Supplementary Information

Insights into the Thioamidation of Thiopeptins to Enhance the Understanding of the Biosynthetic Logic of Thioamide-containing Thiopeptides

Jingyu Liu,‡[a,b] Zhi Lin, ‡[a,c] Yuqing Li,[a] Qingfei Zheng,[a] Dandan Chen[a,d] and Wen Liu*[,a,d]

[a] State Key Laboratory of Bioorganic and Natural Products Chemistry, Center for Excellence on Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China. [b] Innovation Research Institute of Traditional Chinese Medicine (IRI), Shanghai University of Traditional Chinese Medicine, 1200 Cailun Road, Shanghai 201203, China.[c] State Key Laboratory of Microbial Metabolism, School of Life Science & Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China. [d] Huzhou Center of Bio-Synthetic Innovation, 1366 Hongfeng Road, Huzhou 313000, China.

‡These Authors equally contributed to this work.

* To whom correspondence should be addressed: Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Rd., Shanghai 200032, China. Wen Liu, Email: wliu@mail.sioc.ac.cn, Tel: 86-21-54925111, Fax: 86-21-64166128.

Experimental

Materials, bacterial strains, and plasmids. Biochemicals and media were purchased from Sinopharm Chemical Reagent Co., Ltd. (China), Oxoid Ltd. (U.K.) or Sigma-Aldrich Co. LLC. (USA) unless otherwise stated. Enzymes were purchased from Takara Biotechnology Co. Ltd. (China). Restriction endonucleases were purchased from Thermo Fisher Scientific Co. Ltd. (USA). Chemical reagents were purchased from standard commercial sources. The bacterial strains and plasmids used in this study are listed in **Supplementary Table 1**.

DNA isolation, manipulation, and sequencing. DNA isolation and manipulation in *E. coli* or actinobacteria were carried out according to standard methods.[1] PCR amplifications were carried out on an Applied Biosystems Veriti Thermal Cycler using either Taq DNA polymerase (Vazyme Biotech Co. Ltd, China) for routine genotype verification or Kod DNA polymerase (Takara Biotechnology Co., Ltd.) or PrimeSTAR HS DNA polymerase (Takara Biotechnology Co., Ltd. Japan) for high fidelity amplification. Primer synthesis was performed at Shanghai Sangon Biotech Co. Ltd. (China). DNA sequencing was performed at Shanghai Majorbio Biotech Co. Ltd. (China). Primers used for diagnostic PCR are listed in **Supplementary Table 2**.

Sequence analysis. Biosynthetic gene clusters (BGCc) were mined from microbial genomes using the AntiSMASH web tool.[2] Open reading frames (ORFs) were identified using the FramePlot 4.0beta program (http://nocardia.nih.go.jp/fp4/).[3] The deduced proteins were compared with other known proteins in the databases using available BLAST methods (http://blast.ncbi.nlm.nih.gov/Blast.cgi).[4] Amino acid sequence alignments were performed using Vector NT1 software and ESPript (http://espript.ibcp.fr/ESPript/ESPript/).[5] Sequence similarity networks (SSNs) were generated using the EFI-EST webtool (http://efi.igb.illinois.edu/efi-est/).[6]

General chemical analysis. High performance liquid chromatography (HPLC) analysis was carried out on an Agilent 1260 HPLC system (Agilent Technologies Inc., USA) equipped with a DAD detector. Semi-preparative HPLC was performed on an Agilent 1100 system. HPLC Electrospray ionization MS (HPLC-ESI-MS) and tandem MS (MS/MS) were performed on a Thermo Fisher LTQ Fleet ESI-MS spectrometer (Thermo Fisher Scientific Inc., USA), and the data were analyzed using Thermo Xcalibur software. ESI-high resolution MS (ESI-HR-MS) analysis was carried out on an instrument consisting of a 1260 HPLC system or a 6538 UHD quadrupole time of flight (QTOF) high resolution mass spectrometry (Agilent Technologies, Santa Clara, USA). NMR data were recorded on an Agilent 500 MHz PremiumCompact+ NMR spectrometer (Agilent Technologies Inc., USA) or on the BrukerDRX400 and Bruker AV500spectrometers (Bruker Co. Ltd, Germany).

Gene Inactivation of $TppX_2$. The genomic DNA of the *Streptomyces tateyamensis* ATCC21389 wild-type strain served as the template for PCR amplification unless otherwise stated. To exclude polar effects on downstream gene expression, gene inactivation in *S. tateyamensis* ATCC21389 was performed by in-frame deletion. For $tppX_2$ deletion, the 2-kb fragment obtained using primers $tppX_2$-L-for and $tppX_2$-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1001. The 2-kb fragment obtained using primer $tppX_2$-R-for and $tppX_2$-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1001 to yield the recombinant plasmid pLL1002, in which a 807-bp in-frame coding region of $tppX_2$ was deleted. To transfer pLL1002 into the thiopeptin-producing strain S. *tateyamensis*, conjugation between *E. coli* ET12567-Streptomyces was carried out following the standard procedure. The colonies that were apramycin resistant at 37 ℃ were identified as integrating mutants, in which a single-crossover homologous recombination event took place. These mutants were cultured for several rounds in the absence of apramycin, and the resulting apramycin-sensitive isolates were subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 0.7-kb product when using primers $tppX_2$-DC-for and $tppX_2$-DC-rev. Further sequencing of this PCR product confirmed the genotype of SL101, in which $tppX_2$ was in-frame deleted.

For heterologous complementation of $tppX_1$, $tppX_2$ *and* $tppX_1$&$tppX_2$ in TSR-producing strain *S. laurentii*, the related gene-containing fragments were amplified by PCR using primer pairs $tppX_1$-C-for/ $tppX_1$-C-rev, $tppX_2$-C-for/ $tppX_2$-C-rev, $tppX_2$-C-for/ $tppX_1$-C-rev, respectively, and then cloned into *PermE**-containing pSET152to yield pLL1003, pLL1004 and pLL1005, respectively, in which each gene was under the control of the constitutive promoter *PermE**, which is a constitutive promoter responsible for expressing the erythromycin-resistance gene *ermE* in *Saccharopolyspora erythraea*. pLL1003, pLL1004 and pLL1005 were then introduced into *S. laurentii* by conjugation, generating the corresponding recombinant strain SL102, SL103 and SL104 that expressing $tppX_1$, $tppX_2$ *and* $tppX_1$&$tppX_2$ *in trans*.

For homologous complementation of $tppX_1$, $tppX_2$ *and* $tppX_1$&$tppX_2$ in TSR analog-producing strain *ΔtsrB*, the related gene-containing plasmids were constructed as above. pLL1003, pLL1004 and pLL1005 were then introduced into *ΔtsrB* by conjugation, generating the corresponding recombinant strain SL105, SL106 and SL107 that expressing $tppX_1$, $tppX_2$ *and* $tppX_1$&$tppX_2$ *in trans*.

Fermentation, examination, chemical feeding and isolation of Products. The *S. tateyamensis* wild type strain or its mutant derivative was spread on PS5 agar plates that contain the medium composed of 20 g of soluble starch, 5 g of pharmamedia, and 20 g of agar per liter (pH 7.0), and incubated at 28 ℃ for sporulation and growth. The sporulated *S. tateyamensis* was inoculated into 30 mL of the

solid fermentation medium, which was composed of 40 g of soluble starch, 5 g of glucose, 5 g of soybean meal, 14 g of , 1 g of $CaCO_3$ and 20 g of agar per liter. After incubation at 28 ℃ for 96 h, solid fermentation media were cut, chopped and stored at -80 ℃ before methanol extraction.

For product examination, 2 mL of each solid fermentation medium was chopped and soaked by 2 mL of methanol for 30 min. After centrifugation to remove the residue, the supernatant was evaporated before re-dissolution in 200 μL of methanol. The methanol sample was subjected to HPLC and HPLC-MS analysis on an Agilent Zorbax column (SB-C18, 4.6×250 mm, 5μm, Agilent Technologies Inc., USA) by gradient elution of solvent A (H_2O + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) with a flow rate of 1 mL/min over a 35 min period as follows: T = 0 min, 15% B; T = 3 min, 15% B; T = 6 min, 40% B; T = 12 min, 40% B; T = 19 min, 55% B; T = 22 min, 85% B; T = 28 min, 85% B;T = 30 min, 15% B; and T = 35 min, 15% B. Related data were analyzed using Thermo Xcalibur software.

The *S. laurentii* wild type strain or its mutant derivative was spread on ISP2 agar plates that contain the medium composed of 4 g of yeast extract, 10 g of malt extract, 4 g of glucose, and 20 g of agar per liter (pH 7.2), and incubated at 28 ℃ for sporulation and growth. Approximately $1cm^2$ of the sporulated agar of *S. laurentii* or its derivative was cut, chopped, and inoculated into 25ml of seed medium composed of 15 g of tryptic soy broth, 15 g of soluble starch, and 50 g of glucose per liter and then incubated at 28 ℃ and 220 rpm for 36 h. 5ml of the resulting primary culture broth was added into 100ml of the fermentation medium composed of 15 g of tryptic soy broth, 11 g of yeast extract, 50 g of glucose, 15 g of $CaSO_4$ and 1ml of 2×Trace elements solution per liter for scale-up and then further incubated at 28 ℃ and 220 rpm for 3 days. The mycelia were then harvested by centrifugation at 4 ℃ and 5,000 rpm for 15min.

For compound **6** and **7** detection, 20 mL of each fermentation broth was centrifuged, and collected mycelia were soaked by 1 mL of acetone for 5 min, while the supernatants were extracted 3 times with ethyl acetate (EA). After centrifugation to remove the residue, the acetone-supernatants and EA-extracts were evaporated before re-dissolution in 200 μL of methanol. The methanol sample was subjected to HPLC-MS analysis on an Agilent Zorbax column (SB-C18, 4.6×250 mm, 5μm, Agilent Technologies Inc., USA) by gradient elution of solvent A (H_2O + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) with a flow rate of 1 mL/min over a 35 min period as follows: T = 0 min, 15% B; T = 3 min, 15% B; T = 6 min, 40% B; T = 12 min, 40% B; T = 19 min, 55% B; T = 22 min, 85% B; T = 28 min, 85% B;T = 30 min, 15% B; and T = 35 min, 15% B. Related data were analyzed using Thermo Xcalibur software.

