Electronic Supplementary Information

Divergent Biosynthesis of Indole Alkaloids FR900452 and Spiro-maremycins

Yingyi Duan,[a] Yanyan Liu,[a] Tao Huang,[b] Yi Zou,[c] Tingting Huang,[a] Kaifeng Hu,[b] Zixin Deng[a] and Shuangjun Lin*[a]

[a] State Key Laboratory of Microbial Metabolism, Joint International Laboratory of Metabolic & Developmental Sciences, School of Life Sciences & Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China.

[b] Kunming Institute of Botany, Chinese Academy of Science, Kunming, P. R. China.

[c] College of Pharmaceutical Science and Chinese Medicine, Southwest University, Chongqing, P. R. China.

*Correspondence to S.L.: linsj@sjtu.edu.cn

Contents

General

Strains and plasmids/cosmids used in this study are listed in Table S1. DNA isolation and manipulation were carried out following standard procedures for *E.* coli and *Streptomyces.* Primers were synthesized by Sangon Biotech Co., Ltd (Shanghai, China) and DNA sequencing was accomplished at Shanghai Majorbio Biotech Co., Ltd (Shanghai, China). PCR amplifications were performed on a Veriti Thermal Cycler (Applied Biosystems, Carlsbad, CA) using high fidelity PCR DNA-polymerase or Taq DNA polymerase (Vazyme Biotech Co. Ltd., Nanjing, China). EasyPure RNA Kit and TransScript One-Step RT-PCR SuperMix kit used in RT-PCR were purchased from TransGen Biotech Co., Ltd (Beijing, China). All the restriction enzymes and T4 DNA ligase were purchased from Thermo Fisher Scientific. Agents and chemicals were purchased from Sigma Aldrich, TCI (Shanghai) or Sangon Biotech.

E. coli were cultured at 37 °C in LB liquid/solid medium for 12 - 15 h. *Streptomyces* were cultured in TSB media (3% tryptone soy broth) at 30 °C for 24 h to prepare the seed culture. 4 mL of the seed cultures were then transferred into 50 mL ISP2 liquid fermentation medium (0.4% glucose, 0.4% yeast extract, 1% malt extract, pH 7.0) and incubated on a rotary shaker (220 rpm) at 30 °C for 24 - 96 h. MS agar plates (2% mannitol, 2% soybean meal and 2% agar) was used for sporulation of *Streptomyces* and conjugation. The media were supplemented with appropriate antibiotics for selection: thiostrepton 25 μg/mL, kanamycin 50 μg/mL and apramycin 30 μg/mL for *Streptomyces*; apramycin 30 μg/mL, kanamycin 50 μg/mL, ampicillin 100 μg/mL, spectinomycin 50 μg/mL and chloramphenicol 25 μg/mL for *E. coli.*

For compound analysis, fermentation broth was extracted three times with equal volume of ethyl acetate. Combined extract was concentrated and re-dissolved in methanol, then centrifuged (13000 rpm, 8 min), filtered, and subjected to HPLC and liquid chromatography/MS (LC-MS) analysis. HPLC analysis was performed on a ZORBAX Eclipse-XDB-C18 column (250 x 4.6 mm, 5 μm, Agilent) and carried out with an Agilent 1260 HPLC system. The column was equilibrated with 95% solvent A (H_2O) and 5% solvent B (CH_3CN). The program of HPLC analysis was developed as follows: 0-15 min, 5% to 40% B; 15-22 min, 40%-70% B; 22-27 min, 70%-100% B; 27-30 min, 100% B; 30-33 min, 100%-5% B; 33-40 min, 5% B. This process was carried out at a flow rate of 0.6 mL/min and UV detection at 210 nm, 254 nm and 347 nm. LC-MS was performed under the same conditions (except that the flow rate is 0.4 mL/min for LC-MS) on an Agilent 6120 Quadruple LC/MS system operated in positive ion electrospray mode.

For the structure elucidation, the purified compounds were subjected to HRMS and NMR analysis. NMR spectra were recorded on Bruker AV-III 600 (600 MHz) spectrometers. LC-QTOF-MS analysis was performed on a 6530 Accurate-Mass Q-TOF spectrometer coupled to an Agilent HPLC 1200 series (Agilent Technologies).

Experimental Procedures

Construction of gene inactivation and complementation mutants.

The genomic library of *Streptomyces* sp. B9173 was constructed with the vector pJTU2463 (*Apr*^{R}) [1]. In the library, cosmid 20G3 contains *marB* ~ *marQ* and the downstream ~20 kb sequences, while cosmid 24B6 contains *marA* ~ *marP* and the upstream ~25 kb sequences (see Fig. S5). Cosmid 20G3 and 24B6 were subcloned to shuttle vector pJTU1289 (*Amp*^{R}) to generate plasmids pJL0a and pLS0b as the template for gene inactivation.

Strategy of PCR-targeting based on the λ-Red recombination functions was performed to disrupt the genes of interest (*marA, marB, marC, marD, marL marM, marP* and *marQ*). Each target gene in the template plasmids was replaced by the apramycin resistance gene *aac(3)IV* + *oriT* (RK2). Recombination plasmids for gene replacement were further introduced into *Streptomyces* sp. B9173 by conjugation from *E. coli* ET12567/pUZ8002. The gene inactivation mutant strains were selected by apramycin resistance and trimethoprim (50 μg/mL), and the genetic phenotype of the mutants was further confirmed by PCR (see Fig. S2).

For the gene complementation of *mar* mutants, the corresponding gene was cloned into pJTU2170 (*Km*^{R}) *E. coli–Streptomyces* expression shuttle vector which contains the *ermE** strong promoter, generating plasmids pJL1008 (Δ*marP::marP*), pJL1009 (Δ*marA::marA*), pJL1010 (Δ*marD::marD*), pJL1011 (Δ*marL::marL*) and pJL1012 (Δ*marM::marM*), respectively. The constructed plasmids was further introduced into corresponding *Streptomyces* mutant strains by conjugation from *E. coli* ET12567/pUZ8002 to afford the JL1008 ~ JL1012. Exconjugants were selected by kanamycin and trimethoprim (50 μg/mL).

Overexpression of *marP* gene in *Streptomyces* sp. B9173.

The constructed pJL1008 was introduced into *Streptomyces* sp. B9173 wild type (WT) to afford the JL1008a, which was used for the overexpression of *marP* in WT strain. Kanamycin-sensitive screening was used to identify the mutant strain. Transcription level of *marP* in B9173 WT and JL1008a were analyzed by semiquantative RT-PCR with the following procedure. Mycelia were collected from the culturing broth of *Streptomyces* and applied to the isolation of total RNA. The total RNA was extracted and purified using

EasyPure RNA Kit. The strand cDNA was synthesized from the purified RNA by using TransScript One-Step RT-PCR SuperMix kit. Amplification of cDNA was carried out to perform semi-quantitative RT-PCR analysis. Corresponding primers are listed in Table S2.

Real-time quantitative RT-PCR.

Real-time quantitative RT-PCR was carried out to analysis the transcription level of *marP* in different growth periods of *Streptomyces* sp. B9173. The isolation of total RNA, the synthesis and amplication of corresponding cDNA followed the procedure described above. The amplified products of the *marP* gene and *hrdb* gene (house-keeping gene) were less than 200 bp, and they were further applied to quantitative real-time PCR. Reactions were performed by using SYBR Green qPCR Master Mix (Thermo Fisher Scientific Inc.) and ABI7500 Fast Real Time System (Applied Biosystems). Reaction conditions followed a reported method from Dai.[2]

Cosmid combination and heterologous expression of *mar* cluster in *S. lividans* TK24.

The procedure shown in Fig. S6 was used to combine the cosmid 20G3 and cosmid 24B6 for heterologous expression. (1) Using the PCR targeting method, spectinomycin resistance cassette was inserted into cosimid 20G3 at the position between B9173 genome segment and vector pJTU2463 (Apr^R), generating a new cosmid 20G3 ($Spec^R$). (2) The cosmid 20G3 ($Spec^R$) was digested by enzyme DraI and SpecI to get a fragment that included *marB* ~ *marQ* (16.5 kb) and vector (5.6 kb) segment. (3) With *marB* ~ *marQ* (16.5 kb) and vector (5.6 kb) segment serving as left and right homologous flanks respectively, this fragment was further recombined with cosmid 24B6 with PCR targeting method, resulting in a recombined plasmid pJL1001 ($Spec^RApr^R$) which covers the complete maremycin gene cluster (~17 kb), the upstream genes (~25 kb) and downstream genes (~20 kb).

In order to generate a plasmid with only the *mar* gene cluster, the upstream and downstream fragment were both deleted, following the procedure: (1) By using PCR targeting system, the downstream (~20 kb) fragment in pJL1001 was replaced with the kanamycin resistance gene to generate the plasmid pLJ1001D (Km^RApr^R). (2) The kanamycin resistance gene in pLJ1001D was excised by enzyme NdeI, and then the digested pJL1001D was self-ligased by T4 DNA ligase, generating a plasmid pJL1001Ds (Apr^R). (3) The upstream (~25 kb) genes in pJL1001Ds were then replaced by kanamycin resistance gene via PCR-targeting, generating plasmid pJL1001C (Km^RApr^R) which contained the *mar* gene cluster.

The constructed plasmids pJL1001C were further introduced into *S. lividans* TK24 by conjugation from *E. coli* S17-1. Vector pJTU2463 was also introduced into *S. lividans* TK24

to afford *S. lividan*/pJTU2463 as control. Exconjugants were selected by nalidixic acid (25 μg/mL) and apramycin.

Feeding experiment with FR900452 (1)

For the feeding experiment with FR900452 (**1**), 0.6 mg **1** was dissolved in 100 μL DMSO. Δ*marM* and Δ*marQ* mutants were cultured for 48 h, then the dissolved **1** was added into the culture broth, and the fermentation continued until 96 h. Metabolites from the feeding experiment were extracted, and then analyzed by HPLC and LC-MS analysis as mentioned above.

Isolation of FR900452s from *Streptomyces* sp. B9173.

For large-scale fermentation of *Streptomyces* and isolation of the metabolites, 2 L Erlenmeyer flasks containing 500 mL ISP2 medium with 8% (v/v) seed culture were incubated at 30 °C for 48 h (for **1** and **2**) and 72 h (for **3**). The culture broth was extracted three times with equal volume ethyl acetate and the combined extracts were concentrated in vacuo. The crude extract was fractioned by silica gel with ethyl acetate/methanol solvent system. The fraction washed with ethyl acetate: methanol (20:1) was collected, and then purified on reverse phase silica gel (C18, YMC, Japan) using a gradient of MeOH in water, yielding **1** (78 mg), **2** (8 mg) and **3** (12 mg).

Tables

Table S1. Strains and plasmids used in this study.

