Supplementary Information

Cyclization of Fungal Nonribosomal Peptides by a Terminal Condensation-Like Domain

Xue Gao,[1] *Stuart W. Haynes*[2]*, Brian D. Ames,*[2] *Peng Wang,*[1] *Linda P. Vien*[1] *Christopher T. Walsh*[2]*, *Yi Tang*[1,3]*

[1] Department of Chemical and Biomolecular Engineering, [3] Department of Chemistry and Biochemistry, University of California Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095.

[2] Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115.

*Correspondence e-mail: yitang@ucla.edu, christopher_walsh@hms.harvard.edu

Table of Contents

SUPPLEMENTARY METHODS

1. Molecular cloning

E. coli XL1-Blue (Stratagene) and *E. coli* TOPO10 (Invitrogen) were used for DNA manipulations following standard techniques. DNA restriction enzymes were purchased from New England Biolabs and used as recommended by the manufacturer. PCR was performed using Phusion® DNA Polymerase (New England Biolabs), Platinum Pfx DNA polymerase (Invitrogen) and GoTaq polymerase (Promega, Madison, WI). The constructs of pCR-Blunt vector (Invitrogen) containing desired PCR products were confirmed by DNA sequencing (Retrogen, CA). *Saccharomyces cerevisiae* strain BJ5464-NpgA (*MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1Δ1.6R can1 GAL*) was used as the yeast expression host[1, 2].

2. Protein expression and purification

2.1. Expression and purification of proteins from *E. coli*

The expression plasmids were transformed into *E. coli* BL21 (DE3) strain through electroporation for protein expression. The cells were grown at 37 °C in 0.5 L Luria–Bertani (LB) medium with 35 mg l^{-1} kanamycin to an OD600 of 0.4~0.6 and then isopropylthio-β-D-galactoside (IPTG) was added to a final concentration of 0.1 mM to induce protein expression for 16 hours at 16 °C. The cells were harvested by centrifugation (3750 rpm, 15 mins, 4 °C), resuspended in 30 ml lysis buffer (50 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 500 mM NaCl, 5 mM imidazole, pH 7.9) and lysed through sonication on ice. Cellular lysate was centrifuged at 14,000 rpm for 30 min at 4 °C to remove cellular debris. Ni-NTA agarose resin was added to the soluble fraction and the solution was incubated at 4°C for at least 2 hours. The protein-resin

mixture was loaded into a gravity flow column and proteins were eluted with an increasing gradients of imidazole in buffer A (50 mM Tris-HCl, 500 mM NaCl, pH 7.9).

2.2. Expression and purification of proteins from *S. cerevisiae*

For TqaA, AnaPS, TqaA-ΔC_T, TqaA $H^{3766}A$, TqaA $H^{2658}A$ and AnaPS-E^0 enzymes, which all contain *N*-terminal FLAG tag, the yeast strain BJ5464-NpgA was used as an expression host. Two vacuolar proteases PEP4 and PRB1 were inactivated in this host, which is critical to minimize proteolysis of the large recombinant proteins. The 2 μm expression plasmids were transformed into *S. cerevisiae* BJ5464-NpgA by using S. c. EasyComp™ Transformation Kit (Invitrogen). For 1 L of yeast culture, the cells were grown in YPD media with 1% dextrose for 72 hours at 25 °C. The cells were harvested by centrifugation (2500 g, 20 min, 4 °C), resuspended in 20 ml lysis buffer (50 mM NaH_2PO_4 pH 8.0, 0.15 M NaCl, 10 mM imidazole) and lysed using sonication on ice. Lysate was centrifuged at 35,000 g for 60 min at 4 °C to remove cellular debris. FLAG-tagged proteins were purified by using ANTI-FLAG® M1 Agarose Affinity Gel (Sigma-Aldrich), following the supplied protocols. The cleared cell lysate was applied onto a gravity flow column with packed ANTI-FLAG Agarose Affinity Gel. After washing steps as standard protocols, the protein was eluted with the FLAG peptide elution buffer (0.5 mg ml^{-1} FLAG peptide, 50 mM Tris-HCl, pH 7.4, 100 mM NaCl). After affinity column, TqaA/AnaPS protein was loaded onto a Superdex 200 (GE Healthcare) column and eluted with a flow rate of 0.5 ml min^{-1} in 50 mM Tris pH 8.0 and 100 mM NaCl buffer. A major peak eluting at approximately over 600 kDa was collected for further analysis.

3. ATP-[^{32}P]PP_i exchange assay for TqaA and AnaPS[3]

This assay was used to monitor the substrate-dependent exchange of the ^{32}P label of [^{32}P]PP_i into ATP from the adenylation reactions catalyzed by the A-domains of TqaA or AnaPS . Reactions (100 µl) contained 2 mM ATP, 2 mM $MgCl_2$, 3 mM Na_4[^{32}P]PP_i (0.12 µCi), 0.2 µM enzyme, and 2 mM amino acid substrate in Tris reaction buffer (50 mM Tris-HCl [pH 7.5], 100 mM NaCl, 5 mM TCEP, and 5% glycerol). Reactions were initiated by addition of enzyme, incubated at 25 °C for 1.5 hours, and then quenched by addition of a charcoal solution (1.6% w/v activated charcoal, 100 mM sodium pyrophosphate, 3.5% perchloric acid in water). The charcoal was pelleted by centrifugation and washed with a solution containing 100 mM sodium pyrophosphate and 3.5% perchloric acid, and the charcoal-bound radioactivity detected by liquid scintillation counting.

4. Chemical synthesis of peptidyl-SNAC, peptidyl-*S*-CoA substrates

Ant-D-Trp, Ant-L-Trp, Ant-D-Trp-L-Ala, Ant-L-Trp-L-Ala, Ant-L-*N*-Me-Trp-L-Ala, benzoate (benz)-D-Trp-L-Ala and salicylate (sal)-D-Trp-L-Ala are purchased from RS Synthesis®. The purity of all the peptides are >95% determined by HPLC. The traces provided by the vendors are shown in **Supplementary Figure 18.**

4.1. BOC protection of dipeptides

Ant-Trp (D or L) **4** (50 mg, 0.155 mmol, 1.0 equiv) and BOC anhydride (169 mg, 0.773 mmol, 1.0 equiv) were dissolved in ethanol (10 ml) and heated at 60 °C for 5 hours. The solvent was then removed *in vacuo* and the residue was dissolved in 50% acetonitrile/water and purified by preparative HPLC on a Beckmann Coulter Gold system equipped with a reverse phase C18 column (Phenomenex Luna, 250 x 21.2 mm, 10 micron) with detection at 280 nm. Solvent

system A (water plus 0.1% TFA) and B (acetonitrile plus 0.1% TFA) held at 35% B for 1 min and then run over a linear gradient of 35-75% over 30 min, before increasing from 75-95% B over 1 min and a holding at 95% B for 5 min. The column was then equilibrated back to initial conditions by returning to 35% B and holding for 5 min. The peak with retention time of 29.5 min was collected.

^{1}H-NMR (DMSO-d_6, 400 MHz) δ ppm: 1.45 (s, 9 H), 3.13 - 3.35 (m, 3 H), 4.55 - 4.69 (m, 1 H), 6.93 - 7.01 (m, 1 H), 7.05 (q, *J*=6.91 Hz, 2 H), 7.21 (s, 1 H), 7.33 (d, *J*=7.83 Hz, 1 H), 7.46 (t, *J*=7.83 Hz, 1 H), 7.59 (d, *J*=7.43 Hz, 1 H), 7.72 (d, *J*=7.83 Hz, 1 H), 8.20 (d, *J*=8.61 Hz, 1 H), 8.91 (d, *J*=7.83 Hz, 1 H), 10.47 (s, 1 H), 10.83 (br. s., 1 H).

^{13}C-NMR (DMSO-d_6, 100 MHz) δ: 26.39, 27.93, 53.71, 79.84, 110.25, 111.46, 118.07, 118.32, 118.40, 118.60, 120.98, 121.22, 123.61, 127.08, 128.43, 132.37, 136.12, 139.69, 152.02, 168.46, 173.18

HRMS: *m/z* calculated for $C_{23}H_{25}N_3O_5$: 424.1867 $[M+H]^+$. Found: 424.1869

4.2. Synthesis of dipeptide SNACs (4-SNAC)

Diisopropylethylamine (3.3 µl, 18.9 µmol, 4.0 equiv) was added to a stirred solution of *N*-BOC-Ant-Trp (D or L) (2 mg, 4.73 µmol, 1.0 equiv) and PyBOP (7.4 mg, 14.18 µmol, 3.0 equiv) in DCM (2 ml). To the resulting clear and colorless solution was added *N*-acetylcysteamine (0.75 µl, 7.09 µmol, 1.5 equiv) and the reaction was monitored by LC-MS and stirred until completion (circa 1 hour). The reaction was cooled to 5 °C and TFA (2 ml) was added in order to remove the BOC protecting group. The reaction was again monitored by LC-MS until deprotection was complete (circa 1 hour). The solvent was then removed *in vacuo* and the residue was dissolved in 50% acetonitrile/water and purified by preparative HPLC on a Beckmann Coulter Gold system equipped with a reverse phase C18 column (Phenomenex Luna,

250 x 21.2 mm, 10 micron) with detection at 275 nm. Solvent system A (water plus 0.1% TFA) and B (acetonitrile plus 0.1% TFA) held at 20% B for 1 min and then run over a linear gradient of 20-30% over 10 min, followed by a gradient of 30-40% B over 20 min, before increasing from 40-95% B over 1 min and a holding at 95% B for 5 min. The column was then equilibrated back to initial conditions by returning to 20% B and holding for 5 min. The peak with retention time of 31 min was collected. HPLC assessments of the purity of D or L **4-SNAC** can be seen in **Supplementary Fig. 5**.

4-SNAC: ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm: 2.84 - 2.95 (m, 2 H), 3.10 - 3.18 (m, 2 H), 3.19 - 3.24 (m, 1 H), 3.30 (dd, *J*=14.87, 4.30 Hz, 1 H), 4.72 - 4.82 (m, 1 H), 6.52 (t, *J*=7.43 Hz, 1 H), 6.67 (d, *J*=8.22 Hz, 1 H), 6.94 - 7.02 (m, 1 H), 7.03 - 7.10 (m, 1 H), 7.15 (t, *J*=7.63 Hz, 1 H), 7.19 (s, 1 H), 7.32 (d, *J*=7.83 Hz, 1 H), 7.54-7.58 (m, 1 H), 7.56-7.60 (m, 1 H), 8.03 (br. s., 1 H), 8.70 (d, *J*=7.83 Hz, 1 H), 10.82 (br. s., 1 H).