For compound **5** and **8** feeding experiments, The *S. tateyamensis* wild type was grown in tube

containing 3 ml fermentation medium composed of 40 g of soluble starch, 5 g of glucose, 5 g of soybean meal, 14 g of and 1 g of $CaCO_3$ per litre at 28 °C and 220 rpm. The medium was fed with **5** or **8** (in DMSO) at a final concentration of 1mM and incubated at 28 °C and 220 rpm at the second day for another 2 days. Each fermentation broth was centrifuged, and collected mycelia were soaked by 1 mL of acetone for 5 min, while the supernatants were extracted 2 times with ethyl acetate (EA). After centrifugation to remove the residue, the acetone-supernatants and EA-extracts were evaporated before re-dissolution in 50 μL of methanol. The methanol sample was subjected to HPLC-MS analysis on an Agilent Zorbax column (SB-C18, 4.6×250 mm, 5μm, Agilent Technologies Inc., USA) by gradient elution of solvent A (H_2O + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) with a flow rate of 1 mL/min over a 35 min period as follows: T = 0 min, 15% B; T = 3 min, 15% B; T = 6 min, 40% B; T = 12 min, 40% B; T = 19 min, 55% B; T = 22 min, 85% B; T = 28 min, 85% B;T = 30 min, 15% B; and T = 35 min, 15% B. Related data were analyzed using Thermo Xcalibur software.

For compound **4** isolation, 7 L of the solid fermentation broth was cut, chopped and extracted with 3 L of methanol two times. After filtration and concentration, the extract was loaded onto a silica gel column, which was treated with dichloromethane–methanol (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 15% MeOH). After HPLC analyses, the fraction containing compound **4** was combined and concentrated to ~45 mg of dried extract. Further purification by RP-HPLC on an Xselect CSH C18 column (250 × 10 mm, 5μm, Waters Technology Co., Ltd., USA) by gradient elution of solvent A (H_2O + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) with a flow rate of 3 mL/min over a 33 min period as follows: T = 0 min, 15% B; T = 3 min, 15% B; T = 25 min, 85% B; T = 30 min, 15% B; and T = 33 min, 15% B.

Protein expression and purification. The genes *tppX₁* and *tppX₂* were amplified from the genome of *S. tateyamensis* by PCR using the primer pairs *tppX₁*-P-for/ *tppX₁*-P-rev and *tppX₂*-P-for/*tppX₂*-P-rev, respectively (**Table S2**). While *tppX₁* was cloned into EcoRI and HindIII digested-pRSFDeut to generate pLL1006, *tppX₂* was cloned into NdeI and XhoI digested-pLL1006 to generate pLL1007 through homologous recombination for the expression of the recombinant $TppX_1$ protein that is tagged by 6×His at N-terminus and f the protein $TppX_2$, which is not tagged by 6×His at N- or C-terminus. The sequences of *tppX₁* and *tppX₂* are listed in **Supplementary Table 3**. The above plasmid pLL1007 were introduced into *E. coli* BL21(DE3). The culture of each resulting recombinant *E. coli* strain was incubated in Luria-Bertani (LB) medium (5 g of yeast extract, 10 g of tryptone and 10 g of NaCl per liter) containing 50 μg/mL kanamycin at 37°C and 250 rpm until the cell density reached 0.6-0.8 at OD_{600}. Protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG).

The cells were harvested by centrifugation at 3000 × g for 20 min, flash-frozen and then stored at -80 ℃.

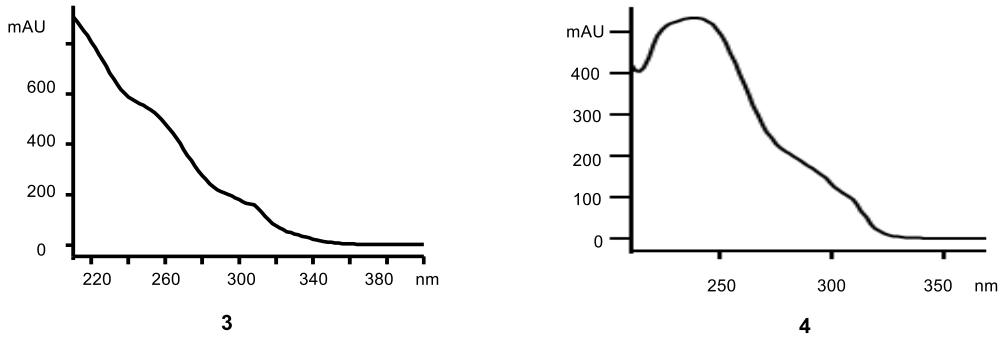
E. coli cells were re-suspended in lysis buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10% glycerol and 5 mM imidazole, pH 8.0). After disruption by FB-110X Low Temperature Ultra-Pressure Continuous Flow Cell Disrupter (Shanghai Litu Mechanical Equipment Engineering Co., Ltd, China), soluble fractions were collected by centrifugation. Recombinant proteins that contain a 6×His were purified on a HisTrap HP column (GE Healthcare, USA), which was pre-treated with 10 column volumes (CVs) of lysis buffer followed by 10 CVs of wash buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10% glycerol and 40 mM imidazole, pH 7.4), using elution buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10% glycerol and 250 mM imidazole, pH 7.4). Desired protein fractions were concentrated (to 500 μM-1 mM) using Amicon® Ultra-15 Centrifugal Filter Devices (MILLIPORE, USA) and desalted using a PD-10 Desalting Column (GE Healthcare, USA) according to the manufacturer's protocols, and then quantified in concentration by Bradford assay using bovine serum albumin as the standard. The purity of recombinant proteins was determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (**Figure S2**).

In vitro assays of $TppX_1$ and **$TppX_2$ activity.** Each conversion was conducted at 30 ℃ for 8 h in 100 μL of the reaction mixture that contained 200 μM substrates and 50 μM $TppX_1$ and $TppX_2$, DMSO 5%, 10 mM ATP, 1 mM Na_2S, 5 mM $MgCl_2$ along with 50 mM TrisCl (pH 7.5). Conversions were quenched by adding equal volumes of acetonitrile, and after centrifugation, reaction mixtures were subjected to HPLC, HPLC-MS and HR-MS analyses. HPLC was conducted on a reversed-phase Agilent ZORBAX column (SB-C18, 4.6×250 mm, 5 μm, Agilent Technologies Inc., USA) by gradient elution of solvent A (H_2O + 0.1% formic acid) and solvent B (acetonitrile + 0.1% formic acid) with a flow rate of 1 mL/min over a 35 min period as follows: T = 0 min, 15% B; T = 3 min, 15% B; T = 6 min, 40% B; T = 12 min, 40% B; T = 19 min, 55% B; T = 22 min, 85% B; T = 28 min, 85% B;T = 30 min, 15% B; and T = 35 min, 15% B. Related data were analyzed using Thermo Xcalibur software and Agilent ChemStation.

