Supporting Information

Genome Mining of *Streptomyces atratus* SCSIO ZH16: Discovery of Atratumycin and Identification of Its Biosynthetic Gene Cluster

Changli Sun,[†] Zhijie Yang,[†,§] Chunyan Zhang,[†,§] Zhiyong Liu,[‡] Jianqiao He,[†,§] Qing Liu,[†,∥] Tianyu Zhang,[‡,§] Jianhua Ju,*[,†,§] and Junying Ma*[,†]

[†]CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, Guangdong Key Laboratory of Marine Materia Medica, RNAM Center for Marine Microbiology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

[‡]Tuberculosis Research Laboratory, State Key Laboratory of Respiratory Disease, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

[§]University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China

[∥]State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China

Corresponding Author

* E-mail: jju@scsio.ac.cn

* E-mail: majunying@scsio.ac.cn

Table of Contents

1. General materials and experimental procedures.

Optical rotation was obtained with an MPC-500 polarimeter (Anton Paar). UV spectrum was measured with a UV-2600 spectrometer (Shimadzu) and IR spectrum was obtained using an IR Affinity-1 spectrophotometer (Shimadzu). High-resolution mass spectral data were determined using a MaXis quadrupole-time-of-flight mass spectrometer (Bruker). X-ray crystallographic data was collected on an XtaLAB PRO MM007HF X-ray diffractometer equipped with an APEX II CCD using Cu Kα radiation. 1D and 2D NMR spectra were acquired with an AVANCE Ⅲ HD 700 spectrometer (Bruker) with TMS as the internal standard. Chemical shifts (δ) were expressed in ppm with reference to the solvent signals. Column chromatography (CC) was performed on silica gel (100–200 mesh, Yantai Jiangyou Silica Gel Development Co., Ltd.), RP-18 (40–63 μm, Merck). Semi-preparative HPLC was run on Agilent 1260 liquid chromatograph with diode array detector (DAD) and a YMC-Pack ODS-A column (250 × 20 mm, 5 μm). Single-crystal data were collected on an Xcalibur Onyx Nova diffractometer (Oxford) using Cu Kα radiation. All chemicals and solvents were of analytical or chromatographic grade.

Antibiotics were purchased from Sangon Biotech Co., Ltd. (Shanghai, China). The fish meal was purchased from Weihai Xiangrun Aquatic Product Co. Ltd. (Shandong, China), and all other media or media regents were purchased from Huankai Microbial SCI. & Tech, Co. Ltd. (Guangdong, China), and the organic solvent regents for analysis were purchased from Fisher-Scientific (Waltham, MA, USA). All chemicals and solvents were of analytical or chromatographic grade.

Primers were synthesized by Sangon Biotech Co., Ltd. (Shanghai, China) and DNA sequencing PCR amplifications were performed on a Veriti Thermal Cycler (Applied Biosysntems, Carlsbad, CA) using high fidelity DNA-polymerase (Takara Biotech. Co. Ltd., Dalian, China) or Taq polymerase (TransGen Biotech Co., Ltd., Beijing, China). Gel pure DNA Kit were purchased from Omega Biotech. Co. Ltd. (Beijing, China).

2. Bacterial strains and plasmids.

S. atratus SCSIO ZH16 NS was an engineered strains of *S. atratus* SCSIO ZH16, which was isolated form a sediment sample collected from the South China Sea at a depth of 3536 m,[1] and the construction procedure had been described before.[2] DNA isolation and manipulation were carried out following standard procedures for *E. coli* and *Streptomyces*. *E. coli* were cultured at 37 ℃ or

28 ℃ in LB liquid/solid medium for 12–15 h. *Streptomyces* were inoculated on YMS medium (4 g/L yeast extract, 10 g/L malt extract, 4 g/L soluble starch, 7.5 g/L oat，2 g/L $CaCO_3$, 20 g/L agar, pH 7.4) or ISP-2 medium (4 g/L glucose, 4 g/L yeast extract, 10 g/L malt extract, 20 g/L agar) for the sporulation. ISP-4 (10 g/L soluble starch, 0.5 g/L yeast extract, 1g/L peptone, 1 g/L K_2HPO_4, 1 g/L $MgSO_4 \cdot 7H_2O$, 2 g/L $(NH_4)_2SO_4$, 2 g/L $CaCO_3$, 1 g/L NaCl, 20 g/L agar, pH 7.4) supplied with 20 mM Mg^{2+} was used for the conjugation. The media was then supplemented with appropriate antibiotics for selection: apramycin 50 μg/mL for *Strptomyces*; apramycin 50 μg/mL, kanamycin 50 μg/mL, ampicillin 100 μg/mL, chloramphenicol 25 μg/mL and trimethoprim for *E. coli*. pIJ773 plasmid was used for the amplification of *aac(3)IV*+*oriT* (RK2) fragment.