HRMS: *m/z* calculated for $C_{22}H_{24}N_4O_3S$: 425.1642 $[M+H]^+$. Found: 425.1645

4.3. Synthesis of tripeptide (Ant-D-Trp-L-Ala) SNAC (3-SNAC)

Potassium carbonate (28 mg, 0.203 mmol, 4.0 equiv) was added to a stirred solution of Ant-D-Trp-L-Ala **3**(20 mg, 0.051 mmol, 1.0 equiv) and PyBOP (53 mg, 0.152 mmol, 3.0 equiv) in 50% THF/water (4 ml). To the resulting clear and colorless solution was added *N*-acetylcysteamine (11 µl, 0.101 mmol, 2.0 equiv) and the reaction was monitored by LC-MS and stirred for 2 hours. The solvent was then removed *in vacuo* and the residue was dissolved in 50% acetonitrile/water and purified by preparative HPLC on a Beckmann Coulter Gold system equipped with a reverse phase C18 column (Phenomenex Luna, 250 x 21.2 mm, 10 micron) with detection at 275 nm. Solvent system A (water plus 0.1% TFA) and B (acetonitrile plus 0.1% TFA) held at 5% B for 1 min and then run over a linear gradient of 5-40% over 20 min, followed

by a gradient of 40-60% B over 10 min, before increasing from 60-95% B over 2 min and a holding at 95% B for 5 min. The column was then equilibrated back to initial conditions by returning to 5% B and holding for 5 min. The peak with retention time of 27.5 min was collected. HPLC assessment of the purity of **3-SNAC** can be seen in **Figure 4b** and **Supplementary Figure 11**.

3-SNAC: ^{1}H-NMR (600 MHz, MeCN-d_3) δ ppm: 1.27 (d, *J*=7.04 Hz, 3 H), 1.84 (s, 3 H), 2.91 (t, *J*=6.75 Hz, 2 H), 3.20-3.28 (m, 2H), 3.26-3.32 (m, 1H), 3.41 (dd, *J*=14.67, 5.28 Hz, 1 H), 4.50 (quin, *J*=7.34 Hz, 1 H), 4.78 - 4.87 (m, 1 H), 6.60 - 6.72 (m, 1 H), 6.73 - 6.82 (m, 1 H), 7.02 - 7.08 (m, 1 H), 7.13 (t, *J*=7.63 Hz, 1 H), 7.19 (s, 1 H), 7.21 - 7.25 (m, 1 H), 7.26-7.30 (m, 2 H), 7.31 - 7.36 (m, 1 H), 7.39 (d, *J*=8.22 Hz, 1 H), 7.66 (d, *J*=7.63 Hz, 1 H), 9.25 (br. s., 2 H).

HRMS: *m/z* calculated for $C_{25}H_{29}N_5O_4S$: 496.2013 $[M+H]^+$. Found: 496.2013

4.4. Synthesis of tripeptide-*S*-CoAs: Ant-D-Trp-L-Ala-*S*-CoA (3-*S*-CoA), Ant-L-Trp-L-Ala-*S*-CoA (*epi*-3-*S*-CoA), Ant-L-*N*-Me-Trp-L-Ala-*S*-CoA, Benz-D-Trp-L-Ala-*S*-CoA and Sal-D-Trp-L-Ala-*S*-CoA

Potassium carbonate (7 mg, 50.93 μmol, 4.0 equiv) was added to a stirred solution of Ant-D-Trp-L-Ala **3,** Ant-L-Trp-L-Ala *epi*-**3,** Ant-L-*N*-Me-Trp-L-Ala*,* Benz-D-Trp-L-Ala and Sal-D-Trp-L-Ala (12.73 μmol, 1.0 equiv) and PyBOP (19 mg, 38.20 μmol, 3.0 equiv) in 50% THF/water (2 ml). To the resulting clear and colorless solution was added coenzyme A trilithium salt (10 mg, 12.73 μmol, 1.0 equiv) and the reaction was monitored by LC-MS and stirred for 2 hours. The solvent was then removed *in vacuo* and the residue was dissolved in 50% acetonitrile/water and purified by preparative HPLC on a Beckmann Coulter Gold system equipped with a reverse phase C18 column (Phenomenex Luna, 250 x 21.2 mm, 10 micron) with detection at 275 nm. Solvent system A (water plus 0.1% TFA) and B (acetonitrile plus 0.1% TFA) held at 5% B for 1 min and

then run over a linear gradient of 5-30% over 20 min, followed by a gradient of 30-45% B over 10 min, before increasing from 45-95% B over 2 min and a holding at 95% B for 5 min. The column was then equilibrated back to initial conditions by returning to 5% B and holding for 5 min. The peak with the desired product was collected. HPLC assessment of the purity of the various tripeptidyl-CoA samples are shown in **Supplementary Figure 15**.

3-*S*-CoA: ^{1}H-NMR (DMSO-d_6, 600 MHz) δ ppm: 0.73 (s, 3 H), 0.94 (s, 3 H), 1.23 (d, *J*=7.04 Hz, 3 H), 2.25 (t, *J*=6.75 Hz, 2 H), 2.79-2.90 (m, 2 H), 3.09-3.13 (m, 2H), 3.16-3.22 (m, 1H), 3.22-3.28 (m, 2H), 3.30-3.35 (m, 1H), 3.43-3.56 (m, 1H), 3.74-3.81 (m, 1H), 3.87-3.98 (m, 2H), 4.16 (d, *J*=8.22 Hz, 2H), 4.37 (t, *J*=7.04 Hz, 1H), 4.72 (d, *J*=4.70 Hz, 1H), 4.73-4.78 (m, 1H), 4.78-4.83 (m, 1H), 5.96 (d, *J*=5.28 Hz, 1H), 6.48 (t, *J*=7.34 Hz, 1 H), 6.65 (d, *J*=8.22 Hz, 1 H), 6.95 - 7.00 (m, 1 H), 7.05 (t, *J*=7.63 Hz, 1 H), 7.11 (t, *J*=7.63 Hz, 1 H), 7.21 (s, 1 H), 7.30 (d, *J*=8.22 Hz, 1 H), 7.48 (d, *J*=7.63 Hz, 1 H), 7.67 (d, *J*=7.63 Hz, 1 H), 7.75 (t, *J*=5.87 Hz, 1 H), 8.10-8.11 (m, 1 H), 8.11-8.13 (m, 1 H), 8.24 (s, 1 H), 8.48 (s, 1 H), 8.71 (d, *J*=7.63 Hz, 1 H), 10.79 (br. s., 1 H).

HRMS: *m/z* calculated for $C_{42}H_{56}N_{11}O_{19}P_3S$: 1144.2760 $[M+H]^+$. Found: 1144.2761

***epi*-3**-*S*-CoA: ^{1}H-NMR (DMSO-d_6, 600 MHz) δ ppm: 0.72 (s, 3 H), 0.94 (s, 3 H), 1.33 (d, *J*=7.04 Hz, 3 H), 2.25 (t, *J*=6.75 Hz, 2 H), 2.81-2.91 (m, 2 H), 3.09-3.13 (m, 2H), 3.13-3.21 (m, 1H), 3.22-3.28 (m, 2H), 3.33-3.41 (m, 1H), 3.43-3.56 (m, 1H), 3.74-3.81 (m, 1H), 3.87-3.98 (m, 2H), 4.10-4.19 (m, 2H), 4.46 (quin, *J*=7.19 Hz, 1H), 4.69-4.73 (m, 1H), 4.73-4.78 (m, 1H), 4.78-4.82 (m, 1H), 5.96 (d, J=5.28 Hz, 1H), 6.46 (t, *J*=7.34 Hz, 1 H), 6.63 (d, *J*=8.22 Hz, 1 H), 6.99 (t, *J*=7.34 Hz, 1 H), 7.03 – 7.07 (m, 1 H), 7.09 (t, *J*=7.63 Hz, 1 H), 7.25 (s, 1 H), 7.31 (d, *J*=8.22 Hz, 1 H), 7.46 (d, *J*=7.63 Hz, 1 H), 7.74 (t, *J*=5.58 Hz, 1 H), 7.77 (d, *J*=8.22 Hz, 1 H), 8.12 (t,

J=5.58 Hz, 1 H), 8.14 (d, *J*=8.22 Hz, 1 H), 8.23 (s, 1 H), 8.47 (s, 1 H), 8.84 (d, *J*=7.04 Hz, 1 H), 10.79 (br. s., 1 H).

HRMS: *m/z* calculated for $C_{42}H_{56}N_{11}O_{19}P_3S$: 1144.2760 $[M+H]^+$. Found: 1144.2759

Ant-L-*N*-Me-Trp-L-Ala-*S*-CoA:

HRMS: *m/z* calculated for $C_{43}H_{58}N_{11}O_{19}P_3S$: 1158.2917 $[M+H]^+$. Found: 1158.2917

Benz-D-Trp-L-Ala-*S*-CoA:

HRMS: *m/z* calculated for $C_{42}H_{55}N_{10}O_{19}P_3S$: 1129.2651 $[M+H]^+$. Found: 1129.2653

Sal-D-Trp-L-Ala-*S*-CoA:

HRMS: *m/z* calculated for $C_{42}H_{55}N_{10}O_{20}P_3S$: 1145.2600 $[M+H]^+$. Found: 1145.2601

4.5. Synthesis of *R*-2 and *S*-2

Isatoic anhydride (1.0 g, 6.13 mmol, 1.0 equiv) and tryptophan (L or D) (1.25 g, 6.13 mmol, 1.0 equiv) were suspended in DMSO (7 ml). The resulting mixture was refluxed for 1.5 hours, before being cooled to room temperature and partitioned between ethyl acetate (15 ml) and water (80 ml). The aqueous phase was then extracted with ethyl acetate (2 x 15 ml). The combined organic fractions were then washed with water (15 ml), 5% $NaHCO_3$ (2 x 15 ml), water (15 ml) and brine (15 ml). The organic layer was then dried over magnesium sulphate and the solvent removed *in vacuo*. The resulting residue was purified by flash column chromatography over silica eluting with 50-100% ethyl acetate/hexane to give the desired product.

A portion of the resulting white *R*-**2** or off-white *S*-**2** solids were then dissolved in acetonitrile and further purified by preparative HPLC on a Beckmann Coulter Gold system equipped with a reverse phase C18 column (Phenomenex Luna, 250 x 21.2 mm, 10 micron) with

detection at 254 nm. Solvent system A (water plus 0.1% TFA) and B (acetonitrile plus 0.1% TFA) held at 25% B for 1 min and then run over a linear gradient of 25-55% over 20 min, before increasing from 55-95% B over 1 min and a holding at 95% B for 5 min. The column was then equilibrated back to initial conditions by returning to 25% B and holding for 8 min. The peak with retention time of 18 min was collected.