Strains/Plasmids	**Characteristics**	**Source/Ref.**
Streptomyces		
Streptomyces sp. B9173	Native producer of maremycins and FR900452s	[3]
S. lividans TK24	heterologous host for the expression of maremycin gene cluster	Laboratory stock
LS24	*Streptomyces* sp. B9173 (Δ*marC::aac(3)IV*)	[4]
LS29	Streptomyces sp. B9173 (Δ*marM::aac(3)IV*)	[4]
JL1002	*Streptomyces* sp. B9173 (Δ*marA::aac(3)IV*)	This work
JL1003	*Streptomyces* sp. B9173 (Δ*marB::aac(3)IV*)	This work
JL1004	*Streptomyces* sp. B9173 (Δ*marD::aac(3)IV*)	This work
JL1005	*Streptomyces* sp. B9173 (Δ*marL::aac(3)IV*)	This work
JL1006	*Streptomyces* sp. B9173 (Δ*marP::aac(3)IV*)	This work
JL1007	*Streptomyces* sp. B9173 (Δ*marQ::aac(3)IV*)	This work
JL1008a	*Streptomyces* sp. B9173 (WT::*marP*)	This work
JL1008b	*Streptomyces* sp. B9173 (Δ*marP::marP*)	This work
JL1009	*Streptomyces* sp. B9173 (Δ*marA::marA*)	This work
JL1010	*Streptomyces* sp. B9173 (Δ*marD::marD*)	This work
JL1011	*Streptomyces* sp. B9173 (Δ*marL::marL*)	This work
JL1012	*Streptomyces* sp. B9173 (Δ*marM::marM*)	This work
S.lividans. JL1001C	Heterologous expression strain of pJL1001C	This work
*S.lividans/*pJTU2463	*Streptomyces lividans* TK24 containing the vector pJTU2463	This work
E.coli		
DH10B	Host for gene cloning	Laboratory stock
BW25113/pIJ790	Host for PCR-targeting system	[5]
DH5α/BT340	Host for in-frame deletion of genes	[6]
ET12567/pUZ8002	Donor strain for *E.coli-Streptomyces* exconjugation	[6]

S17-1	Strain used for *E.coli-Streptomyces* exconjugation	Laboratory stock
Plasmids		
pJTU2463	Shuttle vector used for construction of genomic cosmid library, *ori* (ColE1), *intattpP*ϕC31, *oriT* (RK2), *cos, aac(3)IV*	Laboratory stock
pJTU2170	Shuttle vectors for gene complementation and overexpression	Laboratory stock
pJTU1278	Shuttle vectors for gene replacement, *oriT* (RP4), *ori* (ColE1), *bla, tsr, cos, rep* (pIJ101), *ori* (pIJ101), *lacZ*	Laboratory stock
pJTU1289	oriT region of pJTU1278 deleted for gene replacement in PCR-targeting system	Laboratory stock
pJTU4659	Kan FRT ligased to pBSK+（EcoRI-SmaI）	Laboratory stock
Cosmid 20G3	A cosmid which contains *marB* ~ *marQ* and the downstream ~20 kb genes	[4]
Cosmid 24B6	A cosmid which contains *marA* ~ *marP* and the upstream ~25 kb genes	[4]
Cosmid 20G3 (*Spec*^{R})	Cosimid 20G3 inserted with spectinomycin resistance gene	This work
pJL0a	An 9 kb fragment from cosmid 24B6 (digested by KpnI) was cloned into the corresponding site of pJTU1289 and used for *marA*, *marB* and *marD* gene replacement	This work
pLS0b	An 18 kb fragment from cosmid 20G3 (digested by XbaI and HindIII) was cloned into the corresponding site of pJTU1289 and used for *marL, marP* and *marQ* gene replacement	[4]
pJL1002	*marA* in the pJL0a was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1003	*marB* in the pJL0a was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1004	*marD* in the pJL0a was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1005	*marL* in the pLS0b was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1006	*marP* in the pLS0b was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1007	*marQ* in the pLS0b was substituted by the *aac(3)IV+oriT* cassette using the PCR targeting strategy	This work
pJL1008	The complementary *marP* constructed on vector pJTU2170	This work
pJL1001	Recombined cosmid of cosmid 20G3 and cosmid 24B6	This work
pJL1009	The complementary *marA* constructed on vector pJTU2170	This work
pJL1010	The complementary *marD* constructed on vector pJTU2170	This work
pJL1011	The complementary *marL* constructed on vector pJTU2170	This work
pJL1012	The complementary *marM* constructed on vector pJTU2170	This work
pJL1001D	pJL1001 with downstream ~20 kb genes substituted by kanamycin resistance gene	This work

pJL1001Ds	pJL1001D with kananmycin resisitance gene excised	This work
pJL1001C	pJL1001 with both upstream and downstream genes excised	This work

Amp^R, ampicillin resistance; *Km*^{R}, kanamycin resistance; *Apr*^{R}, apramycin resistance; *Spec*^{R}, spectinomycin resistance.

Table S2. PCR primers used in this study.

Primer name	Sequence 5′-3′	Description
marATarF	gtgtgcatcaccggaatgggctggaccacagcactgggtattccgggg atccgtcgacc	Primers for *marA* inactivition
marATarR	tcacgcggcactcccggtcacggtccgcggttcgccggctgtaggct ggagctgcttc	
marAVtarF	tcggaacaccgggggattta	Primers for the verification of *marA* inactivition
marAVtarR	gggcatgggcagcatcgacac	
marBTarF	gtccccggggtctacgccgatccggtggcctggctgatgattccgggg atccgtcgacc	Primers for *marB* inactivition
marBTarR	gcgaggtggtgctctgcggggtggtgttctgcgaggtcatgtaggctg gagctgcttc	
marBVtarF	gtgatcaccggctccgccct	Primers for the verification of *marB* inactivition
marBVtarR	gacgacaccggggtcgagaa	
marDTarF	gtgacgggccccgaaccgctcaccctgttctgcgtcccgattccgggg atccgtcgacc	Primers for *marD* inactivition
marDTarR	tcaccgcctcgaagcgttcgtggaggaagaagtggccgctgtaggct ggagctgcttc	
marDVtarF	atccagccgctgcgcgcgggc	Primers for the verification of *marD* inactivition
marDVtarR	tggacggcagtgtcctcggct	
marLTarF	gacaccctggaagcgactctcaaggagatcctcgtcgagattccgggg atccgtcgacc	Primers for *marL* inactivition
marLTarR	ggcgcccgtgcgcagggcgcgcacgagatcggccaccgttgtaggc ctggagctgcttc	
marLVtarF	atgacagccggttcgggcga	Primers for the verification of *marL* inactivition
marLVtarR	ccagacgtccaactgctc	
marPTarF	ccgcaccgcttacgacgccctaaccgaggagcacgcatgattccgg ggatccgtcgacc	Primers for *marP* inactivition
marPTarR	ccctccttcagcggttcgcggatgcgtgggcaggggtcatgtaggctg gagctgcttc	
marPVtarF	cgaccgtcagttctgcttgt	Primers for the verification of *marP* inactivition
marPVtarR	tgtcgcgtcatctgctctg	
CmarPF	ggaattccatatgtccgtcaccgccgacctctac	Primers for the complementation of *marP*
CmarPR	ggaattctcagatcacccgtccgtac	
hrdB-F	gcggtggagaagttcgacta	Primers for semi-quantitative and real-time quantitative RT-PCR of *hrdb* gene
hrdB-R	ttgatgacctcgaccatgtg	
marP-RTF	gtcaccgccgacctctacat	Primers for semi-quantitative and real-time quantitative RT-PCR of marP gene
marP-RTR	tcttggtacgcatcgcctc	

marQTarF	atgacgcgacagccggcagagcgaccgacgcggcaaccgattccg gggatccgtcgacc	Primers for *marQ* inactivition
marQTarR	cgttcacccccccgggcgttcacccccctcggcgttcatgtaggctg gagctgcttc	
marQVtarF	atgacgcgacagccggcagag	Primers for the verification of *marQ* inactivition
marQVtarR	gccccccccgggcgttca	
CmarAF	ggaattccatatggtgtgcatcaccggaatggg	Primers for the complementation of *marA*
CmarAR	ggaattctcacgcggcactcccggtcacggtc	
CmarDF	ggaattccatatggtgacgggccccgaaccgct	Primers for the complementation of *marD*
CmarDR	ggaattctcagcggcccgctcgccgct	
CmarLF	ggaattccatatgacagccggttcgggcgagaca	Primers for the complementation of *marL*
CmarLR	ggaattctcactgcgggaccgccaggtgctc	
CmarMF	ggaattccatatgagggaacaggcctcctcc	Primers for the complementation of *marM*
CmarMR	ggaattctcaggagtgtgcggggag	
20G3SpecF	gcccagttcaagccggaacggttcaccctctacacgatgattccgggg atccgtcgacc	Primers for the spectinomycin resistance gene insertion
20G3SpecR	tcttcggcccgacgacgctgaggccgacgctcagcgtcatgtaggct ggagctgcttc	
20G3SpecVF	tcatggcggacaacgaga	Primers for the verification of spectinomycin resistance gene insertion
20G3SpecVR	ggcgtcgtgtaacccaccca	
upstreamTarF	caggcgaccgtcacccggtcggcgtgccgcagcagcgcgatggtta acagctattccagaagtagt	Primers for ~25 kb upstream genes deletion in pJL1001Ds
upstreamTarR	gtgcgcccgaacgggggtttcccctatagagggtttgtctcagttaact ggatgccgacggatttg	
upstreamVtarF	gaagaacgtcgtctgccgc	Primers for the verification of upstream genes deletion
upstreamVtarR	caggtgggccgggcggggtg	
downstreamTarF	tcgatgagttcgttcatctggttgcggctgagcttctccagccatatgag ctattccagaagtagt	Primers for ~20 kb downstream genes deletion in pJL1001
downstreamTarR	ggatcttcggcccgacgacgctgaggccgacgctcagcgtcacatat gtggatgccgacggatttg	
downstreamVtarF	cgaccaggctggtcttcttg	Primers for the verification of downstream genes deletion
downstreamVtarR	ggatcttcggcccgacgac	
marA PCR-F	gctcagagcgtttatcggaac	PCR primers for *marA* verification in pJL1001, pJL1001D and pJL1001C
marA PCR-R	gggcggagccggtgatcac	
MarQinnerPCR-F	cgcagatcgacacctatg	primers for the PCR of 750 bp segment of *marQ* in pJL1001, pJL1001D and pJL1001C
MarQinnerPCR-R	ccactggatgacgttgac	

Table S3. Proposed functions of proteins encoded by genes in the maremycin biosynthetic gene cluster

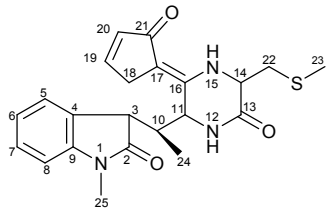
ORF	Size (aa)	Proposed function, Homologous protein	Identity/ Similarity	Sequence similarity origin
marA	369	Coronafacic acid beta-keto acyl synthetase component, cfa3	90%/91%	*Pseudomonas syringae pv. aesculi str. 2250*
marB	171	Coronafacic acid synthetase component, cfa4	39%/55%	*Pseudomonas syringae*
marC	512	Coronafacic acid synthetase, ligase component, cfa5	42%/54%	*Bacillus cereus* NVH0597-99
marD	240	Thioesterase	30%/49%	*Bacillus cereus* AH1134
marE	284	Tryptophan 2,3-dioxygenase	27%/41%	*Ktedonobacter racemifer* DSM 44963
marF	347	*O*-Methyltransferase	42%/59%	*Streptomyces* sp. e14
marG	352	Histidinol-phosphate aminotransferase	64%/73%	AFW04575.1, *S. flocculus*
marH	129	Cupin protein	87%/90%	KHK89141.1, *S. pluripotens*
marI	345	Methyltransferase	29%/43%	*S. hygroscopicus* NRRL 30439
marJ	84	PCP domain	51%/71%	*Chitinophaga pinensis* DSM 2588
marK	349	2-Dehydropantoate 2-reductase	38%/47%	*S. avermitilis* MA-4680
marL	102	Coronafacic acid synthetase, acyl carrier protein component, cfa1	49%/76%	*Bacillus cereus* NVH0597-99
marM	1383	Nonribosomal peptide synthetase, C-A-PCP	88%/90%	*Actinobacteria bacterium* OV320
marN	337	Dehydrogenase	44%/54%	*Amycolatopsisjaponica*
marO	413	Putative beta-ketoacyl synthase, KS	37%/55%	*Photobacterium profundum* SS9
marP	142	SnoaL_4 super family (cl09109)	43%/55%	*Prauserella marina*
marQ	852	NRPS, A domain	43%/55%	*Micromonospora*sp. ML1