3. Activation of the atratumycin (1) biosynthetic gene cluster (BGC).

The *S. atratus* SCSIO ZH16 NS mutant strain was inoculated into 50 mL batches of ten different media (Table S1) in 250 mL Erlenmeyer flasks and then cultivated at 28 ℃ on a rotary shaker at 200 rpm for 7 d. The culture were extracted by butanone separately, and the extract were analyzed by HPLC. One small peak was observed in the extract got from FYG medium. LC-MS analysis of the extract exhibited the existence of peptide fragment ions. In order to enhance the production of the compound, the mutant strain was dealt by atmospheric and room temperature plasma (ARTP) with different time[3]. The monoclonal strains grown on ISP-2 (4 g/L glucose, 4 g/L yeast extract, 10 g/L malt extract, 20 g/L agar) plates were inoculated into 50 mL FYG medium in 250 mL Erlenmeyer flasks respectively, and then cultivated at 28℃ on a rotary shaker at 200 rpm for 7 d. By HPLC analysis of the fermentation extract, a new mutant generated by ARTP for 80 s with high production of the compound, named *S. atratus* SCSIO ZH16 NS-80S, was obtained. In order to repeat the results and establish the stable fermentation conditions for the newly discovered compound, the FYG media supplemented to be 2% XAD-16 were used. The analytical procedure is as follows: solvent system (solvent A, 15 % acetonitrile in water supplied with 0.1% acetic acids; solvent B, 85% acetonitrile in water supplied with 0.1% acetic acids); 20% B to 80% B (linear gradient, 0-20 min), 80% B -100% B (linear gradient, 20-21.5 min), 100% B (21.5-27.0 min), 100% to 0 % B (27.0-27.1min), 0% B (27.1-30.0min); flow rate was set as 1.0 mL/min. A reversed phase column (SB-C18, 5μm, 4.6 × 150 mm) was used for analysis.

4. Large scale fermentation and isolation of atratumycin (1).

To isolate the aforementioned compound, *S. atratus* SCSIO ZH16-80S was used for the large scale fermentation. The *S. atratus* SCSIO ZH16-80S mutant strain grown on YMS (4 g/L yeast extract, 10 g/L malt extract, 4 g/L soluble starch, 7.5 g/L oatmeal，2 g/L $CaCO_3$, 20 g/L agar, pH 7.4) plates was inoculated into 50 mL of FYG medium in 250 mL Erlenmeyer flasks (15 flasks were used). The media in the flasks were incubated at 28 ℃ on a rotary shaker at 200 rpm for 2 d to be used as seed culture. Then the prepared seed culture was successively transferred into 1.0 L Erlenmeyer flasks containing 200 mL of FYG medium and 2% XAD-16 macroporous resin, each for about 15 mL seed culture (50 flasks were used). The flasks were incubated at 28 ℃ on rotary shakers at 200 rpm for another 7 d.

The entire culture broth (10 L) was harvested and filtered to yield the mixture of mycelium and macroporous resin. The mixture was extracted by ultrasonication using acetone three times. The acetone layers were combined and evaporated to dryness to yield a residue (4.8 g). The residue was subjected to silica gel CC using gradient elution with $CHCl_3$ and MeOH mixtures (100:0, 95:5, 90:10, 85:15, 80:20, 70:30, 60:40, 50:50) to give eight fractions (Fr.1–Fr.8). Fr.6 was subjected to Rp-18 CC, eluted with MeCN-H_2O (30:70→100:0, 0→120 min) to afford five fractions (Fr.2-1–Fr.2-5). Fr.2-4 were separated by semi-preparative HPLC (MeCN-H_2O, 50:50) to yield analytically pure atratumycin (**1**) (21.0 mg, t_R = 17.1 min).

Atratumycin (**1**): colorless crystal (MeOH); $[\alpha]_D^{25}$ +60 (*c* 0.05, MeOH); UV (MeOH): 204 (4.99), 220 (4.86), 283 (4.46); IR (film) ν_{max}: 3289, 2955, 1755, 1693, 1628, 1516, 1443, 1277, 1242, 1223, 1066 cm^{-1}; ^1H and ^{13}C NMR spectroscopic data (Table S2); (+)-HR-ESI-MS *m/z* 1325.6193 $[M+H]^+$ (calcd for $C_{68}H_{85}N_{12}O_{16}$, 1325.6201), 1347.6007 $[M+Na]^+$ (calcd for $C_{68}H_{84}N_{12}NaO_{16}$, 1347.6020).