R-**2** or *S*-**2**: ^{1}H-NMR (400 MHz, MeOD-d_4) δ ppm: 3.14 (dd, *J*=14.87, 9.00 Hz, 1 H), 3.39 (dd, *J*=14.87, 5.87 Hz, 1 H), 4.09 (dd, *J*=8.61, 5.87 Hz, 1 H), 6.89 - 6.98 (m, 1 H), 7.05 (t, *J*=7.43 Hz, 1 H) 7.09 - 7.16 (m, 1 H), 7.13 (s, 1H), 7.22 (t, *J*=7.83 Hz, 1 H), 7.30 (d, *J*=7.83 Hz, 1 H), 7.43 (d, *J*=7.83 Hz, 1 H), 7.48 - 7.55 (m, 1 H), 7.78 (dd, *J*=7.83, 1.17 Hz, 1 H).

^{13}C-NMR (MeOD-d_4, 100 MHz) δ ppm 25.10, 54.70, 110.71, 112.50, 118.99, 119.98, 122.42, 122.62, 124.95, 125.94, 127.21, 128.52, 131.81, 134.28, 138.19, 138.22 171.10, 173.86. These NMR signals matched with those previously reported[4].

HRMS: *m/z* calculated for $C_{18}H_{15}N_3O_2$: 306.1237 $[M+H]^+$. Found 306.1234

5. Chiral analysis of *R*-2 and *S*-2 standards and product of AnaPS

200 μM solutions of *R*-**2** and *S*-**2** were prepared by dissolving a known quantity of synthetic stndards in 50% acetonitrile/water. The product of AnaPS and AnaPS-E^0 was assessed by preparing a 50 μl reaction containing 1 μM AnaPS, 3 mM ATP, 2 mM $MgCl_2$, 0.5 mM TCEP, and 1 mM amino acid substrates (Ant and L-Trp) in NaP_i reaction buffer (50 mM NaP_i [pH 7.4], 100 mM NaCl, and 5% glycerol), which was incubated at 25 °C for 22 hours after addition of enzyme. The sample was then quenched by adding an equal volume of MeCN and removing the observed precipitate by centrifugation. 20 μl of each sample (*R*-**2** standard, *S*-**2** standard, product of AnaPS incubation, product of AnaPS mutant incubation) were injected onto a ChiralCel OD-

RH cellulose tris column (150 × 4.6 mm) for homochiral HPLC analysis (with diode-array detection). The injected sample was separated at a flow rate of 0.5 ml min^{-1} using a linear gradient of 25-65% MeCN in water over 30 min, followed by a ramp up to 95% MeCN over 1 min and a hold at 95% MeCN for 2.5 min.

6. HPLC-based time course study of product formation by TqaA and AnaPS

Reactions (300 µl) contained 1 µM enzyme, 3 mM ATP, 2 mM $MgCl_2$, 0.5 mM TCEP, and 1 mM amino acid substrates (Ant, L-Trp, and L-Ala for TqaA or Ant and L-Trp for AnaPS) in NaP_i reaction buffer. Reactions were incubated at 25 °C and 50 µl aliquots were taken at 1, 2.5, 4.5, 6, and 22 hours after addition of enzyme and quenched by adding an equal volume of MeCN. Precipitate was removed by centrifugation and 20 µl samples were injected onto an Alltima C18 column (150 × 4.6 mm) for HPLC analysis (with diode-array detection). The injected sample was separated at a flow rate of 1 ml min^{-1} using a linear gradient of 25-55% MeCN in water (including 0.1% TFA) over 20 min, followed by a ramp up to 95% MeCN over 1 min and a hold at 95% MeCN for 5 min. Under these conditions **1** eluted at 16.2 min and *R*-**2** eluted at 11.6 min. Integration of the product peaks (**1** at 276 nm and *R*-**2** at 254 nm) was used to generate a plot of product peak area vs. time in order to approximate enzymatic rate. Initial rate data (obtained as integration area per hour) was converted to µM per hour using standard curves generated from 20 µl injections of **1** or *R*-**2** samples of known concentration.

7. Amino acid loading assay for TqaA and AnaPS

The loading of [^{14}C] labeled amino acids onto the T-domains of TqaA or AnaPS was performed to monitor thiolation activity and estimate percentage of the *holo* proteins (as

expressed from *S. cerevisiae* BJ5464-NpgA). For TqaA, reactions contained 2 μM enzyme, 3 mM ATP, 2 mM $MgCl_2$ and 50 μM amino acid substrate (either [carboxy-^{14}C]Ant (0.13 μCi), L-[3-^{14}C]Trp(0.13 μCi), L-[3-^{14}C]Ala (0.14 μCi), or L-[U-^{14}C]Ala (0.32 μCi)) in Tris reaction buffer. The reaction parameters were identical for AnaPS except that [^{14}C] was omitted from testing. Reactions were incubated at 25 °C and 50 μl of the reaction was quenched with 10% TCA (containing 50 μg BSA) at 15 and 30 min after addition of substrate. Precipitated protein was collected by centrifugation and the protein pellet washed twice with 10% TCA and dissolved in 80% formic acid for liquid scintillation counting. A ratio of nmoles radioactivity counted to nmoles protein was used to calculate % loading based on 1 equivalent of [^{14}C] amino acid labeling 1 equivalent of NRPS protein.

SUPPLEMENTARY RESULTS

Supplementary Table 1 Sequences of primers used in this paper

Primer	5′-Sequence-3′
TqaA_P1_F:	ATGGCTAGCGATTATAAGGATGATGATGATAAGACTAGTATGACGGTCCAATCTCTCGAG
TqaA_P1_R:	CCAACAAAGCAGGATGGTCACC
TqaA_P2_F:	CCACAATCTCTCCAACAGCCAG
TqaA_P2_R:	GATCAATTATCACGCCCTTGGG
TqaA_P3_F:	TACATCTGGCTCGACCGGCACA
TqaA_P3_R:	CATCAATGGGAAGGTCACGGCC
TqaA_P4_F:	TTGCTTCGGCTATGTCAAGTCT
TqaA_P4_R:	TCATTTAAATTAGTGATGGTGATGGTGATGCACGTGGAATGTAATATCCAATAATCTTGC
TqaA-ΔC_T _P4 _R:	TCATTTAAATTAGTGATGGTGATGGTGATGCACGTGATCTTCCAATGTGCTTGTGAC
C_T-BspHI-F:	AAAAAA**TCATGA**TGGCAGTTGCCAAAGACAGCACC
C_T-EcoRI-R:	TTTTTT**GAATTC**TCAGTGGTGGTGGTGGTGGTGGAATGTAATATCCAATAATCT
AnaPS_P1F:	ATGGCTAGCGATTATAAGGATGATGATGATAAGACTAGTATGACAATCACTACAATGACT
AnaPS_P1R:	TCCGTTGCTCAAGTTCGCAG
AnaPS_P2F:	TGTTGGCTGAACTCGCTTCTG
AnaPS_P2R:	CATCGGGAAGCCAATGTCCC
AnaPS_P3F:	AGGTCTCGGATGTCGAGCAAA
AnaPS_P3R:	TCATTTAAATTAGTGATGGTGATGGTGATGCACGTGAGCAAAACCGCTGGCAAGAT
C_T_H173A-R:	AGAGACTCCATCCCACTGGGAGGCTGAAATCTGGATCGTGAGAAC
C_T_H173A-F:	GTTCTCACGATCCAGATTTCAGCCTCCCAGTGGGATGGAGTCTCT
TqaA S^{3765}A _F:	GTCGTTCTCACGATCCAGATTGCACACTCCCAGTGGGATGGAGTC
TqaA S^{3765}A _R:	GACTCCATCCCACTGGGAGTGTGCAATCTGGATCGTGAGAACGAC
TqaA S^{3765}H _F:	GTCGTTCTCACGATCCAGATTCACCACTCCCAGTGGGATGGAGTC
TqaA S^{3765}H _R:	GACTCCATCCCACTGGGAGTGGTGAATCTGGATCGTGAGAACGAC
AnaPS-E^0_F:	TCCCTTTTCATGATTGGACATGCCCTGGTACTTGATCTTGTCTCG
AnaPS-E^0_R:	CGAGACAAGATCAAGTACCAGGGCATGTCCAATCATGAAAAGGGA
TqaA H^{2658}A _F:	ATGATTCGCCTCGACATGAGTGCCGCAATTATGGACCGAACAACC
TqaA H^{2658}A _R:	GGTTGTTCGGTCCATAATTGCGGCACTCATGTCGAGGCGAATCAT
T_3_BspHI_F:	AAAAAA**TCATGA**TGAAGCAGCTCTCCACCGACGC
T_3_EcoRI_R:	TTTTTT**GAATTC**TCAGTGGTGGTGGTGGTGGTGATCTTTGATTGGGCTGAGGGA

*Nucleotides underlined and in bold represent restriction enzyme recognition sites.

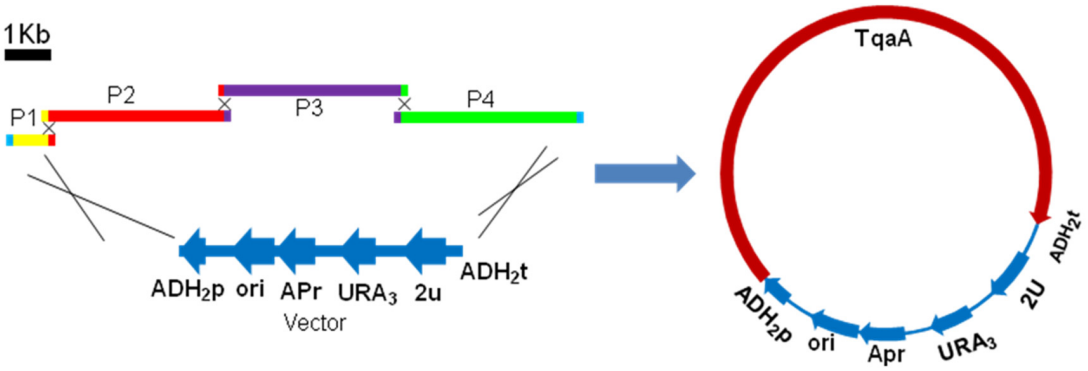
Supplementary Table 2 Fungal NRPSs domain architecture.