Table S4. HR-ESI-MS data of the compounds

Compound	Calcd. $[M+H]^+$	Found $[M+H]^+$
FR900452 (**1**)	412.1650	412.1655
FR900452 B (**2**)	414.1867	414.1894
FR900452 C (**3**)	428.1599	428.1607
Maremycin G (**4**)	408.1376	408.1388
Maremycin F (**5**)	424.1331	424.1329
Maremycin A (**6**)	364.1286	364.1340
Maremycin B (**7**)	364.1286	364.1341
*Compound (*)	412.1650	412.1648

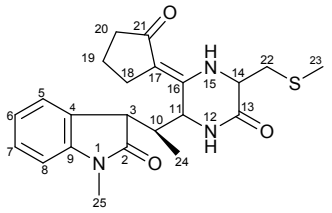
*Compound is the isomer of **1**

Table S5. NMR data of compound **1** ($CDCl_3$; δ_H 7.26 ppm, δ_C 77.0 ppm)



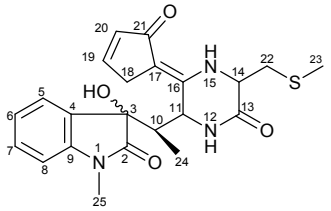
No.	δ_H	δ_C	No.	δ_H	δ_C
2		176.1 (s)	14	4.33 (m)	56.7 (d)
3	3.96 (d, 3.0)	44.2 (d)	15	10.15 (d, 4.3)	
4		127.2 (s)	16		150.1 (s)
5	7.13 (d, 7.6)	123.2 (d)	17		104.9 (s)
6	7.09 (t, 7.6)	123.1 (d)	18	3.18 (br d, 22.4)	34.5 (t)
				3.26 (br d, 22.4)	
7	7.31 (t, 7.6)	128.4 (d)	19	7.22 (br d, 5.8)	151.1 (d)
8	6.87 (d, 7.6)	108.5 (d)	20	6.34 (br d, 5.8)	136.3 (d)
9		144.3 (s)	21		196.8 (s)
10	3.00 (m)	45.9 (d)	22	3.08 (dd, 14.0, 7.3)	39.2 (t)
				3.15 (dd, 14.0, 3.4)	
11	4.49 (dd, 9.6, 4.3)	54.8 (d)	23	2.27 (s)	16.9 (q)
12	7.05 (br s)		24	0.72 (d, 7.0)	13.2 (q)
13		168.1 (s)	25	3.23 (s)	26.2 (q)

Table S6. NMR data of compound **2** (DMSO-d_6; δ_H 2.49 ppm, δ_C 39.5 ppm)



No.	δ_H	δ_C	HMBC	^{1}H,^{1}H-COSY
2		175.5 (s)		
3	3.93 (d, 2.2)	44.8 (d)	C-2, C-4, C-9, C-10, C-11, C-25	H-10
4		127.5 (s)		
5	7.24 (d, 7.6)	123.1 (d)	C-3, C-7, C-9	H-6
6	7.06 (t, 7.6)	122.4 (d)	C-4, C-8	H-5, H-7
7	7.30 (t, 7.6)	128.1 (d)	C-5, C-9	H-6, H-8
8	7.01 (d, 7.6)	108.5 (d)	C-4, C-6	H-7
9		144.3 (s)		
10	2.47 (m)	45.5 (d)	C-2, C-11, C-25	H-3, H-11, H-25
11	4.73 (dd, 9.6, 4.5)	52.3 (d)	C-10, C-13, C-16, C-17	H-10, H-12
12	9.02 (d, 4.5)		C-11, C-13, C-14, C-16	H-11
13		168.0 (s)		
14	4.19 (m)	55.4 (d)	C-13, C-16, C-22	H-15, H-22
15	10.30 (d, 4.9)		C-11, C-13, C-14, C-17	H-14
16		150.8 (s)		
17		103.2 (s)		
18	2.46 (m), 2.56 (m)	28.0 (t)	C-16, C-17, C-19, C-20, C-21	H-19
19	1.72–1.82 (m)	20.8 (t)	C-17, C-18, C-20, C-21	H-18, H-20
20	2.14–2.25 (m)	38.6 (t)	C-18, C-19, C-21	H-19
21		203.7 (s)		
22	2.86 (dd, 14.0, 9.3)	38.8 (t)	C-13, C-14, C-24	H-14
	3.02 (dd, 14.0, 3.6)			
23	2.16 (s)	15.6 (q)	C-22	
24	0.54 (d, 7.1)	11.0 (q)	C-3, C-10, C-11	H-10
25	3.13 (s)	25.8 (q)	C-2, C-9	

Table S7. NMR data of compound **3** ($CDCl_3$; δ_H 7.26 ppm, δ_C 77.0 ppm)



No.	δ_H	δ_C	HMBC	^{1}H,^{1}H-COSY
2		176.3 (s)		
3		78.7 (s)		
3-OH	4.36 (s)		C-2, C-3, C-4, C-10	
4		130.3 (s)		
5	7.35 (d, 7.6)	123.0 (d)	C-3, C-7, C-9	H-6
6	7.15 (t, 7.6)	123.7 (d)	C-4, C-8	H-5, H-7
7	7.36 (t, 7.6)	130.3 (d)	C-5, C-9	H-6, H-8
8	6.84 (d, 7.6)	108.5 (d)	C-4, C-6	H-7
9		143.1 (s)		
10	2.73 (m)	50.2 (d)	C-2, C-3, C-4, C-11, C-25	H-11, H-25
11	5.19 (dd, 8.0, 4.1)	52.3 (d)	C-3, C-10, C-13, C-16, C-17	H-10, H-12
12	7.32 (d, 4.1)		C-13, C-14, C-16	H-11
13		169.0 (s)		
14	4.16 (m)	55.8 (d)	C-13, C-22	H-15, H-22
15	10.28 (d, 4.8)		C-11, C-13, C-14, C-17	H-14
16		149.1 (s)		
17		104.7 (s)		
18	3.12 (ddd, 22.4, 2.2, 1.9)	34.8 (t)	C-17, C-19, C-20, C-21	H-19, H-20
	3.31 (ddd, 22.4, 2.2, 1.9)			
19	7.23 (dt, 5.8, 2.2)	151.3 (d)	C-17, C-18, C-20, C-21	H-18, H-20
20	6.28 (dt, 5.8, 1.9)	136.1 (d)	C-17, C-18, C-19, C-21	H-18, H-19
21		196.7 (s)		
22	2.86 (dd, 13.9, 8.2)	39.3 (t)	C-13, C-14, C-24	H-14
	3.01 (dd, 13.9, 3.6)			
23	2.21 (s)	16.7 (q)	C-22	
24	0.56 (d, 7.4)	11.9 (q)	C-3, C-10, C-11	H-10
25	3.20 (s)	26.1 (q)	C-2, C-9	

Results Figures

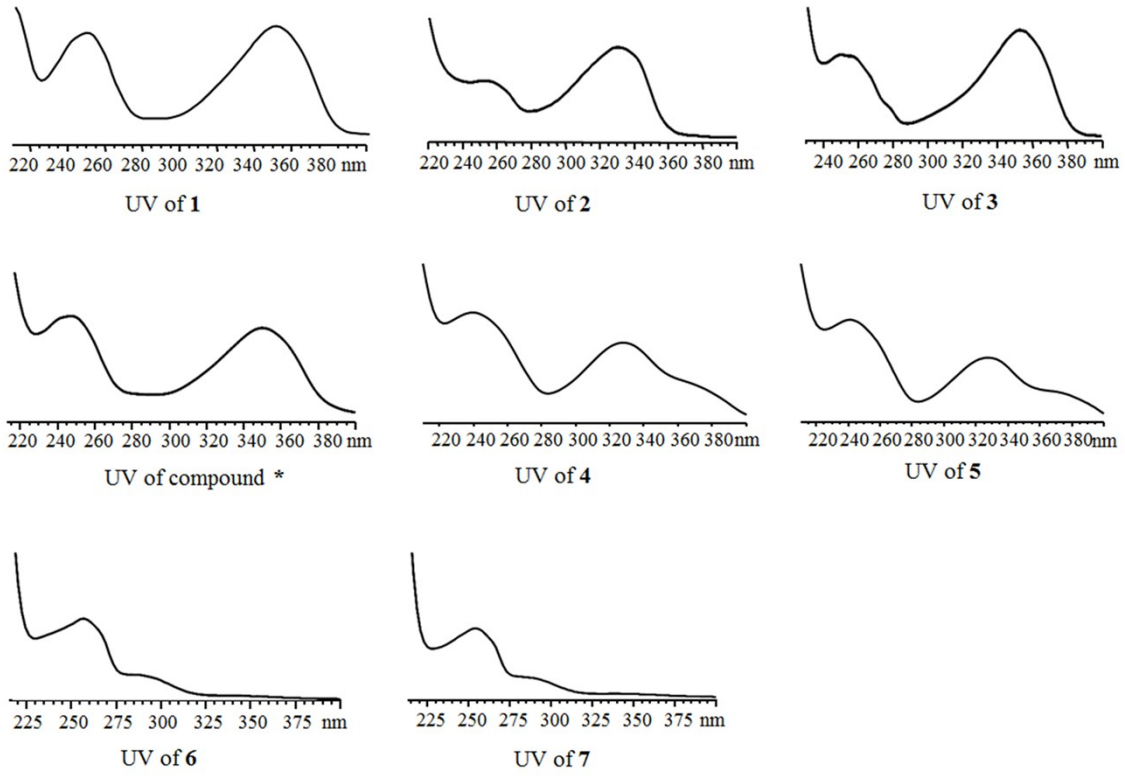


Figure S1. UV spectra of the metabolites of *Streptomyces* sp. B9173. **1**, **4** and **5** show UV absorption at both 254 nm and 347 nm; **2** and **3** show strong UV absorption at 347 nm but weak absorption at 254 nm; **6** and **7** exhibit no absorption at 347 nm but strong absorption at 254 nm. Thus we used 347 nm to detect **1-5** and 254 nm to detect **6** and **7** in this work.

