Supplementary Information

Heterologous expression and *in vivo* characterization of the longipeptin biosynthetic gene cluster

Yuwei Duan[1], Menghan Shi[1], Weifeng Sun[1], Youming Zhang[1]*, Xiaoying Bian[1]*, Guannan Zhong[1]*

1, State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China.

* To whom correspondence should be addressed. Email: zhonggn@sdu.edu.cn, bianxiaoying@sdu.edu.cn, zhangyouming@sdu.edu.cn

Methods

General materials and methods

Chemical compounds and reagents were purchased from Sinopharm Chemical Reagent Co., Ltd. (China), Bide Pharmatech Ltd. (China), Macklin Biochemical Technology Co., Ltd. (China) or J&K Scientific Ltd. (China). Components of culture media were purchased from Oxoid Ltd. (UK) or Thermo Fisher Scientific Inc. (USA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) and antibiotics were purchased from Shanghai Sangon Biotech Co. Ltd. (China). Restriction endonucleases were purchased from New England Biolabs Ltd. (UK) or Monad Biotech Co., Ltd. (China). High-fidelity DNA polymerase was purchased from Accurate Biology (China). DNA gel extraction and plasmid extraction kits were purchased from Tiangen Biotech Co., Ltd. (China). Gibson assembly cloning kit was purchased from ABclonal Biotechnology Co., Ltd. (China). Primer synthesis and DNA sequencing were performed by Shanghai Sangon Biotech Co. Ltd. (China) or Tsingke Biotechnology Co., Ltd. (China). Gene synthesis and codon optimization were performed at GENEWIZ, Inc. (China). Polypeptides were synthesized by GenScript Biotech Corporation (China). Primers used in this study are summarized in Table S3. Plasmids and bacterial strains used in this study are listed in Tables S4 and S5, respectively.

Analysis

HPLC analysis was carried out on Thermo Fisher Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific Inc., USA). Low resolution MS analysis was performed on Thermo Fisher LTQ XL™ Linear Ion Trap Mass Spectrometer (Thermo Fisher Scientific Inc., USA), and the data were analyzed by Thermo Xcalibur software. HRMS analysis was carried out on Bruker High Resolution Q-TOF mass spectrometry (impactHD) (Bruker Co. Ltd, Germany) and the data were analyzed by Bruker Daltonics Data Analysis. Monoisotopic masses are commonly used in this article unless stated otherwise. Semi-preparative HPLC was performed on Shimadzu LC-20AT system (Shimadzu, Japan).

Heterologous expression of *lop* cluster in *S. coelicolor* A3(2)

The putative *lop* cluster along with flanking genes (i.e., from *-4* to *+2*) was divided into three parts. Primers lop-fra1-F/R, lop-fra2-F/R and lop-fra3-F/R were used individually for the amplification of these three parts from genomic DNA (gDNA) extracted from *Longimycelium tulufanense* TRM 46004 (CGMCC 4.5737). In addition, p15A vector with *φ*C31 integrase gene and *attP* (phage attachment site) was linearized using p15A-amp-attP-F/R primers. The four amplified DNA fragments were purified and assembled using Gibson method to give *lop* heterologous expression plasmid p15A-amp-attP-lop.[1] The plasmid was certified by restriction analysis with different enzymes and DNA sequencing.

The bidirectional promoter P_{apra}-apra-SP44 was then inserted between *lop-1* and *lopA* genes via Red/ET recombineering. Briefly, the plasmid p15A-amp-attP-lop was electroporated into *E. coli* GB05-red which carries a recombineering operon composed of *redα/redβ/redγ/recA* genes under

P_{BAD} promoter in its chromosome.[2] The bidirectional promoter P_{apra}-apra-SP44 was amplified using lop-SP44-apra-F/R primers and electroporated into *E. coli* GB05-red cells that had been cultured with L-arabinose for the inducible expression of Redα/Redβ/Redγ/RecA proteins. After selection by apramycin resistance, the obtained recombinant product p15A-amp-attP-P_{apra}-SP44-lop was subsequently verified by DNA sequencing.

The resulted plasmid p15A-amp-attP-P_{apra}-SP44-lop was electroporated into *E. coli* ET12567 (pUZ8002), conjugated into *S. coelicolor* A3(2) and finally integrated at the *φ*C31 *attB* (bacterial attachment site) locus of the *S. coelicolor* A3(2) chromosome. The spore of *S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop were inoculated into 50 mL TSB medium and shaken at 30 °C, 200 rpm for about 4-5 days. Then the seed culture was transferred into secondary flask-fermentation containing 50 mL R4 modified medium (glucose 0.5%, yeast extract 0.1%, $MgCl_2 \cdot 6H_2O$ 0.5%, $CaCl_2 \cdot 2H_2O$ 0.2%, K_2SO_4 0.1%, casamino acids 0.05%, L-proline 0.07%, L-valine 0.12%, TES (N-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid) 0.28%, and 50 μL trace elements solution ($ZnCl_2$ 40 mg/L, $FeCl_3 \cdot 6H_2O$ 20 mg/L, $CuCl_2 \cdot 2H_2O$ 10 mg/L, $MnCl_2 \cdot 4H_2O$ 10 mg/L, $Na_2B_4O_7 \cdot 10H_2O$ 10 mg/L and $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ 10 mg/L)).[3] After incubation at 30 °C and 200 rpm for 7 days, Amberlite XAD-16 polymeric adsorbent resin (Aladdin Scientific, USA) was added to the flask and further incubated for another 1 day. The resin and cells were collected and eluted with methanol. The methanolic elution was concentrated in vacuo and centrifugated to remove the residue. The supernatant was subjected to HPLC-(HR)MS analysis on an Acclaim™ RSLC 120 C18 column (2.1 × 100 mm, 2.2 μm, 120 Å, Thermo Fisher Scientific Inc., USA) by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) at a flow rate of 0.3 mL/min over a 25-min period as follows: T = 0 min, 5% B; T = 3 min, 5% B; T = 18 min, 95% B; T = 22 min, 95% B; T = 23 min, 5% B; and T = 25 min, 5% B.

Scarless gene deletions

The scarless gene deletions were performed by in-frame deletion to exclude polar effects on downstream gene expression using RedEx strategy.[4,5] In brief, p15A-amp-attP-P_{apra}-SP44-lop was electroporated into *ccdB* resistant *E. coli* strain GBred-gyrA462 and the expression of Redαβ recombinases was induced by L-arabinose. The fragment containing homology arm (HA) 1 (40 bp), *Pac*I, chloramphenicol resistance gene (Cm), *ccdB* gene, *Pac*I, 3'-terminal HA1 (25 bp) and HA2 (40 bp) was amplified by polymerase chain reaction (PCR) and electroporated into the above *E. coli* GBred-gyrA462 strain for recombineering. The recombinant plasmid carrying Cm-*ccdB* cassette in the target gene was identified by chloramphenicol resistance and confirmed by DNA sequencing. Then Cm-*ccdB* cassette was removed by *Pac*I digestion and Gibson assembly to stitch the incision. The circular recombined plasmid was electroporated into the *ccdB* sensitive *E. coli* strain GB2005 that with neither RecET nor Redαβ recombinase expression to minimize unexpected recombination. The in-frame deletion of the plasmid extracted from *E. coli* GB2005 was validated by DNA sequencing and the correct plasmid was conjugated into *S. coelicolor* A3(2) for fermentation.

