



Supplementary Materials for

Use of a Scaffold Peptide in the Biosynthesis of Amino Acid Derived Natural Products*

Authors: Chi P. Ting¹, Michael A. Funk^{2#}, Steve L. Halaby^{4,5}, Zhengan Zhang,¹ Tamir Gonen^{4,5†} and Wilfred A. van der Donk^{1,2,3†}

Affiliations:

¹ Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

² Department of Chemistry, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

³ Howard Hughes Medical Institute, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

⁴ Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles CA 90095, USA

⁵ Departments of Biological Chemistry and Physiology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles CA 90095, USA

Present address: American Association for the Advancement of Science, Washington, DC, USA.

† Correspondence addressed to T.G (tgonen@ucla.edu) or W.A.V. (vddonk@illinois.edu)

*This manuscript has been accepted for publication in Science. This version has not undergone final editing. Please refer to the complete version of record at <http://www.sciencemag.org/>. The manuscript may not be reproduced or used in any manner that does not fall within the fair use provisions of the Copyright Act without the prior, written permission of AAAS.

This PDF file includes:

Materials and Methods
Figs. S1 to S11
Table S1-S3
Scheme S1

Materials and Methods

Figure preparation

Gene cluster diagrams (Fig. 1B, fig. S8 and S11) were made with EasyFig (29). Mass spectra were plotted in Matlab R2015b. All figures were constructed with Adobe Illustrator. NMR spectra were analyzed in Mnova (Mestrelab Research). Sequence alignment (Fig. 4B) was made with Geneious 11.0.5 using Geneious Alignment.

Materials

Materials were obtained from Sigma-Aldrich unless otherwise noted. All water was deionized and purified using a Milli-Q IQ 7000 purifier. MALDI-TOF MS was performed on a Bruker Daltonics UltrafleXtreme MALDI TOF/TOF instrument. ESI-MS analyses were performed using a SYNAPT ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters). ¹H NMR and ¹³C NMR spectra were recorded on an Agilent 600 MHz spectrometer for ¹H (150 MHz for ¹³C) in D₂O. Chemical shifts are reported relative to the residual solvent signal (¹H NMR: $\delta = 4.79$). UV-Vis spectroscopy was performed on an Agilent Cary 4000 spectrophotometer. gBlocks were codon optimized and ordered from Twistbio and Integrated DNA Technologies (Table S2).

Cloning and cell culture. DNA was prepared using MiniPrep kits (Qiagen) using the manufacturer's instructions from *E. coli* DH10b cells (New England Biolabs) made chemically competent by the KCM method (30). Genomic DNA from *P. syringae* pv. *maculicola* ES4326 (Pma) was prepared using an UltraClean Microbial DNA isolation kit (Mo Bio Laboratories) according to the manufacturer's instructions; Pma cells were grown in liquid KB medium with vigorous aeration. Plasmids were constructed with type-II restriction enzymes using the New England Biolabs Golden Gate Assembly Tool (<http://goldengate.neb.com/editor>) or manually designed and prepared by Gibson assembly. For plasmids assembled by Golden Gate Assembly, two primer pairs were designed for PCR to prepare both the gene and the vector with complementary sticky ends for type II's restriction digest. BsaI and T4 DNA Ligase were used for single step DNA ligation. For plasmids assembled by Gibson Assembly, one primer pair was designed for PCR amplification of the gene with an overhang complementary to the vector. The vector was linearized by restriction digest with NdeI and XhoI and ligated with the gene of interest using NEBuilder® HiFi DNA Assembly Master Mix. Restriction enzymes and PCR polymerases were obtained from New England Biolabs. Genes for each of the *tgl* gene cluster products and the *P. syringae* cysteine tRNA synthetase (CysRS) were identified using the Joint Genome Initiative Integrated Microbial Genomes and Microbiomes webtool (<http://img.jgi.doe.gov>). Primers for cloning and site-directed mutagenesis were obtained from Integrated DNA Technologies. Mutagenesis was accomplished using QuikChange II Site-Directed Mutagenesis Kit according to manufacturer's instructions. Mutagenesis primers were designed using the QuikChange Primer Design webtool. https://www.genomics.agilent.com/primerDesignProgram.jsp?_DARGS=/primerDesignProgram.jsp. Primers are listed in Table S1.

Peptide expression and purification. *E. coli* BL21(DE3) cells (New England Biolabs) expressing N-terminal His₆-TglA from pRSF-Duet were grown with 50 µg/mL kanamycin in autoinduction (AI) media (for 1 L total: 10 g bacto tryptone, 5 g yeast extract, 5 g NaCl, 3 g KH₂PO₄, 6 g Na₂HPO₄) with 1X AI sugar solution containing 0.5% (vol/vol) glycerol, 0.05% (w/vol) glucose, and 0.2% (w/vol) lactose (20 mL of 50X stock). Cultures were shaken at 37 °C for 6-7 h following inoculation with 1 mL saturated culture. Cells were harvested and lysed by sonication. Peptides were purified by immobilized metal affinity chromatography (IMAC). The lysate was clarified by centrifugation at 29,000 rcf and applied to a 5 mL NiNTA-agarose column (GE Healthcare) using a peristaltic pump. The immobilized peptide was washed with 5 column volumes (CV) of 90% lysis buffer (50 mM HEPES, 100 mM NaCl, pH 7.5), 10% elution buffer (50 mM HEPES, 100 mM NaCl, 500 mM imidazole, pH 7.5; final imidazole concentration in wash: 50 mM) and eluted with 100% elution buffer. The elution fraction was concentrated using a 3 kDa MWCO Amicon spin filter and washed with 10-20 CV of deionized water to remove imidazole. Crude peptide was desalted with a VYDAC[®] Bioselect C4 cartridge. Peptide elution fractions were lyophilized.

