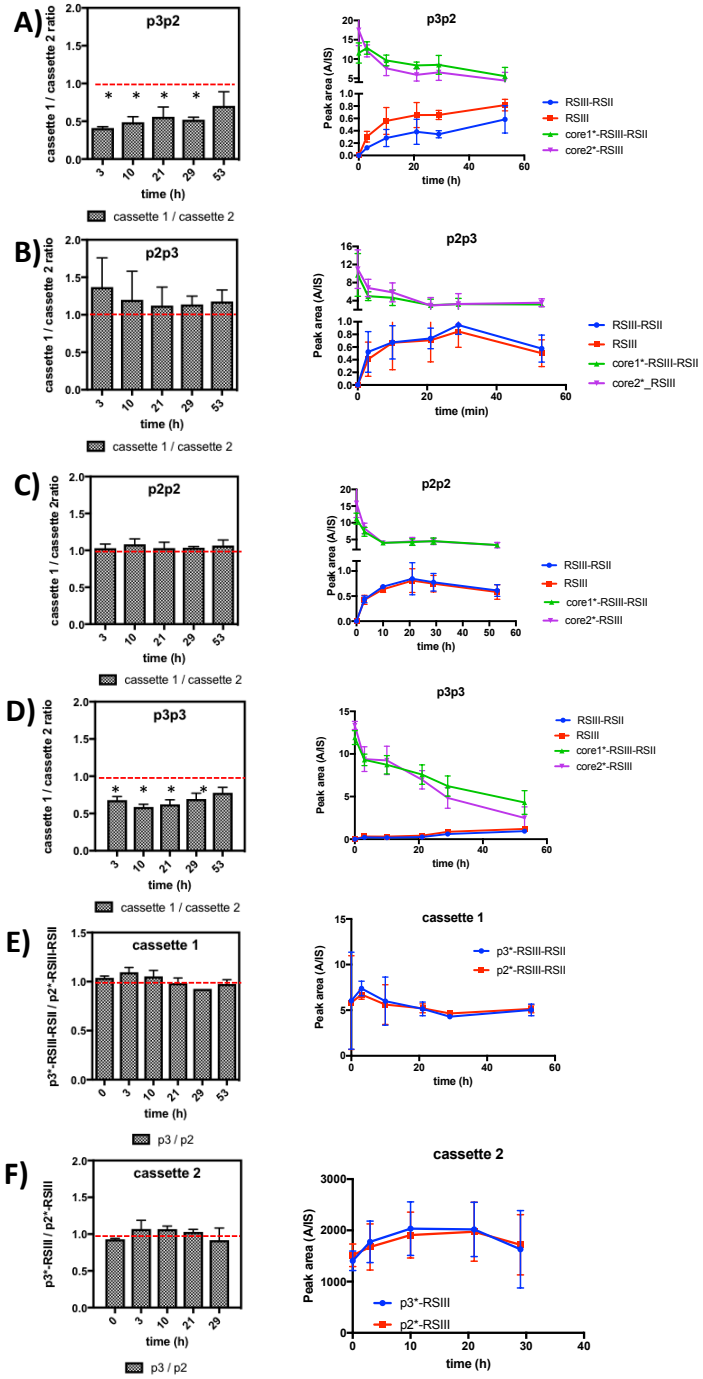


**Figure S14.** Alternative biosynthetic pathway routes for a dual cassette precursor peptide. **A)** Routes also observed in this study. **B)** Routes not measured in this experimental design but should occur under the postulated biosynthetic model.