5. Marfey's Analysis of atratumycin (1).

Atratumycin (**1**) (0.5 mg) was hydrolyzed in 2.0 mL of 6 N HCl at 115 ℃ for 24 h. The HCl was removed in vacuo and the dry material was resuspended in 100 μL of H_2O. It was reacted with 100 μL of 1 N $NaHCO_3$, and 50 μL 1% (w/v) 1-fluoro-2, 4-dinitrophenyl-5-L-alaninamide (FDAA) in acetone. The reaction mixtures were incubated in a water bath at 40 °C for 1 h. 5 μL of 1 N HCl was added to quench the reaction by neutralization. MeOH was then added to the quenched reaction to afford a total volume of 500 μL; 10 μL of hydrolysate derivatization reaction was used for HPLC analysis using a Kromasil 100-5C_{18} column (5 μL, 4.6 × 250 mm) with a gradient from 20% to 50%

solvent B (CH_3CN/ddH_2O/TFA, A: 0/100/0.1, B: 100/0/0.1) over the course of 42 min，followed with 50% to 100% solvent B from 42.01 to 46 min, and UV detection at 340 nm at a flow rate of 1 mL/min. Similarly, 10 μL of the standard amino acids in H_2O (10 μM) was added to 1 N $NaHCO_3$ (5 μL) and each mixture was treated with 1% (w/v) FDAA (20 μL) for 1 h at 40 °C. Derivatization reactions were terminated with 1 N HCl (5 μL) and diluted to a total volume of 500 μL with MeOH. Of these standard amino acid derivatization reactions, 10 μL was subjected to HPLC analysis and used as structural standards in the elucidation of structure **1** (Figure S4).

6. X-Ray crystallography of atratumycin (1).

Colorless flaky crystals of atratumycin (**1**) were obtained in MeOH. A suitable crystal ($0.10 \times 0.10 \times 0.05$ mm^3) was selected and data recorded with an XtaLAB PRO MM007HF X-ray diffractometer equipped with an APEX II CCD using Cu Kα radiation (Table S3). The structure was solved with the ShelXT structure solution program using the dual solution method and refined using full-matrix least-squares difference Fourier techniques. Crystallographic data for **1** has been deposited with the Cambridge Crystallographic Data Centre with the deposition number CCDC 1883521. Copies of the data can be obtained, free of charge, on application to the Director, CCDC, 12 Union Road, Cambridge CB2 1EZ, UK [fax: +44(0)-1233-336033 or e-mail: deposit@ccdc.cam.ac.uk].

7. Bioinformatics analyses of the atratumycin BGC.

The genomic sequence of *S. atratus* SCSIO ZH16 and the gene cluster analysis by online anti-SMASH software (http://antismash secondary metabolites.org/) had been reported before.[1] ORFs were analyzed using the online Frameplot 4.0Beta software (http://nocardia.nih.go.jo/fp4/), and their functional predictions were obtained using online BLAST software (http://blast. ncbi.nlm.nih.gov/) (Table S4). The NRPS domains were predicted using the PKS/NRPS analysis website (http://nrps.igs.umaryland.edu/nrps/). The gene cluster for atratumycin (*atr*) was deposited in GenBank under the accession number MK370905.

8. Construction of gene inactivation mutants.

The construction of genomic library of *S. atraus* SCSIO ZH16 with the vector of SupercosI

(Amp^R, Kan^R) has been previously described.[3] Seven pairs of primers were used for the screen of the genomic library (Table S7). In the library, ten cosmids (56G, 102C, 177G, 189A, 73C, 84E, 183H, 198F, 54D and 1610C) cover the whole gene cluster and the up/down stream sequences. Cosmids 198F, 102C, 56G and 1610C were used for genes inactivation in this study (Figure S19). A summary of all strains and plasmids generated for this work is shown in Table S8.

λ-Red recombination mediated PCR-targeting strategy was employed to disrupt the genes of interest (*orf(-1), orf(-2), atr23, atr27, orf(+1), orf(+2))*. Each target gene in the template plasmids was replaced by the apramycin resistance gene *aac(3)IV+oriT* (RK2). Recombination plasmids for gene replacements were further introduced into *S. atraus* SCSIO ZH16NS-80S by conjugation from *E.coli* ET12567/pUZ8002. The gene inactivated mutant strains were selected on the basis of apramycin resistance and trimethoprim (50 μg/mL), and the genetic phenotype of the mutants was further confirmed by PCR.