Construction the plasmids for the heterogenous expression of peptide and proteins

lopA and *lopE* genes were codon optimized (Table S6), commercially synthesized and inserted into the *Bam*HI-*Hin*dIII site of the pRSFDuet-1 vector to afford the plasmids pRSFDuet-1 + *lopA* and pRSFDuet-1 + *lopE*, respectively. The DNA fragment containing optimized *lopE* and *lopA* was amplified by PCR and then inserted into the *Nde*I-*Xho*I site of pRSFDuet-1 + *lopA* and pRSFDuet-1 + *lopE* to yield the recombinant plasmid pRSFDuet-1 + *lopA* + *lopE* and pRSFDuet-1 + *lopE* + *lopA*, respectively.

The DNA fragment of *lopB* containing homology arms of pCold-TF was amplified by PCR. The pCold-TF vector was linearized by *Nde*I and *Xho*I digestion and assembled with *lopB* fragment to yield the recombinant plasmid pCold-TF + *lopB*.

The DNA fragment of *lopF or lopG* containing homology arms of pET28a(+) was amplified by PCR. The pET28a(+) vector was linearized by *Nde*I and *Xho*I digestion and assembled with *lopF* and *lopG* fragment to yield the recombinant plasmid pET28a + *lopF*, pET28a + *lopG*, respectively. The *lopH* gene was codon optimized (Table S6), commercially synthesized and inserted into the *Nde*I-*Xho*I site of the pET28a(+) vector to afford the plasmids pET28a + *lopH.*

Expression and purification of precursor peptide and proteins

The constructed plasmids were electroporated into *E. coli* BL21 (DE3) for heterogenous expressions. The transformed colonies were identified by kanamycin resistance (50 μg/mL) except pCold-TF + *lopB* by ampicillin resistance (100 μg/mL). A single colony was chosen and inoculated into 100 mL of LB with kanamycin or ampicillin respectively, and cultivated at 37 °C and 200 rpm overnight. The seed cultures were utilized for the inoculation of 4 L of LB medium containing kanamycin (50 μg/mL) or ampicillin (100 μg/mL). These cultures were incubated at 37 °C and 200 rpm until OD_{600} reached 0.6-0.8. Then the incubation temperature was reduced to 20 °C except for pCold-TF + *lopB*, which was reduced to 15 °C. IPTG was then added to the cultures to a final concentration of 100 μM. These cultures were further incubated at 20 °C or 15 °C and 200 rpm for 24 h. Cells were harvested by centrifugation at 6,000 rpm for 10 min and stored at -80 °C if not purified immediately. The supernatant was decanted and discarded.

For the purification of precursor peptide (LopA), the cells were resuspended in lysis buffer containing 4 M guanidine hydrochloride, 20 mM NaH_2PO_4 (pH 7.5), 500 mM NaCl and 0.5 mM imidazole. After disruption by high-pressure homogenizer (JN-Mini Pro, Guangzhou Juneng Nano & Bio Technology Co., Ltd., China) at 4 °C, the insoluble material was removed by centrifugation at 4 °C, 12000 rpm for more than 1 h. The supernatant was applied to 5 mL Ni-NTA agarose resin (Bestchrom, China) and washed with elution buffer containing a low concentration of imidazole (25 mM) to remove the impurities. The sample was further eluted with elution buffer containing a high concentration of imidazole (1 M) to obtain purified peptide. The presence and stability of the target peptide were monitored by HPLC-HRMS. The eluted fractions were desalted by reversed phase (RP) HPLC on a Shimadzu LC-20AT system (Shimadzu Corporation, Japan) equipped with a ReproSil 300 C4 column (10 × 250 mm, 10 μm, 300 Å, Dr. Maisch, Germany) using gradient elution of solvent A (H_2O containing 0.1% trifluoroacetic acid (TFA)) and solvent B (CH_3CN containing 0.1% TFA) with a flow rate of 3 mL/min over a 82-min period as follows: T = 0 min, 5 % B; T = 15 min,

5 % B; T = 65 min, 100 % B; T = 72 min, 100 %; T = 77 min, 5 % B; and T = 82 min, 5 % B (mAU at 220 nm). Desalted precursor peptides were lyophilized and stored at -20 °C.

For the purification of proteins (LopE, LopB, LopF, LopG, LopH), the cells were resuspended in lysis buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 5 mM imidazole and 10% (v/v) glycerol. After disruption by high-pressure homogenizer at 4 °C, the insoluble material was removed by centrifugation at 4 °C, 12000 rpm for more than 1 h. The supernatant was applied to 5 mL Ni-NTA agarose resin (Bestchrom, China) and washed with elution buffer by a stepwise gradient containing different concentrations of imidazole (10-500 mM). The elution fractions were determined by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis, and fractions containing the recombinant protein were combined and desalted using a PD-10 Desalting Column (GE Healthcare, USA) into storage buffer containing 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% (v/v) glycerol, and 1 mM DTT. The resulting proteins were concentrated by Amicon Ultra Centrifugal Filter (Millipore Sigma, USA) and stored at -80°C. The purity of proteins was dictated by SDS-PAGE analysis, and the concentrations were measured by protein absorbance at 280 nm. The functional P450 concentrations were determined from the CO-reduced difference spectra using the extinction coefficient of $\epsilon_{450-490}$ = 91,000 M^{-1} cm^{-1}.[6]

Site-directed mutagenesis of LopA

The plasmids containing LopA mutant were generated using the primers listed in Table S3 for PCR amplification with pRSFDuet-1 + *lopA* + *lopE* as the template. Then, 1 μL of *Dpn*I enzyme was added to the PCR system and incubated at 37 °C for 3 h to digest parental plasmid. The digested products were transformed into *E. coli* GB2005 for single colony isolation. After sequencing to validate the fidelity, the resulting mutant precursor peptide was expressed in *E. coli* BL21(DE3), purified to homogeneity, desalted, and lyophilized according to the procedures described above for the native precursor peptides.

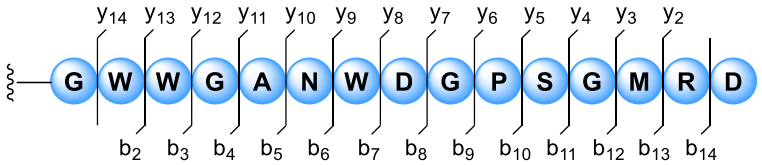
***In vitro* Enzymatic Assay**

The assay for TF-LopB (total volume, 100 μL) was performed at 30 °C for 12 h in 50 mM Tris-HCl buffer (pH 8.0) containing 100 μM LopA, 100 mM NaCl, 2 mM $MgCl_2 \cdot 6H_2O$, 2 mM ATP, 1 mM DTT, 20 μM LopE and 20 μM TF-LopB. ATP was added last to initiate the reaction. The assays were quenched by adding an equal volume of acetonitrile and centrifugated at 12,000 rpm for 15 min to remove the denatured proteins. The supernatant was subjected to HPLC-HRMS analysis on an Acclaim^{TM} RSLC 120 C18 column (2.1 × 100 mm, 2.2 μm, 120 Å, Thermo Fisher Scientific Inc., USA) by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) at a flow rate of 0.3 mL/min over a 25-min period as follows: T = 0 min, 5% B; T = 3 min, 5% B; T = 18 min, 95% B; T = 22 min, 95% B; T = 23 min, 5% B; and T = 25 min, 5% B.