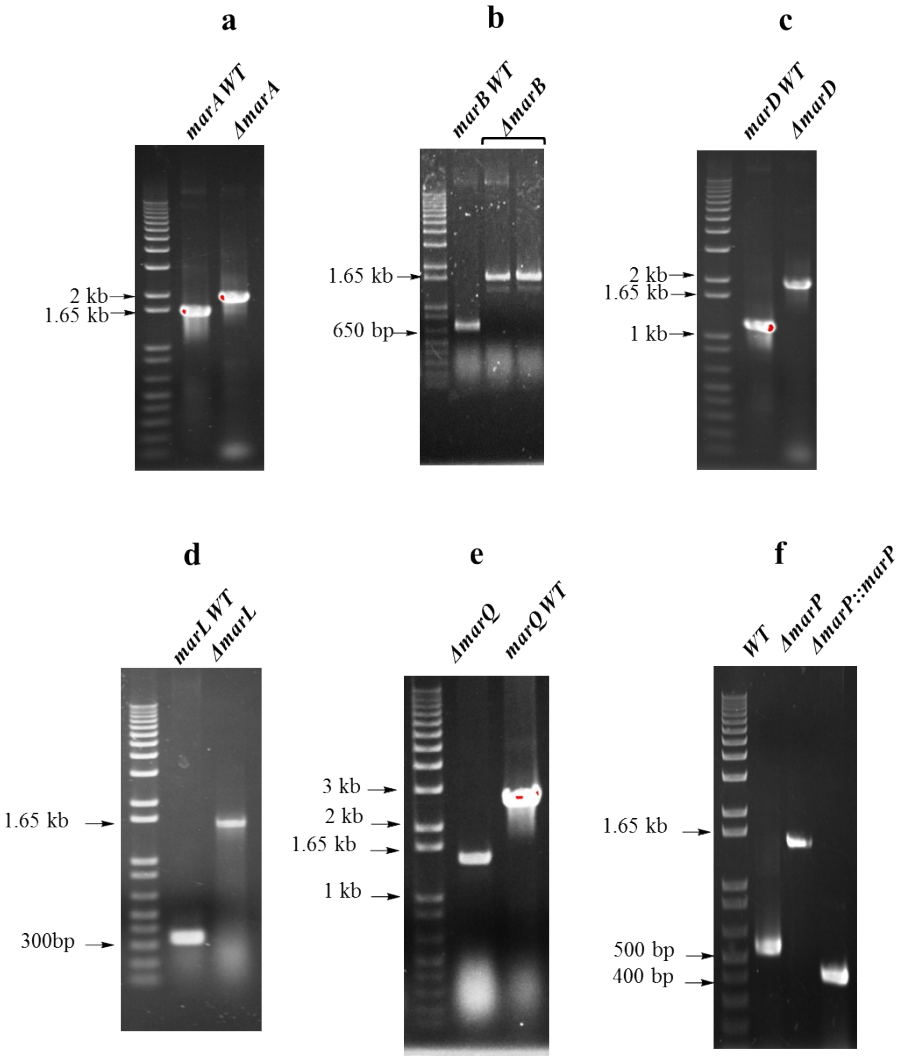


Figure S2. Confirmation of gene-inactivation and complementation mutants, (a-f) PCR is carried out with genomic DNA of *Streptomyces* sp. B9173 wild type (WT) and its mutants as templates. Primer pairs are the same as those used for cloning and expression of the *mar* genes (Table S2). (a) *marA* was replaced with *aac(3)IV*. Primer pair (*marAVtarF* / *marAVtarR*) was used to amplify a 1550 bp fragment from B9173 WT and a 1814 bp fragment from mutant Δ*marA*::*aac(3)IV*. (b) *marB* was replaced with *aac(3)IV*. Primer pair (*marBVtarF* / *marBVtarR*) was used to amplify a 688 bp fragment from B9173 WT and a 1556 bp fragment from mutant Δ*marB*::*aac(3)IV*. (c) *marD* was replaced with *aac(3)IV*. Primer pair (*marDVtarF*/ *marDVtarR*) was used to amplify a 1123 bp fragment from B9173 WT and a 1784 bp fragment from mutant Δ*marD*::*aac(3)IV*. (d) *marL* was replaced with *aac(3)IV*. Primer pair (*marLVtarF* / *marLVtarR*) was used to amplify a 310 bp fragment from B9173 WT and a 1384 bp fragment from mutant *ΔmarL*::*aac(3)IV*. (e) The *marQ* was replaced with *aac(3)IV*. Primer pair (*marQVtarF* / *marQVtarR*) was used to amplify a 2656 bp fragment from B9173 WT and a 1481 bp fragment from mutant Δ*marQ*::*aac(3)IV*. (f) The *marP* was replaced with *aac(3)IV*, and the complement of *marP* to Δ*marP* mutant. Primer pair (*marPVtarF* / *marPVtarR*) was used to amplify a 540 bp fragment from B9173 WT and a 1495 bp fragment from mutant Δ*marP*::*aac(3)IV*. Primer pair (*CmarPF* / *CmarPR*) was used to amplify a 429 bp *marP* fragment from Δ*marP*::*marP*.

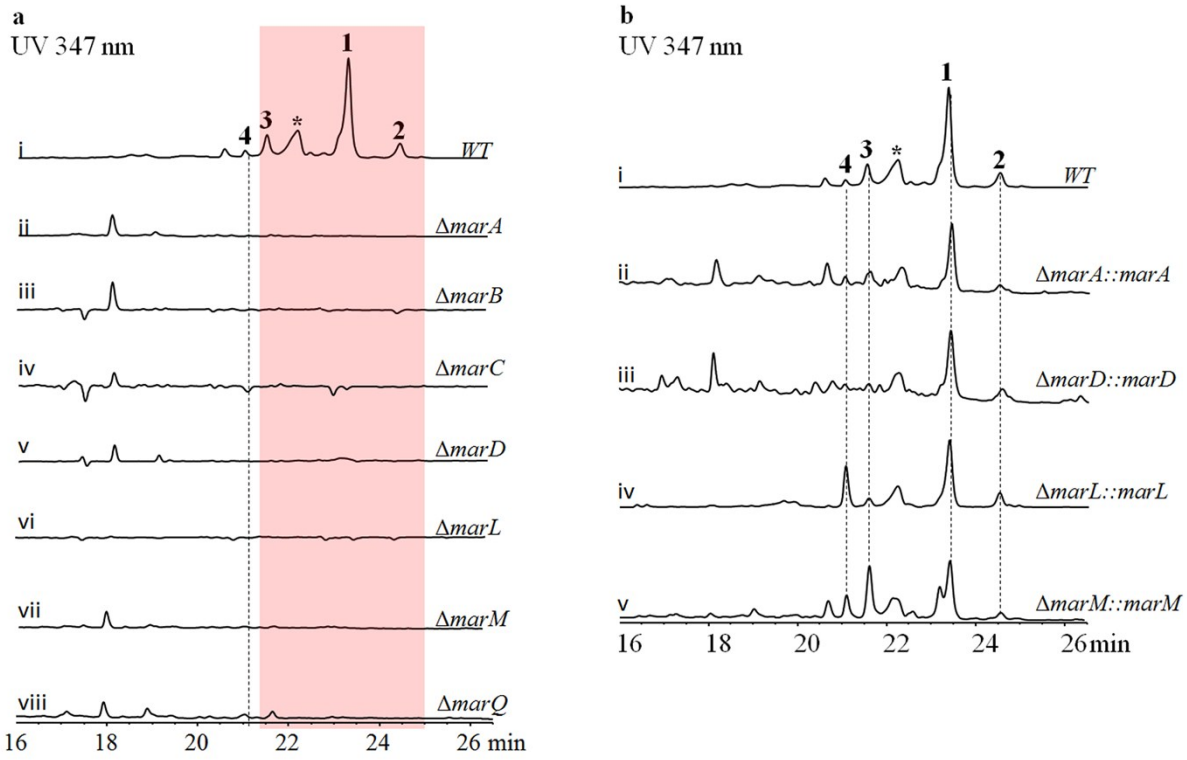


Figure S3. Genetic characterization of the structural genes for the biosynthesis of **1-4** *in vivo*. (a) HPLC analysis of **1-4** (UV at 347 nm) from (a-i) *Streptomyces* sp. B9173 wild type, (a-ii) JL1002 (Δ*marA* mutant strain), (a-iii) JL1003 (Δ*marB* mutant strain), (a-iv) LS24 (Δ*marC* mutant strain), (a-v) JL1004 (Δ*marD* mutant strain), (a-vi) JL1005 (Δ*marL* mutant strain), (a-vii) LS29 (Δ*marM* mutant strain), (a-viii) JL1007 (Δ*marQ* mutant strain). (b) HPLC analysis of **1-4** (UV at 347 nm) from the gene complementary strains of (b-ii) JL1009 (Δ*marA::marA* complementary strain), (b-iii) JL1010 (Δ*marD::marD* complementary strain), (b-iv) JL1011 (Δ*marL::marL* complementary strain), (b-v) JL1012 (Δ*marM::marM* complementary strain). The peak marked by an asterisk (*) was identified as an isomer of **1** by HRMS.

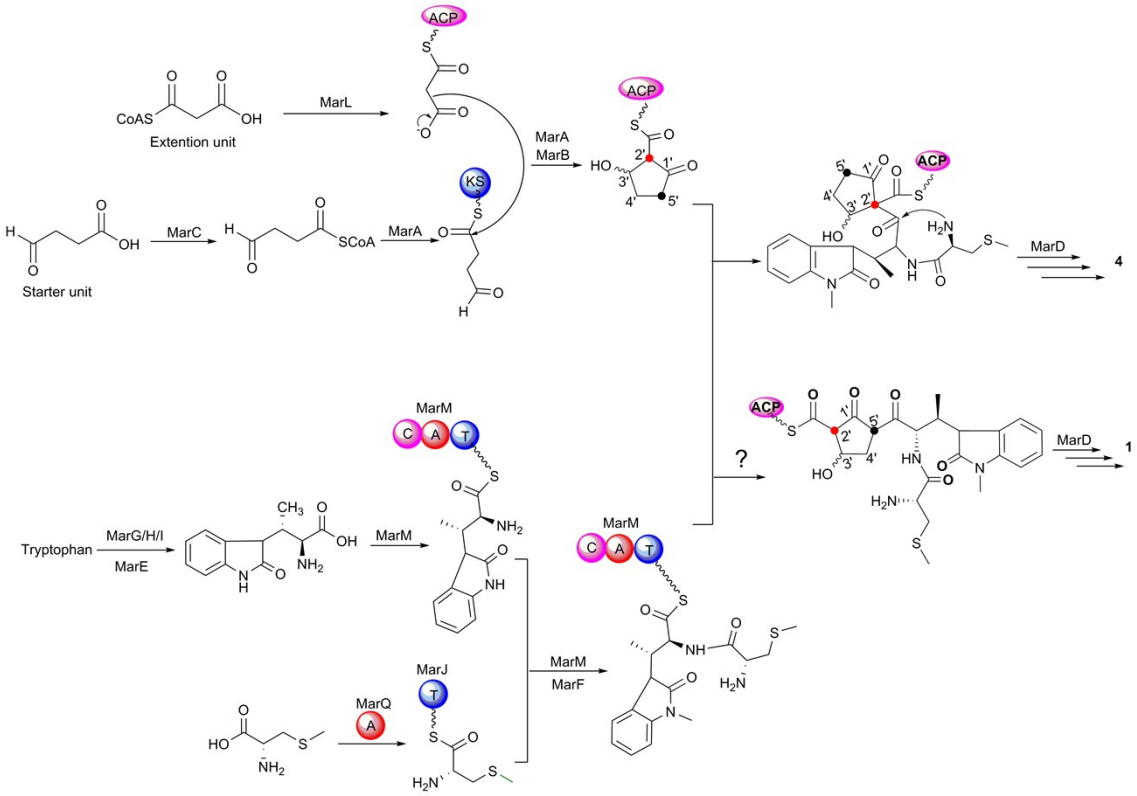


Figure S4. Proposed biosynthetic pathway of **1** and **4**.