9. Anti-bacterial assays.

The anti-bacterial activities were measured using the method previously described[3]. Eight kinds of Gram positive (*Enterococus faecalis* ATCC29212, *Straphylococcus aureus* ATCC29213, MRSA ssha-01, *Micrococcus luteus* SC-1, *Mycobacteria Tuberculosis* H37Ra and H37Rv) and Gram negative bacteria (*Escherichia coli* ATCC25922, *Acinetobater baumanii* ATCC19606) were used for antimicrobial bioassays. A data summary is presented in Table S5.

10. Cytotoxicity assays

Cytotoxicity evaluations were performed using the MTT method using human gastric carcinoma cell line MGC803, human lung cancer cell line A549, human cervical carcinoma cell line HeLa, human hepatocellular carcinoma cell line HepG2, human breast adenocarcinoma cell line MCF7, human colorectal carcinoma cell line RKO as described in our previous publication.[3] The human normal hepatic cell line L02 and human normal umbilical vein endothelial cell line Huvec-12 cell lines were used as normal cell line controls. Atratumycin (1) with different concentrations (0 μM, 2 μM, 4 μM, 8 μM, 16 μM, 32 μM, 64 μM) was tested in the experiment. Data for these cytotoxicity assays are shown in Table S6. All the cell lines were obtained from American Type Culture Collection (ATCC) and cultured according to ATCC recommendations. All experiments were performed in triplicate with doxorubicin and *cis*-platin agents serving as controls.

11. Construction of plasmid for the production of Atr27 protein.

The *atr27* sequence was amplified from 198F cosmid DNA by PCR using primers Exp-*atr27*F and Exp-*atr27*R (Table S8). The amplified fragment was digested with *Nde*I and *Hin*dIII enzymes and then was cloned into pCWori-HisN vector, which was also treated with the same two enzymes to generate the Atr27 expression plasmid pCWori-HisN-*atr27*. Then, it was introduced into *E.coli* BL21(DE3) for protein expression.

12. Atr27 protein expression, purification and enzyme assays.

A single colony of *E.coli* BL21(DE3) harboring pCWori-HisN-*atr27* was inoculated into LB medium containing 100 g/mL ampicillin and was cultivated at 37℃ for overnight. The seed culture was used for 1:100 inoculation of 1 liter of TB broth [4] with 100 μg/mL ampicillin cultivating for 2 hours until OD600 reached at 0.6-1.0. After incubation on ice for 10 mins, heterologous gene expression was induced by adding 0.1mM isopropyl-beta-d-thiogalactopyranoside (IPTG), 1mM thiamine and 1mM 5-aminolevulinic acid to the culture. The cells were further incubated with shaking for 14-16 hours at 25℃, and they were harvested by centrifugation. The pellet was resuspended in binding buffer (50 mM Tris-HCl, 500mM NaCl, 10 mM imidazole and 10% glycerol, pH 8.0). After lysis by sonication, the cell debris was removed by centrifugation. His60 Ni^{2+} resin (GE Health care, USA) was added to the clear lysate and incubated at 4℃ for 1h. Then, the resins were loaded onto an open column and washed with 50 mL binding buffer. Then the absorbed protein was eluted stepwise with 20 mL wash buffer containing different concentrations of imidazole (10, 20, 30, 40, 50, 60, 100 and 125 mM), finally it was washed with 3 mL elution buffer (50 mM Tris-HCl, 500mM NaCl, 250 mM imidazole and 10% glycerol, pH 8.0). The fractions containing the recombinant protein abundantly were desalted and concentrated with an Amicon Ultra centrifugal with 30 kDa molecular mass cutoff (Millipore, Billerica, USA). The protein concentrations were determined using NanoDrop (thermo Fisher Sceintific, Waltham, USA) with the extinction coefficient of 47565 M^{-1}. Cm^{-1} for Atr27. The concentrations of redox partner proteins were determined by a Bradford assay [5], using bovine serum albumin (BSA) as a standard.