Table S1. Proteins encoded in the *lop* cluster and its flanking regions as well as their putative functions.

Protein	Amino acids	Proposed function	Accession number
Lop(-4)	657	Maleylpyruvate isomerase N-terminal domain-containing protein	WP_189061734.1
Lop(-3)	876	Transposase	-
Lop(-2)	1371	Dihydrolipoyl dehydrogenase family protein	WP_189061733.1
Lop(-1)	1422	DEAD/DEAH box helicase	WP_189061732.1
LopA	117	Precursor peptide	WP_189061731.1
LopC	1613	ATP-dependent macrolactam synthetase	WP_194500064.1
LopF	1215	Cytochrome P450	WP_189061729.1
LopD	1791	ABC transporter	WP_189061728.1
LopG	1221	Cytochrome P450	WP_189061727.1
LopH	576	DUF6919 domain-containing protein	WP_189061726.1
LopE	258	PqqD family protein	WP_189061725.1
LopB	414	ATP-dependent cysteine protease	WP_189061724.1
Lop(+1)	90	Hypothetical protein	WP_189061723.1
Lop(+2)	159	Hypothetical protein	WP_189061722.1

Table S2. MS/MS fragmentation analysis of LopA core peptide generated by TF-LopB and LopE.



Ions	Calc.	Obs.	Er. (ppm)	Ions	Calc.	Obs.	Er. (ppm)
b1	58.0288	-	-	**y14**	1634.6805	817.8468 ($[M+2H]^{2+}$)	3.6
b2	244.1081	244.1081	0.2	**y13**	1448.6012	1448.6029	1.2
b3	430.1874	430.1869	1.1	**y12**	1262.5219	1262.5198	1.6
b4	487.2088	487.2083	1.1	**y11**	1205.5004	1205.4966	3.2
b5	558.2459	558.2454	1.0	**y10**	1134.4633	1134.4607	2.3
b6	672.2889	672.2879	1.4	**y9**	1020.4204	1020.4185	1.8
b7	858.3682	858.3676	0.7	**y8**	834.3410	834.3398	1.5
b8	973.3951	973.3941	1.1	**y7**	719.3141	719.3130	1.5
b9	1030.4166	1030.4159	0.7	**y6**	662.2926	662.2920	1.0
b10	1127.4694	1127.4720	2.4	**y5**	565.2399	565.2384	2.6
b11	1214.5014	1214.4992	1.8	**y4**	478.2078	478.2072	1.3
b12	1271.5229	1271.5293	5.1	**y3**	421.1864	421.1858	1.4
b13	1402.5633	1402.5625	0.6	**y2**	290.1459	290.1455	1.4
b14	1558.6644	779.8367 ($[M+2H]^{2+}$)	1.1	**y1**	134.0448	-	-

Table S3. Primers used in this study. The restriction enzyme cutting sites are highlighted in bold font. The mutation sites are colored in red.

Primer	Sequence (5'-3')
lop-fra1-F	CGACCTTGCCCCTCCAACGTCATCTCGTTCTCCGCTCAGGAAGTGACGTGGAGATGTGC
lop-fra1-R	CGACGCCGGAACCACCGACGTGATGGCTACCCTGCAGGCCGTCGACCTATGTGAGTGGA
lop-fra2-F	CCCCGACCGGAGTGTCGAGTCCCTCCGGTAGTGAGTCGATCCACTCACATAGGTCGACG
lop-fra2-R	CGGATCTGGTCCATATTTCGTAAGCGGGGCCAGAACCCACAACCCACGTGTGAAGTCAG
lop-fra3-F	CGGGTTGCGATGGTCGAGAACTTCTTCGCGCAGGGGTGTCTGACTTCACACGTGGGTTG
lop-fra3-R	GCCTTTCGTTTTATTTGATGCCTGGAGATCCTTAAGATCAGTTTTCCGCCACGTTGGGA
p15A-amp-attP-F	CGGGACGGCTGCCATCTCGTCCCAACGTGGCGGAAAACTGATCTTAAGGATCTCCAGGC
p15A-amp-attP-R	CCAACCGAGGGACTCCCATAGCACATCTCCACGTCACTTCCTGAGCGGAGAACGAGATG
lop-SP44-apra-F	CCCACCAGCCCTGGGTGTCACACTCGAAGGTGCCCACCTCAACCAGCGCCGGCGACTCGTAGGCAGTCTGCTCCTGCATCAGCCTACTCCTTACTTAGA
lop-SP44-apra-R	AAGGACAAGGACGAGGTGGCAGGGGCAGAAGAGGTGTGCAGGGGGCGGCTGACCGCCATAAATTCTCCGAGAGGTAGCATCAGCCAATCGACTGGCGAG
lop-delta(-1)-F	AGGCAAGACCTATGCGTTCGTCCTGCCGGTCCTGGCCCGG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-delta(-1)-R	TGGGCGGCGGGGTGAAGGTGCGTTCACCGGGAGCGAGTTCCCGGGCCAGGACCGGCAGGACGAAC**TTAATTAA**TTTGTTCAAAAAAAAGCCCGC
lop-delta(-2)-(-4)-F	CGGCCATGATCGTTGCACCCAGTACCCGTCCGCCCCCCAG**TTAATTAA**TTTGTTCAAAAAAAAGCCCGC
lop-delta(-2)-(-4)-R	GCAGCCGTCCGGCTGTACCGGCTGGCTCGTGCGGGACCTGCTGGGGGGCGGACGGGTACTGGGTG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-delta(-1)-(-4)-F	AGGCAAGACCTATGCGTTCGTCCTGCCGGTCCTGGCCCGG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-delta(-1)-(-4)-R	GCAGCCGTCCGGCTGTACCGGCTGGCTCGTGCGGGACCTGCCGGGCCAGGACCGGCAGGACGAAC**TTAATTAA**TTTGTTCAAAAAAAAGCCCGC
lop-delta(+1)-(+2)-F	CGGAGGGCTGACCTTGACGGCGACCGTGGCCAAGTTTGAG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-delta(+1)-(+2)-R	CCTGAGAGAGCGAGCCCGACGCTCTTCAGGGCCACGGCGACTCAAACTTGGCCACGGTCGCCGTC**TTAATTAA**TTTGTTCAAAAAAAAGCCCGC
lop-deltaA-F	TCCGATCAGTTAATGCGGGTCAGTCGCGCATTCCACTCGG**TTAATTAA**TTTGTTCAAAAAAAAGCCCGC