Strain	Broad designation	Domain Architecture	% of NRPSs ending with C_T domain
A. terreus NIH2624	ATEG_08448.1	C-A-T-C-A-T-C_T	66%
	ATEG_04322.1	A-T-C-C-A-C-A-T-C_T	
	ATEG_02944.1	C-A-C-A-T-C-Lanthionine synthetase C-like protein	
	ATEG_03576.1	A-T-C-A-T-C_T	
	ATEG_08427.1	A-T-C-A-T-C_T	
	ATEG_01002.1	A-T-C-C-A-T-C-A-C-A-R	
	ATEG_03528.1	A-T-C-A-T-C_T	
	ATEG_05073.1	A-T-C-A-T-C-A-T-C-T-C-T-C_T	
	ATEG_07488.1	A-T-C-T-C_T	
A. fumigatus Af293	Afu1g10380	A-T-C-C-A-A-T-C-A-T-C-C-T-C-T	62.5%
	Afu1g17200	A-T-C-A-T-C-A-T-C-T-C-T-C_T	
	Afu3g03350	A-T-C-A-T-C_T	
	Afu3g03420	A-T-C-A-T-C_T	
	Afu3g12920	A-T-C-A-T-C-T	
	Afu3g13730	A-T-C_T	
	Afu3g15270	A-C-A-T-C_T	
	Afu5g10120	A-T-R	
	Afu5g12730	A-T-C-A-T-E-C-A-T-C-E-T-C-A-T-C-A-T-C-A-T-E-C_T	
	Afu6g08560	A-T-R	
	Afu6g09610	C-A-T-C_T	
	Afu6g09660	A-T-C-A-T-C-T	
	Afu6g12050	A-T-C_T	
	Afu6g12080	A-T-C-A-T-C-E-A-T-C_T	
	Afu8g00170	A-C-A-T-C_T	
	Afu8g01640	A-T-R	
A. clavatus NRRL 1	ACLA_017890	A-T-C-A-T-E-C-A-T-C_T	90%
	ACLA_095980	A-T-C-A-T-E-C-A-T-C_T	
	ACLA_076770	A-T-C-A-T-C-A-T-C_T	
	ACLA_093780	A-T-C-A-T-C-A-T-E-C-A-T-C-A-T-C_T	
	ACLA_098420	A-T-C-A-T-E-C-A-T-E-C-A-T-E-C-A-T-C-A-T-C_T	
	ACLA_025160	A-T-E-C-A-T-C-A-T-C-A-T-E-C-T-C-T	
	ACLA_059530	A-T-E-C-A-T-C-A-T-C-A-T-C_T	
	ACLA_061000	A-T-C-A-T-C_T	
	ACLA_017900	A-T-C_T	
	ACLA_044390	C-A-T-C-A-T-C_T	
	CHGG_00041	A-T-R	
C. globosum CBS 148.51	CHGG_02251	A-T-C-A-T-C-A-T-C-T-C-T-C_T	75%
	CHGG_02283	T-C-C-A-T-C_T	

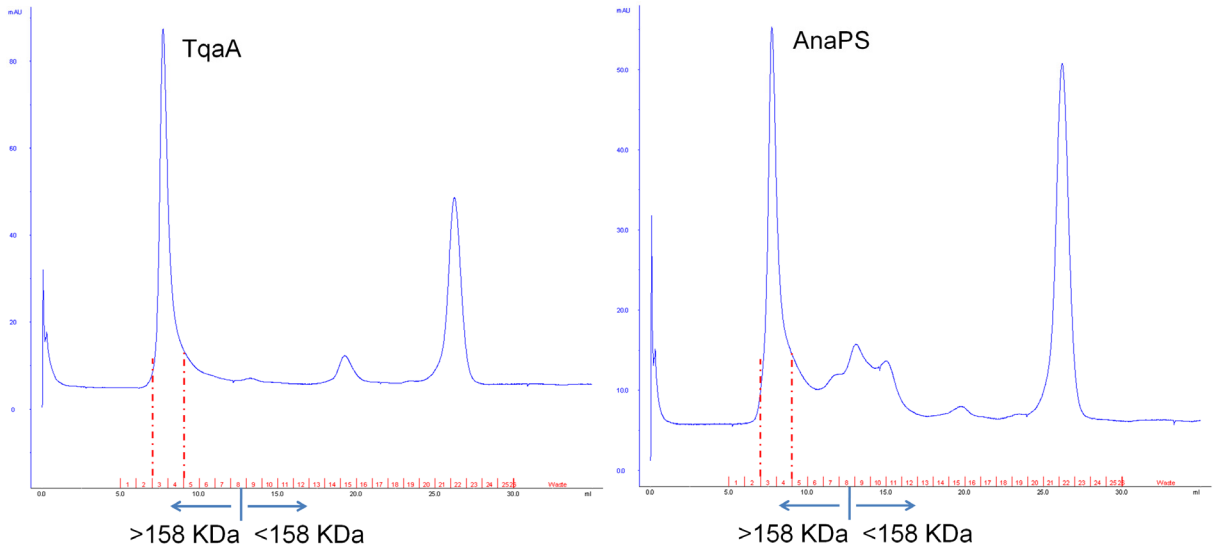
	CHGT_03491	C-A-T-C-A-T-C_T	
	CHGG_04477	T-C-T-C-A-T-C-A-T-C_T	
	CHGG_06052	C-A-T-R	
	CHGG_06789	C- A- T- E-C- A-T-C- A-T-E-C- A-T-C-A-T-E-C-T-C_T	
	CHGG_08540	A-T-C-A-T-C-A-T-E-C-A-T-E-C-A-T-E-T-C_T	
	CHGG_09475	A-T-C-A-T-C_T	
	CHGG_09543	A-T-C-A-T-C-T-C-A-T-C-T-C-T-C_T	
	CHGG_10057	A-T-C-A-T-C-A-T-C-A-T-E-C-A-T-E-C-A-T-E-C-A-T-C_T	
	CHGG_10135	A-T-R	

.

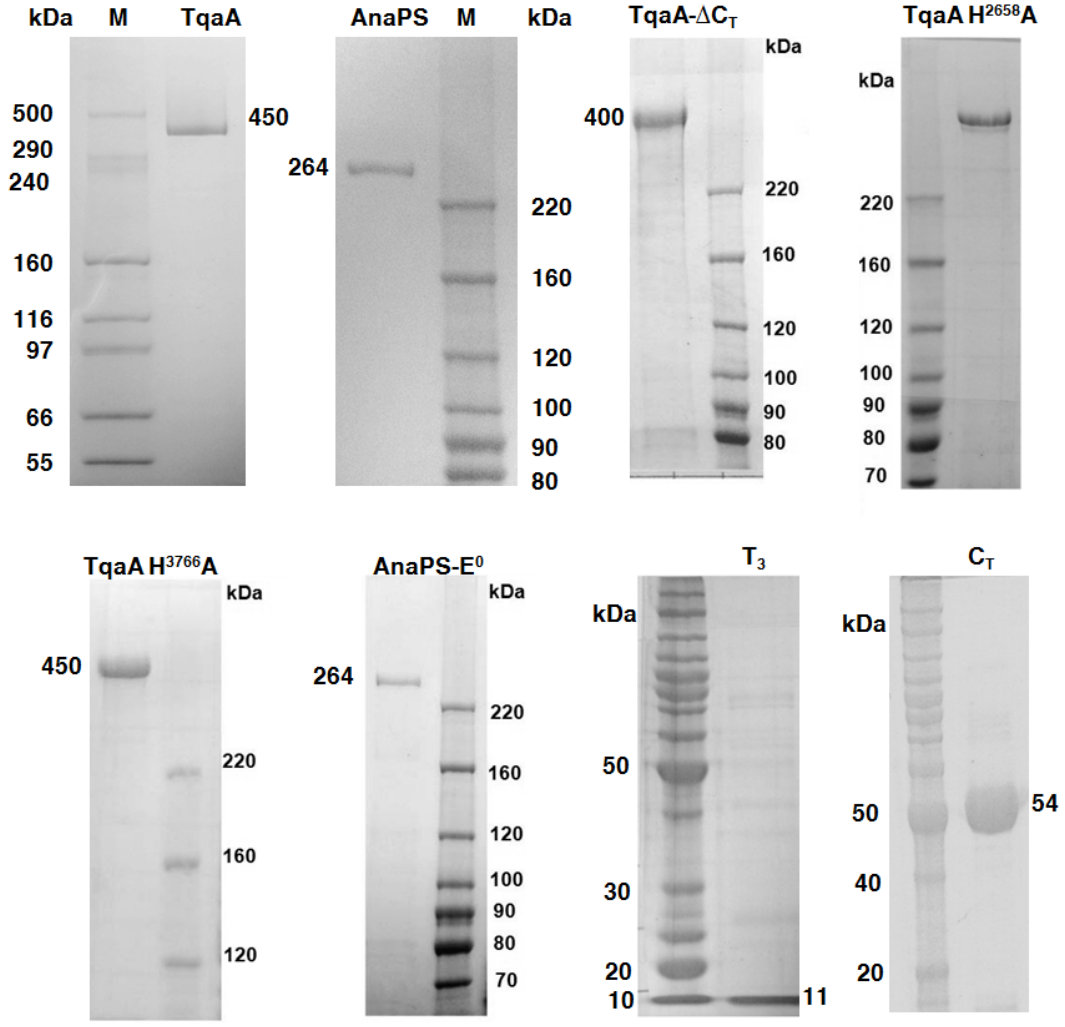
Supplementary Fig. 1 Scheme of *in vivo* homologous recombination for assembly of *tqaA* gene into the 2-*μ*m vector for *S. cerevisiae* expression. The sizes of the overlapping regions between two flanking pieces are within 30 ~ 40 bps.