The standard reaction mixture contained 10 μM Atr27, 10 μM Fdx (20 μg/ml for *sp*Fdx) plus 10 μM FdR (0.2 U/mL for *sp*FdR), 0.5 mM L-Phe, and 1mM NADPH in 100 μL of 8 different reaction

buffer at 30℃ for overnight. The reaction was quenched by added 5 μL trifluoroacetic acid (TFA) and centrifuged to generate the supernatants, which were subject to HPLC analysis. HPLC analysis was normally performed with a linear gradient of 2 to 20% CH_3CN (v/v) over 20 min, 20% to 90% CH_3CN (v/v) over 5 min, 90% CH_3CN (v/v) over 5min, 90-2% CH_3CN (v/v) over 0.1min, and 2% CH_3CN for a further 4.9 min in H_2O supplemented with 0.1% (v/v) TFA at a flow rate of 1 mL/min. The detection wavelengths were set at 210 nm and 254 nm.

Table S1. The media used in the optimization of *S. atratus* SCSIO ZH16NS.

The name of the media	The composition of the media (gram per liter)
N4	fishmeal (8), soluble starch (4), corn flour (8), peptone (5), glycerol (6), $CaCO_3$ (2), KBr (0.2), Sea Salt (30);
N4a1	fishmeal (10), glycerol (10), $CaCO_3$(2), KBr (0.2), sea salt (30);
N4b1	fishmeal (8), soluble starch (15), glycerol (6), $CaCO_3$ (2), KBr (0.2), sea salt (30);
N4c1	corn flour (10), peptone (5), glycerol (6), $CaCO_3$ (2), KBr (0.2), sea salt (30);
N4d1	soy flour (10), glycerol (6), $CaCO_3$ (2), KBr (0.2), sea salt (30);
Am6-1	soluble starch (20), glycerol (10), yeast extract (5), $CaCO_3$ (5), sea salt (30);
Am2ab	soluble starch (5), glucose (20), yeast extract (2), peptone (2), soy flour (5), $MgSO_4.7H_2O$ (0.5), KH_2PO_4 (0.5), NaCl (4), $CaCO_3$ (2), sea salt (30);
Am3	soy flour (5), bacteria peptone (15), soluble starch(15), glycerol (15), $CaCO_3$ (2), sea salt (30);
FYG	fish meal (10), yeast extract (5), glycerol(2), calcium carbonate (5);
ISP-2	yeast extract (4), malt extract (10), glucose (4), sea salt (30);

Table S2. ^{1}H and ^{13}C NMR data for atratumycin (**1**) in DMSO-d_6.

Position	δ_C (ppm)	δ_H (ppm)	position	δ_C (ppm)	δ_H (ppm)
1	169.8		35	73.6	4.52, overlap
2	123.1	7.33, overlap	36	141.9	
3	138.2	7.89 d (15.7)	37	126.8	7.57, d (7.1)
4	137.3		38	128.2	7.31, overlap
5	126.5	7.48, d (7.7)	39	127.8	7.30, overlap
6	126.7	7.10, overlap	40	128.2	7.31, overlap
7	130.2	7.16, m	41	126.8	7.57, d (7.1)
8	131.2	7.00, d (7.7)	42	170.7	
9	133.6		43	60.9	4.46, overlap
10	18.9	1.76, s	44	30.6	2.29, m
11	174.9				1.72, m
12	61.2	5.13, d (2.4)	45	25.2	1.92, m
12-NH		9.29, brs			1.85, m
13	69.1	5.23, m	46	48.1	3.92, overlap
14	18.0	1.38, d (6.7)			3.68, m
15	171.0		47	172.5	
16	49.1	3.17	48	58.3	4.31, m
16-NH		9.20, s	48-NH		8.37, d (9.6)
17	60.2	3.25, overlap	49	32.1	1.82, overlap
		3.10, overlap	50	19.7	0.82, d (6.0)
18	171.7		51	19.3	0.82, d (6.0)
19	54.7	4.47, m	52	171.0	
19-NH		7.98, d (9.8)	53	40.5	4.72, m
20	28.1	3.80, br. D (13.4)	53-NH		8.43, d (9.3)
		3.23, overlap	54	171.0	
21	111.4		55	55.2	4.23, dd (13.4, 6.9)
21-NH		10.75, s	55-NH		8.88, s
22	123.4	7.08, s	56	40.5	3.55, m
23	127.8		57	24.4	1.54, m
24	118.6	7.43 d (7.8)	58	21.9	0.79, d (6.5)
25	118.8	6.88, dd (7.8, 7.4)	59	22.6	0.91, d (6.5
26	121.2	6.93, dd (7.4, 7.1)	60	169.8	
27	127.8	7.22, d (7.1)	61	53.0	4.83, m
28	136.4		61-NH		8.08, d (7.7)
29	171.4		62	34.1	3.11, overlap
30	50.9	4.52, m			2.88, m
30-NH		7.95, d (7.4)	63	128.3	
31	38.6	2.53, dd (15.4, 10.3)	64	130.3	6.98, d (8.3)
32	168.5		65	115.3	6.64, d (8.3)
33	169.9		66	156.3	
34	57.7	4.53, overlap	67	115.3	6.64, d (8.3)
34-NH		7.34，br. s	68	130.3	6.98, d (8.3)

Table S3. Crystal structure data and structure refinement for atratumycin (**1**).