lop-deltaA-R	CTAAGTAAGGAGTAGGCTGATGCAGGAGCAGACTGCCTACCCGAGTGGA ATGCGCGACTGACCCG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaE-F	ATACAGTCACCAGCTTAGCTCGCCGGAGGTGCTCCACCAG**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaE-R	GCCGTTCACCCTGCGTCGGGACGTATCGATGACGACGGCTCTGGTGGAG CACCTCCGGCGAGCTA**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaB-F	GGTAGTAGCCGGCACCGATCGGCTCGCCGACCGGCTGGCC**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaB-R	CCCCAAGCCCCGGCTGTACCGCCGCCTCGTCGCGCGTGTGGGCCAGCCG GTCGGCGAGCCGATCG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaC-F	CAATTCCTCGCATCGCCTCGACCAGGAGCGGCTTGTACTG**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaC-R	CCGGCTCGCCGCGACCCTGCCGGGCAGCTTCCACCTGATCCAGTACAAG CCGCTCCTGGTCGAGG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaD-F	CATCGAGAATCAGGACATCCGCTTGGCGCAGCAGTGCTCT**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaD-R	GGAACGATTTGGCCACCGCGTGATCTTCGGAGTCCGTCAAAGAGCACTG CTGCGCCAAGCGGATG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaF-F	GGGCGCCGTGGCCGAACGCCACGTGCGAGGTGTCCTGCCG**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaF-R	GTCCAAAGCCCTTCTGCAGGACAGCCGCCTGCAACGCAACCGGCAGGAC ACCTCGCACGTGGCGT**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaG-F	GGGTACCGAGGGCGAGCTCGGCTTCCATCTTGGCCAGGCG**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaG-R	CACCGATCCACGGCTCAGCAACAAGTCTTATGCCGCAGCGCGCCTGGCC AAGATGGAAGCCGAGC**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaH-F	CCCACACCGGATCCACGAGGGTGGGCTGGTCTGCCGAGCA**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaH-R	TGAACTGACCGCCCACTGGCTGACGGGTGACCTCGCCTCCTGCTCGGCA GACCAGCCCACCCTCG**TTAATTAA**GACGTTGATCGGCACGTAAG
lop-deltaGH-F	CCCACACCGGATCCACGAGGGTGGGCTGGTCTGCCGAGCA**TTAATTAA**T TTGTTCAAAAAAAAGCCCGC
lop-deltaGH-R	CACCGATCCACGGCTCAGCAACAAGTCTTATGCCGCAGCGTGCTCGGCA GACCAGCCCACCCTCG**TTAATTAA**GACGTTGATCGGCACGTAAG
LopE-F	TAGTATATTAGTTAAGTATAAGAAGGAGATATA**CATATG**CCGTTTACCC TGCGCCGCGA
LopE-R	TTCAAATTTCGCAGCAGCGGTTTCTTTACCAGA**CTCGAG**TTACACGGTC ACCAGTTTCG
LopA-F	TAGTATATTAGTTAAGTATAAGAAGGAGATATA**CATATG**CAAGAACAGA CCGCGTATGA
LopA-R	TTCAAATTTCGCAGCAGCGGTTTCTTTACCAGA**CTCGAG**TTAATCGCGC ATGCCGCTCG
LopF-F	TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC**CATATG**ACCGAT ACCCTACCGT

LopF-R	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTG**CTCGAG**TCAGCCAAG CTGGACGGGA
LopG-F	TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC**CATATG**GACACG GATGCCGACC
LopG-R	TAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTG**CTCGAG**TCATCCTCG ATCGTCCCCT
LopA-C(-4)A-F	GTGGGCACCTTTGAA**GCA**GATACCCAAGGCTGG
LopA-C(-4)A-R	CCAGCCTTGGGTATC**TGC**TTCAAAGGTGCCCAC

Table S4. Plasmids used in this study.

Plasmid	Characteristic(s)	Source/ Reference
p15A-amp-attp	Heterologous expression vector with φC31 integrase gene and *attP*, amp^R	[7]
p15A-amp-attP-lop	p15A-amp-attp derivative containing *lop* cluster and the flanking regions	This study
p15A-amp-attP-P_{apra}-SP44-lop	p15A-amp-attP-lop derivative inserted P_{apra}-apra-SP44 bidirectional promoter between *lop-1* and *lopA*, apra^R	This study
p15A-cm-ccdB	PCR template of the *cm-ccdB* cassette	[2]
p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-1)*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *Δlop(-1)*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-2)-(-4)*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *Δlop(-2)-(-4)*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-1)-(-4)*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *Δlop(-1)-(-4)*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(+1)-(+2)*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *Δlop(+1)-(+2)*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopA*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopA*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopE*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopE*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopB*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopB*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopC*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopC*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopD*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopD*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopF*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopF*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopG*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopG*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopH*	p15A-amp-attP-P_{apra}-SP44-lop derivative deleting *ΔlopH*	This study
p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopFGH*	p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopF* derivative deleting *ΔlopGH*	This study
pRSFDuet-1	Protein expression vector used in *E. coli*, containing two multiple cloning sites (MCS) and encoding *N*-terminal 6 × His-tag in the first MCS, kan^R	Novagen
pRSFDuet-1 + *lopA* + *lopE*	pRSFDuet-1 derivative containing *lopA* (*Bam*HI + *Hin*dIII) and *lopE* (*Nde*I + *Xho*I)	This study

pRSFDuet-1 + *lopA* C(-4)A +*lopE*	pRSFDuet-1 + *lopA* + *lopE* derivative containing *lopA* C(-4)A	This study
pRSFDuet-1 + *lopE* + *lopA*	pRSFDuet-1 derivative containing *lopE* (*Bam*HI + *Hin*dIII) and *lopA* (*Nde*I + *Xho*I)	This study
pCold-TF	Protein expression vector used in *E. coli*, encoding *N*-terminal 6 × His-tag, Trigger Factor (TF), amp^R	TaKaRa
pCold-TF + *lopB*	pCold-TF derivative containing *lopB* (*Nde*I + *Xho*I)	This study
pET-28a(+)	Protein expression vector used in *E. coli*, encoding *N*-terminal 6 × His-tag, kan^R	Novagen
pET-28a(+)+*lopF*	pET28a(+) derivative containing *lopF* (*Nde*I + *Xho*I)	This study
pET-28a(+)+*lopG*	pET28a(+) derivative containing *lopG* (*Nde*I + *Xho*I)	This study
pET-28a(+)+*lopH*	pET28a(+) derivative containing *lopH* (*Nde*I + *Xho*I)	This study

Table S5. Strains used in this study.