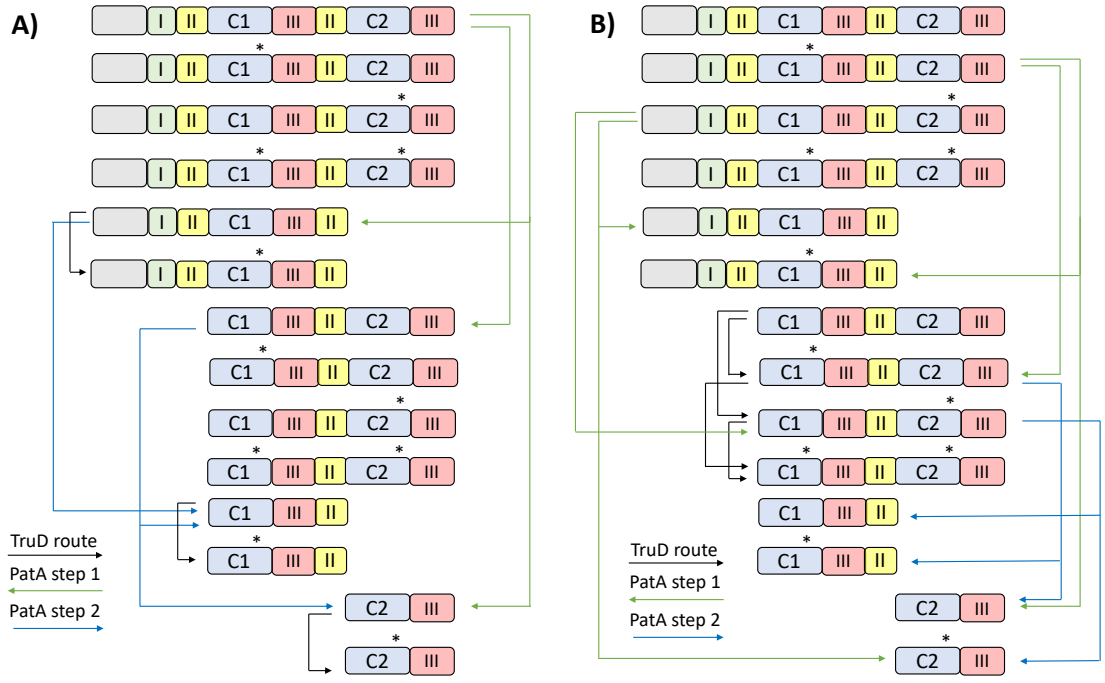


**Figure S15.** SDS-PAGE analysis of precursor peptides. M represents 10-200 kD broad range marker (NEB P7704). **A)** Tris-tricine SDS-PAGE analysis of TruE-p2p2. Elution 1 (E1), elution 2 (E2) and elution 3 (E3) are shown. **B)** Tris-glycine SDS-PAGE analysis of (1) TruE-p3p2, (2)TruE-p3p3, (3) TruE-p2p3, (4) TruE-p23 and (5) TruE-p3.

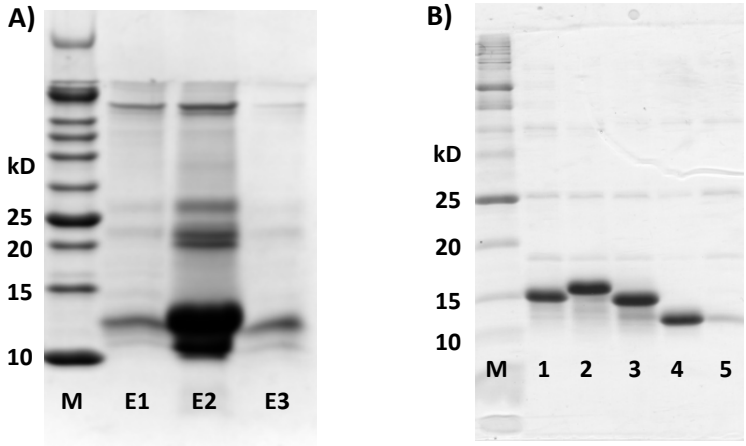
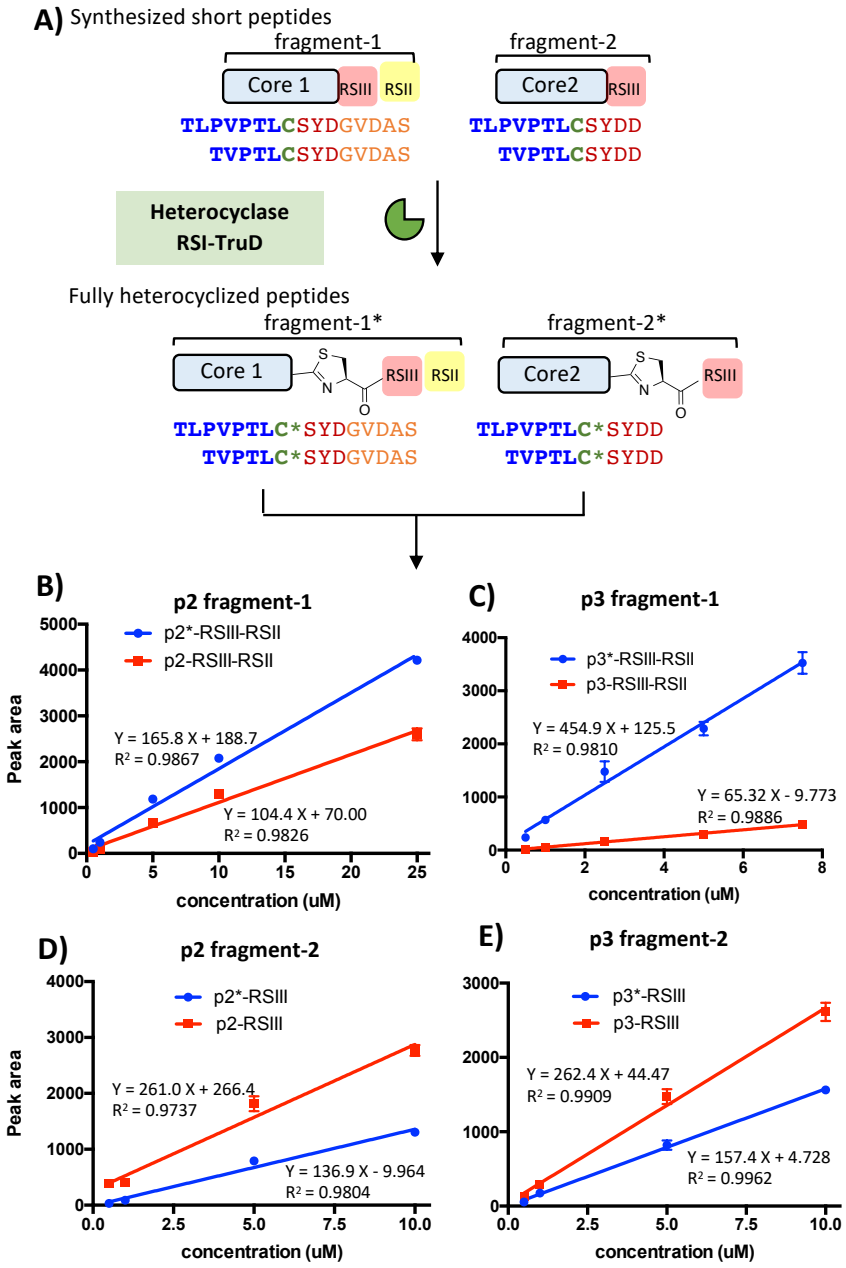