His₆-TglA-Cys and the modified peptide **1** obtained by co-expression of TglA with TglB or TglH/I/B were obtained using the same procedure described above. For analysis of modified TglA peptides, the same procedure was used but the peptide was reacted with 20 mM N-ethylmaleimide (NEM) or 20 mM 2-iodoacetic acid (IAA) prior to concentration by Amicon. To reduce the size of the peptide, the lyophilized product was resolubilized in denaturing buffer (1 mL 6 M guanidinium chloride, 20 mM NaH₂PO₄, 500 mM NaCl, 0.5 M imidazole, pH 7.5), diluted 10-fold and digested by trypsin (0.5 mg/mL) to release the VFAC-NEM, VFAX-NEM (X is the thioaminal in structure **1**), or VFA-thiaGlu fragments. Plasmid DNA encoding N-terminal truncants of TglA-Cys were prepared by Gibson assembly and designed to introduce an insertion of a TEV cleavage site (ENLYFQS) prior to the sequence of interest. Peptides were purified by IMAC, digested by TEV protease and fragments were purified by HPLC (see HPLC peptide purification). Hala2 and ProcA2.8 peptides were prepared as previously reported (31).

His₆-AmmA, His₆-AmmA-ΔTrp56 and His₆-AmmA* peptides were prepared with the same procedure as His₆-TglA with the following deviation. *E. coli* BL21(DE3) cells (New England Biolabs) expressing N-terminal His₆-AmmA, His₆-AmmA-ΔTrp56 or His₆-AmmA* from the pACYCDuet1 vector were grown with 33 µg/mL chloramphenicol in LB media (for 1 L total: 10 g bacto tryptone, 5 g yeast extract, 10 g NaCl). Cells were grown with vigorous shaking at 37 °C for 2-3 h until the optical density (OD) at 600 nm reached 0.6 at which point the culture was cooled to 0 °C and induced with 0.5 mM IPTG. Upon induction, the cell culture was incubated at 37 °C for 5 h with vigorous shaking prior to collection by centrifugation.

Protein expression and purification for TglB, TglF, and CysRS. All heterologously-expressed proteins were obtained from *E. coli* BL21(DE3) cells (New England Biolabs) made chemically competent through the KCM method (30). The general expression protocol for the *tgl* cluster enzymes TglB, and TglF is as follows: A 1 mL inoculum was added to 1 L of AI medium containing 50 µg/mL kanamycin and 1x AI sugars. For N-

terminal His₆-TglB or, His₆-CysRS, cells were grown lysogeny broth (LB) medium and expression was induced with 0.5 mM isopropyl-beta-D-1-thiogalactopyranoside (IPTG). Culture were grown with vigorous shaking at 37 °C for 3-4 h and then shifted to 21 °C for overnight (~10 h) expression. Cells were harvested by centrifugation, collected in 50 mL tubes and frozen in liquid nitrogen. For purification of His₆-TglB and His₆-CysRS, cells were thawed, resuspended in lysis buffer (50 mM HEPES, 100 mM NaCl, pH 7.5; 30 mL per 10 g wet cell paste) and lysed by treatment with lysozyme (100 µg/mL) and sonication (3 min active time; 1 s pulse, 2 s rest at 60 % max amplitude using a 1 cm tip). Proteins were purified by immobilized metal affinity chromatography (IMAC). TCEP (1 mM) was added to maintain reduced thiols. The eluate was concentrated to 2.5 mL in a 30 kDa MWCO centrifuge filter and desalted on a PD-10 size-exclusion column (GE Healthcare Life Sciences). Protein was separated into aliquots and stored at -78 °C. Protein purity was judged by SDS-PAGE (See Supplementary Information).

Protein expression and purification for TglHI. TglHI was obtained from electrocompetent *E. coli* BL21(DE3) cells (New England Biolabs) transformed with pET15b-*tglHI*. A 25 mL inoculum was added to 1 L of LB medium containing 100 µg/mL ampicillin. Cells were grown with vigorous shaking at 37 °C for 3-4 h until the optical density (OD) at 600 nm reached 0.6 at which point the culture was cooled to 0 °C and induced with 0.2 mM IPTG. Upon induction, the cell culture was incubated at 18 °C for 11 h with vigorous shaking. Cells were harvested by centrifugation, collected in 50 mL tubes and suspended in lysis buffer (20 mM Tris, 300 mM NaCl, 10% glycerol, pH 7.6,) supplemented with 1 mg/mL lysozyme, 600 U DNase, and 1 mM TCEP. The cells were lysed by passage through a French pressure cell twice and cell debris was removed by centrifugation (30,000 rcf) for 50 min at 10 °C. The supernatant was loaded onto a column containing 5 mL of Ni-NTA resin previously equilibrated with lysis buffer. After equilibration of the resin with the lysate by orbiting for 30 min, the flow-through was discarded. The resin was washed with 2 x 40 mL of wash buffer (25 mM imidazole, 20 mM Tris, 300 mM NaCl, 10% glycerol, pH 7.6,) followed by elution with elution buffer (250 mM imidazole, 20 mM Tris, 300 mM NaCl, 10% glycerol, pH 7.6.). The eluate was concentrated to 2.5 mL in a 30 kDa MWCO centrifuge filter and desalted on a PD-10 size-exclusion column (GE Healthcare Life Sciences). Protein was separated into aliquots and stored at -78 °C. The protein was judged pure by SDS-PAGE (See Supplementary Information).