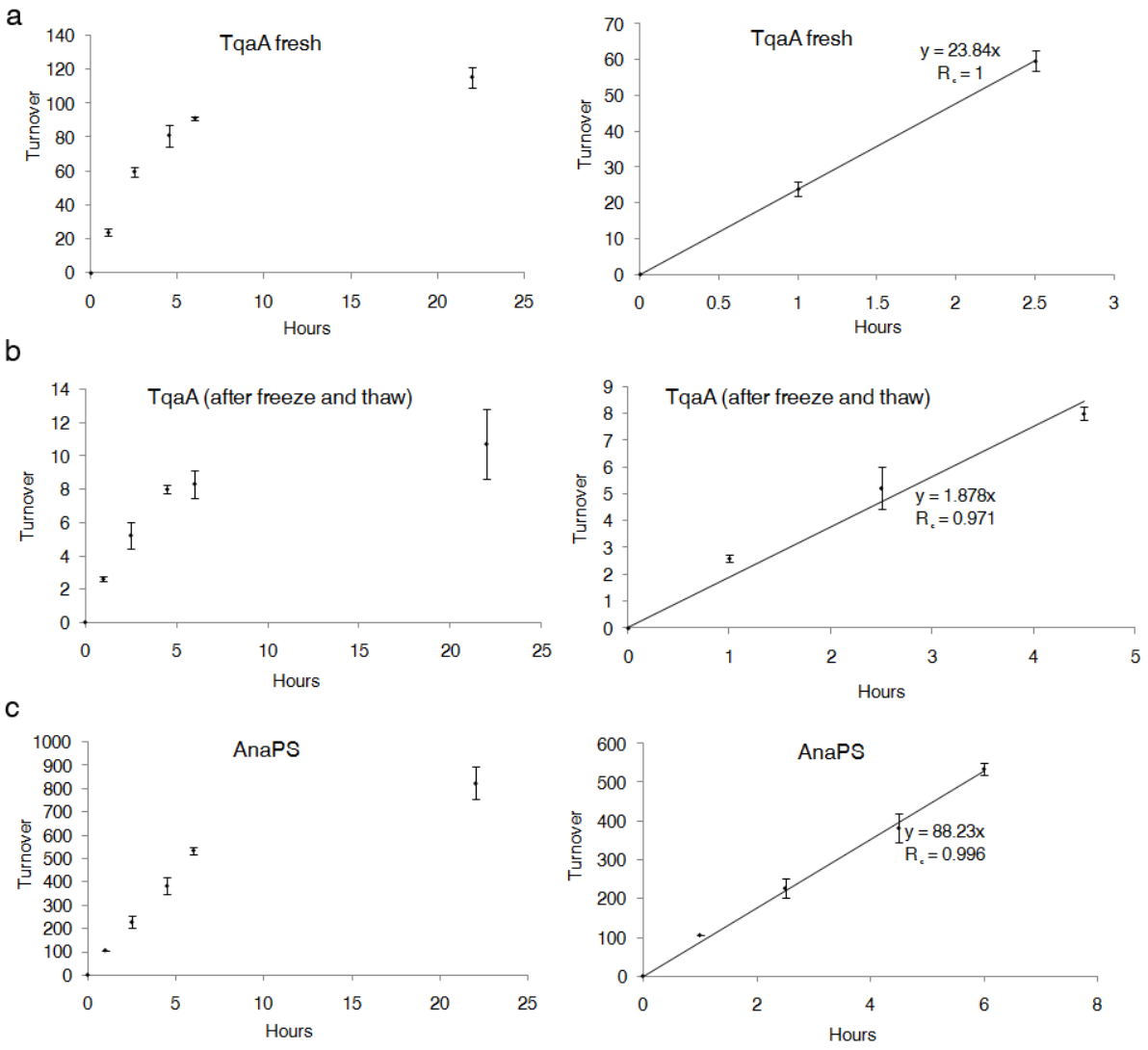
Supplementary Fig. 2 Gel filtration chromatography for TqaA and AnaPS. TqaA/AnaPS purified from ANTI-FLAG® column was loaded onto a Superdex 200 (GE Healthcare) column and eluted with a flow rate of 0.5 ml min^{-1} in 50 mM Tris pH 8.0 and 100 mM NaCl. Fractions 3 and 4 were collected for further assays.



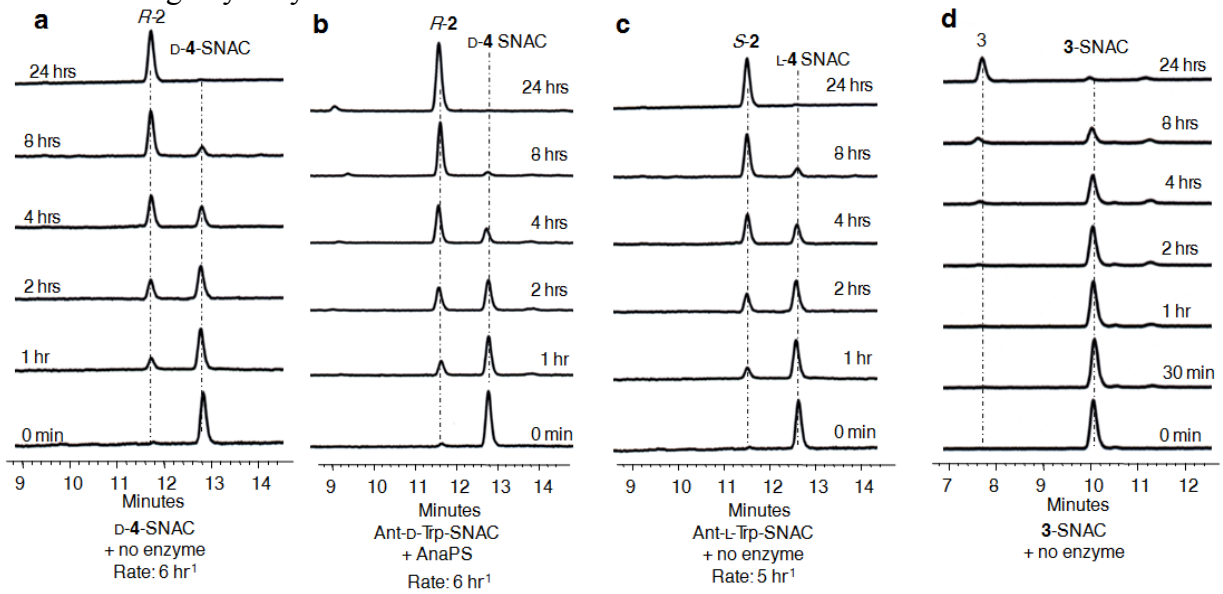
Supplementary Fig. 3 SDS-PAGE of the heterologously expressed proteins in this study. TqaA (450kDa), AnaPS (264kDa), TqaA-ΔC_T (400 kDa), TqaA $H^{3766}A$ (450 kDa), TqaA $H^{2658}A$ (450 kDa) and AnaPS-E^0 (264 kDa) were expressed from BJ5464-NpgA and purified by using ANTI-FLAG® M1 Agarose Affinity Gel; C_T (54 kDa) and T_3 (11 kDa) were expressed from *E. coli* BL21 (DE3) with C-terminal His-tag and purified by using Ni-NTA agarose affinity resin.



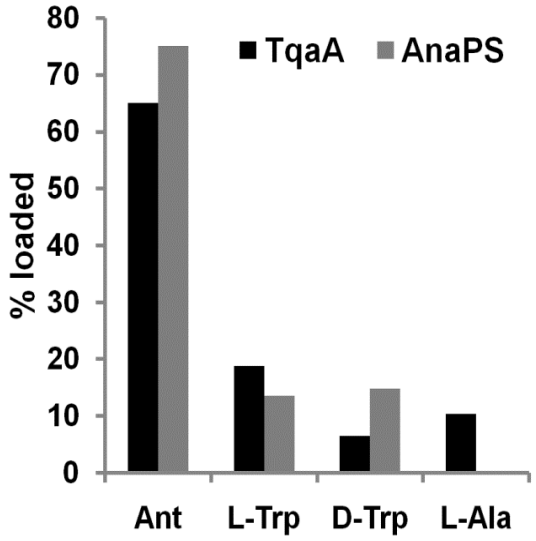
Supplementary Fig. 4 HPLC-based time course study of TqaA and AnaPS in the turnover of **1** and *R*-**2**, respectively. Assays were performed using (**a**) freshly purified TqaA; (**b**) stored (-80 ºC) and thawed TqaA; (**c**) AnaPS (both freshly purified and storage AnaPS displayed nearly the same activity). Each data point was repeated three times. Only data points within the linear range were used to calculate the initial product formation rates. Data represent mean values ± s.d.



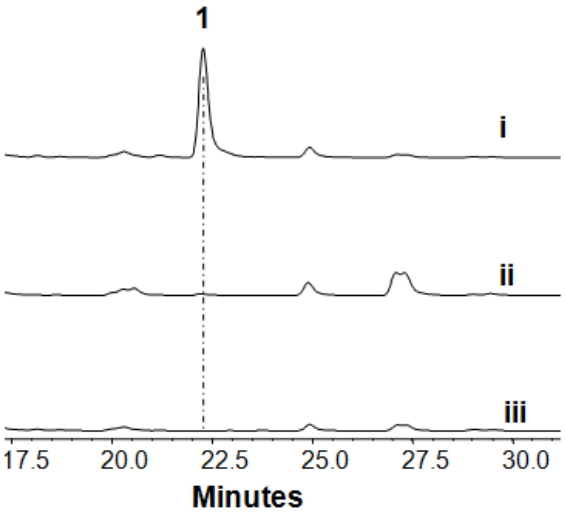
Supplementary Fig. 5 Comparison of the reaction rates of SNACs either with or without enzyme. SNAC reactions (400 µl) contained 0.5 mM TCEP and (**a**) 1 mM D-**4** SNAC without enzyme; (**b**) 1 mM D-**4** SNAC with 1 µM AnaPS; (**c**) 1 mM L-**4** SNAC without enzyme; (**d**) 1 mM **3**-SNAC without enzymes. **3**-SNAC does not undergo spontaneous cyclization to afford **1**, but can undergo hydrolysis to release the free acid **3**. HPLC traces are shown with λ = 272 nm.