Strain	Characteristic(s)	Source/Reference
L. tulufanense TRM 46004 (CGMCC 4.5737)	Original strain for longipeptins	[8]
Streptomyces		
S. coelicolor A3(2)	Heterologous expression host bacteria	[9]
S. coelicolor A3(2)/ lop	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop	This study
S. coelicolor A3(2)/ *Δlop(-1)*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-1)*	This study
S. coelicolor A3(2)/ *Δlop(-2)-(-4)*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-2)-(-4)*	This study
S. coelicolor A3(2)/ *Δlop(-1)-(-4)*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(-1)-(-4)*	This study
S. coelicolor A3(2)/ *Δlop(+1)-(+2)*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*Δlop(+1)-(+2)*	This study
S. coelicolor A3(2)/ *ΔlopA*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopA*	This study
S. coelicolor A3(2)/ *ΔlopE*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopE*	This study
S. coelicolor A3(2)/ *ΔlopB*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopB*	This study
S. coelicolor A3(2)/ *ΔlopC*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopC*	This study
S. coelicolor A3(2)/ *ΔlopD*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopD*	This study
S. coelicolor A3(2)/ *ΔlopF*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopF*	This study
S. coelicolor A3(2)/ *ΔlopG*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopG*	This study
S. coelicolor A3(2)/ *ΔlopH*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopH*	This study
S. coelicolor A3(2)/ *ΔlopFGH*	*S. coelicolor* A3(2) integrated with p15A-amp-attP-P_{apra}-SP44-lop-*ΔlopFGH*	This study
E. coli		
GB2005	*ccdB* sensitive, containing neither RecET nor Redαβ recombinase expression system	[10,11]
GB05-red	GB2005 derivative, carries a recombineering operon composed of *redα*/*redβ*/*redγ*/*recA* genes under P_{BAD} promoter in its chromosome	[10]
GBred-gyrA462	GB05-red derivative, carries the GyrA R462C mutation to confer *ccdB* resistance	[7]

ET12567(pUZ8002)	Conjugative transfer donor bacteria	[12]
BL21 (DE3)	Heterologous host for protein expression	NEB
BL21 (DE3) pRSFDuet-1+*lopA*+*lopE*	BL21 (DE3) derivative, containing pRSFDuet-1 + *lopA* +*lopE* for the expression of LopA	This study
BL21 (DE3) pRSFDuet-1+*lopA* C(-4)A +*lopE*	BL21 (DE3) derivative, containing pRSFDuet-1 + *lopA* C(-4)A +*lopE* for the expression of LopA C(-4)A	This study
BL21 (DE3) pRSFDuet-1+*lopE*+*lopA*	BL21 (DE3) derivative, containing pRSFDuet-1 + *lopE* +*lopA* for the expression of LopE	This study
BL21 (DE3) pCold-TF + *lopB*	BL21 (DE3) derivative, containing pCold-TF + *lopB* for the expression of TF-LopB	This study
BL21 (DE3) pET-28a(+) + *lopF*	BL21 (DE3) derivative, containing pET-28a(+) + *lopF* for the expression of LopF	This study
BL21 (DE3) pET-28a(+) + *lopG*	BL21 (DE3) derivative, containing pET-28a(+) + *lopG* for the expression of LopG	This study
BL21 (DE3) pET-28a(+) + *lopH*	BL21 (DE3) derivative, containing pET-28a(+) + *lopH* for the expression of LopH	This study

Table S6. Sequences of synthetic genes optimized for *E. coli* expression.

Gene Name	Sequence
lopA	ATGCAAGAACAGACCGCGTATGAAAGCCCGGCGCTGGTGGAAGTGGGCACCTTT GAATGCGATACCCAAGGCTGGTGGGGCGCGAACTGGGATGGCCCGAGCGGCATG CGCGATTAA
lopE	ATGCCGTTTACCCTGCGCCGCGATGTGAGCATGACCACCGCGGATGATGCGACC GTGCTGCTGGATGAACGCACCGGCACCTATTGGCAGCTGAACCCGAGCGGCAGC CTGGTTCTGAGCACCCTGCTGAGTGGCGGCAGCCCGCAGCAAGCGACCGAACAG CTGGTGGAACGCTATGCGGTGGAACCGGCGCGCGCGGCGGATGATGTGACGAGC CTGGTGGAACATCTGCGCCGCGCGAAACTGGTGACCGTGTAA
lopH	ATGAACGAATATGAAGTGCAGCAGTGGCGCACCGCGCGCGGCCTGGGCGATCTG GGCGAACTGACCGCGCATTGGCTGACCGGCGATCTGGCGAGCCGCCCGGGTTAT CCGCCGGGCGAAGGCCCGTATGTGGAAACCCATGAACTGATTGGCACCCTGGCG GCGTGCAACCGCGGCGGCTTTGTGACCAACGCGAGTCAGCCGGGCTTTCCGGAA AGCGCGGGCCCGGATGGCGTGGTTTGGGTGCAGCGCGCGGCGGTGACCGGCTTT GCGGAAGAGGAAGTGCTGGAACGCATTCGCGCCGCGGTGGCGGGCACCGAACTG GTGCTGCTGACCCCGACCGCGACCGATCGCAACCCGGATGGCATTGTGGTGACC ACCCGCAACGGCAAACAGCATACCACCTTTGGCACCTGGCTGAGCCGCCGCGCG GTGGCGGATAGCTTTGATGTGTGCAAAAGCGCGGCGATTGATGCGCTGTGCAGC GCGGATCAGCCGACCCTGGTGGATCCGGTGTGGGGCCGCAACACCGTGCTGTGG CCGTTTCTGGATGATCTGTTTGGCACCACGACCTAA

Figure S1. HPLC-HRMS analysis of the fermentation for *Longimycelium tulufanense* TRM 46004 (CGMCC 4.5737) in modified R4 medium.