Protein expression and purification for AmmB1, AmmB2, AmmB3, AmmB4 and TrpRS. AmmB1, AmmB2, AmmB3, AmmB4, and TrpRS were obtained from electrocompetent *E. coli* BL21(DE3) cells (New England Biolabs) transformed with pET28a-AmmB1, pET28a-AmmB2, pET28a-AmmB3, pET28a-AmmB4 and pET28a-TrpRS respectively. A 25 mL inoculum was added to 1 L of LB medium containing 50 µg/mL kanamycin. Cells were grown with vigorous shaking at 37 °C for 3-4 h until the optical density (OD) at 600 nm reached 0.6 at which point the culture was cooled to 0 °C and induced with 0.5 mM IPTG. Upon induction, the cell culture was incubated at 18 °C for 14 h with vigorous shaking. Cells were harvested by centrifugation, collected in 50 mL tubes and suspended in lysis buffer (50 mM HEPES, 100 mM NaCl, 10 mM imidazole, 10% glycerol, pH 7.6,) supplemented with 1 mg/mL lysozyme, 600 U DNase, and 1 mM TCEP. The cells were

lysed by passage through a French pressure cell twice and cell debris was removed by centrifugation (30,000 rcf) for 50 min at 10 °C. The supernatant was loaded onto a column containing 5 mL of Ni-NTA resin previously equilibrated with lysis buffer. After equilibration of the resin with the lysate by orbiting for 30 min, the flow-through was discarded. The resin was washed with 2 x 40 mL of wash buffer (25 mM imidazole, 50 mM HEPES, 100 mM NaCl, 10% glycerol, pH 7.6,) followed by elution with elution buffer (500 mM imidazole, 50 mM HEPES, 100 mM NaCl, 10% glycerol, pH 7.6,). The eluate was concentrated to 2.5 mL in a 30 kDa MWCO centrifuge filter and desalted on a PD-10 size-exclusion column (GE Healthcare Life Sciences). Protein was separated into aliquots and stored at -78 °C. The protein was judged pure by SDS-PAGE (See Supplementary Information).

Peptide purification and LC-MS. Following IMAC and C4 desalting, TglA was purified using an Agilent 1200 HPLC. HPLC buffers were 0.1% TFA in water (HPLC buffer A) and 0.1% TFA in acetonitrile (HPLC buffer B). The desalted material was lyophilized, resuspended in buffer A, and applied to a Phenomenex Luna C18 column equilibrated with 95% HPLC buffer A, 5% HPLC buffer B. A linear gradient to 100% HPLC buffer B was run over 20 min at 0.1 mL/min. TglA eluted near the end of the gradient. TglA-Cys behaved similarly and eluted slightly later. Full-length and AspN or trypsin-digested peptides were analyzed by LC-MS/MS on a Waters Synapt Q-TOF equipped with a Waters UPLC and Phenomenex C18 Luna or C4 Jupiter columns. Buffers for LC-MS were 0.1% formic acid in water (LC-MS buffer A) or 0.1% formic acid in LC-MS grade acetonitrile (LC-MS buffer B).

Analysis of metal concentration. TglHI was purified by size-exclusion chromatography (SEC) using a Superdex 200 column. The protein complex eluted with tris buffer (20 mM Tris, 300 mM NaCl, 10% glycerol, pH 7.6) at 1 mL/min flow rate. TglHI was then reconstituted in an anaerobic chamber with 1 equivalent $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2$ in (20 mM Tris, pH 7.6) with incubation for 10 min on ice. Excess iron was removed from TglHI by SEC using a PD10 column prior to iron analysis. Iron quantification of TglHI was determined using Ferene S as a spectrophotometric dye as reported by Hennessy and co-workers (32). A standard curve was generated using an iron standard in 2% HNO_3 solution (Claritas PPT).

Chemical synthesis of CxSAM. CxSAM was synthesized and purified as previously described (9). In a 2.5 mL tube, S-adenosylhomocysteine (3 mg) and iodoacetic acid (0.1 g) were dissolved in 150 mM ammonium bicarbonate (0.5 mL) and incubated at 37 °C for 12 h with shaking. The resulting mixture was analyzed by LC-MS and purified by analytical C18 HPLC. CxSAM is poorly retained on C18 and eluted from the column at 5 min in 0% HPLC buffer B (~3.5 min dead time). HPLC fractions were analyzed by LC-MS and had characteristic fragments for S-carboxymethyl thioadenosine.

In vitro assay of TglB. In a 2.5 mL centrifuge tube, reaction components were combined to the following optimized final concentrations: TglA (5 μM), TglB (0.5 μM), CysRS (0.5 μM), ATP (5 mM), L-cysteine (2 mM), TCEP (1 mM) *E. coli* cell extract (1:20 dilution), HEPES assay buffer (50 mM HEPES pH 7.5, pH adjusted with NaOH; 10 mM

MgCl₂, 100 mM NaCl). The assay mixture was incubated at 30 °C for 2 h, desalted and concentrated by ZipTip, eluted in 80% acetonitrile, 20% water, 0.1% TFA, and analyzed by MALDI-TOF MS. H₂¹⁸O reactions were performed in a similar manner with the following changes: all enzyme components were initially withheld from the reaction mixture, the incomplete mixture was lyophilized to remove unlabeled water, the proper volume was reconstituted with H₂¹⁸O, and the enzymes were added at the end to initiate the reaction. The residual unlabeled H₂O was estimated to be ~15 %. For ³¹P NMR analysis, reaction components were combined to final concentrations as follows: TglA (50 μM), TglB (5 μM), Cys-tRNA^{Cys} (50 μM), ATP (0.5 mM), Tris assay buffer (50 mM Tris pH 7.5, 5 mM MgCl₂, 100 mM NaCl). Cys-tRNA^{Cys} was obtained using a similar procedure for preparation of Glu-tRNA^{Glu} reported previously (6).