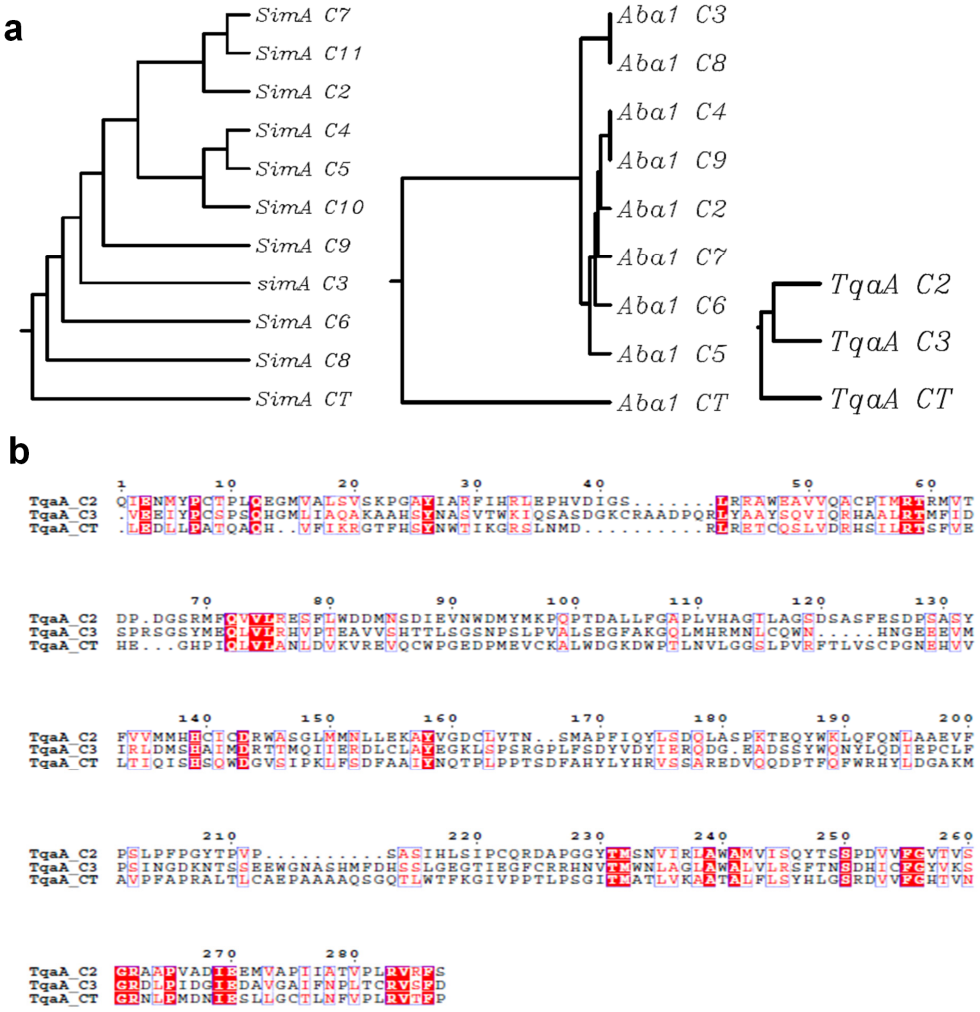
Supplementary Fig. 6 Characterization of TqaA and AnaPS T domain loading assay with [^{14}C] labeled amino acid.



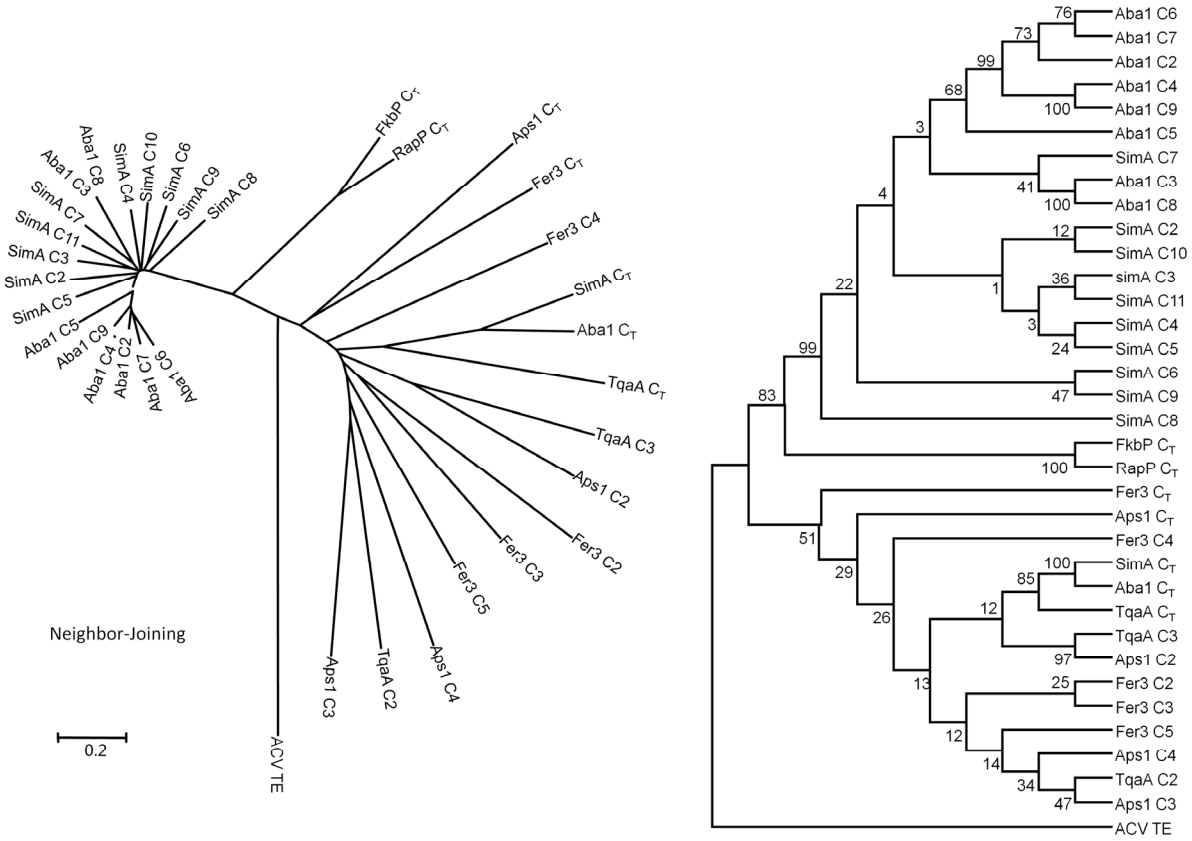
Supplementary Fig. 7 *R*-**2** is not produced by TqaA under *in vitro* assay conditions. HPLC traces (λ = 272 nm) shown here are organic extracts of the following *in vitro* assays: i) 2 mM of Ant, L-Trp, L-Ala and 10 μM TqaA; ii) 2 mM of Ant, L-Trp, L-Ala and 10 μM TqaA C_3^0 (containing the inactivating $H^{2658}A$ mutation); iii) 2 mM of Ant, L-Trp and 10 μM TqaA (no L-Ala added). The inability to produce *R*-**2** in ii and iii suggests C_3 may structurally protect the dipeptide from being cyclized into *R*-**2** spontaneously.



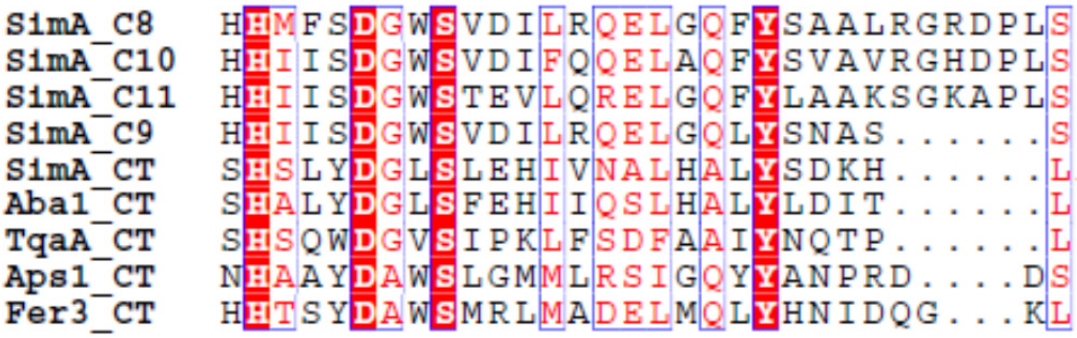
Supplementary Fig. 8 (a) Neighbor-joining phylogenetic analyses of C domains and C_T domain in SimA, Aba1 and TqaA NRPSs; (**b**) Sequences alignment of TqaA C domains and C_T domain. Sequence alignment and tree construction were performed using ClustalW[4].



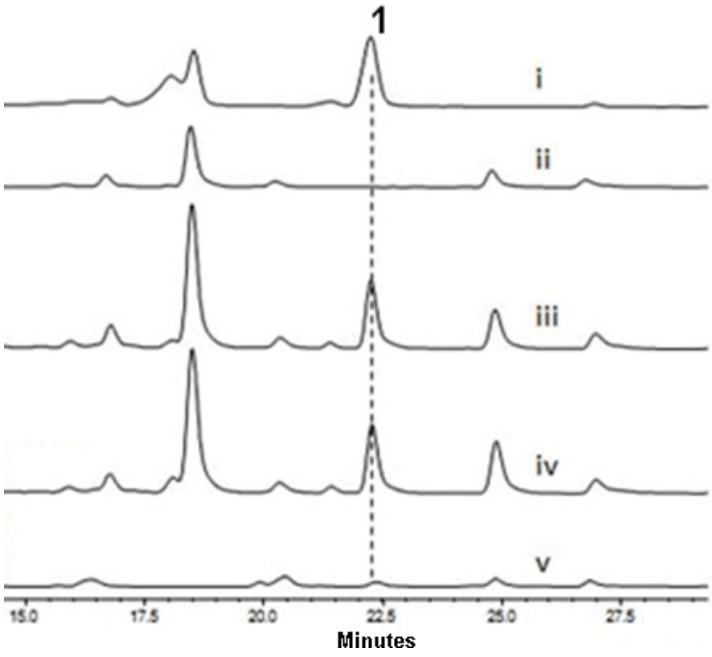
Supplementary Figure 9 Neighbor-joining phylogenetic tree of select fungal C-domains and ACV synthetase terminal TE domain. Sequence alignment and tree construction were performed using MEGA[5].