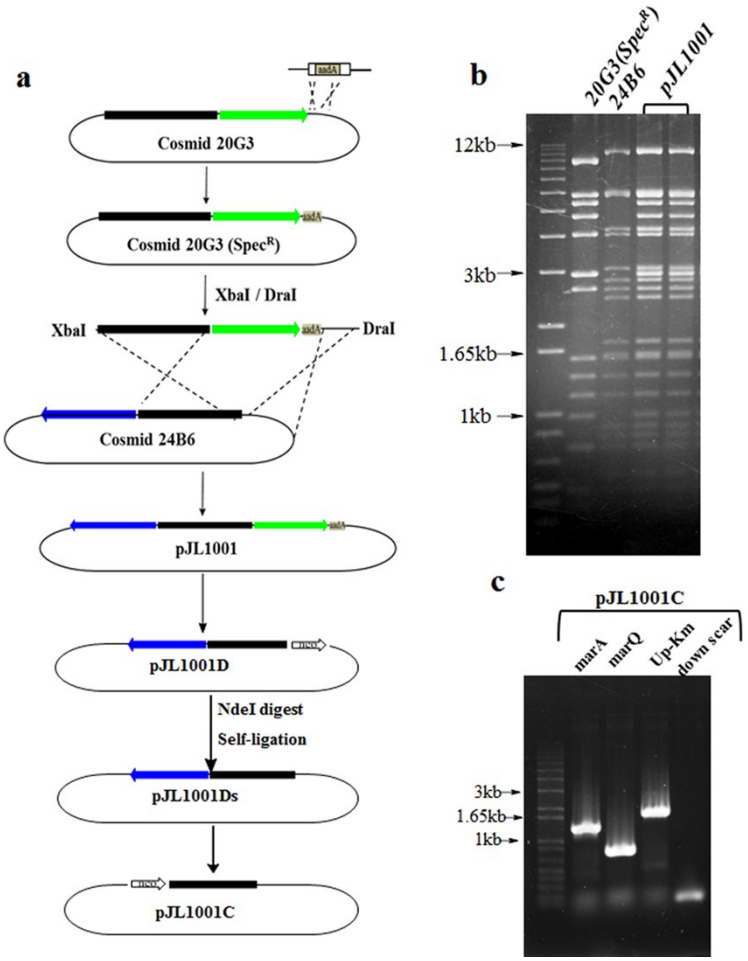


Figure S5. Cosmids combination and fragment deletions to construct pJL1001 and pJL1001C. (a) Schematic diagram for the cosmids combination and fragment deletions. (b) Restriction enzyme digestion maps of the cosmids (pJL1001, cosmid 20G3 and cosmid 24B6) by BamHI enzyme. (c) The upstream genes (~25 kb) out of the *mar* cluster were replaced with *neo*, and the *neo* in downstream genes was cut by NdeI and then self-ligased. PCR is carried out with DNA of pJL1001C as template. Primer pair (*upstreamVtarF*/*upstreamVtarR*) was used to amplify the 1.5 kb kanamycin resistance cassette in upstream genes of pJL1001C, primer pair (*downstreamVtarF*/*downstreamVtarR*) was used to amplify the 200 bp scar in downstream genes of pJL1001C, primer pair (*marA PCR-F/marA PCR-R*) and primer pair (*marQ innerPCR-F/marQ innerPCR-R*) were used to amplify *marA* gene and 742 bp fragment of *marQ* gene so as to confirm the integrality of *mar* cluster in pJL1001C.

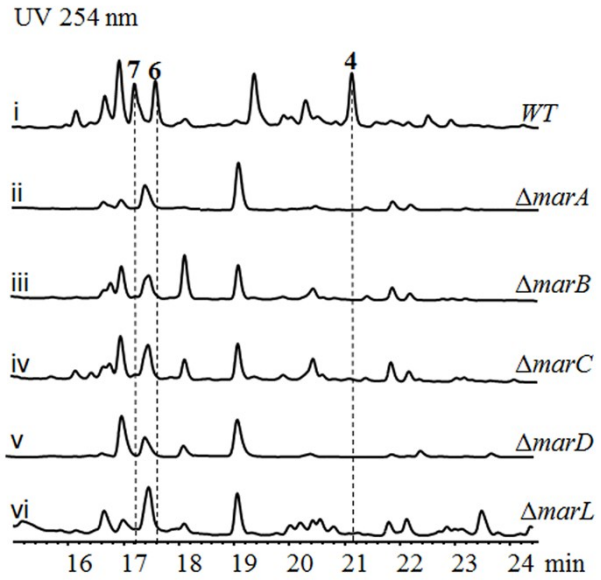


Figure S6. Characterization of the PKS genes (Δ*marA* ~ Δ*marD*，Δ*marL*) for the biosynthesis of **4** and **6**-**7** *in vivo*. HPLC analysis the production of f **4** and **6**-**7** (UV 254 nm) from (i) *Streptomyces* sp. B9173 wild type, (ii) JL1002 (Δ*marA* mutant strain), (iii) JL1003 (Δ*marB* mutant strain), (iv) LS24 (Δ*marC* mutant strain), (v) JL1004 (Δ*marD* mutant strain), (vi) JL1005 (Δ*marL* mutant strain).

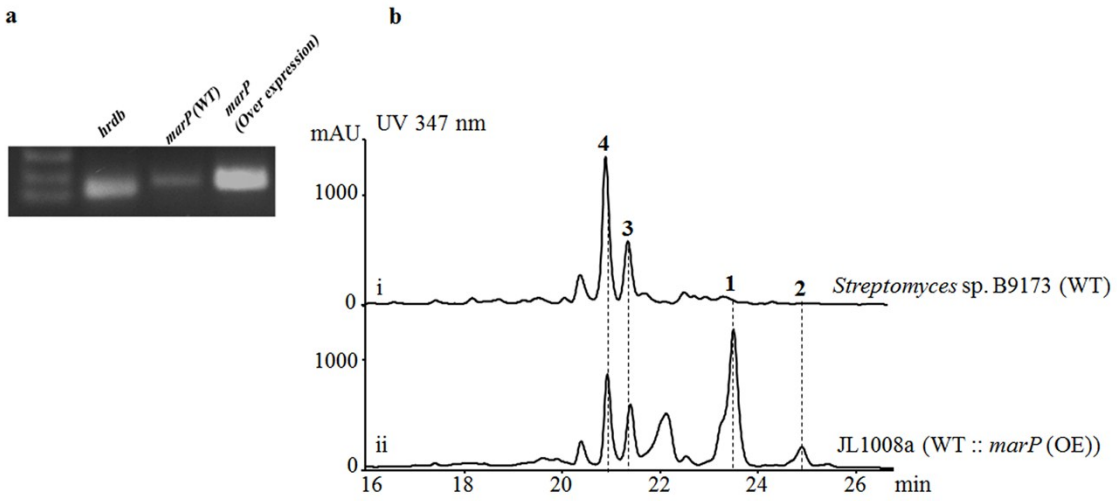


Figure S7. Characterization of **1-3** and **4** production in *Streptomyces* sp. B9173 and the recombinant strain JL1008a fermented for 72 h. (a) Semi-quantitative RT-PCR analysis of the over-expression of *marP* in *Streptomyces* sp. B9173, using *hrdb* gene as a house-keeping gene. (b) HPLC analysis of the metabolites from (i) *Streptomyces* sp. B9173 wild type strain and (ii) recombination strain for *marP* overexpression.

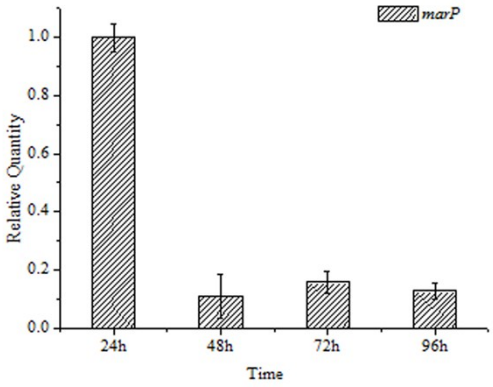
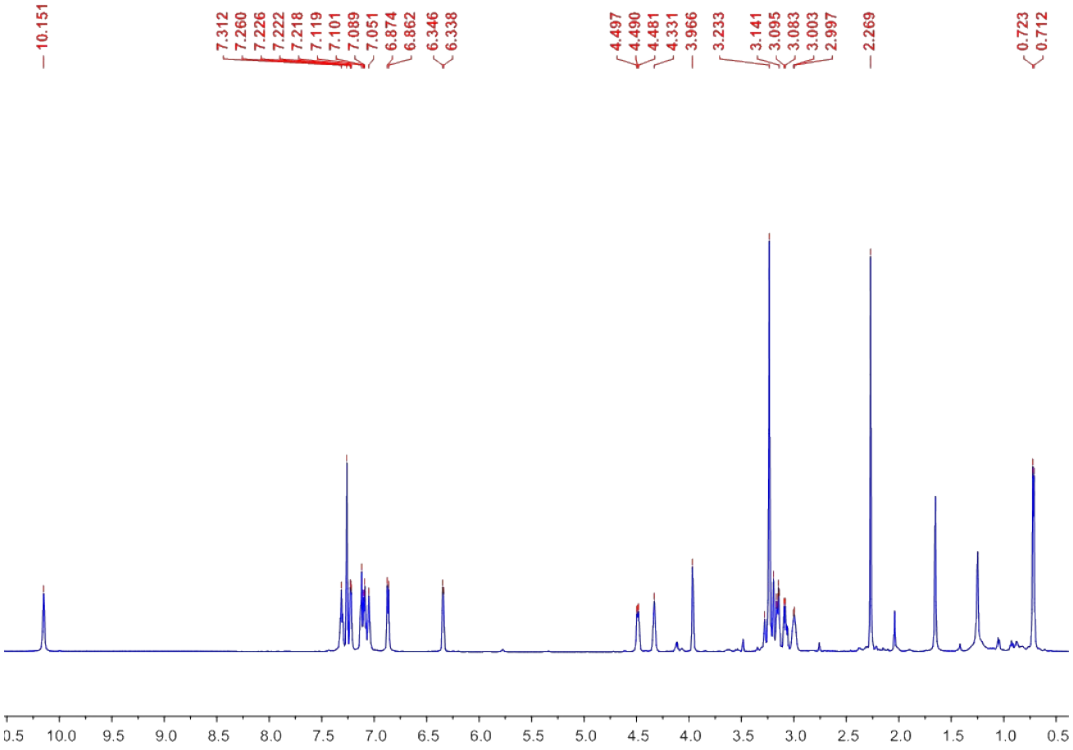


Figure S8. Transcription levels of the *marP* gene in different growth stages of *Streptomyces* sp. B9173. The cDNA quantity of *marP* gene from the mycelia of 24 h culture broth was used as relative control.

NMR Figures

^{1}H NMR spectrum of **1** (600 MHz, $CDCl_3$).



^{13}C NMR spectrum of **1** (150 MHz, $CDCl_3$).