To determine the specific activity of TglB, the enzyme (0.1 μM) was incubated with TglA (50 μM), CysRS (10 μM), tRNA^{Cys} (10 μM), ATP (5 mM), L-cysteine (5 mM), and TCEP (1 mM) in reaction buffer (50 mM HEPES, pH 7.5). The reaction was allowed to proceed at 25 °C, and aliquots were quenched by addition of acetonitrile to 50% (v/v) at set time points (30, 60, 90, and 120 s). At 120 s the conversion was about 11%. Product formation over time was linear with a velocity of 28.2 min⁻¹. To assure that under these conditions the reaction catalyzed by CysRS to form Cys-tRNA^{Cys} was not rate limiting, the amounts of CysRS and tRNA^{Cys} were doubled, and no change in velocity was observed. Similarly, to assure that the substrate was saturating, the concentration of TglA was doubled without observing a change in velocity.

The hydroxylamine (NH₂OH) quenching assay was performed by incubating 50 μM TglB and 50 μM TglA in the presence or absence of ATP (5 mM) for 5 min at 30 °C. Then NH₂OH was added to the reaction to a final concentration of 1 M and the mixture was incubated at 30 °C for another 20 min. The assay was then analyzed by both MALDI-TOF-MS and ESI-MS.

In vitro assay of TglHI. In a 2.5 mL centrifuge tube, TglA-Cys (100 μM) was added to TglHI (10 μM) in phosphate buffer (0.2 mL, 50 mM Na₂HPO₄, 300 mM NaCl, 10% glycerol, pH 7.6). The reaction vessel was left open to air at 24 °C for 16 h. At this time, the reaction was directly desalted and concentrated by ZipTip, eluted in 80% acetonitrile, 20% water, 0.1% TFA, and analyzed by MALDI-TOF MS (Fig. 2B). TglA-CysAla, TglA-Cys truncants, Hala2 and ProcA2.8 were subjected to identical reaction conditions as mentioned above. For the reaction at low oxygen concentrations, TglHI was subjected to buffer exchange with degassed phosphate buffer ten times using a 10 kDa MWCO Amicon 0.5 mL spin filter in an anaerobic chamber maintained at < 1.0 ppm oxygen. TglA-Cys was also subjected to buffer exchange with degassed phosphate buffer using a 3 kDa MWCO Amicon 0.5 mL spin filter in the anaerobic chamber. TglA-Cys (100 μM) was added to TglHI (10 μM) in phosphate buffer (0.2 mL, 50 mM Na₂HPO₄, 300 mM NaCl, 10% glycerol, pH 7.6) and incubated in the anaerobic chamber. After 1 h, the enzyme was first deactivated with 1 M formic acid before exposing the reaction mixture to air. A separate TglHI reaction under aerobic conditions was set up in parallel as a control experiment. The reaction mixtures were analyzed as described above (fig. S6).

To determine the specific activity of TglHI, a coupled assay with formate dehydrogenase from *Candida boidinii* was developed. The TglHI enzyme (5 μM) was

incubated with TglA-Cys (100 μM), $\beta\text{-NAD}^+$ (200 μM), and formate dehydrogenase (8 U) in reaction buffer (20 mM Tris, pH 7.5). The reaction was allowed to proceed at 25 $^{\circ}\text{C}$, and reaction progress was monitored by UV-Vis spectroscopy with continuous monitoring of absorbance at 340 nm which corresponds to production of NADH. Product formation over time was linear with a velocity of 1.1 min^{-1} . To assure that under these conditions the reaction catalyzed by formate dehydrogenase was not rate limiting, the amounts of formate dehydrogenase were doubled, and no change in velocity was observed. Similarly, to assure that the substrate was saturating, the concentration of TglA-Cys was doubled without observing a change in velocity.

In vitro assay of TglF. In a 2.5 mL centrifuge tube, peptide 1 (40 μM) and Cx-SAM (120 μM) was added to TglF (10 μM) in HEPES buffer (0.5 mL, 50 mM HEPES, 100 mM NaCl, pH 7.6). After 13 h, the reaction was directly desalted and concentrated by ZipTip, eluted in 80% acetonitrile, 20% water, 0.1% TFA, and analyzed by MALDI-MS (Fig. 2C).

Chemical analysis of modified TglA peptide. TglBHI-modified TglA was first treated with trypsin as described under *Peptide expression and purification*. Peptide fragments were esterified using methanolic HCl(33). First 160 μL of acetyl chloride was added dropwise to 1 mL of anhydrous methanol while stirring in an ice bath. Then 100 μL of this mixture was added to dry peptide and incubated at room temperature for 4 h. The esterification mixture was diluted with 400 μL of water, frozen in liquid nitrogen, and lyophilized. The modified peptides were analyzed by LC-MS as described above.