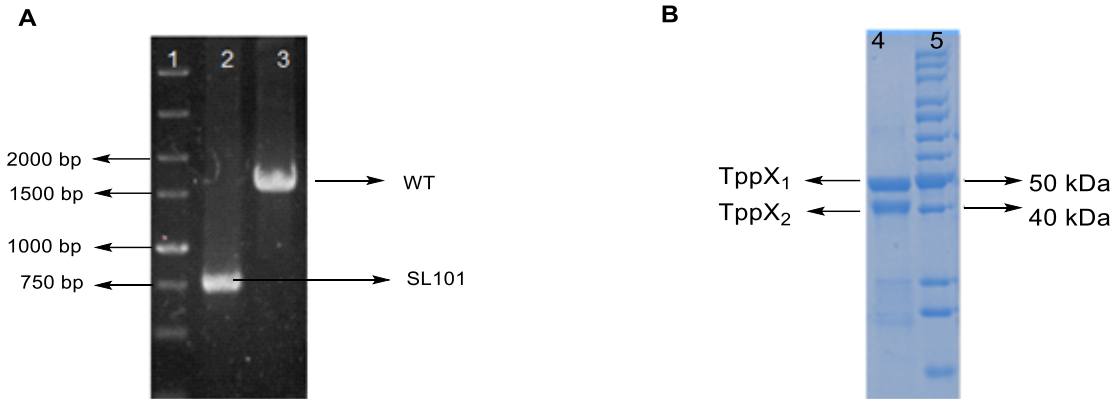
Table of Contents

1. Supplementary Figures

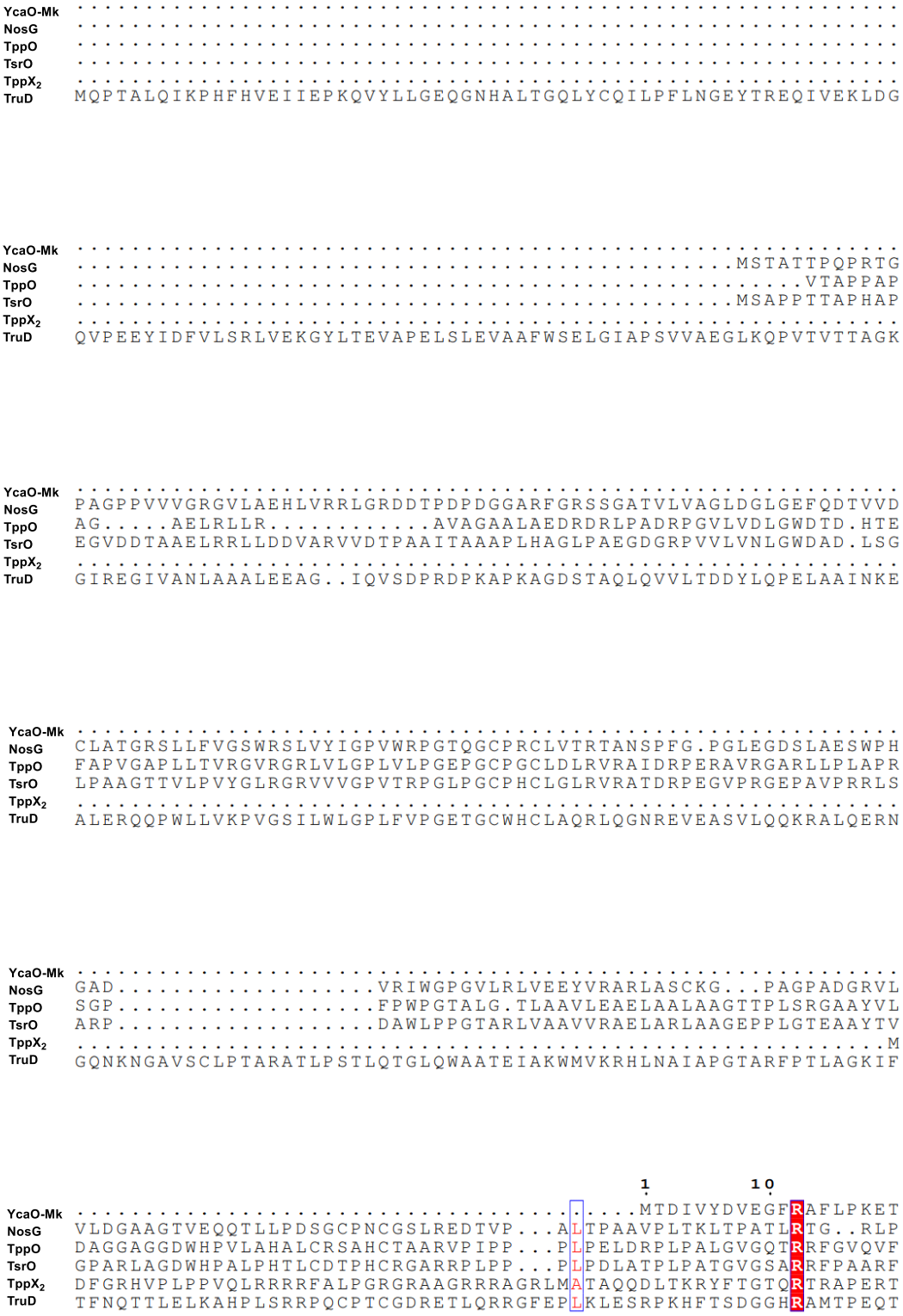
Supplementary Fig.1. The UV spectra of **3** and **4**.

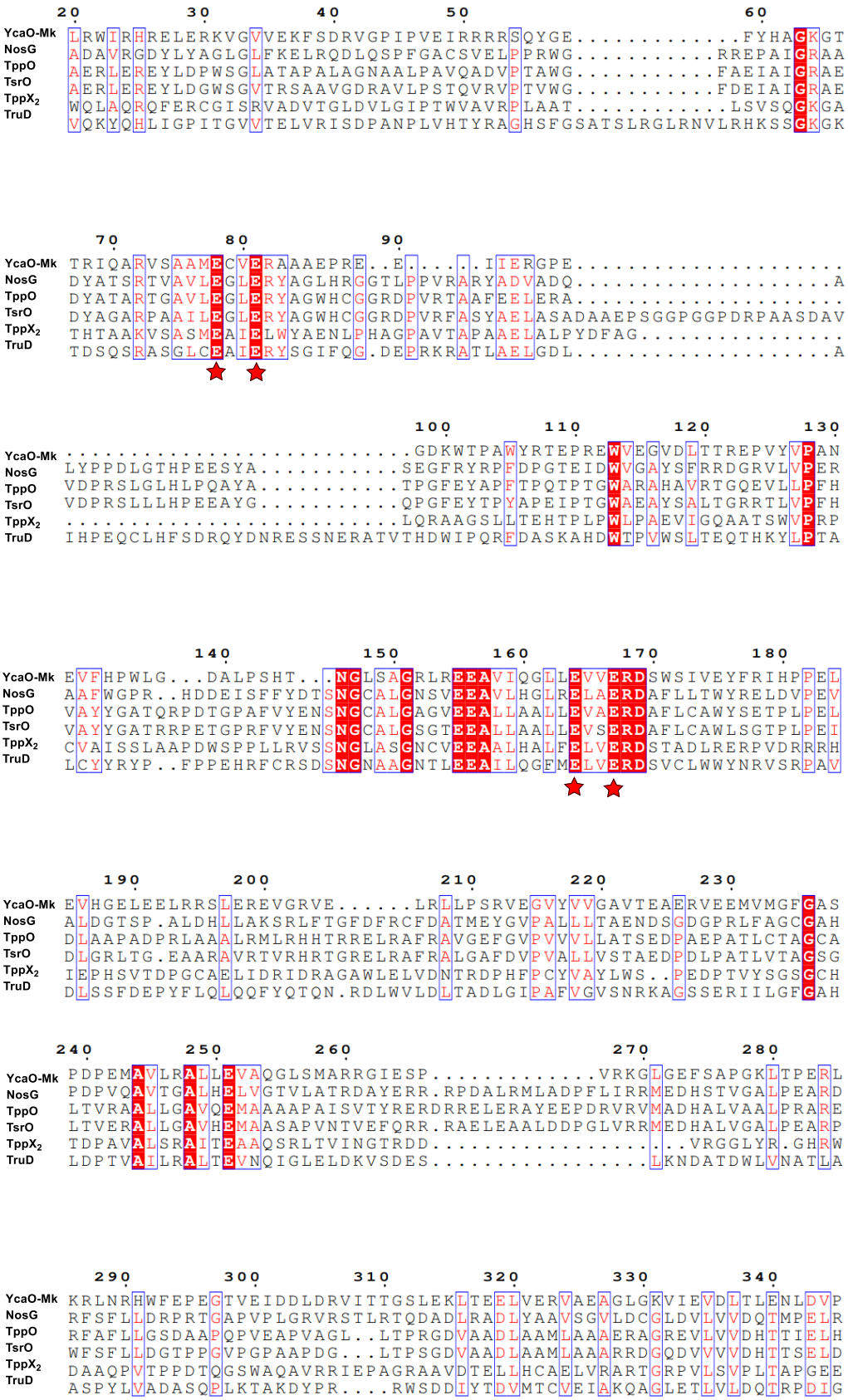


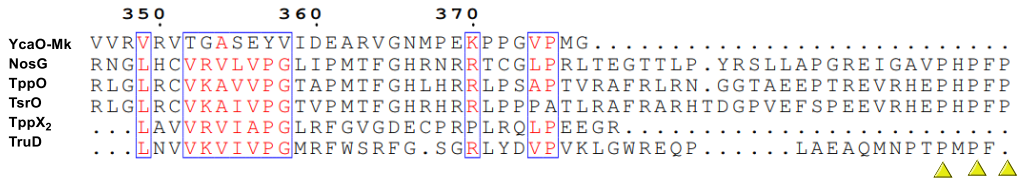
Supplementary Fig.2. DNA and protein examination. DNA examination by agarose gel electrophoresis and protein examination by SDS-PAGE analysis. Each gel contains the DNA/protein ladder (left) and the purified recombinant protein (right). **A**. DNA ladder (Lane1), DNA examination of mutant SL101(729 bp)(Lane2） and *S. tateyamensis* (WT) (1536 bp) (Lane3). **B.** Protein ladder (Lane5) and protein examination of $TppX_1$ (48 kDa) and $TppX_2$ (43 kDa) (Lane4).



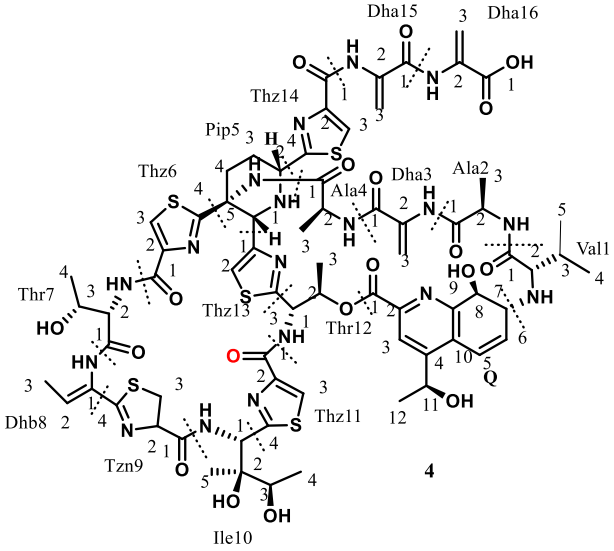
Supplementary Fig.3. Sequence alignment of some YcaO proteins. Red stars indicate the conserved residues for binding ATP and Mg^{2+}. Yellow triangles indicate the Pro-rich C-termini. $TppX_2$ is a protein from *Streptomyces tateyamensis*. YcaO-Mk is a protein from *Methanosarcina kandleri*. TppO is a protein from *Streptomyces tateyamensis*. TsrO is a protein from *Streptomyces laurentii*. NosG is a protein from *Streptomyces actuosus*. TruD is a protein from uncultured Prochloron sp. 06037A.