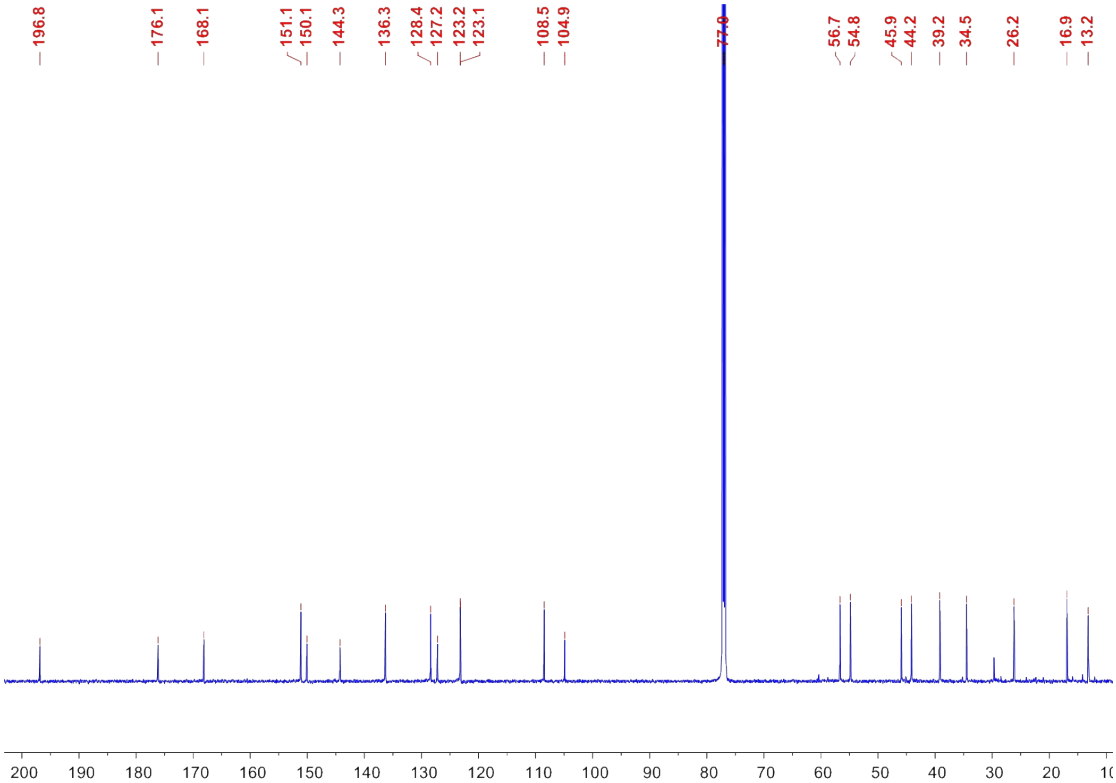
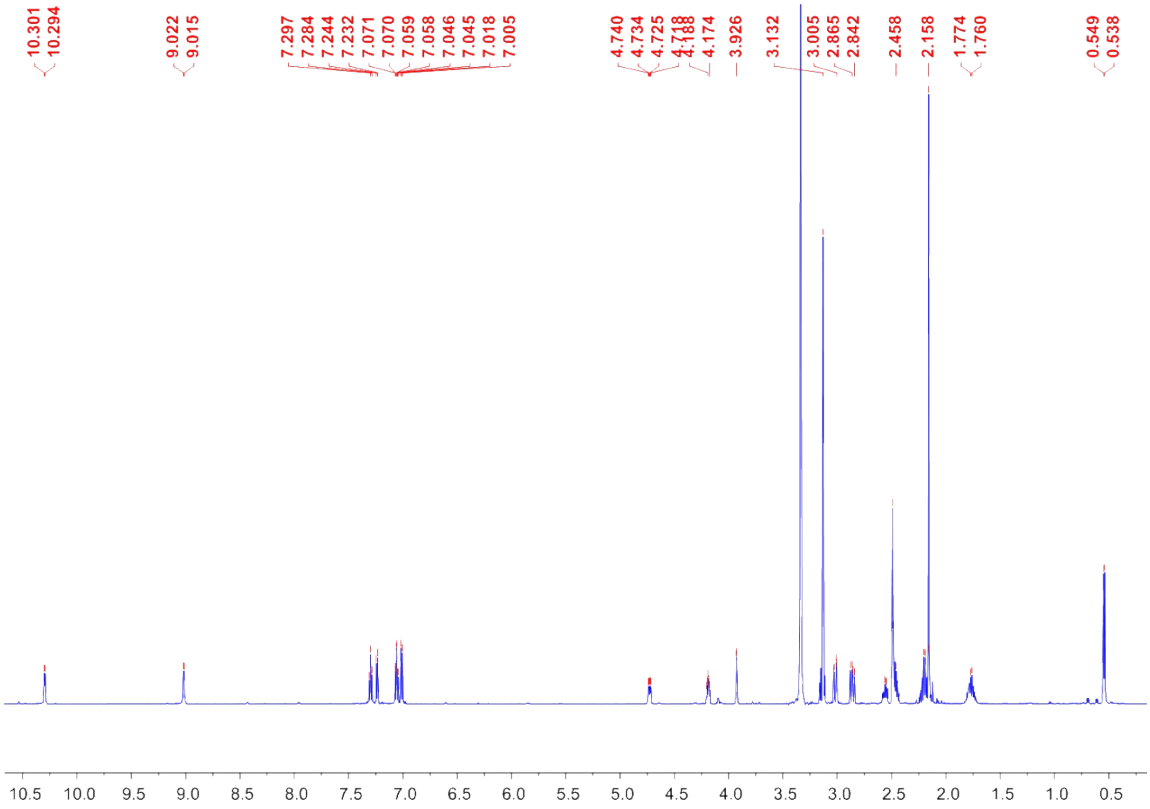


Figure S9. ^{1}H and ^{13}C NMR spectra of **1**

^{1}H NMR spectrum of **2** (600 MHz, DMSO-d_6).



^{13}C NMR spectrum of **2** (150 MHz, DMSO-d_6)

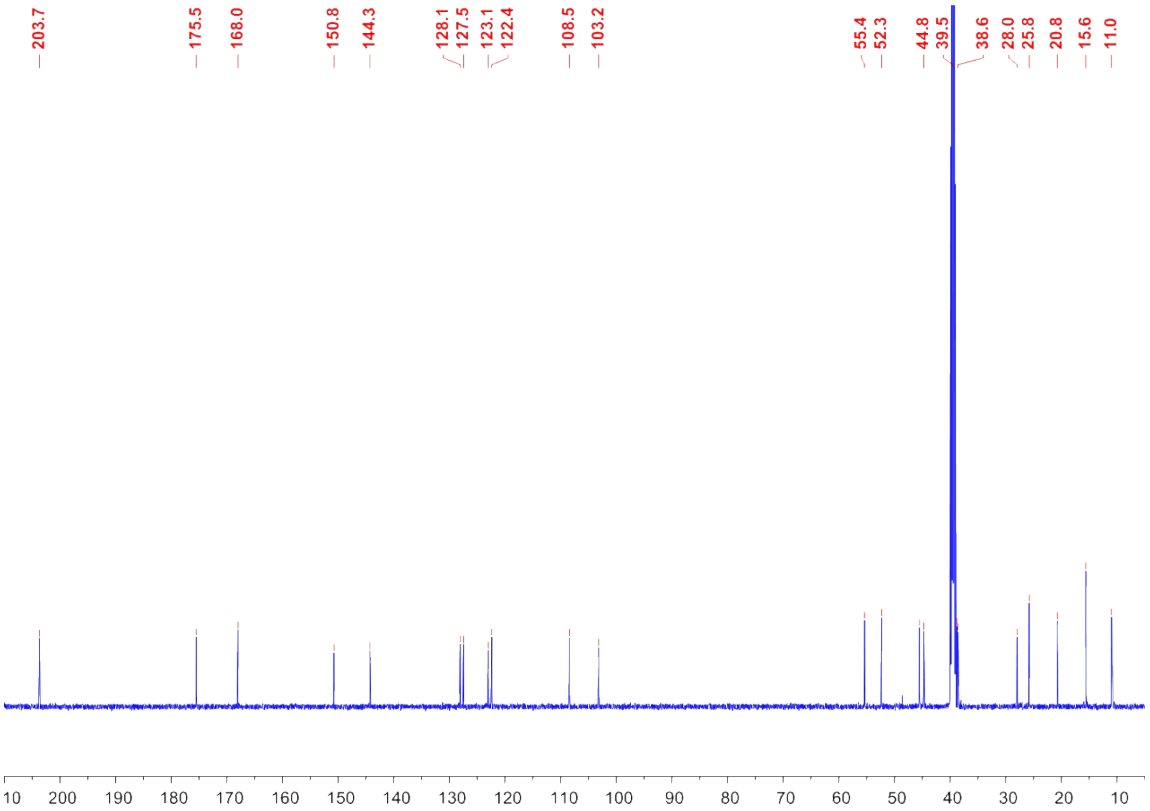
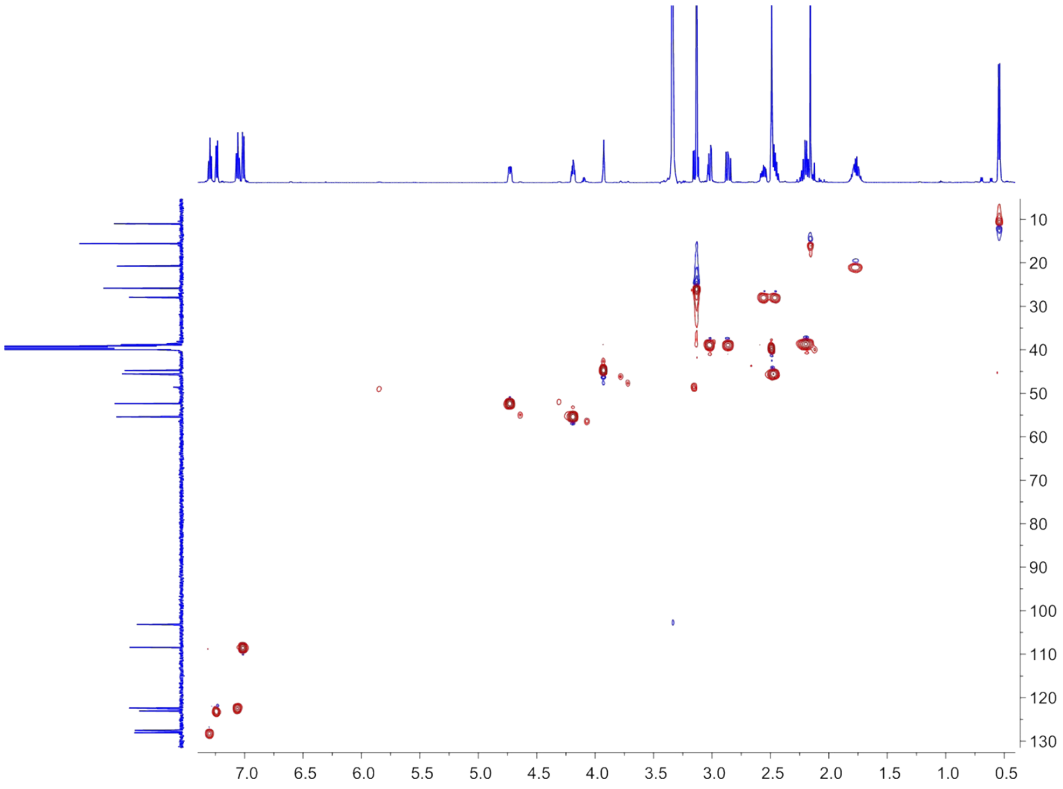