Proteolytic activity of GFP-TglG membrane fraction. For protein expression of GFP-TglG, *E. coli* BL21(DE3) cells (New England Biolabs) transformed with pet28b-GFP-tglG were grown with 50 $\mu\text{g}/\text{mL}$ kanamycin in LB media (for 1 L total: 10 g bacto tryptone, 5 g yeast extract, 10 g NaCl). A 25 mL starting culture was inoculated with a single colony and incubated at 37 $^{\circ}\text{C}$ for 14 h followed by 40x dilution into 1 L of LB containing 50 $\mu\text{g}/\text{mL}$ kanamycin. The culture was kept at 37 $^{\circ}\text{C}$ until optical density at 600 nm reached 0.6-0.8. Then, the flask was placed in an ice/water bath for 30 min before the addition of IPTG to a final concentration of 1 mM. The culture was then incubated for an additional 14 h at 18 $^{\circ}\text{C}$. At this time, the cells were collected by centrifugation at 11,270 rcf for 20 min. Cells were harvested, resuspended in 25 mL of Tris lysis buffer (20 mM Tris, 300 mM NaCl, 10 mM imidazole, 10% glycerol, pH 7.6), and 10 mg of lysozyme, TCEP (1 mM) and 600 U DNase were added. Cells were lysed by sonication, and the cell lysate (0.4 mL) was added to TglA-thiaGlu (2, prepared separately) in 0.1 mL HEPES buffer (50 mM HEPES, 100 mM NaCl, pH 7.6) at rt. After 16 h, the reaction was centrifuged at 16,100 rcf for 5 min to pellet cell debris, and the supernatant was directly desalted and concentrated by ZipTip, eluted in 80% acetonitrile, 20% water, 0.1% TFA, and analyzed by MALDI-MS (fig. S7). TglA-Glu and TglA-GluAla were subjected to identical reaction conditions as mentioned above.

Location of TglG proteolysis. PmaG is a member of the family of site-2 proteases (S2P), which are zinc-dependent metalloproteases containing six transmembrane helices. Notably, the structure of an active core fragment of a related S2P protease, mjS2P, from

Methanocaldococcus jannaschii was previously determined by X-ray diffraction (20). The active site of mjs2P consists of a zinc atom coordinated by two histidines and one aspartate residue and appears accessible only from the cytosol (20). Sequence alignment between PmaG and mjs2P reveals that the active site residues are conserved in PmaG and are located on the same transmembrane helices. This finding suggests that the active site of PmaG is also on the cytosolic side. This conclusion is also supported by transmembrane prediction analysis using TMHMM 2.0 (34, 35).

Isotope Labeling Experiments. ^{13}C -cysteine labeled His₆-TglA-Cys was expressed using *E. coli* strain JW3582-2. *E. coli* JW3582-2, an auxotrophic strain for cysteine, was purchased from the Coli Genetic Stock Center at Yale University, <https://cgsc2.biology.yale.edu/Strain.php?ID=108920>. A lysogenization step was performed to JW3582-2 so that the host strain could be used to express target genes cloned in T7 expression vectors. Lysogenization was performed using the λ DE3 Lysogenization Kit (Novagen) as per the manufacturer's instructions. The lysogenized JW3582-2 was transformed with a pACYCDuet-1 plasmid encoding His₆-TglA-Cys, TglH and TglI. Expression of ^{13}C labeled, modified His₆-TglA-Cys was performed in modified M9 minimal media. A starter culture of *E. coli* BL21(DE3) cells containing a pACYCDuet-1 plasmid encoding His₆-TglA-Cys, TglH and TglI were grown overnight at 37 °C in LB containing 25 $\mu\text{g}/\text{mL}$ kanamycin and 20 $\mu\text{g}/\text{mL}$ chloramphenicol. After harvesting the cells, the supernatant was discarded and the cells were washed with 5 mL of wash buffer (22 mM KH_2PO_4 , 42 mM Na_2HPO_4 , 8.5 mM NaCl, pH 7.4). After washing, the cells were resuspended in wash buffer and used to inoculate (1:200) modified M9 minimal media, with the following composition per 100 mL: 10 mL of a 10x minimal media (220 mM KH_2PO_4 , 420 mM Na_2HPO_4 , 85 mM NaCl, pH 7.4), 0.3 mL of 40% aqueous $(\text{NH}_4)_2\text{SO}_4$, 2 mL of 20% aqueous glucose, 0.1 mg of FeSO_4 , 10 μg of thiamine, 200 μL of 1 M MgSO_4 , 10 μL of 1 M CaCl_2 , and 75 μL of a trace element solution (5 mM CaCl_2 , 1.25 mM ZnCl_2 , 260 μM $\text{CuCl}_2 \cdot \text{H}_2\text{O}$, 252 μM $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 250 μM $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, pH 7.4). L-cysteine or U- ^{13}C -L-cysteine or 3- ^{13}C -L-cysteine (1 mM) was added as the sole cysteine source. The media also contained 10 $\mu\text{g}/\text{mL}$ chloramphenicol and 25 $\mu\text{g}/\text{mL}$ kanamycin. The cells were grown at 37 °C and induced at $\text{OD}_{600} = 0.6\text{-}0.8$ by the addition of IPTG to a final concentration of 1 mM and grown for an additional 3 h at 37 °C before harvesting. Peptide 1 was purified as described above. After modification with iodoacetic acid (20 mM) and trypsin digestion, the VFA-thiaGlu tetrapeptide was analyzed by LC/MS (fig. S4). For preparation of ^{13}C -enriched TglA-Cys for *in vitro* TglHI reaction and ^{13}C NMR spectroscopy (Fig. 3), the above procedure was modified to use pACYC-Duet1 plasmid encoding only His₆-TglA-Cys, and in the purification phosphate buffer (50 mM Na_2HPO_4 , 300 mM NaCl, pH 7.5) was used instead of HEPES. For preparation of TglA-d₃-Cys for *in vitro* TglHI reaction, the above procedure was modified to use pACYC-Duet1 plasmid encoding only His₆-TglA-Cys and d₃-L-cysteine (1 mM) was added as the sole cysteine source. The deuterated peptide was modified by TglHI and analyzed under conditions described above.