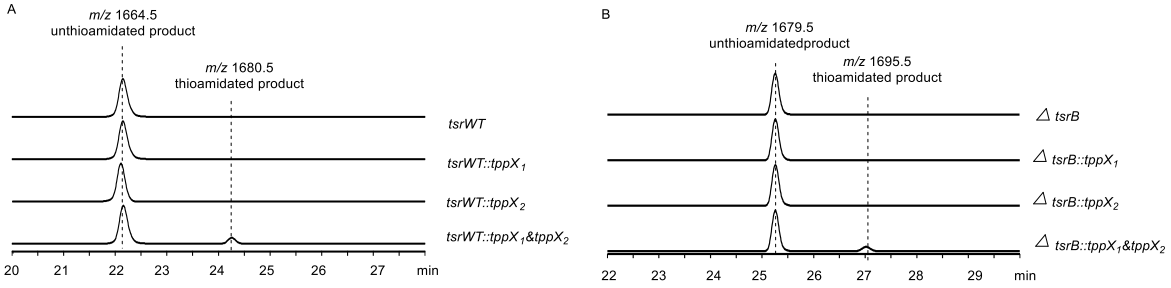

20 30 40 50 60
YcaO-Mk LRWIRHRELERKVGVVEKFSDRVGPIPVEIRRRRSQYGE.....FYHAGKGT
NosG ADAVRGDYLYAGLGLFKELRQDLQSPFGACSVELPPRWG.....RREPAIGRAA
TppO AERLERFYLDPWSGLATAPALAGNAALPAVQADVPTAWG.....FAEIAIGRAE
TsrO AERLERFYLDGWSGVTRSAAVGDRAVLPSTQVRVPTVWG.....FDEIAIGRAE
TppX2 WQLAQRQFERCGISRVADVTGLDVLGIPTWVAVRPLAAT.....LSVSQGKGA
TruD VQKYQHLIGPITGVVTELVRISDPANPLVHTYRAGHSFGSATSLRGLRNVLRHKSSGKGK
70 80 90
YcaO-Mk TRIQARVSAAMECVERAAEPRE..E....I.IERGPE.....
NosG DYATSRTVAVLEGLERYAGLHRGGTLPPVRARYADVADQ.....A
TppO DYATARTGAVLEGLERYAGWHCGGRDPVRTAAFEELERA.....
TsrO DYAGARPAAILEGLERYAGWHCGGRDPVRFASYAELASADAAEPSGGPGGPDRPAASDAV
TppX2 THTAAKVSASMEAIELWYAENLPHAGPAVTAPAAELALPYDFAG.....
TruD TDSQSRASGLCEAIERYSGIFQG.DEPRKRATLAELGDL.....A
100 110 120 130
YcaO-Mk .....GDKWTPAWYRTEPREWVEGVDLTTREPVYVPAN
NosG LYPPDLGTHPEESYA.....SEGFRYRPFDPGTEIDWVGAYSFRRDGRVLVPER
TppO VDPRSLGLHLPQAYA.....TPGFEYAPFTPQTPTGWARAHAVRTGQEVLLPFH
TsrO VDPRSLLLHPEEAYG.....QPGFEYTPYAPEIPTGWAEAYSALTGRRTLVPFH
TppX2 .....LQRAAGSLLTEHTPLPWLPAEVIGQAATSWVPRP
TruD IHPEQCLHFSDRQYDNRESSNERATVTHDWIPQRFDASKAHDWTPVWSLTEQTHKYLPTA
140 150 160 170 180
YcaO-Mk EVFHPWLG...DALPSHT..NGLSAGRLREEAVIQGLLEVVERDSWSIVEYFRIHPPEL
NosG AAFWGPR..HDDEISFFYDTSNGCALGNSVEEAVLHGLRELAERDAFLLTWYRELDVPEV
TppO VAYYGATQRPDTGPAFVYENSNGCALGAGVEEALLAALLEVAERDAFLCAWYSETPLPEL
TsrO VAYYGATRRPETGPRFVYENSNGCALGSGTEEALLAALLEVSERDAFLCAWLSGTPLPEI
TppX2 CVAISSLAAPDWSPPLLRVSSNGLASGNCVEEAALHALFELVERDSTADLRERPVDRRRH
TruD LCYYRYP..FPPEHRFCRSDSNGNAAGNTLEEAILQGFMELVERDSVCLWWYNRVSRPAV
190 200 210 220 230
YcaO-Mk EVHGELEELRRSLEREVGRVE.....LRLLPSRVEGVYVVGAVTEAERVEEMVMGFGAS
NosG ALDGTSP.ALDHLLAKSRLFTGFDFRCFDATMEYGVPALLLTAENDSGDGPRLFAGCGAH
TppO DLAAPADPRLAAALRMLRHHTRRELRAFRAVGEFGVPVVVLLATSEDPAEPATLCTAGCA
TsrO DLGRLTG.EAARAVRTVRHRTGRELRAFRALGAFDVPVALLVSTAEDPDLPATLVTAGSG
TppX2 IEPHSVTDPGCAELIDRIDRAGAWLELVDNTRDPHFPCYVAYLWS..PEDPTVYSGSGCH
TruD DLSSFDEPYFLQLQQFYQTQN.RDLWVLDLTADLGIPAFVGVSNRKAGSSERIILGFGAH
240 250 260 270 280
YcaO-Mk PDPEMAVLRALLEVAQGLSMARRGIESP.....VRKGLGEFSAPGKLTPERL
NosG PDPVQAVTGALHELVGTVLATRDAYERR.RPDALRMLADPFLIRRMEDHSTVGALPEARD
TppO LTVRAALLGAVQEMAAAAPAISVTYRERDRRELERAYEEPDRVRVMADHALVAALPRARE
TsrO LTVERALLGAVHEMAASAPVNTVEFQRR.RAELEAALDDPGLVRRMEDHALVGALPEARP
TppX2 TDPAVALSRAITEAAQSRLTVINGTRDD.....VRGGLYR.GHRW
TruD LDPTVAILRALTEVNQIGLELDKVSDES.....LKNDATDWLVNATLA
290 300 310 320 330 340
YcaO-Mk KRLNRHWFEPEGTVEIDDLDRVITTGSLEKLTEELVERVAEAGLGKVIEVDLTLENLDVP
NosG RFSFLLDRPRTGAPVPLGRVRSTLRTQDADLRADLYAAVSGVLDCGLDVLVVDQTMPELR
TppO RFAFLLGSDAAPQPVEAPVAGL..LTPRGDVAADLAAMLAAAERAGREVLVVDHTTIELH
TsrO WFSFLLDGTPPGVPGPAAPDG...LTPSGDVAADLAAMLAAARRDGQDVVVVDHTTSELD
TppX2 DAAQPVTPPDTQGSWAQAVRRIEPAGRAAVDTELLHCAELVRARTGRPVLSVPLTAPGEE
TruD ASPYLVADASQPLKTAKDYPR....RWSDDIYTDVMTCVEIAKQAGLETLVLDQTRPDIG



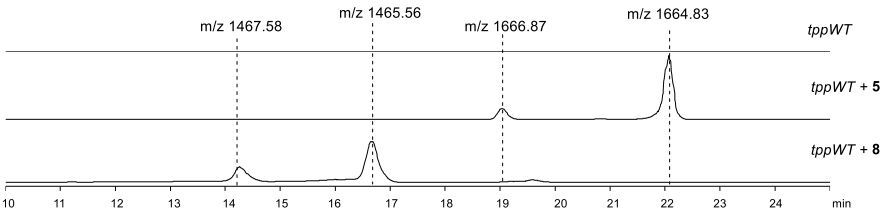
Supplementary Fig.4. Structure of **4**.



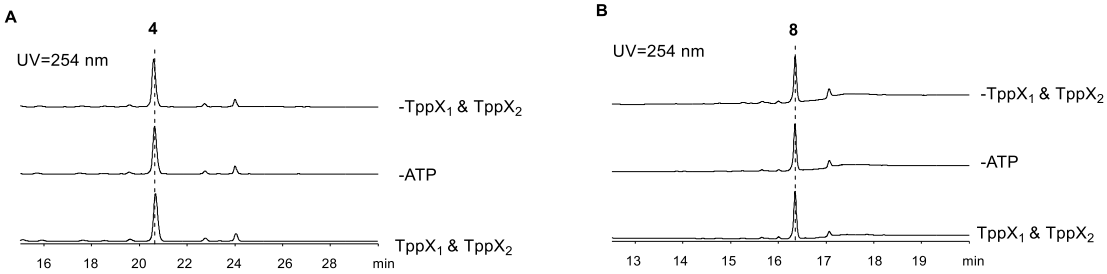
Supplementary Fig.5. Comparation of thioamidated and unthioamidated products in gene complementation experiments. (A) HPLC-MS analyses of the culture extracts from the wild-type *S. laurentii*, the wild-type *S. laurentii* with the integrative plasmid containing $tppX_1$, the wild-type *S. laurentii* with the integrative plasmid containing $tppX_2$, the wild-type *S. laurentii* with the integrative plasmid containing $tppX_1$ and $tppX_2$; (B) HPLC-MS analyses of the culture extracts from the mutant *ΔtsrB* with the integrative plasmid containing $tppX_1$, the mutant *ΔtsrB* with the integrative plasmid containing $tppX_2$, the mutant *ΔtsrB* with the integrative plasmid containing $tppX_1$ and $tppX_2$.



Supplementary Fig.6. Feeding experiments. HPLC-MS analyses of the fermentation cultures of the *tpp* wild type strain *tppWT* in the absence or presence of the exogenous TSR (**5**) or TSR-analog **8**.