Figure S10. ^{1}H and ^{13}C NMR spectra of **2**

HSQC spectrum of **2**



HMBC spectrum of **2**

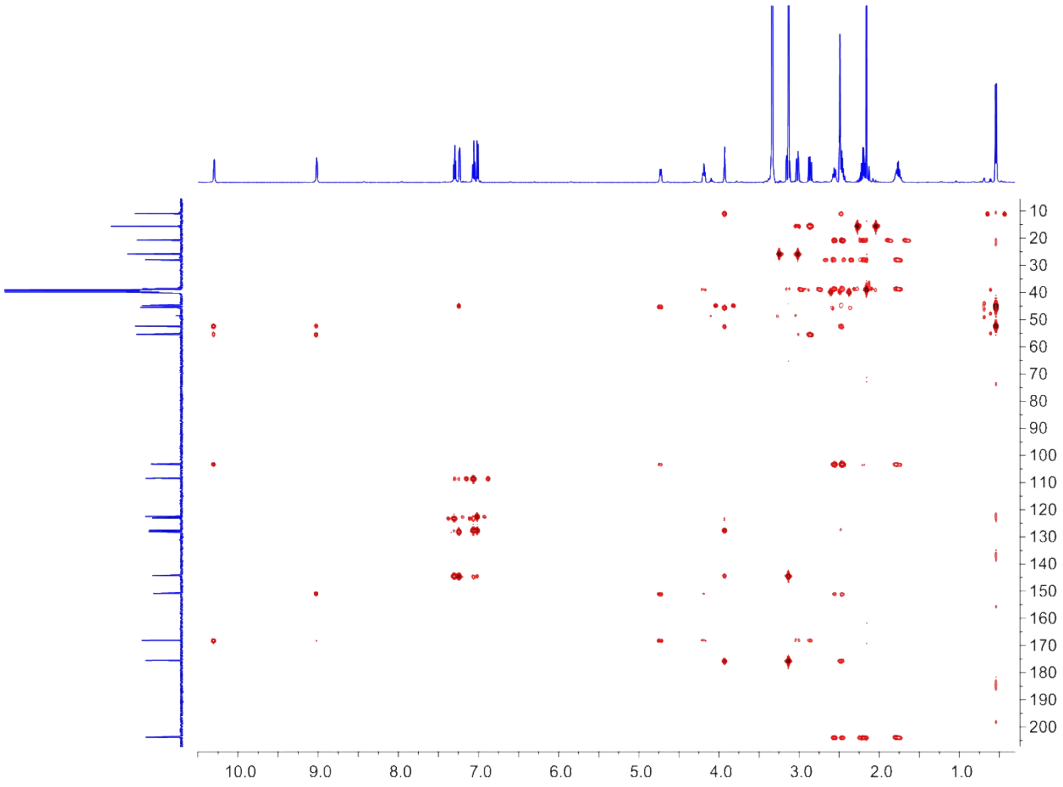
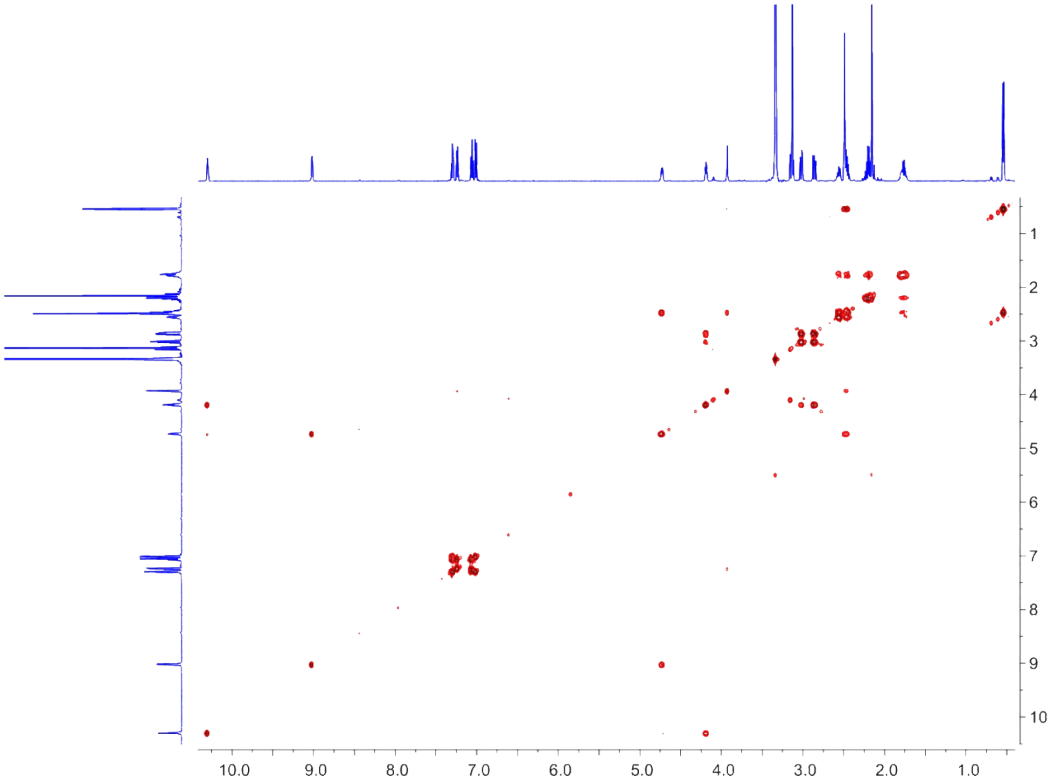


Figure S11. HSQC and HMBC NMR spectra of **2**

^{1}H,^{1}H-COSY spectrum of **2**



NOESY spectrum of **2**

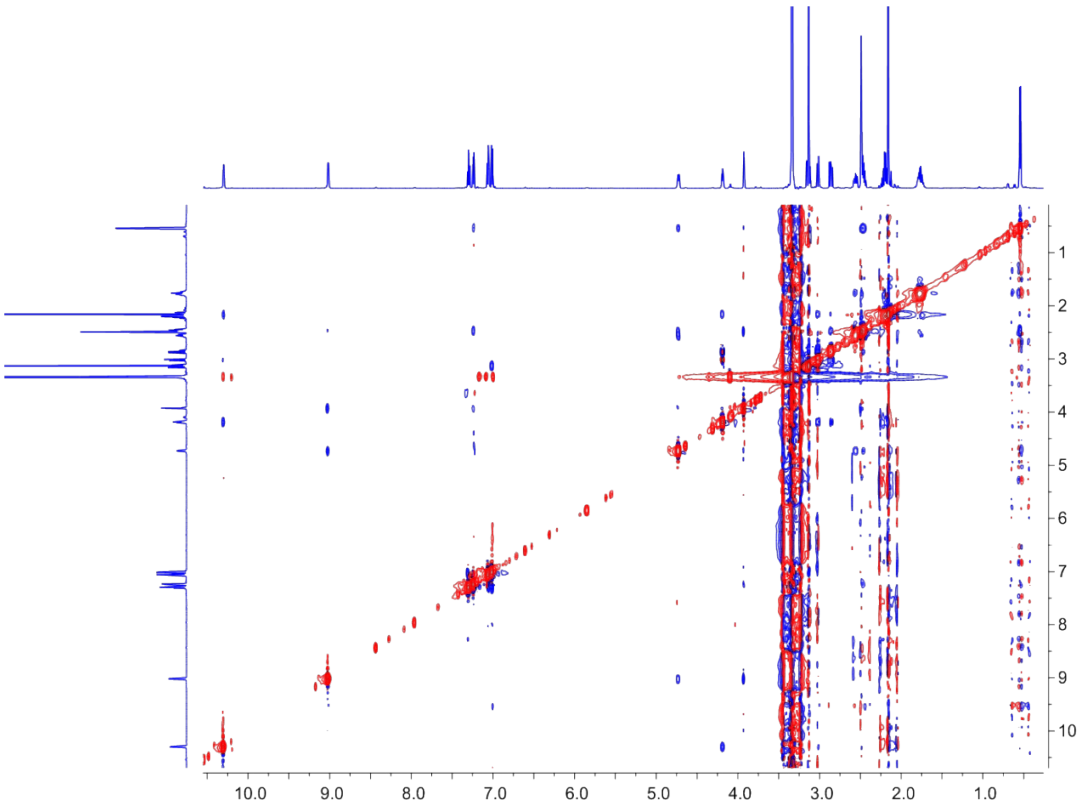
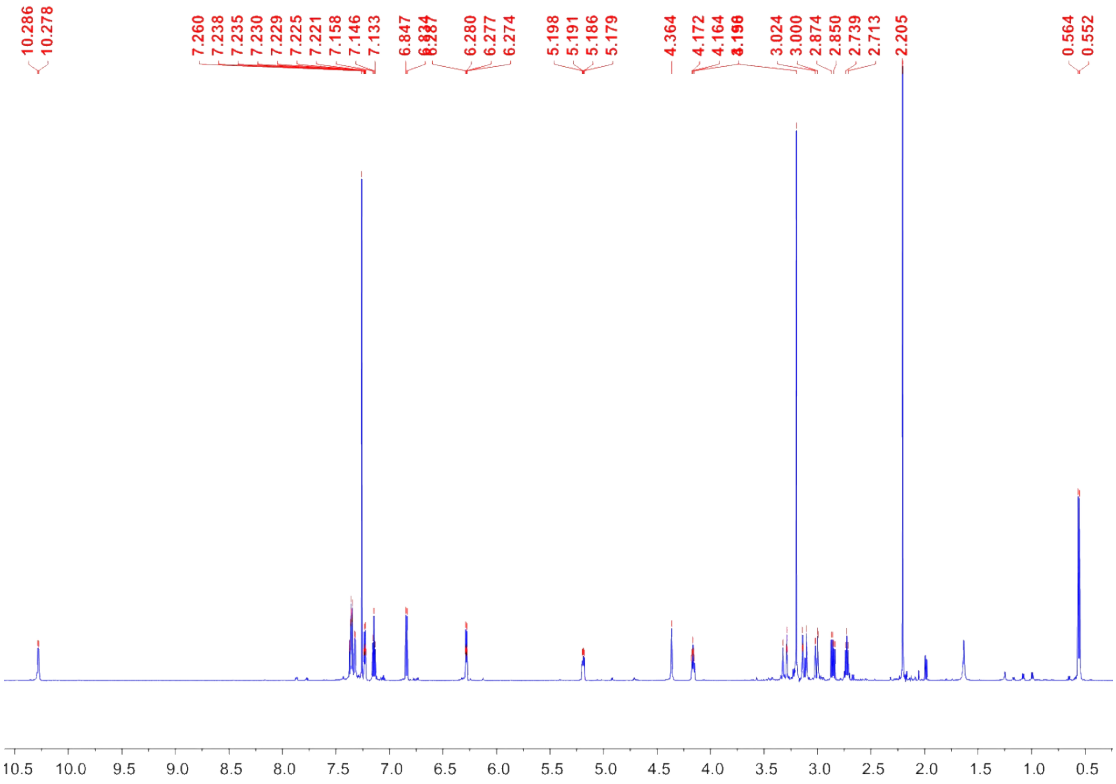