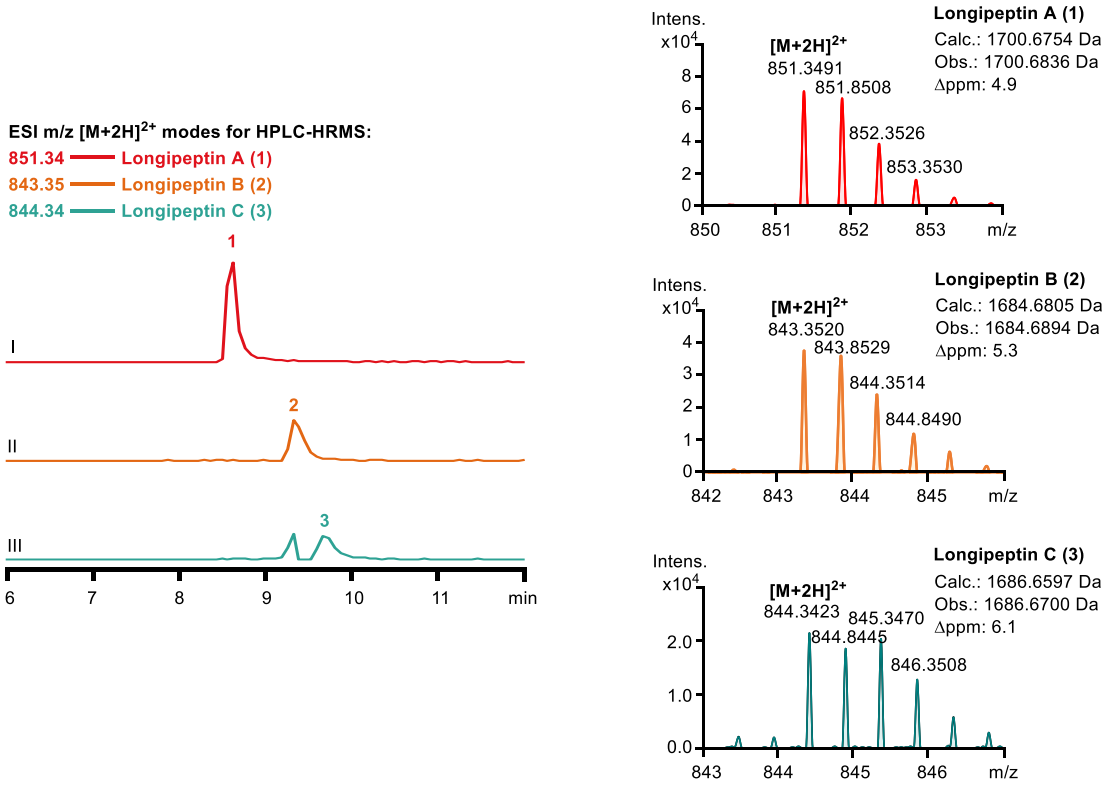


Figure S2. The map of the heterologous expression plasmid p15A-amp-attP-P_{apra}-SP44-lop. The bidirectional promoter P_{apra}-apra-SP44 was inserted between *lop(-1)* and *lopA*.

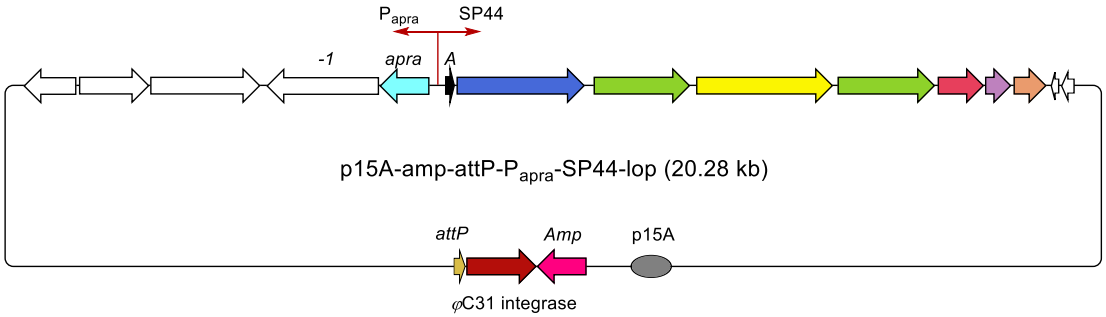


Figure S3. Denaturing SDS-PAGE analyses of proteins. Lane 1, Protein standard; Lane 2, LopE (10.80 kDa); Lane 3, Protein standard; Lane 4, TF-LopB (66.84 kDa), Lane 5, LopF (47.69 kDa); Lane 6, Protein standard; Lane 7, LopG (45.86 kDa); Lane 8, Protein standard; Lane 9, LopH (22.71 kDa); Lane 10, Protein standard; Lane 11, *Sel*Fdx (12.82 kDa); Lane 12, Protein standard; Lane 13, Protein standard; Lane 14, *Sel*FdR (46.60 kDa); Lane 15, CamB (11.80 kDa); Lane 16, Protein standard; Lane 17, CamA (47.96 kDa); Lane 18, Protein standard; Lane 19, PuxB (13.55 kDa); Lane 20, Protein standard; Lane 21, PuR (45.71 kDa); Lane 22, Protein standard.

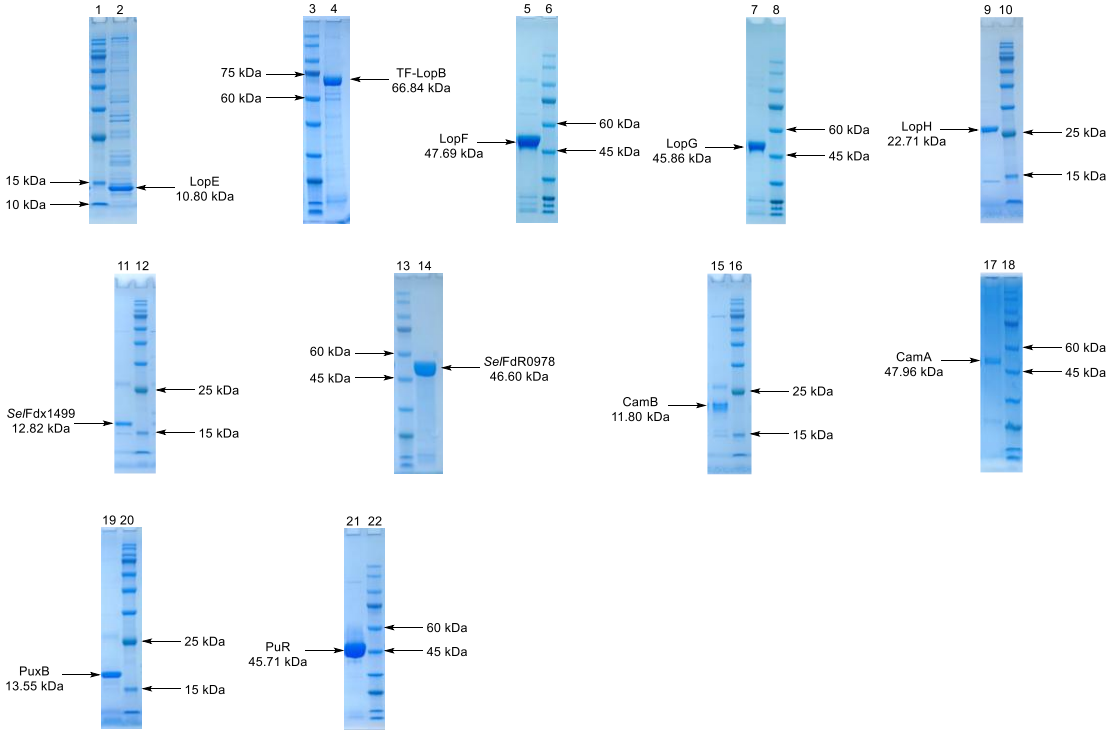


Figure S4. CO-bound reduced difference spectra of purified LopF and LopG. Orange line, P450 absorbance spectra in ferrous CO-complexed state. Cyan line, P450 absorbance spectra in ferric state.