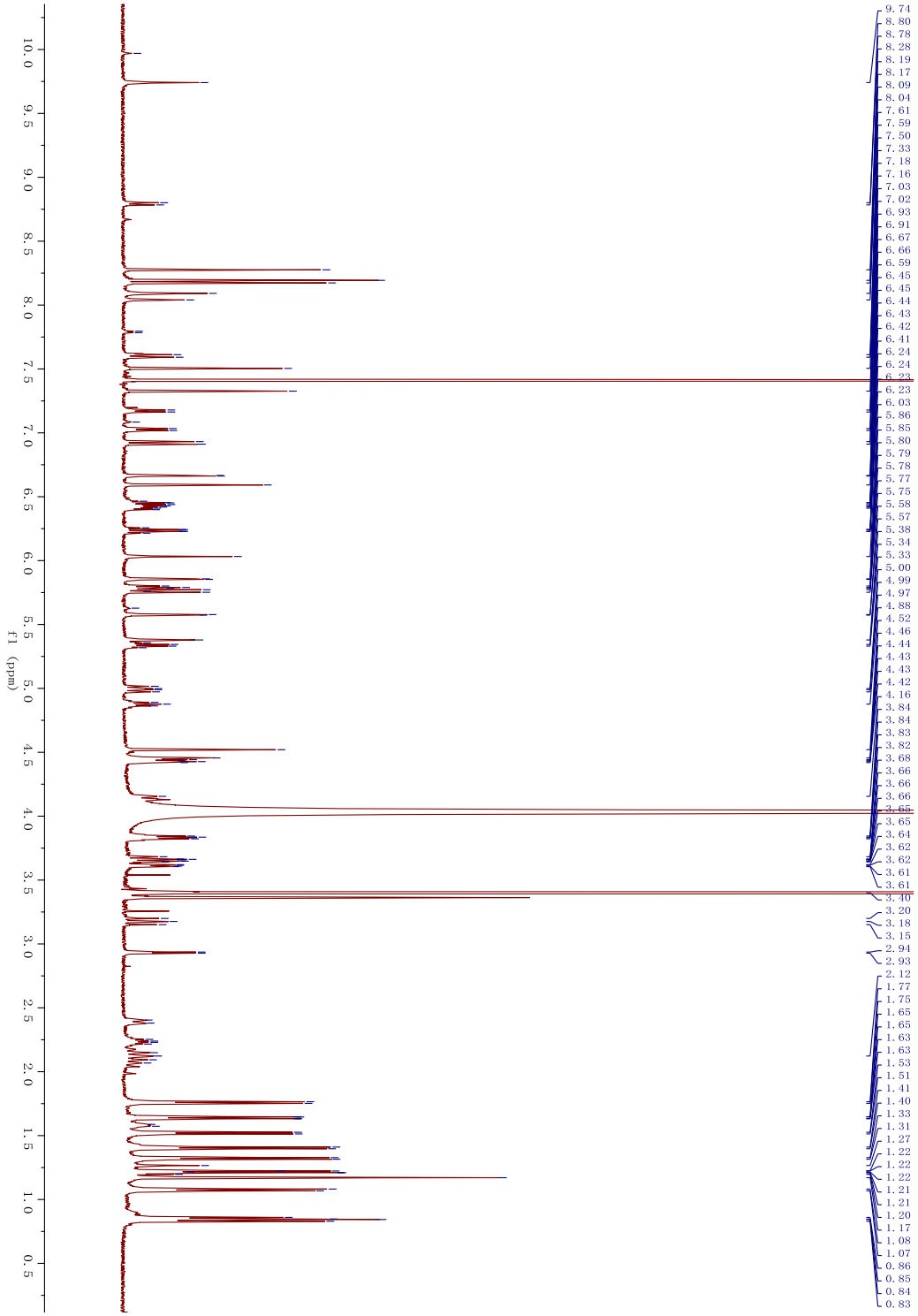


Supplementary Fig.7. HPLC analyses of in vitro bioactivity assays of $TppX_1$ and $TppX_2$ using **4** or **8** as the substrates. The control group is with absence of enzymes or ATP.

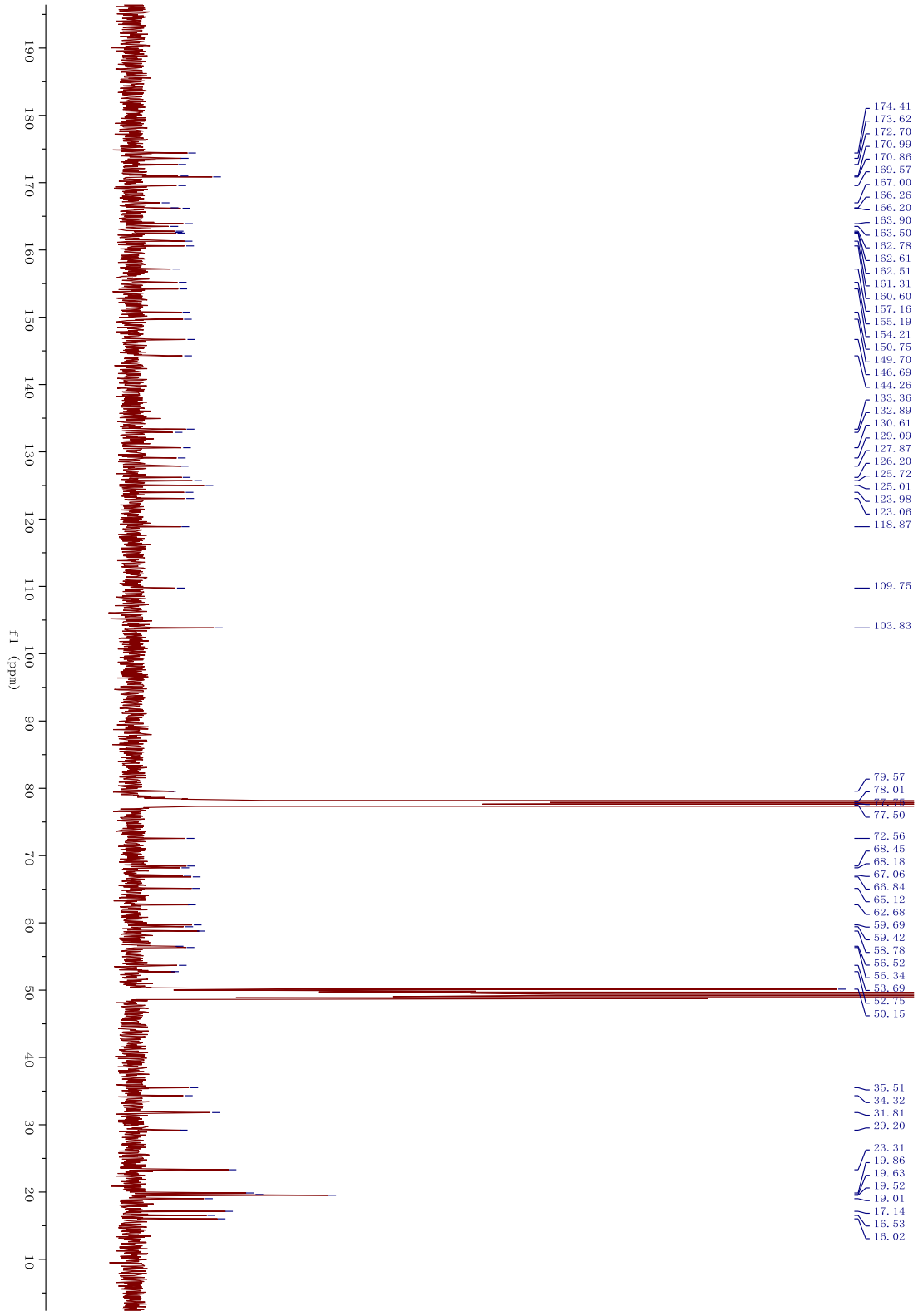


Supplementary Fig. 8. NMR spectra of **4**. **A**. ^{1}H spectrum. **B.** ^{13}C spectrum. **C.** COSY spectrum. **D.** HSQC spectrum. **E.** HMBC spectrum. The compound is dissolved in a mixed solvent containing $CDCl_3$ and CD_3OD with a ratio of 4:1.