Identification code	atratumycin
Empirical formula	$C_{68}H_{84}N_{12}O_{16}$·MeOH·3H_2O
Formula weight	1411.56
Temperature/K	293(2)
Crystal system	monoclinic
Space group	$P2_1$
a/Å	12.41130(10)
b/Å	21.0957(2)
c/Å	15.1196(2)
α/°	90
β/°	106.5030(10)
γ/°	90
Volume/Å3	3795.61(7)
Z	2
ρcalcg/cm^3	1.235
μ/mm^{-1}	0.760
F(000)	1504.0
Crystal size/mm^3	0.1 × 0.1 × 0.05
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	7.4 to 133.744
Index ranges	-13 ⩽ h ⩽ 14, -24 ⩽ k ⩽ 25, -18 ⩽ l ⩽ 17
Reflections collected	38877
Independent reflections	13094 [R_{int} = 0.0359, R_{sigma} = 0.0338]
Data/restraints/parameters	13094/1/930
Goodness-of-fit on F^2	1.067
Final R indexes [I>=2σ (I)]	R_1 = 0.0474, wR_2 = 0.1335
Final R indexes [all data]	R_1 = 0.0511, wR_2 = 0.1364
Largest diff. peak/hole / e Å^{-3}	0.48/-0.32
Flack parameter	0.04(6)

Table S4. Deduced functions of ORF in the *atr* gene cluster in *S. atratus* SCSIO ZH16.

ORF	Size[a]	Proposed functions	SI/ID[b]	Protein homologue and origin
orf(-2)	516	histidine ammonia-lyase	98/97	WP_072486511.1, *Streptomyces atratus*
orf(-1)	134	IS481 family transposase	89/85	WP_093735660.1, *Streptomyces* sp. DvalAA-14
atr1	304	LuxR family transcriptional regulator	81/72	WP_016824036.1, *Streptomyces viridosporus*
atr2	325	LuxR family transcriptional regulator	66/54	WP_012379778.1, *Streptomyces griseus*
atr3	296	ACP S-malonyltransferase	88/83	WP_004981554.1, *Streptomyces viridosporus*
atr4	255	Thioesterase	90/82	WP_081240493.1, *Streptomyces viridosporus*
atr5	84	Putative acyl carrier protein (ACP)	92/87	WP_100109525.1, *Streptomyces peucetius*
atr6	420	3-oxoacyl-ACP synthase	90/85	WP_100109526.1, *Streptomyces peucetius*
atr7	381	3-oxoacyl-ACP synthase	95/91	WP_004981562.1, *Streptomyces viridosporus*
atr8	317	3-oxoacyl-ACP synthase	93/92	WP_081236343.1, *Streptomyces viridosporus*
atr9	283	alpha/beta hydrolase	94/93	WP_016824042.1, *Streptomyces viridosporus*
atr10	323	hypothetical protein (hot-dog family)	97/95	WP_081236344.1, *Streptomyces viridosporus*
atr11	366	3-oxoacyl-ACP synthase	91/90	WP_081236346.1, Streptomyces viridosporus
atr12	79	Putative ACP protein	93/87	WP_004981588.1, *Streptomyces viridosporus*
atr13	152	dehydratase	70/59	WP_060952244.1, *Streptomyces hygroscopicus*
atr14	159	beta-hydroxyacyl-ACP dehydratase	96/94	WP_004981594.1, *Streptomyces viridosporus*
atr15	248	3-oxoacyl-ACP-reductase	97/92	WP_081236349.1, *Streptomyces viridosporus*
atr16	206	isomerase	92/90	EFE66248.1, *Streptomyces ghanaensis* ATCC 14672
atr17	70	MbtH family protein	95/91	WP_081236351.1, *Streptomyces viridosporus*
atr18	162	condensation	92/90	EFE66250.1, *Streptomyces ghanaensis* ATCC 14672
atr19	187	DUF3995 domain-containing protein	64/48	WP_098889550.1, *Bacillus cereus*
atr20	63	MbtH family protein	71/56	WP_071803045.1, *Couchioplanes caeruleus*
atr21	3932	CAT-CATE-CAT-Te	70/59	AEA30274.1, *Streptomyces sp. Acta* 2897
atr22	3692	CAT-CAT-CATE	68/57	AEA30273.1, *Streptomyces sp. Acta* 2897
atr23	4855	CAT-CAT-CAT-CATE	70/59	EFE66253.1, *Streptomyces ghanaensis* ATCC 14672
atr24	79	IS3/IS911 transposase	77/71	ADW07812.1, *Streptomyces pratensis* ATCC 33331
atr25	323	transposase	87/87	WP_016824053.1, *Streptomyces viridosporus*
atr26	108	non-ribosomal peptide synthetase	75/70	WP_016824053.1, *Streptomyces viridosporus*