In vitro assay of AmmB2. In a 2.5 mL centrifuge tube, reaction components were combined to the final concentrations: AmmA* (50 μM), AmmB2 (5 μM), TrpRS from *Streptomyces sp.* CNR698 (50 μM), ATP (5 mM), L-tryptophan (2 mM), TCEP (1 mM),

tRNA^{Trp} from *Streptomyces sp.* CNR698 (20 μ M), HEPES assay buffer (50 mM HEPES pH 7.5, pH adjusted with NaOH; 10 mM MgCl₂, 100 mM NaCl). The assay mixture was incubated at room temperature for 2 h, desalted and concentrated by ZipTip, eluted in 80% acetonitrile, 20% water, 0.1% TFA, and analyzed by MALDI-TOF MS. tRNA^{Trp} was obtained using a similar procedure for preparation of tRNA^{Glu} reported previously (6). AmmA*W was digested with LysC at 37 °C for 2 h, and the fragments were analyzed by LC-MS/MS.

Bioinformatics for operon prediction. The operon containing the *tgl* gene cluster was predicted by the Softberry FGENESB program (Softberry, Inc., Mount Kisco, NY) (<http://www.softberry.com/>) (36).

MicroED structure determination:

Sample Preparation for MicroED. ~1-5 mg of the D-3 peptide was turned into a fine powder by placing the sample in between two 25-mm glass slide cover slips and rubbing them together. A pre-clipped Quantifoil R1.2/1.3 Cu300 mesh grid was then gently placed on the fine powder, tapped very gently to ensure the sample bound to the grid, and then tapped to remove any excess before placing the grid in liquid nitrogen and inserting it in the sample cartridge followed by transfer to the microscope.

Data collection. Data sets were collected as previously described (11, 12) with minor changes to settings and the use of a faster, and more sensitive direct electron detector (15). Briefly, data was collected using a Thermo-Fischer Talos Artica electron cryomicroscope operating at an acceleration voltage of 200 keV at a wavelength of 0.0251 Å. Crystals were identified in overfocus diffraction mode and data collected from each nanocrystal using an ultra-low dose of less than 0.01 e⁻/Å². The stage was continuously rotated at a rate of ~0.9 °/s and ~50° worth of data collected from each nanocrystal at a rate of 1 second per frame on a bottom mount Falcon III direct electron detector with a diffraction length of 1100mm. Data from 7 nanocrystals was collected on the Falcon III.

The images were collected as MRC files and later converted to SMV format using an in-house developed software (15). Frames were indexed and integrated in XDS, five datasets were then scaled and merged using XSCALE as described before (37, 38). The intensities were converted to SHELX format using XDSCONV (37, 38). The structure was then solved by direct methods using SHELXT, and refined in SHELXL as previously described (39, 40). MicroED data, crystallographic processing and refinement statistics are indicated in the crystallographic Table (Table S3).

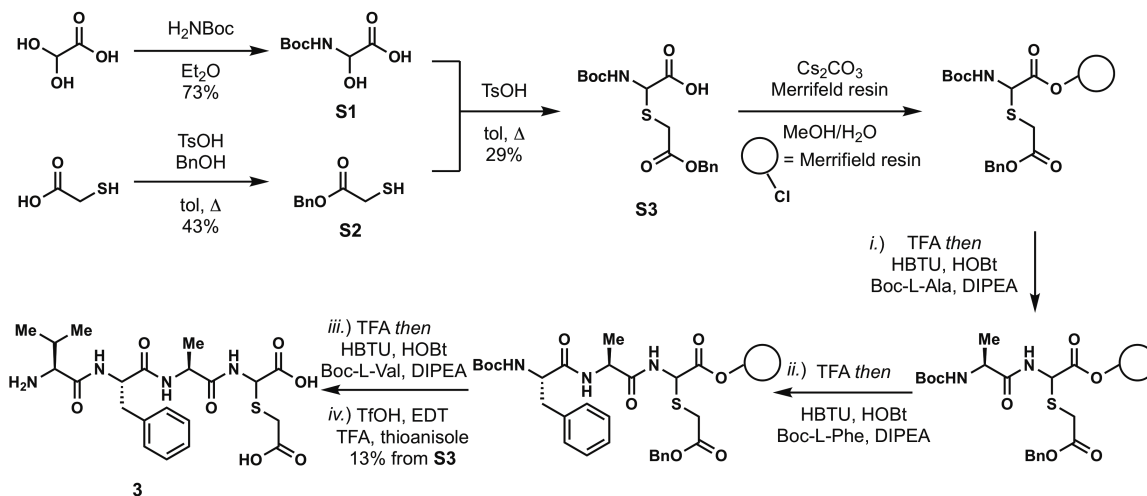
General procedures for chemical synthesis

Unless stated otherwise, all reactions were performed in flame-dried glassware under an atmosphere of dry nitrogen or argon. Dry dichloromethane, and dimethylformamide (DMF) were obtained by passing these degassed solvents through activated alumina columns. All other reagents were used as received from commercial sources, unless stated otherwise. Reactions were monitored by thin layer chromatography (TLC) on Silicycle Siliaplate glass-backed TLC plates (250 μ m thickness, 60 Å porosity, F-254 indicator) and visualized by UV irradiation or development with anisaldehyde or

phosphomolybdic/cerium sulfate stain. Volatile solvents were removed under reduced pressure with a rotary evaporator. All flash column chromatography was performed using Silicycle SiliaFlash F60, 230-400 mesh silica gel (40-63 μm).