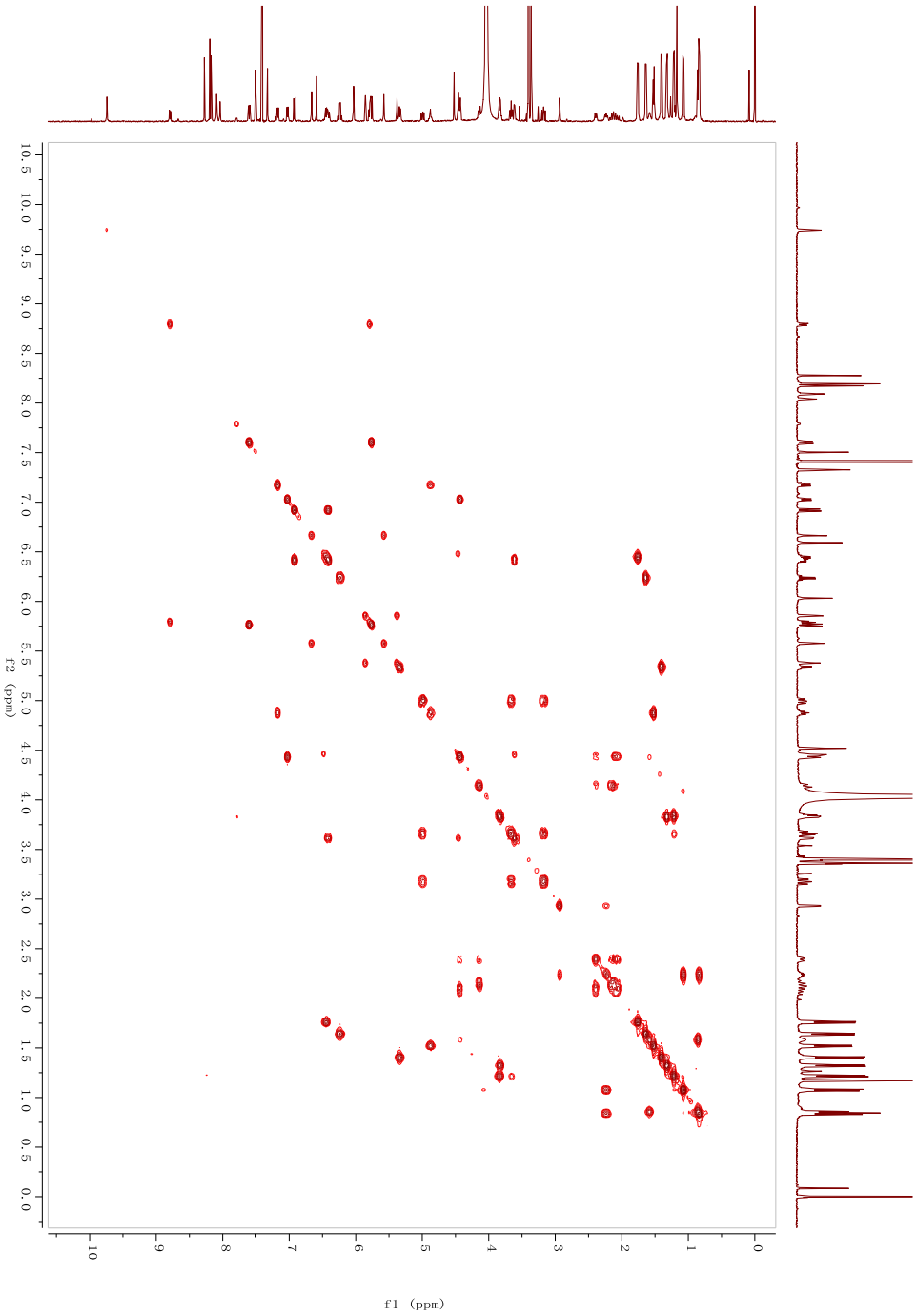
A



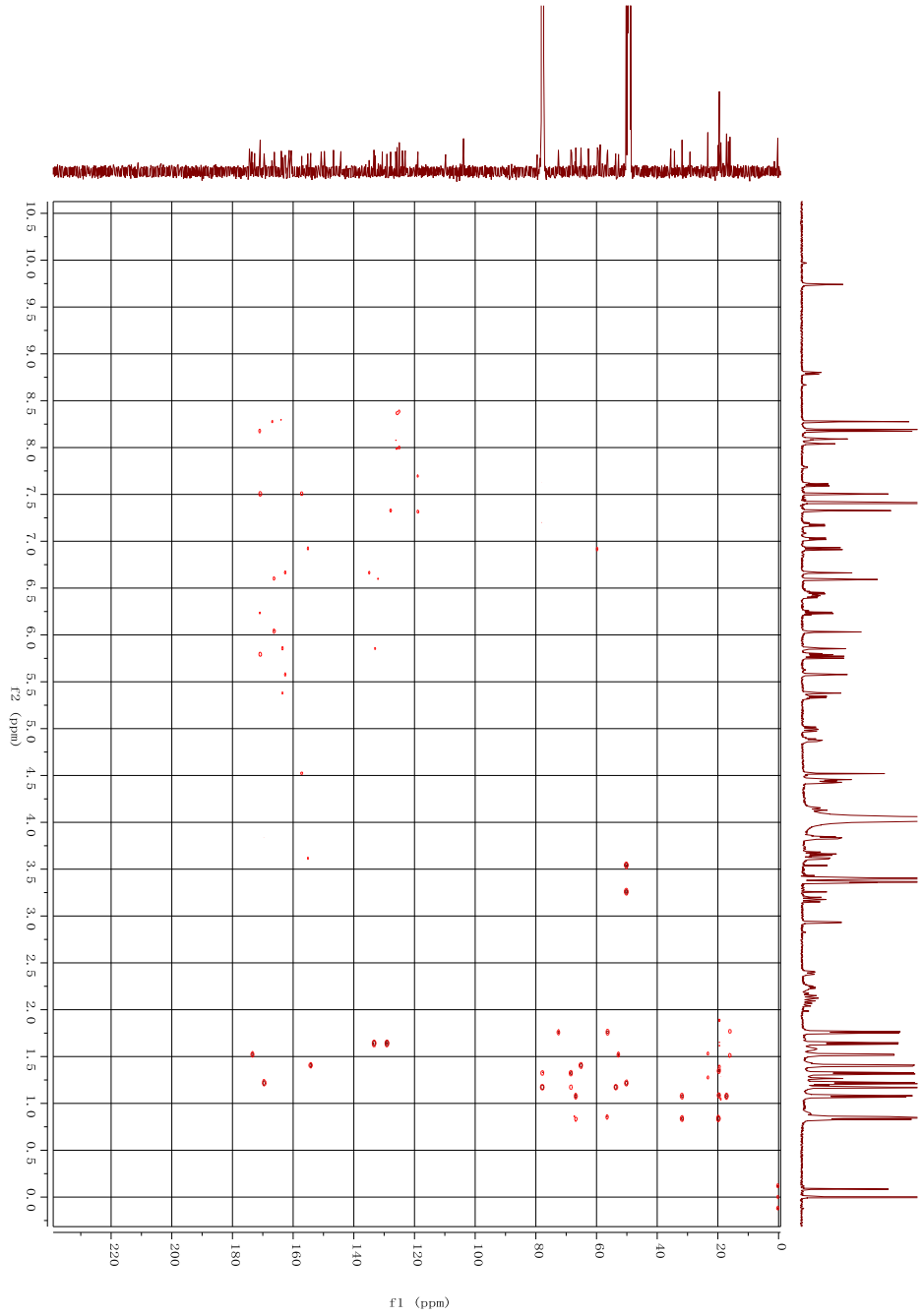
B


f1 (ppm)
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
174.41
173.62
172.70
170.99
170.86
169.57
167.00
166.26
166.20
163.90
163.50
162.78
162.61
162.51
161.31
160.60
157.16
155.19
154.21
150.75
149.70
146.69
144.26
133.36
132.89
130.61
129.09
127.87
126.20
125.72
125.01
123.98
123.06
118.87
109.75
103.83
79.57
78.01
77.75
77.50
72.56
68.45
68.18
67.06
66.84
65.12
62.68
59.69
59.42
58.78
56.52
56.34
53.69
52.75
50.15
35.51
34.32
31.81
29.20
23.31
19.86
19.63
19.52
19.01
17.14
16.53
16.02