atr27	423	cytochrome P450	95/60	WP_043493936.1, *Streptomyces viridosporus*
atr28	154	DUF3995 domain-containing protein	61/58	WP_004981609.1, *Streptomyces viridosporus*
atr29	318	ABC transporter	97/95	WP_004981611.1, *Streptomyces viridosporus*
atr30	262	ABC transporter permease	97/93	WP_081238217.1 ,*Streptomyces viridosporus*
atr31	341	polyprenyl synthetase family protein	87/82	WP_043493938.1 , *Streptomyces viridosporus*
atr32	667	SARP family transcriptional regulator	66/54	WP_044581768.1, *Streptomyces iranensis*
atr33	127	RIP metalloprotease RseP	46/29	WP_047909685.1, *Paenibacillus sp.* TCA20
orf(+1)	95	Hypothetical protein	88/85	WP_093900098.1, *Streptomyces* sp. Ncost-T10-10d
orf(+2)	122	Hypothetical protein	95/93	WP_093900098.1, *Streptomyces* sp. Ncost-T10-10d
orf(+3)	414	ErfK/YbiS/YcfS/YnhG family protein	90/87	ADW03751.1, *Streptomyces pratensis* ATCC 33331

[a]: amino acids, [b]: similarity/identity

Table S5. The antibacterial activities of **1** against eight bacteria (MICs, μM)[a].

	E.coli ATCC25922	*A.baumanii* ATCC19606	*E.faecalis* ATCC29212	*S.aureus* ATCC29213	MRSA shha-01	*M. Luteus* SC-1	*M. tuberculosis* H37Ra	*M. tuberculosis* H37Rv
1	>75.5	>75.5	>75.5	>75.5	>75.5	>75.5	3.8	14.6
Kan[b]	1.7	3.4	/	/	/	0.85	/	/
Rif[c]	/	/	/	/	/	/	0.1	0.6

a: *Escherichia coli ATCC25922*, *Acinetobacter baumannii* ATCC19606, *Enterococcus faecalis* ATCC29212, *Staphylococcus aureus* ATCC29213, MRSA shha-01, *Micrococcus luteus* SC-1; *Mycobacteria tuberculosis* (H37Ra and H37Rv)

b: kanamycin

c: rifampicin

/: not tested

Table S6. The cytotoxic activities of compound **1** (IC_{50}, μM)[a].

	MGC803	A549	HeLa	HepG2	MCF-7	RKO	L02	Huvec-12
1	>50	>50	>50	>50	>50	>50	>50	>50
Dox[b]	3.64	2.08	0.75	0.72	0.81	1.53	8.57	7.16
Cis[c]	12.92	9.31	6.07	10.27	6.39	19.96	21.61	23.08

[a]MGC803, human gastric carcinoma cell line; A549, human lung cancer cell line; HeLa, human cervical carcinoma cell line; HepG2, human hepatocellular carcinoma cell line; MCF7, human breast adenocarcinoma cell line; RKO, human colorectal carcinoma cell line; L02, human normal hepatic cell line; Huvec-12, human normal umbilical vein endothelial cell line.

[b] Doxorubicin

[c] Cisplatin

Table S7. The primers used in this study.