^1H NMR and ^{13}C NMR spectra were recorded on an Agilent 600 MHz spectrometer for ^1H (150 MHz for ^{13}C) in D_2O . Chemical shifts are reported relative to the residual solvent signal (^1H NMR: $\delta = 4.79$). NMR data are reported as follows: chemical shift (multiplicity, coupling constants where applicable, number of hydrogens). Splitting is reported with the following symbols: s = singlet, bs = broad singlet, d = doublet, t = triplet, app t = apparent triplet, dd = doublet of doublets, ddd = doublet of doublet of doublets, dt = doublet of triplets, hept = heptet, m = multiplet. Infrared spectra (IR) were recorded as a thin film on a Perkin-Elmer FT-IR system and peaks were reported in cm^{-1} . Mass spectrometry analyses were performed using a SYNAPT ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters).

Scheme S1. Chemical synthesis of VFA-thiaGlu.



Hemiaminal S1. **S1** was prepared on multi-gram scale according to the procedure reported by Rivier and co-workers (41). An oven-dried 250 mL round-bottom flask was charged with glyoxylic acid monohydrate (3.6 g, 39 mmol, 1.1 equiv) and *tert*-butyl carbamate (4.1 g, 35 mmol, 1.0 equiv). The reaction vessel was evacuated and backfilled with nitrogen and this process repeated for a total of three times. Diethyl ether (25 mL) was added and the reaction mixture was stirred for 14 h at rt. The reaction mixture was filtered and concentrated *in vacuo*. The crude residue was dissolved in EtOAc and triturated with hexanes to afford **S1** (4.9 g, 73% yield) as a yellow-brown solid. Spectral data were in agreement with that previously reported by Rivier and co-workers (41).

Ester S2. A 250 mL round-bottom flask equipped with a Dean-Stark trap and reflux condenser was charged with thioglycolic acid (3.0 g, 33 mmol, 1.0 equiv), benzyl alcohol (5.3 g, 49 mmol, 1.5 equiv) and *p*-toluenesulfonic acid monohydrate (0.63 g, 3.3 mmol, 0.1 equiv). The vessel was evacuated and filled with nitrogen and this process repeated for a total of three times. Toluene (100 mL) was added and the reaction mixture was

heated to reflux and maintained at this temperature for 14 h. The reaction mixture was cooled to rt, and then concentrated *in vacuo*. The crude residue was purified by column chromatography (5% → 7.5% EtOAc in hexanes) to afford ester **S2** (2.6 g, 43% yield) as a colorless oil. Spectra data were in agreement with that previously reported (42).

Thioether S3. A 250 mL round-bottom flask equipped with a Dean-Stark trap and reflux condenser was charged with hemiaminal **S1** (0.5 g, 2.6 mmol, 1.0 equiv), ester **S2** (1.2 g, 6.6 mmol, 2.5 equiv) and *p*-toluenesulfonic acid monohydrate (25 mg, 0.13 mmol, 0.05 equiv). The reaction vessel was evacuated and backfilled with nitrogen and this process repeated for a total of three times. Toluene (100 mL) was added and the reaction mixture was heated to reflux and maintained at this temperature for 6 h. The reaction mixture was cooled to rt, and then concentrated *in vacuo*. The crude residue was purified by column chromatography (10% → 20% MeOH in CH₂Cl₂) to afford thioether **S3** (269 mg, 29% yield) as a brown oil: ¹H NMR (600 MHz, D₂O) δ 7.52 – 7.42 (m, 5H), 5.26 (d, J = 12.3 Hz, 1H), 5.23 (d, J = 12.3 Hz, 1H), 5.18 (bs, 1H), 3.52 (bs, 2H), 1.43 (s, 9H); ¹³C NMR (150 MHz, D₂O) δ 173.3, 172.5, 156.1, 135.2, 128.8, 128.6, 128.2, 81.6, 67.8, 58.8, 31.9, 27.5; IR (thin film) 3415, 3055, 2984, 1717, 1496, 1455 cm⁻¹; HRMS (ESI) calcd for[C₁₆H₂₂NO₆S]⁺ (M+H)⁺: m/z 356.1168 found 356.1156.