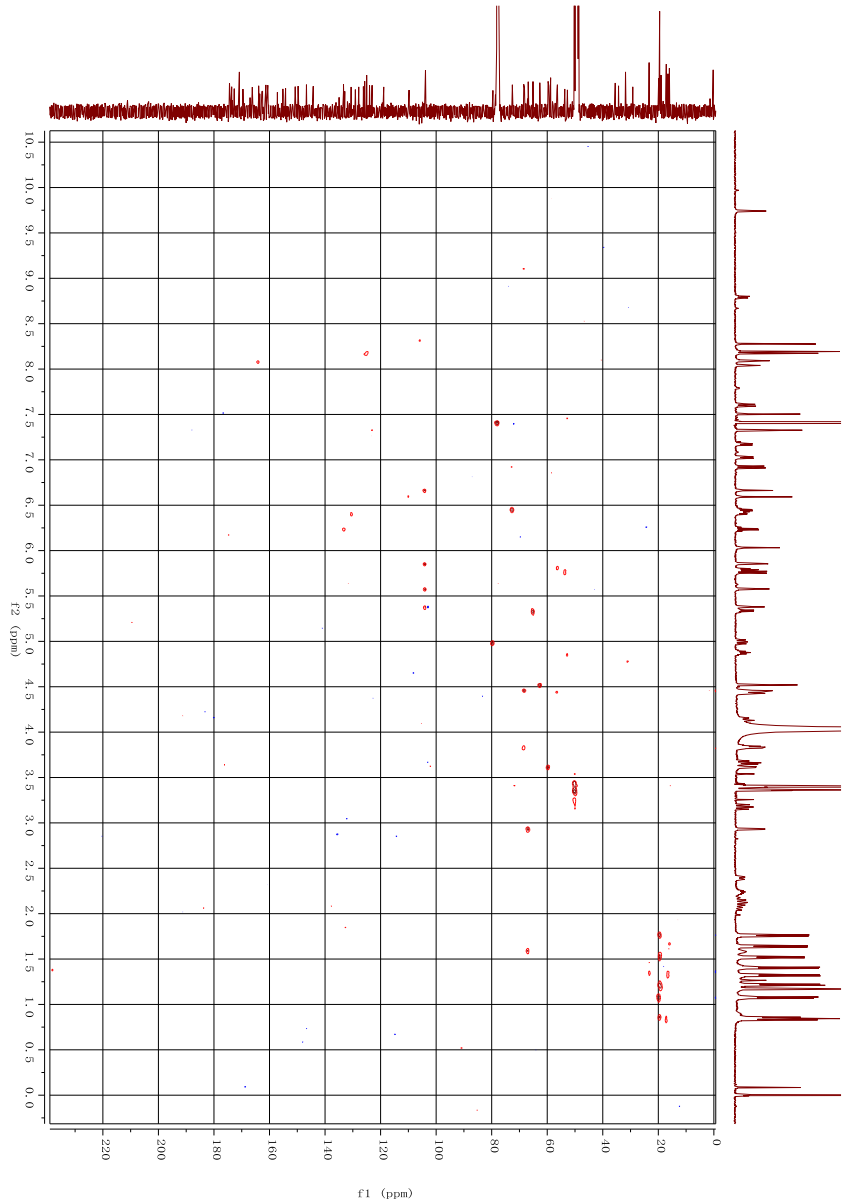
C



D



E



2. Supplementary Tables

Supplementary Table 1. Strains and plasmids used in this study

Strains/Plasmids	Description	Source / Reference
Strains		
Escherichia coli		
DH5α	Host for general cloning	Transgen
ET12567 (pUZ8002)	Donor strain for conjugation between *E. coli* and *Streptomyces*	1
BL21(DE3)	Host for protein expressing	Transgen
SL108	BL21 (DE3) derivative, containing pLL1007for heterogenously producing $TppX_1$ and $TppX_2$	This study
Streptomyces		
tateyamensis ATCC21389	Wild type strain, thiopeptin-producing strain	ATCC
SL101	The *tppX*$_2$ in-frame deletion mutant of *S. tateyamensis*	This study
laurentii	Wild type strain, thiostrepton-producing strain	ATCC
ΔtsrB	Thiostrepton derivative-producing strain	7
ΔtsrP	Thiostrepton derivative-producing strain	8
SL102	The *Streptomyces laurentii* derivative carrying *tppX*$_1$ in trans	This study
SL103	The *Streptomyces laurentii* derivative carrying *tppX*$_2$ in trans	This study
SL104	The *Streptomyces laurentii* derivative carrying *tppX*$_1$ and *tppX*$_2$ in trans	This study
SL105	The *ΔtsrB* derivative carrying *tppX*$_1$ in trans	This study
SL106	The *ΔtsrB* derivative carrying *tppX*$_2$ in trans	This study
SL107	The *ΔtsrB* derivative carrying *tppX*$_1$ and *tppX*$_2$ in trans	This study
Plasmids		
pKC1139	*E.coli-Streptomyces* shuttle vector for gene inactivation	1
pRSFDeut-1	Protein co-expression vector used in *E.coli*, encoding N-terminal His-tag for one sequence	Novagen
pSET152-E*	*E.coli-Streptomyces* shuttle vector for gene complementation with *PermE**	9
pLL1001	pKC1139 derivative containing partial *tppX*$_2$ in-frame deletion fragment	This study
pLL1002	pKC1139 derivative containing total *tppX*$_2$ in-frame deletion fragment	This study
pLL1003	pSET152-E* derivative, containing gene *tppX*$_1$	This study
pLL1004	pSET152-E* derivative, containing gene *tppX*$_2$	This study
pLL1005	pSET152-E* derivative, containing gene *tppX*$_1$ and *tppX*$_2$	This study
pLL1006	pRSFDeut-1 derivative, containing a 1347 bp PCR product that encodes *tppX*$_1$	This study
pLL1007	pLL1006 derivative, containing a 1200 bp PCR product that encodes *tppX*$_2$	This study

Supplementary Table 2. Primers used in this study.

Primer	**Sequence**
TppX$_2$-L-for	ACCGCAAGCTTCCGGCTGGCCGCCGCCTGCGGGCAG
TppX$_2$-L-rev	CTAGTCTAGACAGTACGTCCAGTCCGGTGACGTCG
TppX$_2$-R-for	CTAGTCTAGAGAGCCCGCCGGGCGGGCCGCGGTCG
TppX$_2$-R-rev	ACGGAATTCCCACCGTGCCCGGGTCCCGGCCGC
TppX$_2$ –DC-for	GCTGTTCGACGCCAAGTAC
TppX$_2$-DC- rev	GGAACAGCCCGTCCACGATC
TppX$_1$-C-for	AGAGGATCCATGACCACGCACGTGTTCGC
TppX$_1$-C-rev	GACTCTAGAATATCATGACGGCTCCCCTTCG
TppX$_2$-C-for	AGAGGATCCATGGCCACGGCCCAACAGGAC
TppX$_2$-C-rev	GACTCTAGAATATCATCGGCCCTCCTCGGGC
TppX$_2$-P-for	ATAAGAAGGAGATATACATATGGCCACGGCCCAACAG
TppX$_2$-P-rev	AGCGGTTTCTTTACCAGACTCGAGCATCGGCCCTCCTCGGG
TppX$_1$-P-for	TCATCACCACAGCCAGGATCCAATGACCACGCACGTGTTCG
TppX$_1$-P-rev	GCATTATGCGGCCGCAAGCTTTCATGACGGCTCCCCTTCG

Supplementary Table 3. Gene sequences of *tppX*$_1$ and *tppX*$_2$.

Name	sequence
tppX$_1$	ATGACCACGCACGTGTTCGCCGGGCCGACCATCGGCCCGGACCGGGTGG CCGAGCTACTGCCCGGCGCCGTGCTGCACCCGCCGGTGCAGCACGGCGA CCTGCTGCGGCTGCCGGTGGCCGCCGGGGACACCGTGCTGATCGTGGAC GGGCTGTTCCAACAGGCCCCCGCGGTGCGGCACAAGGAGATCCTGCACC TGGTGCACGAGGGCGTCCGGGTGGCCGGCGCGAGCAGCATGGGCGCGCT GCGCGCGGCCGAGCTGCACCGGTTCGGGATGCTCGGGCTGGGCCAGGTG TTCCGCTGGTACGCCGACGGCACGGTGACCGCCGACGACGAGGTGGCAG TGGCCCACCTGGGCGAGGAGGACGGCTACCGCCAGCTCTCCGACGCCCT GGTCTCGGTGCGCTACGGGCTGGGCCGGGCAGTCGAGGCTGGCGTGCTG AACGCCGCCGAGCAGGCCGGGCTGCTTGCCGCACTGGCCGAGCTGCCGT TCCCGCAGCGCAGTTGGCGCAACCTGTGGCGGATCACCGGGCAGACCGA CCTGGCGGCGGCGGCCGCCCGGGTCCGCGCCCACCTGGCGGTGCGCCCC GCGGACGCCGACGTGAAGCGCCTGGACGCCGAGACCGCGCTGCGGGCGC TGCGCCGGGAGGCCGCTCCCGCGGCGGCCCCCGGGCCCCGCCCGGCCAC CCTGGACACCGTCTACCTGGCCGACTGGCGGTTCGAGCACTCCGCCACCG ACCCGGTGAGCGACTTCCACACCCTCGGCTTCCTCCAGCTGTTCCTGCCG GCTTACCCGCAGCTGAACCGCCATCAGGCACTGGCCCGGATCGGCGGCG GCGAGGCGGGAGCCCTCGCGGCGGCCCGGGCCGCGGGCCTGCTGGGCGC CGGCGACGAGCCGGGGAGCGGGATGCGGGCCTGGCTCACCGAGCGGGA ACTGGCCGAGCTGCCCGGCCGCGAGCTGGCGCTCACCGCCCTGGTGCGG TCCTTCCGCACCGCGCCCGGGGTGCGCACCGGCCACCGGCTGCCCGAGC CGCTCCTGGCCGCCGGGCCGCTGCTGCGCATGGCCCGCAGCTGCGCCGCG GCCGCGGCCGCACTGAACGCGGCCCGCCGGGCCAGGCACCCGGAGTTCC AGGTCGAGCACGTCCGCACCGACCTGGTGGAGGAGTTCTTCGCCGCCCG CTGGCAGTGCGCCGACCTGGTCACCGCCTGCTGGGACCGCGGGCTCACG GGCCTTGGCCAACTGCACGAGCTGGGCCAGTACTTCCTGTTGCTCGGCCG GTCCGGCCGACTGCCGGAACCGCAACTCGCGGCGGTCAGCTTCGCCGCC GCCGAAGGGGAGCCGTCATGA
tppX$_2$	ATGGCCACGGCCCAACAGGACCTCACCAAAAGGTACTTCACCGGCACCC AGCGGACCCGGGCGCCGGAGCGGACCTGGCAGTTGGCGCAGCGTCAGTT CGAGCGGTGCGGGATCAGCCGGGTGGCCGACGTCACCGGACTGGACGTA CTGGGCATCCCGACCTGGGTCGCGGTCCGCCCGCTGGCCGCCACGCTGAG CGTCAGTCAGGGCAAGGGCGCCACCCACACCGCCGCGAAGGTCTCCGCG AGCATGGAGGCGATCGAGCTCTGGTACGCCGAGAACCTGCCGCACGCCG GCCCCGCGGTCACCGCGCCGGCTGCCGAACTCGCCCTGCCCTACGACTTC GCCGGCCTGCAGCGGGCCGCGGGCAGCCTGCTGACCGAGCACACCCCGC TGCCGTGGCTGCCGGCCGAGGTGATCGGCCAGGCGGCCACCAGCTGGGT GCCCCGGCCCTGCGTGGCGATCTCCTCGCTGGCCGCCCCGGACTGGTCGC CGCCGCTGCTGCGGGTCTCCAGCAACGGCCTGGCCAGCGGGAACTGTGT CGAGGAGGCGGCCCTGCACGCGCTGTTCGAACTGGTCGAGCGGGACAGC ACCGCCGACCTGCGGGAGCGCCCGGTGGACCGGCGCCGGCACATCGAGC CGCACAGCGTCACCGACCCCGGCTGCGCCGAGCTGATCGACCGGATCGA CCGGGCCGGGGCCTGGCTGGAGCTGGTGGACAACACCCGCGACCCGCAC TTCCCCTGCTACGTCGCCTACCTGTGGTCCCCGGAGGACCCGACCGTCTA CTCGGGCTCCGGCTGCCACACCGACCCGGCGGTGGCGCTCTCCCGGGCG ATCACCGAGGCCGCGCAGAGCCGGCTCACGGTGATCAACGGCACCCGCG ACGACGTGCGCGGCGGCCTCTACCGCGGCCACCGCTGGGACGCCGCGCA GCCCGTCACACCCCCGGACACGCAGGGCAGTTGGGCGCAGGCGGTGCGG CGGATCGAGCCCGCCGGGCGGGCCGCGGTCGACACTGAACTGCTGCACT GCGCCGAGCTGGTGCGGGCCCGGACCGGCCGACCGGTGCTGAGCGTGCC

	GCTGACCGCCCCCGGGGAGGAGCTCGCGGTGGTCCGGGTGATCGCCCCC GGCCTGCGCTTCGGCGTCGGCGACGAGTGCCCGCGCCCGCTCCGCCAAC TGCCCGAGGAGGGCCGATGA

Supplementary Table 4. *Orfs* in thiopeptin and thiostrepton biosynthetic gene clusters.