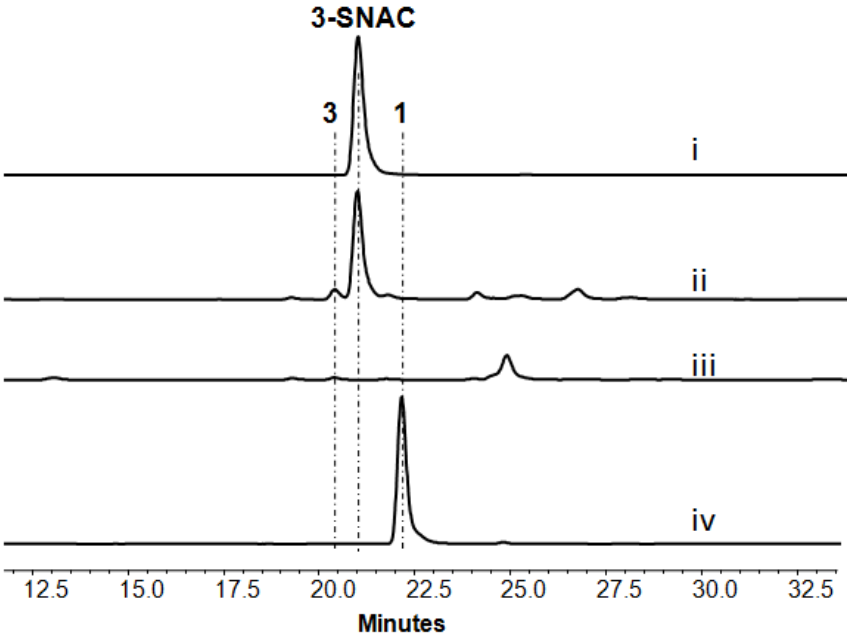
Supplementary Fig. 10 Sequence alignment of the active sites of C_T domains with other fungal NRPS C domains (using the canonical NRPS C domains from the cyclosporine synthetase SimA). All of the chain-extending C domains in SimA containing the "**HH**XXX**D**XX**S/T**" motif. In contrast, TqaA Aba1 and SimA C_T domains have another amino acid (**S** or **N**) in place of the first **H**. Sequence alignment was performed using ClustalW[4].



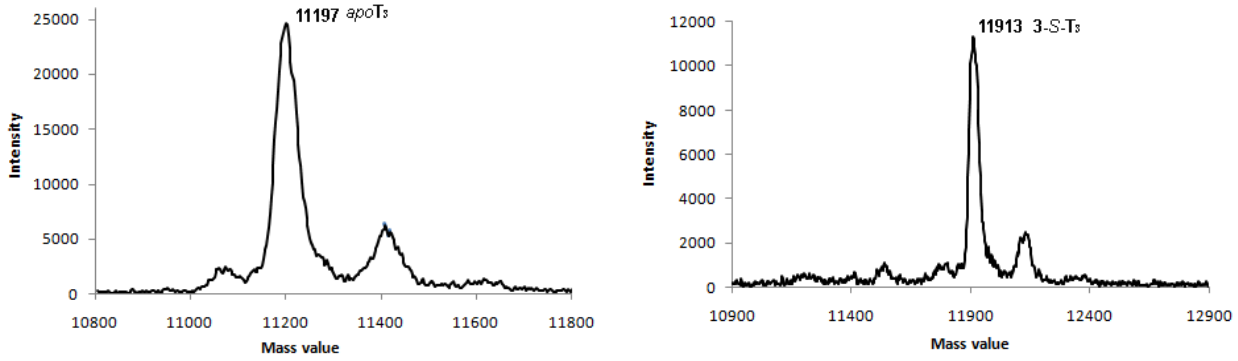
Supplementary Fig. 11 Analysis of TqaA C_T active site mutants by *in vivo* product formation and *in vitro* assay. The assays show TqaA H^{3766} is essential for cyclization of **1**, while S^{3765} in not. Traces (i-iv) shown here are HPLC analyses (λ = 272 nm) of metabolites extracted from 3-day cultures of i) BJ5464-NpgA expressing TqaA; ii) BJ5464-NpgA expressing TqaA H^{3766}A; iii) BJ5464-NpgA expressing TqaA S^{3765}A; and iv) BJ5464-NpgA expressing TqaA S^{3765}H. v) TqaA H^{3766}A was purified and subjected to *in vitro* assay. HPLC analysis of the organic extract from the reactions containing: 2 mM of each amino acid building block and 10 μM TqaA H^{3766}A.



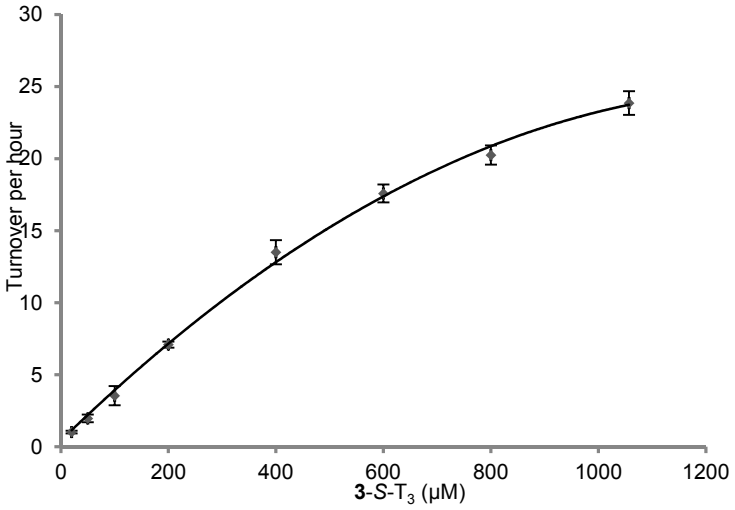
Supplementary Fig. 12 Assay of TqaA C_T domain using **3** attached to small molecule thioester carriers. The assays show that C_T cannot cyclize **3**-SNAC or **3**-*S*-CoA to product **1**. 20 μM C_T was used in all assays. HPLC traces (λ = 272 nm) shown here are extracts from *in vitro* reaction containing ii) 200 μM **3**-SNAC and iii) 200 μM **3**-*S*-CoA. Trace i and iv are the standards of **3**-SNAC and **1**, respectively. Hydrolysis of **3** can be detected in both reactions.



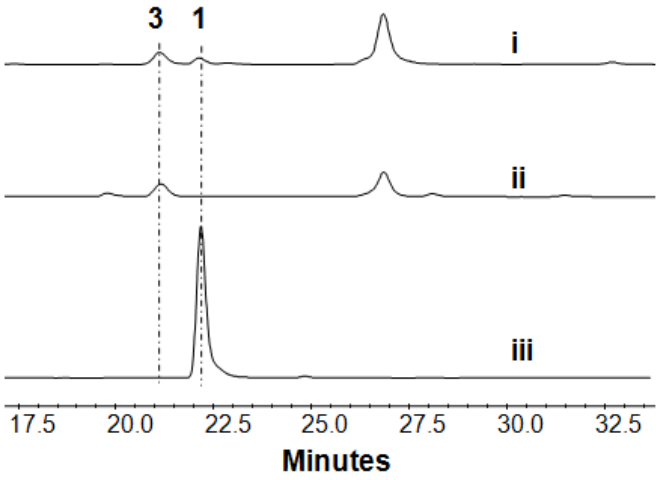
Supplementary Fig. 13 MALDI-TOF mass analyses of the *apo* T_3 and **3**-*S*-T_3. **3**-*S*-T_3 was prepared from an i*n vitro* reaction containing *apo* T_3, Sfp and **3**-*S*-CoA. *Apo* T_3 has a $[M+H]^+$ mass value of 11197 and **3**-*S*-T_3 has a $[M+H]^+$ mass value of 11913.



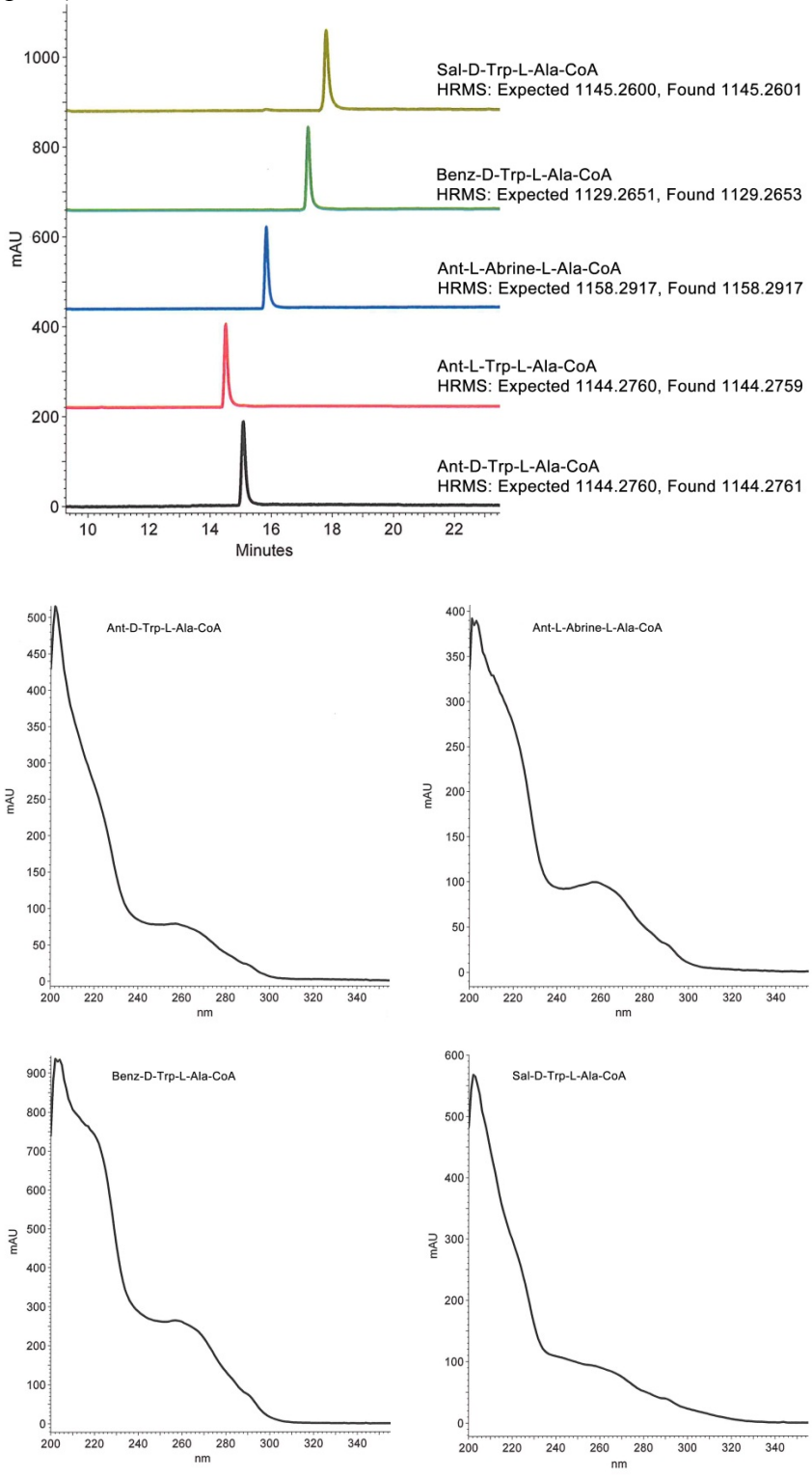
Supplementary Fig. 14 Kinetic properties of the C_T domain in cyclizing **1** from **3**-*S*-T_3. Initial turnover rates (V/E_o) as a function of **3**-*S*-T_3 concentration were measured. The concentration of C_T domain was kept at 2 μM. The data were fitted to the Michaelis-Menten equation using nonlinear least squares regression. Each data point was repeated three times. Data represent mean values ± s.d.