Figure S12. COSY and NOESY NMR spectra of **2**

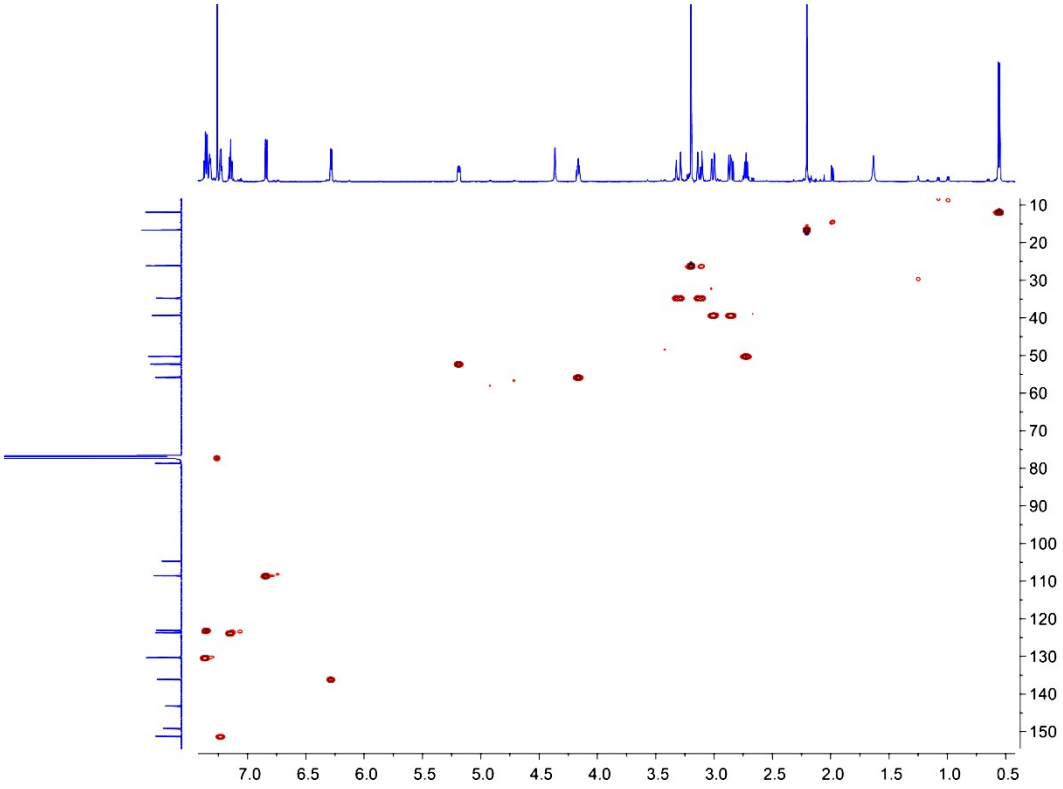
^{1}H NMR spectrum of **3** (600 MHz, $CDCl_3$)



^{13}C NMR spectrum of **3** (150 MHz, $CDCl_3$)

Figure S13. ^{1}H and ^{13}C NMR spectra of **3**

HSQC spectrum of **3**



HMBC spectrum of **3**

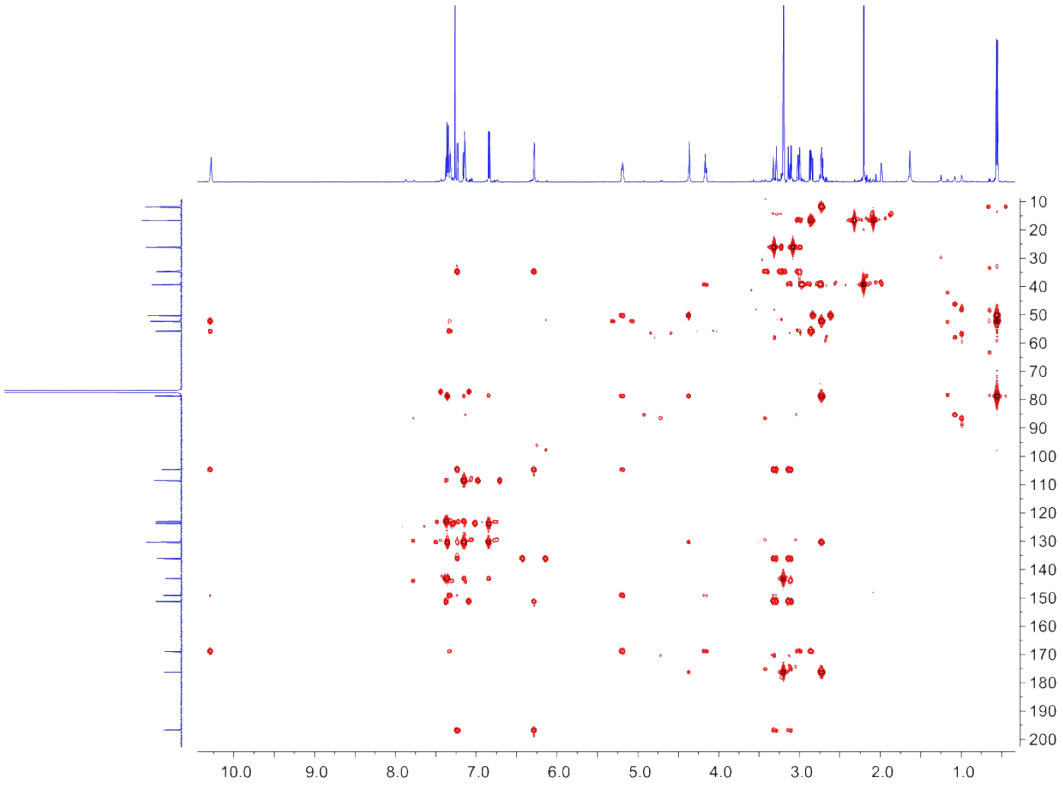
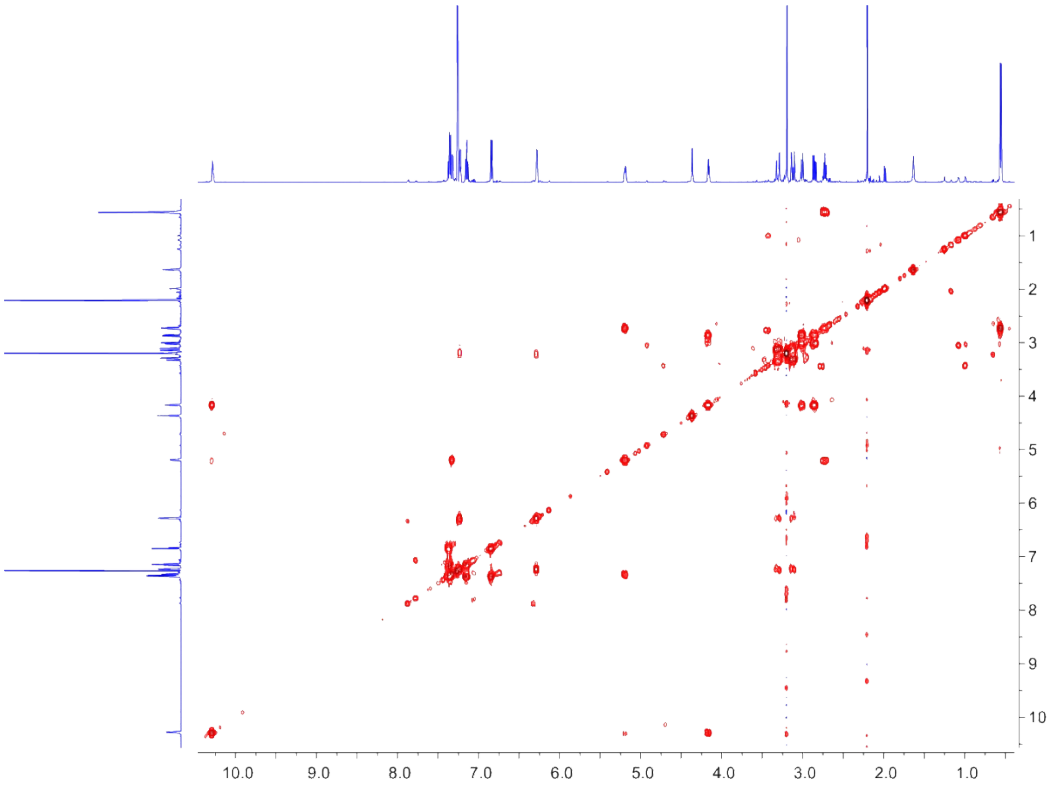


Figure S14. HSQC and HMBC NMR spectra of **3**

^{1}H,^{1}H-COSY spectrum of **3**



NOESY spectrum of **3**

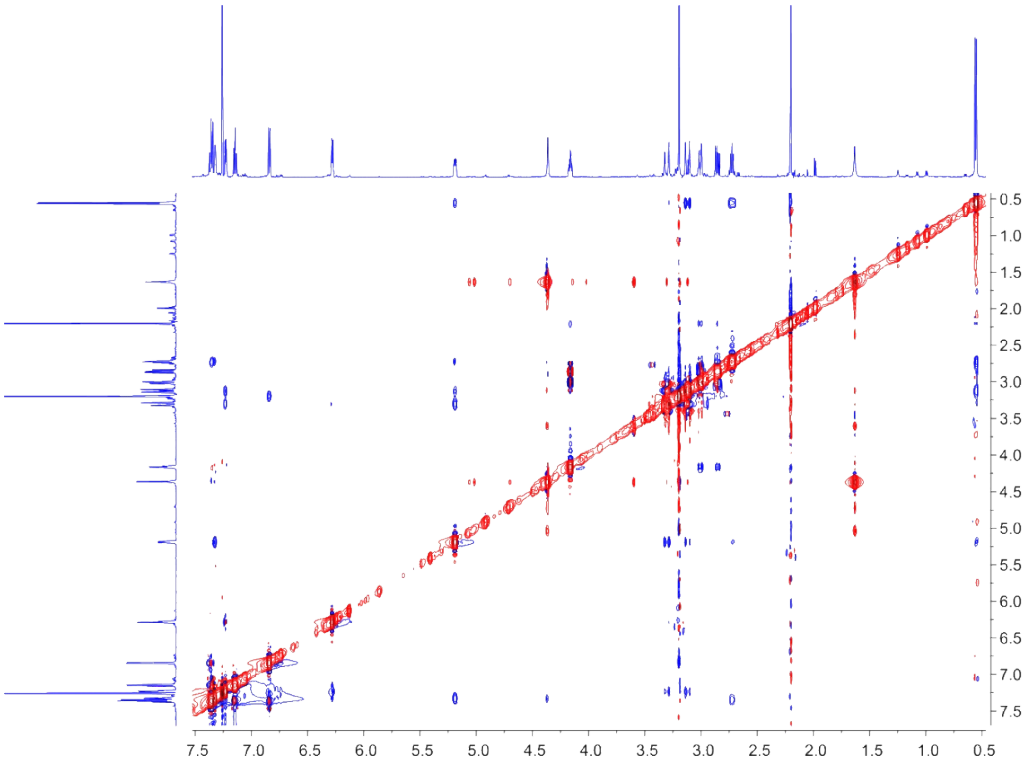


Figure S15. COSY and NOESY NMR spectra of **3**

References

1 W. Chen, D. Qu, L. Zhai, M. Tao, Y. Wang, S. Lin, N. P. Price and Z. Deng, *Protein & cell*, 2010, **1**, 1093-1105.

2 L. Bown, Y. Li, F. Berrue, J. T. P. Verhoeven, S. C. Dufour and D. R. D. Bignell, *Appl. Environ. Microbiol.*, 2017, **83**, e01169-01117.

3 W. Balk-Bindseil, E. Helmke, H. Weyland and H. Laatsch, *Liebigs Annalen*, 1995, **0**, 1291-1294.

4 Y. Lan, Y. Zou, T. Huang, X. Wang, N. L. Brock, Z. Deng and S. Lin, *Sci. China Chem.*, 2016, **59**, 1224-1228.

5 B. Gust, G. L. Challis, K. Fowler, T. Kieser and K. F. Chater, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 1541-1546.

6 T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood, *Practical streptomyces genetics*, John Innes Foundation, Norwich, United Kingdom, 2000.