Protein	Size (aa)	Proposed Function	Protein Homolog Identity/Similarity （%）	GenBank or NCBI accession no.
TppA	355	Putative aminotransferase	TsrA(34/43)	ACN80663.1
TppD	145	Ester cyclase-like	TsrD(58/71)	ACN80666.1
TppE	403	Acyl-CoA dehydrogenase, short-chain specific	TsrE(58/71)	ACN80667.1
TppF	292	Putative methyltransferase	TsrF(51/60)	ACN80668.1
TppG	191	Regulator	TsrG(51/60)	ACN80669.1
$TppX_3$	241	α/β hydrolase	TsrB(14/25)	ACN80664.1
TppH	68	Thiopeptin precursor	TsrH(70/77)	ACN80670.1
TppI	251	Putative hydrolase	TsrI(53/66)	ACN80671.1
TppJ	873	Dehydrotase	TsrJ(51/62)	ACN80672.1
TppK	334	Dehydrotase	TsrK(56/68)	ACN80673.1
TppL	369	[4+2] cycloaddition enzyme	TsrL(74/85)	ACN80674.1
TppM	534	Dehydrogenase	TsrM(59/67)	ACN80675.1
TppN	544	Ocin-ThiF like	TsrN(45/54)	ACN80676.1
TppO	634	Cyclodehydrase	TsrO(59/67)	ACN80677.1
TppP	453	Cytochrome P450-like enzyme	TsrP(63/72)	ACN80678.1
TppQ	432	Acyl-CoA synthetase	TsrQ(62/71)	ACN80679.1
TppR	409	Cytochrome P450	TsrR(52/63)	ACN80680.1
TppS	794	Dehydrotase	TsrS(34/45)	ACN80681.1
$TppX_4$	168	$F_{420}H_2$-dependent reductase	Rv1155(14/26)	O06553.1
$TppX_2$	400	YcaO superfamily	TvaH(24/36)	BAN83923.1
$TppX_1$	449	TfuA-domain	TvaI(17/24)	BAN83924.1
TppT	574	Methyltransferase	TsrT(63/72)	ACN80682.1
TppU	264	Short-chain dehydrogenase	TsrU(37/43)	ACN80683.1

Supplementary Table 5. ^{1}H and ^{13}C NMR spectroscopic data for **4**. ^{13}C NMR spectroscopic data($^{13}C_2$) for **2** is shown for comparation. Each compound is dissolved in a mixed solvent containing $CDCl_3$ and CD_3OD with a ratio of 4:1.

Assigment	δH	δC	δC_2	HMBC	COSY
Val1-1		174.4	172.6		
Val1-2	2.93m	66.9	63.2	Val1-3	Val1-3
Val1-3	2.25 m	31.7	29.8	Val1-4; Val 1-2	Val1-2; Val1-4; Val1-5
Val1-4	0.84d(6.6)	17.2	15.0	Val1-3	Val1-3; Val1-5
Val1-5	1.07d(6.9)	19.9	17.3	Val1-3	Val1-4; Val1-3
Ala2-1		169.5	168.0		
Ala2-2	3.84m	50.1	48.3	Val1-1; Ala2-1; Ala2-3	Ala2-3
Ala2-3	1.18d(6.7)	19.5	17.3	Ala2-1; Ala2-2	Ala2-2
Ala2-NH	7.8br. s				
Dha3-1		163.5	161.9		
Dha3-2		132.8	131.3		
Dha3-3	Ha 5.85bs; Hb 5.37bs	104.1	101.9	Dha3-1; Dha3-2 Dha3-1; Dha3-2	Dha3-3Hb Dha3-3Ha
Dha3-NH					
Ala4-1		174.0	171.8		
Ala4-2	4.85m	52.9	50.9	Ala4-1; Dha3-1	Ala4-3; Ala4-NH
Ala4-3	1.54d(6.5)	19.5	17.7	Ala4-1; Ala4-2	Ala4-2
Ala4-NH	7.18d(7.7)			Dha3-1	Ala4-2
Pip5-1	4.53s	62.7	60.8	Pip5-2; Pip5-5; Thz13-1; Thz13-2; Thz14-4	
Pip5-2	4.42m	58.8	56.9	Pip5-1	
Pip5-3	Ha 2.40m Hb 2.11m	29.4	27.2		Pip5-4- Ha; Pip5-4Hb; Pip5-4- Ha
Pip5-4	Ha 2.13m Hb 4.15m	33.8	32.5	Pip5-3; Pip5-1 Pip5-3; Pip5-5; Pip5-1; Pip5-2; Thz6-4	Pip5-3- Ha; Pip5-3Hb; Pip5-4- Hb; Pip5-3- Ha; Pip5-4Ha
Pip5-NH	9.90br. s				
Pip5-5		58.3	57.6		
Thz6-1		162.4	160.6		
Thz6-2		146.8	144.9		
Thz6-3	8.18s	124.7	123.7	Thz6-1; Thz6-2; Thz6-4	
Thz6-4		170.4	169.4		
Thr7-1		166.1	164.5		
Thr7-2	4.44m	56.6	54.6	Thr7-1; Thr7-3	Thr7-3; Thr7-NH;
Thr7-3	1.59m	67.0	65.1		Thr7-2; Thr7-4
Thr7-4	0.86d(6.6)	16.1	17.3	Thr7-2; Thr7-3	Thr7-3
Thr7-NH	7.19d(7.6)			Thz6-1; Thr7-1	Thr7-3
Dhb8-1		129.1	127.5		
Dhb8-2	6.23q(7.0)	133.4	131.5	Dhb8-1; Dhb8-3; Thr7-1; Tzn9-4	Dhb8-3
Dhb8-3	1.61d(6.8)	16.0	13.7	Dhb8-1; Dhb8-2;	Dhb8-2

				Tzn9-4	
Dhb8-NH	8.69bs				
Tzn9-1		172.8	170.9		
Tzn9-2	4.98m	79.7	77.7	Tzn9-1; Tzn9-3; Tzn9-4	Tzn9-3- Ha; Tzn9-3Hb
Tzn9-3	Ha 3.15m Hb 3.66m	35.4	33.5	Tzn9-1; Tzn9-2; Tzn9-4	Tzn9-2; Tzn9-3Hb Tzn9-2; Tzn9-3Ha
Tzn9-4		170.9	170.9		
Ile10-1	5.75d(9.9)	53.7	52.3	Ile10-2; Tzn9-1; Thz11-4	Ile10-NH
Ile10-2		77.7	75.8		
Ile10-3	3.83m	68.6	66.9	Ile10-1; Ile10-2; Ile10-4	Ile10-4
Ile10-4	1.32d(6.4)	16.6	14.6	Ile10-2; Ile10-3	Ile10-3
Ile10-5	1.17s	19.5	17.5	Ile10-1; Ile10-2; Ile10-3	
Ile10-NH	7.5br s				Ile10-1
Thz11-1		166.9	189.9		
Thz11-2		150.8	154.1		
Thz11-3	8.28s	128.6	126.7	Thz11-1; Thz11-2; Thz11-4	
Thz11-4		166.4	164.8		
Thr12-1	5.81s	56.4	60.2	Thz11-1; Thz13-3	
Thr12-2	6.45m	72.7	70.8	Thr12-1; Thz13-3; Q-1	Thr12-3
Thr12-3	1.76d(6.6)	19.5	17.3	Thr12-1; Thr12-2	Thr12-2
Thr12-NH	8.82m				
Thz13-1		157.2	155.2		
Thz13-2	7.62s	118.9	117.1	Thz13-1; Thz13-3; Pip5-1	
Thz13-3		170.8	166.6		
Thz14-1		160.3	158.8		
Thz14-2		150.5	148.1		
Thz14-3	8.31s	128.3	123.1	Thz14-2; Thz14-4	
Thz14-4		169.0	171.7		
Dha15-1		162.6	160.9		
Dha15-2		134.9	133.4		
Dha15-3	Ha 6.67br. s Hb 5.58br. s	104.2	101.7	Dha15-1; Dha15-2 Dha15-1; Dha15-2	Dha15-3- Hb Dha15-3- Ha
Dha15-NH	10.00br. s				
Dha16-1		166.3	165.0		
Dha16-2		132.9	130.8		
Dha16-3	Ha 6.59br. s Hb 5.79br. s	110.4	107.3	Dha16-1; Dha16-2 Dha16-1; Dha16-2	Dha16-3- Hb Dha16-3- Ha
Q-1		161.4	159.7		
Q-2		144.4	142.5		
Q-3	7.32s	123.0	121.3	Q-1; Q-10; Q-11	
Q-4		154.4	152.5		
Q-5	6.93d(10.1)	124.1	122.2	Q-7; Q-9; Q-10	Q-6
Q-6	6.40m	130.5	128.8	Q-7; Q-8; Q-10	Q-5; Q-7
Q-7	3.61m	59.8	57.6	Q-5; Q-6; Q-8; Q-9; Val1-2	Q-6; Q-8

Q-8	4.46m	68.3	64.8	Q-6; Q-7; Q-9; Q-10	Q-7
Q-9		155.1	153.6		
Q-10		127.9	126.1		
Q-11	5.33m	65.1	66.5	Q-3; Q-12	Q-12
Q-12	1.41d(6.5)	23.2	21.2	Q-4; Q-11	Q-11

3. Supplementary references

1. T. Kieser et al., *Practical Streptomyces Genetics* (John Innes Foundation, **2000**).

2. Weber, T. et al., *Nucleic Acids Res.* 43, W237–W243 (**2015**).

3. States, D. J. & Gish, W., *J. Comput. Biol.* 1, 39–50 (**1994**).

4. Ishikawa J, Hotta K., *FEMS Microbiology Letters*, 174,251-253 (**1999**)

5. Gouet P, Robert X, Courcelle E., *Nucleic Acids Res.*, 31(13): 3320–3323 (**2003**).

6. Gerlt, J. A., et al., *Biochim. Biophys. Acta, Proteins Proteomics*, 1854, 1019−1037 (**2015**).

7. R. Liao & W. Liu, *J. Am. Chem. Soc.*, 133, 2852–2855 (**2011**).

8. Q. Zheng et al., *ACS Chem. Biol.*, 11, 2673–2678(**2016**).

9. H. Guo et al., *Chem. Sci*. 5, 240-246 (**2014)**.