VFA-thiaGlu 3. **S3** was loaded on Merrifield resin using the Gisin Method (43). In a 20 mL scintillation vial, **S3** (77.5 mg, 0.22 mmol, 1 equiv) was dissolved in methanol (5 mL) and water (0.5 mL). The solution was basified to pH 7.0 with 20% aqueous Cs₂CO₃ and concentrated *in vacuo*. To the residue was added DMF (2.5 mL) and then concentrated *in vacuo*, and this process was repeated two times. In a separate 100 mL round-bottom flask, Merrifield resin (442 mg, 1-1.5 mmol/g Cl⁻ loading, 2% cross linked) was added, followed by DMF (8 mL) and the mixture was gently stirred for 10 min. The cesium salt of **S3** was added to the flask with the resin and the mixture was kept at 50 °C for 18 h. The resin was filtered with a fritted glass funnel by vacuum filtration and washed successively with DMF, DMF/H₂O (1:1), MeOH/H₂O (1:1) and then MeOH. After the resin was dry, 50% TFA in CH₂Cl₂ (1 mL) was added on the fritted glass filter. After 3 min, the 50% TFA in CH₂Cl₂ was removed by vacuum filtration, and the resin was washed with CH₂Cl₂ (3 x 2 mL) followed by diisopropylethylamine (3 x 2 mL, 5% v/v in CH₂Cl₂). To the resin was then added Boc-L-Ala-OH (125 mg, 0.66 mmol, 3 equiv) that was premixed with HOBt (89 mg, 0.66 mmol, 3 equiv), HBTU (250 mg, 0.66 mmol, 3 equiv), and diisopropylethylamine (0.12 mL, 0.66 mmol, 3.0 equiv) in DMF (8 mL) for 8 min. The reaction was mixed with a spatula for 15 min or until the Kaiser test was negative. The solution was filtered by vacuum and washed with DMF (3 x 5 mL). The deprotection protocol with TFA was repeated followed by the addition of Boc-L-Phe-OH (175 mg, 0.66 mmol, 3 equiv) using the elongation protocol described above. The deprotection protocol with TFA was repeated followed by the addition of Boc-L-Val-OH (143 mg, 0.66 mmol, 3 equiv) with the elongation protocol. After drying, the resin was transferred to a 100 mL oven-dried flask. The reaction vessel was evacuated and backfilled with nitrogen and this process repeated for a total of three times. After the reaction mixture was cooled to 0 °C, TFA (10 mL) was added followed by the addition of thioanisole (1 mL, 7.6 mmol, 35 equiv), 1,2-ethanedithiol (0.5 mL, 4.7 mmol, 21 equiv), and trifluoromethanesulfonic acid (1 mL, 3.9 mmol, 18 equiv). The reaction mixture was

warmed to rt over the course of 2 h. At this time, the mixture was filtered and the solid was washed with TFA (2 x 3 mL). To the filtrate was added Et₂O (100 mL) and extracted with DI H₂O (3 x 25 mL). The combined aqueous layers were washed with Et₂O (2 x 50 mL), basified with conc. aqueous NH₄OH to pH 4 and then concentrated in *vacuo*. The crude residue was partially purified on a CombiFlash® Rf+ equipped with a 50 g RediSep Rf Gold C18Aq column. Acetonitrile and 0.1% trifluoroacetic acid in H₂O were the mobile phases, and a gradient of 0-100% aq. MeCN was applied over 15 min at 40 mL/min flow rate to afford 14 mg (13% yield) of a white crystalline solid containing a 1:1 mixture of diastereomers. The combined fractions containing **3** were further purified with an XBridge C18 column (250 x 10 mm, 5 μM particle size) to separate the diastereomers. Acetonitrile and water were used as the mobile phases, and a gradient of 5-40% aq. MeCN was applied over 20 min at 5 mL/min flow rate. The diastereomer containing a D-thiaGlu residue as determined by microED (**D-3**) eluted first at 6.8 min followed by elution of the diastereomer containing a L-thiaGlu residue (**L-3**) at 8.0 min. Both diastereomers were obtained as white crystalline solids:

D-3:

¹H NMR (600 MHz, D₂O) δ 7.42 – 7.35 (m, 2H), 7.35 – 7.28 (m, 3H), 5.21 (s, 1H), 4.69 (t, J = 7.7 Hz, 1H), 4.40 (q, J = 7.2 Hz, 1H), 3.38 (d, J = 15.0 Hz, 1H), 3.29 (d, J = 15.0 Hz, 1H), 3.20 – 3.15 (m, 2H), 3.06 (dd, J = 13.8, 8.5 Hz, 1H), 1.92 – 1.82 (m, 1H), 1.38 (d, J = 7.1 Hz, 3H), 0.87 (d, J = 6.9 Hz, 3H), 0.84 (d, J = 6.8 Hz, 3H); ¹³C NMR (125 MHz, D₂O) δ 176.9, 176.6, 173.2, 172.7, 172.5, 136.2, 129.1, 128.8, 127.2, 60.0, 56.6, 54.7, 49.4, 37.1, 35.3, 31.7, 18.4, 16.9; HRMS (ESI) calcd for [C₂₁H₃₁N₄O₇S]⁺ (M+H)⁺: m/z 483.1914 found 483.1898. The structure for the **D-3** was determined by microcrystal electron diffraction (MicroED).

L-3:

¹H NMR (600 MHz, D₂O) δ 7.43 – 7.37 (m, 2H), 7.37 – 7.30 (m, 3H), 5.20 (s, 1H), 4.72 (dd, J = 8.9, 6.5 Hz, 1H), 4.38 (q, J = 7.2 Hz, 1H), 3.39 (d, J = 14.8 Hz, 1H), 3.33 – 3.28 (m, 2H), 3.21 (dd, J = 14.0, 6.5 Hz, 1H), 3.05 (dd, J = 13.9, 9.2 Hz, 1H), 1.98 – 1.90 (m, 1H), 1.38 (d, J = 7.2 Hz, 3H), 0.87 (dd, J = 19.5, 6.8 Hz, 6H); ¹³C NMR (125 MHz, D₂O) δ 177.0, 174.8, 173.3, 172.8, 172.5, 136.3, 129.2, 128.8, 127.2, 59.6, 56.7, 54.7, 49.6, 37.1, 35.5, 31.3, 18.3, 16.8, 16.7; HRMS (ESI) calcd for [C₂₁H₃₁N₄O₇S]⁺ (M+H)⁺: m/z 483.1914 found 483.1913. **L-3** was identical by NMR and HPLC to the tetrapeptide produced by TgIHI and IAA modification of TgIA-Cys.