**Figure S16.** Standard curves used to quantify enzyme reactions. **A)** Heterocyclization of the standard peptides by RSI-TruD. **B)-E)** Standard curves of heterocyclized/unheterocyclized standard peptides and linear regression model fit by Prism GraphPad 7.0.



## References


(1) Tianero, M. D.; Pierce, E.; Raghuraman, S.; Sardar, D.; McIntosh, J. A.; Heemstra, J. R.; Schonrock, Z.; Covington, B. C.; Maschek, J. A.; Cox, J. E.; et al. Metabolic Model for Diversity-Generating Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1772–1777.

(2) Collins, B.; Copley, H.; Emmrich, P.; Handley, W.; Hohmann, A.; Knott, E.; Masset, P.; Reeve, B.; Sanderson, T. BBF RFC 57 : Assembly of BioBricks by the Gibson Method Relation to other BBF RFCs https://openwetware.org/wiki/The_BioBricks_Foundation:RFC#BBF_RFC_57:_Assembly_of_BioBricks_by_the_Gibson_Method.

(3) Mcintosh, J. A.; Donia, M. S.; Schmidt, E. W. Insights into Heterocyclization from Two Highly Similar Enzymes. *J. Am. Chem. Soc.* **2010**, *132*, 4089–4091.

(4) Sardar, D.; Hao, Y.; Lin, Z.; Morita, M.; Nair, S. K.; Schmidt, E. W. Enzymatic N- and C-Protection in Cyanobactin RiPP Natural Products. *J. Am. Chem. Soc.* **2017**, *139*, 2884–2887.

(5) Sardar, D.; Lin, Z.; Schmidt, E. W. Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides. *Chem. Biol.* **2015**, *22* (7), 907–916.

(6) Lee, J.; Mcintosh, J.; Hathaway, B. J.; Schmidt, E. W. Using Marine Natural Products to Discover a Protease That Catalyzes Peptide Macrocyclization of Diverse Substrates. *J. Am. Chem. Soc.* **2009**, *131*, 2122–2124.

(7) Sardar, D.; Pierce, E.; McIntosh, J. A.; Schmidt, E. W. Recognition Sequences and Substrate Evolution in Cyanobactin Biosynthesis. *ACS Synth. Biol.* **2015**, *4*, 167–176.

(8) Schägger, H. Tricine–SDS-PAGE. *Nat. Protoc.* **2006**, *1* (1), 16–22.

(9) McIntosh, J. A.; Robertson, C. R.; Agarwal, V.; Nair, S. K.; Bulaj, G. W.; Schmidt, E. W. Circular Logic: Nonribosomal Peptide-like Macrocyclization with a Ribosomal Peptide Catalyst. *J. Am. Chem. Soc.* **2010**, *132*, 15499–15501.

(10) Oueis, E.; Stevenson, H.; Jaspars, M.; Westwood, N. J.; Naismith, J. H. Bypassing the Proline/Thiazoline Requirement of the Macrocyclase PatG. *Chem. Commun.* **2017**, *53* (D), 12274–12277.

(11) Tianero, M. D.; Pierce, E.; Raghuraman, S.; Sardar, D.; McIntosh, J. A.; Heemstra, J. R.; Schonrock, Z.; Covington, B. C.; Maschek, J. A.; Cox, J. E.; et al. Metabolic Model for Diversity-Generating Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1772–1777.