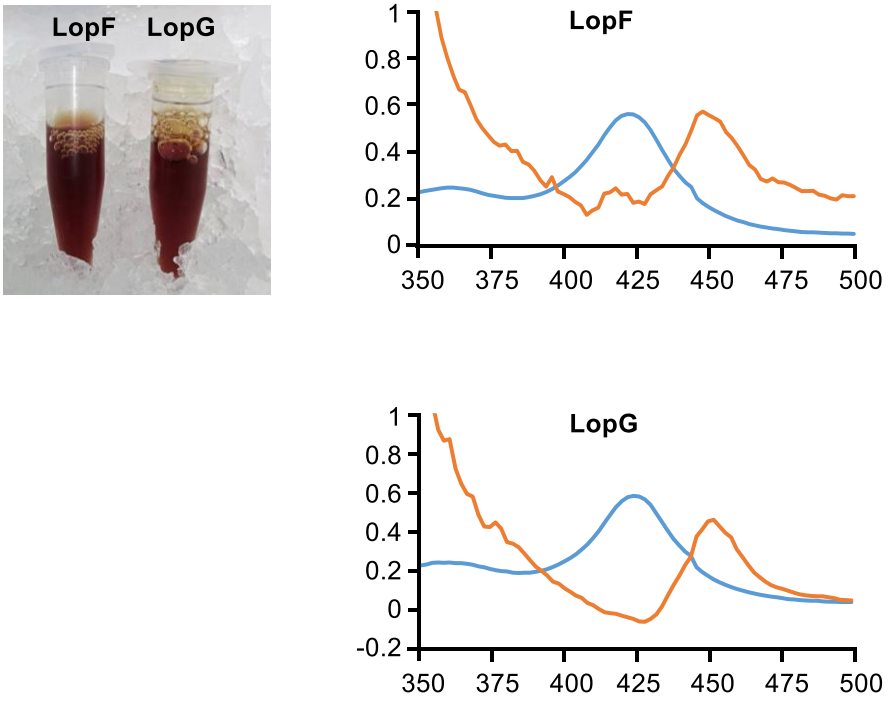


Figure S5. *In vitro* characterization of LopF, LopG and LopH with LopA C(-4)A or LopA-core. (**A**) HPLC-HRMS analysis of LopF and LopG with LopA C(-4)A or LopA-core in the presence of II. CamB/CamA, III. PuxB/PuR, IV. *Sel*Fdx1499 and *Sel*FdR0978 partners. (**B**) HPLC-HRMS analysis of LopH with LopA C(-4)A or LopA-core.

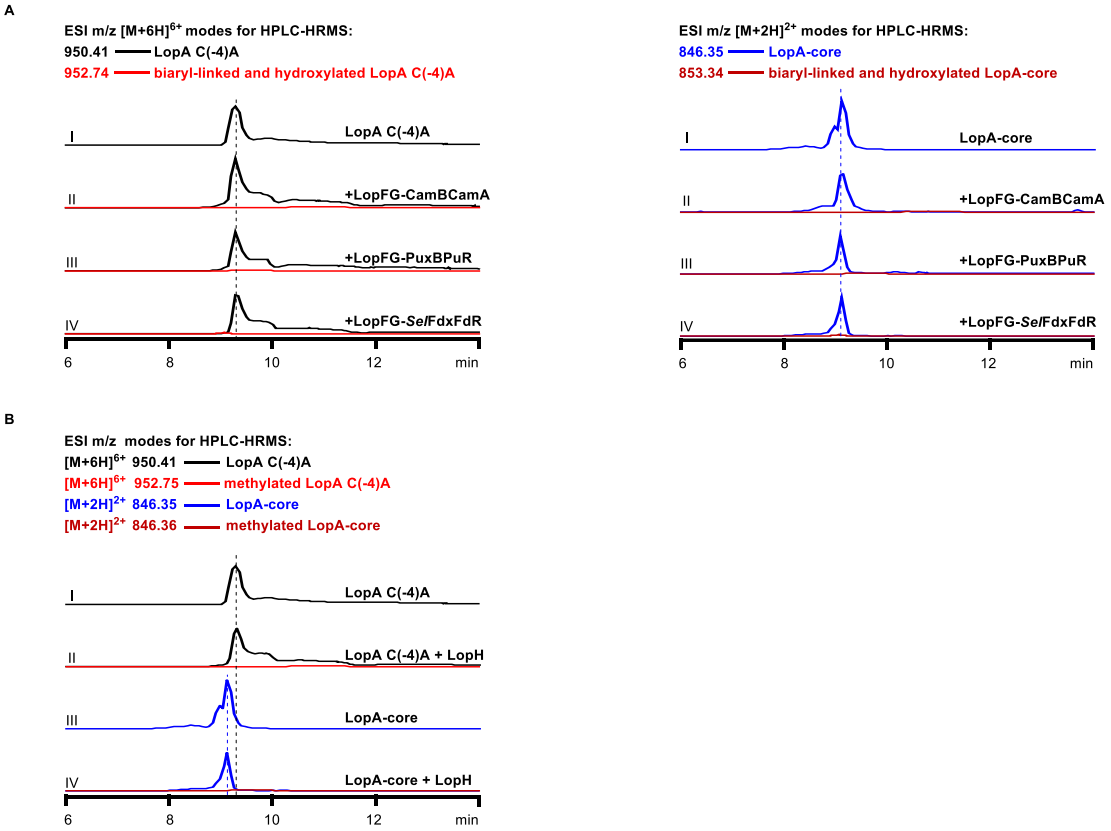


Figure S6. The proposed biosynthetic pathway of longipeptins.

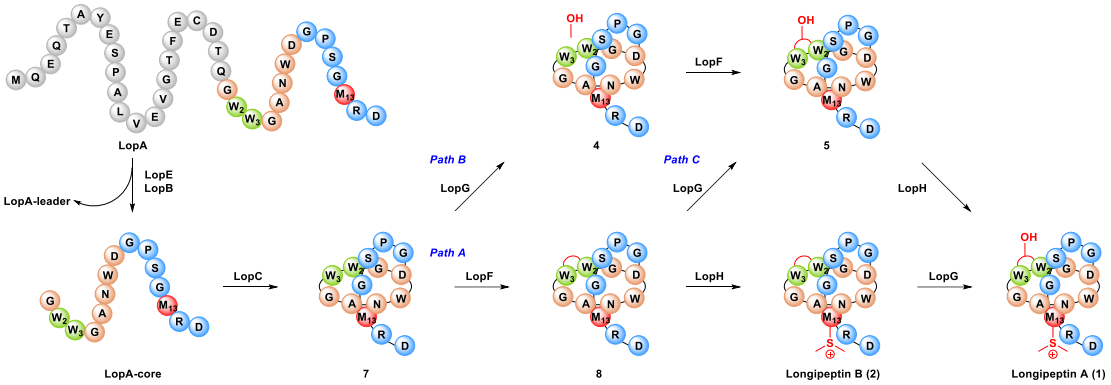


Figure S7. Trivalent sulfonium-containing natural products.