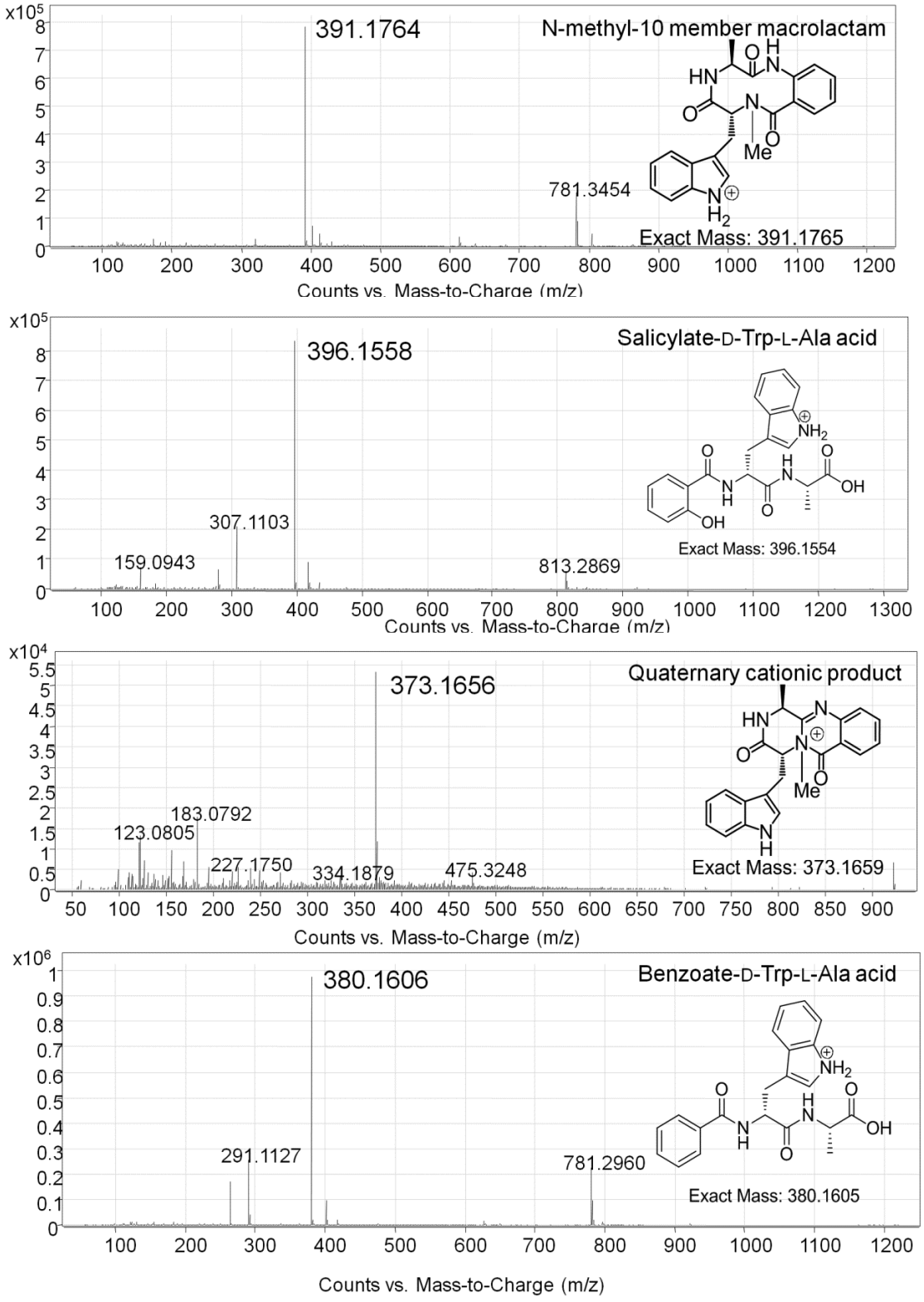
Supplementary Fig. 15 C_T can interact with T_{ApdA} to produce **1** but not with T_{LovF}. In the *in vitro* assays, 200 μM of **3**-*S*-CoA was used as the substrate. All assays contained 20 μM C_T. Traces shown are HPLC traces (λ = 272 nm) of *in vitro* reactions i) with addition of 200 μM T_{ApdA}, C_T was able to cyclize and produce **1** with a relatively low yield, while significant amount of the hydrolysis product **3** was observed; and ii) with addition of 200 μM T_{LovF}, only hydrolysis of **3**-*S*-CoA to **3** free acid can be observed. Trace iii is that of **1** standard.



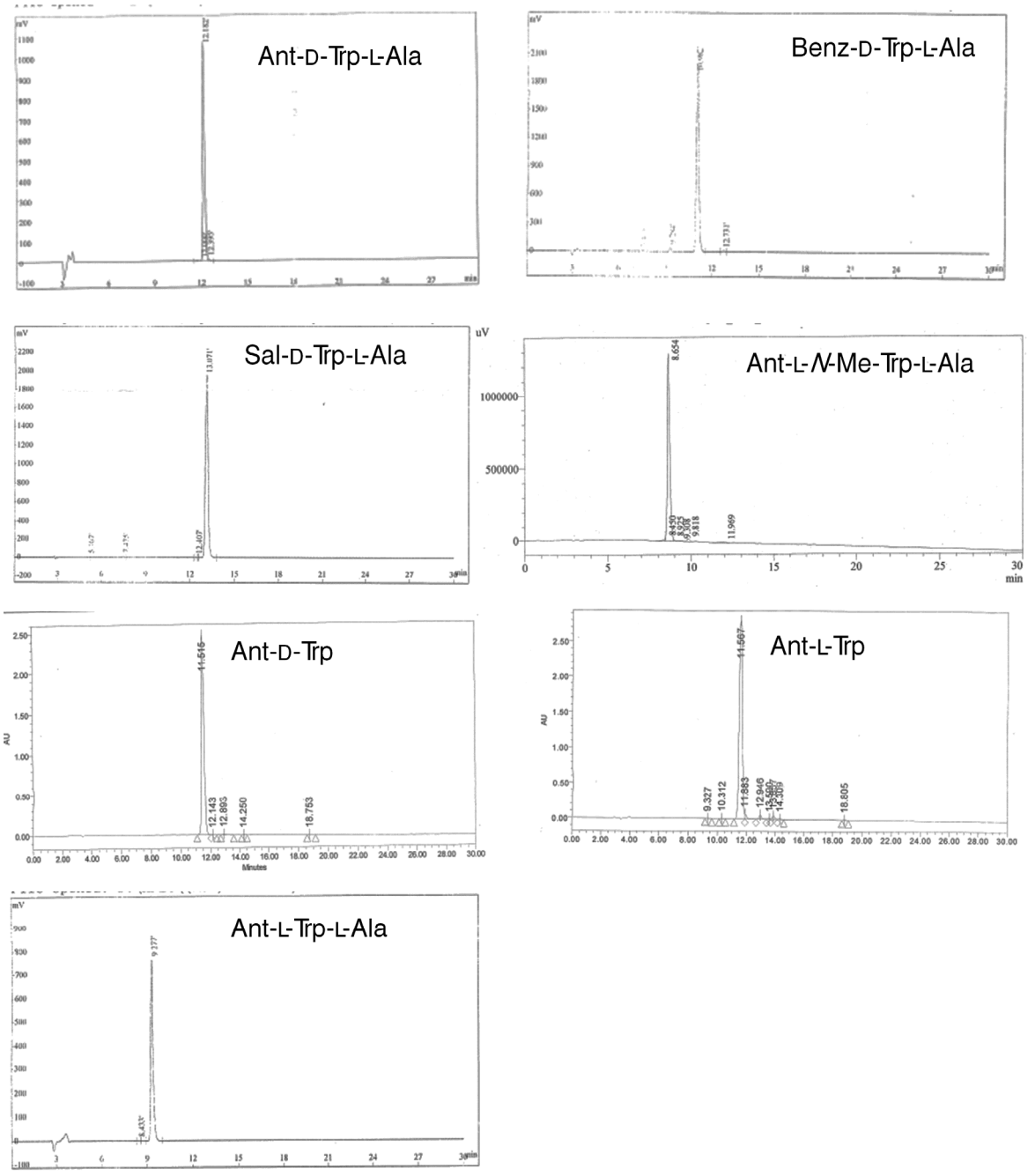
Supplementary Fig. 16 HPLC traces and 2D UV-vis spectra of the tripeptidyl-CoA analogues. High resolution mass (HRMS) analyses showed the measured masses of the synthetic compounds were consistent with the caculated masses. (Sal: Salicylate; Benz: benzoate; abrine: *N*-methyl-tryptophan)



Supplementary Fig. 17 High resolution mass analysis of the products from TqaA C_T domain reactions: *N*-methyl-10 membered macrolactam, quaternary cationic product, salicylate-D-Trp-L-Ala free acid and benzoate-D-Trp-L-Ala free acid.



Supplementary Fig. 18. HPLC assessments of the purity of the peptides as supplied by the vendor (vendor information: RS Synthesis, LLC. PO Box 70301, Louisville, KY 40270, Phone: 502-614-5920, Fax: 801-780-2235)



References

1. Jones, E.W. Tackling the protease problem in *Saccharomyces cerevisiae*. *Methods Enzymol.* **194**, 428-453 (1991).
2. Lee, K.K.M., Da Silva, N.A. & Kealey, J.T. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in *Escherichia coli* and *Saccharomyces cerevisiae*. *Anal. Biochem.* **394**, 75-80 (2009).
3. Rusnak, F., Sakaitani, M., Drueckhammer, D., Reichert, J. & Walsh, C.T. Biosynthesis of the *Escherichia-coli* siderophore enterobactin: sequence of the entf gene, expression and purification of entf, and analysis of covalent phosphopantetheine. *Biochemistry* **30**, 2916-2927 (1991).
4. Yin, W.B., Grundmann, A., Cheng, J. & Li, S.M. Acetylaszonalenin biosynthesis in *Neosartorya fischeri*. Identification of the biosynthetic gene cluster by genomic mining and functional proof of the genes by biochemical investigation. *J. Biol. Chem.* **284**, 100-109 (2009).
5. Larkin, M.A. et al. Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
6. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596-1599 (2007).