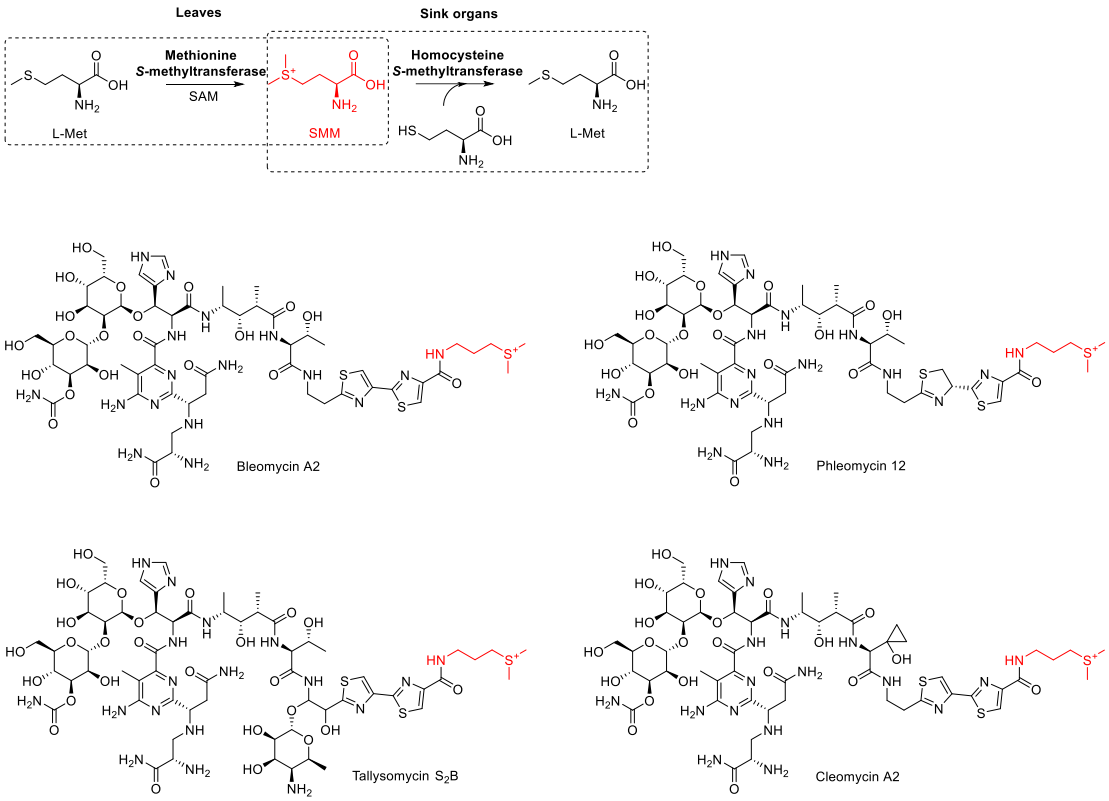
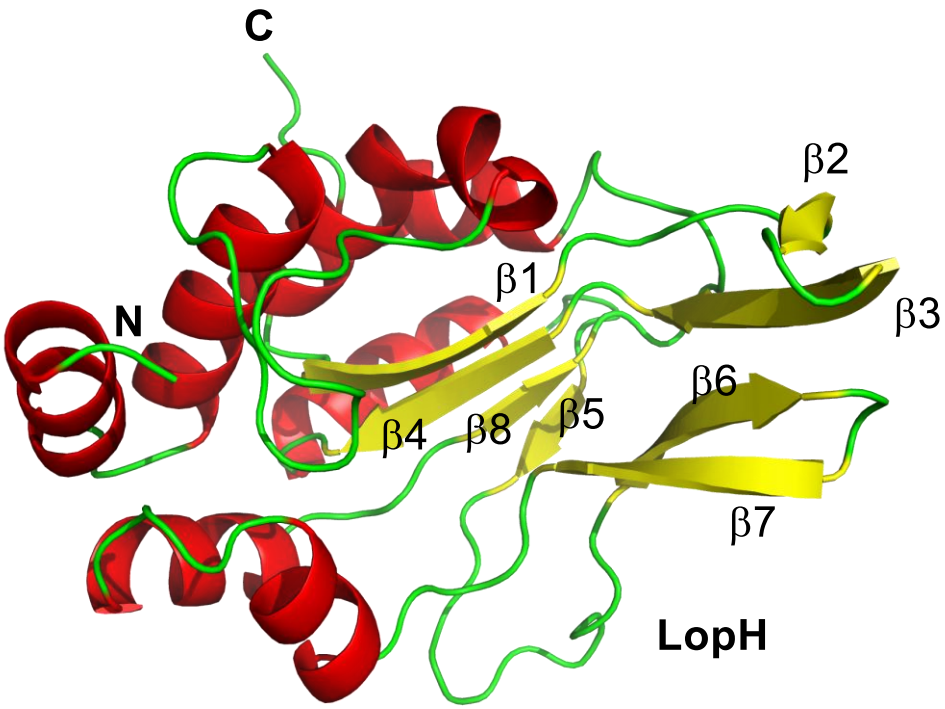


Figure S8. AlphaFold3-predicted structure of LopH.



References

1 D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison III and H. O. Smith, *Nat. Methods*, 2009, **6**, 343–345.

2 H. Wang, X. Bian, L. Xia, X. Ding, R. Müller, Y. Zhang, J. Fu and A. F. Stewart, *Nucleic Acids Res.*, 2014, **42**, e37.

3 H. Saad, T. Majer, K. Bhattarai, S. Lampe, D. T. Nguyen, M. Kramer, J. Straetener, H. Brötz-Oesterhelt, D. A. Mitchell and H. Gross, *Chem. Sci.*, 2023, **14**, 13176-13183.

4 C. Song, J. Luan, R. Li, C. Jiang, Y. Hou, Q. Cui, T. Cui, L. Tan, Z. Ma, Y. J. Tang, A. F. Stewart, J. Fu, Y. Zhang and H. Wang, *Nucleic Acids Res.*, 2020, **48**, e130.

5 J. Luan, C. Song, Y. Liu, R. He, R. Guo, Q. Cui, C. Jiang, X. Li, K. Hao, A. F. Stewart, J. Fu, Y. Zhang and H. Wang, *Nat. Protoc.*, 2024, **19**, 3360-3388.

6 F. P. Guengerich, M. V. Martin, C. D. Sohl and Q. Cheng, *Nat Protoc.*, 2009, **4**, 1245–1251.

7 H. Wang, Z. Li, R. Jia, Y. Hou, J. Yin, X. Bian, A. Li, R. Müller, A. F. Stewart, J. Fu and Y. Zhang, *Nat. Protoc.*, 2016, **11**, 1175–1190.

8 Z. F. Xia, T. W. Guan, J. S. Ruan, Y. Huang and L. L. Zhang, *Int. J. Syst. Evol. Microbiol.*, 2013, **63**, 2813–2818.

9 S. D. Bentley, K. F. Chater, A. M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood, *Nature*, 2002, **417**, 141–147.

10 J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P. M. Seibert, A. Plaza, L. Xia, R. Müller, A. F. Stewart and Y. Zhang, *Nat. Biotechnol*., 2012, **30**, 440–446.

11 J. Fu, M. Teucher, K. Anastassiadis, W. Skarnes and A. F. Stewart, *Method. Enzymol*., 2010, **477**, 125–144.

12 M. Bierman, R. Logan, K. O'Brien, E.T. Seno, R. N. Rao and B. E. Schoner, *Gene*, 1992, **116**, 43–49.