Primer	Sequence (5`-3`)
For the genomic library screening	
Screen 1F	GTGCAGCAGCACGCGCTCG
Screen 1R	GTTCGACCTGTCGATCGACC
Screen 2F	CCGCACCGACGAGCAGGTCAAG
Screen 2R	ACCGCCGTGCCGATGTCCTC
Screen3F	CAGGCTTATCGGAGCAGTAC
Screen3R	TCCAGACCATGCTGACCTTC
Screen4F	TCGGCGAGGACGAACAGCT
Screen4R	GCAGTTGGTGGCCTATCTGG
Screen5F	CGCGTGGTTCTGACGGGACT
Screen5R	ACCTCGGCGTAGATGGTGGC
Screen6F	TGAACGTGTTGAGCGCACC
Screen6R	GCTGATGTCGAGGAGGTCGT
Screen7F	TCGAGTTCAGCAGTTCGCT
Screen7R	TGAGGCGTTCTGCTGCATCT
For construction of gene inactivation mutant	
*atr23*delF	gccggggaactgctggacagcgagcggctgcggtcgctgATTCCGGGGATCCGTCGACC
*atr23*delR	ggccaccaactgcttgaccccgggccggtcctcgcgggcTGTAGGCTGGAGCTGCTTC
atr27delF	gacagcttcgacctgaccgatcccatgacctatcagacg ATTCCGGGGATCCGTCGACC
atr27delR	ccgcaccgcgtcgagcatggcgttgatctcgacacggcc TGTAGGCTGGAGCTGCTTC
*orf(-1)*delF	cagccggtccacgtggacatcaagaagctcggcaacatc ATTCCGGGGATCCGTCGACC
*orf(-1)*delR	ctgggagttactggcgcccacgctgctgccggacgactc TGTAGGCTGGAGCTGCTTC
*orf(-2)*delF	cacatcagtcaggatctccgcgcacagctccagcgcaac ATTCCGGGGATCCGTCGACC
*orf(-2)*delR	ctcggtgatcgaaccgaggtcggccgcgacgatcgccag TGTAGGCTGGAGCTGCTTC
*orf(+1)*delF	gctctccgcccgcttctgaccgccaccgctgcgggcacg ATTCCGGGGATCCGTCGACC
*orf(+1)*delR	gcggcgtgccgagtaggtgacgaacgctgcgccgatgcc TGTAGGCTGGAGCTGCTTC
*orf(+2)*delF	agcggtaccgcactggagtccatcgccgccgagtacggg ATTCCGGGGATCCGTCGACC
*orf(+2)*delR	gctgtgctcgccgtcgtcctcaccattgtgcgggacgtcTGTAGGCTGGAGCTGCTTC
For verification of the mutants	
Id*atr23*F	tacccgggccgcctccttg
Id*atr23*R	gaacgccgagctgctggcc
Id*atr27*F	ctatgaccagcgccaccgc
Id*atr27*R	cgttcggcatcgtattgctcttc

Id*orf(-1)*F	cccatacgccgctacgaacgc
Id*orf(-1)*R	ccgttgcagtggcacgagttc
Id*orf(-2)*F	gaaccggtctacggcgtctcc
Id*orf(-2)*R	ggtgtactgggcgatcatcag
Id*orf(+1)*F	tcagacggacagggtgtcggc
Id*orf(+1)*R	gattccgtggccgctgctctc
Id*orf(+2)*F	gaaatggtgctggaggtcacg
Id*orf(+2)*R	ctagatccgccgctccgcacc
For gene expression	
Exp*atr27*F	aaa*CATATG*atgaccagcgccaccgcgac
Exp*atr27*R	aaa*AAGCTT*tcagctccaccgggcgagcc

Table S8. The strains and plasmids used or constructed in this study.

Strains or plasmids	Description	Reference or source
Strains		
E. coli BW25113/pIJ790	Host strain for PCR-targeting	[6]
E. coli ET12567/pUZ8002	Host strain for conjugation	[7]
S. atratus SCSIO ZH16NS	atratumycin producing strain	[1]
E.coli BL21(DE3)	Expression host	Life Technologies
S. atratus SCSIO ZH16NS-80S	The engineered strain by ARTP	This study
Δatr23	*atr23* inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
Δatr27	*atr27*inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
Δorf(−1)	*orf(−1)* inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
Δorf(−2)	*orf(−2)* inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
Δorf(+1)	*orf(+1)* inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
Δorf(+2)	*orf(+2)* inactivation mutant strain originated from *S. atratus* SCSIO ZH16NS-80S	This study
E. coli BL21(DE3)/*atr27*	*atr27 was expression in E.coli* BL21(DE3)	This study
plasmids		
pIJ790	Cml^r, including *λ-RED (gam, bet, exo)* for PCR-targeting	[8]
pIJ773	Apr^r, *source of acc(3)IV and oriT* fragment	[8]
pUZ8002	Kan^r, including *tra* for conjugation	[9]
p19-8F	Amp^r, Kan^r,used for inactivation of *atr23*	This study
p19-8F	Amp^r, Kan^r,used for inactivation of *atr27*	This study
P16-10C	Amp^r, Kan^r,used for inactivation of *orf(+1), orf(+2)*	This study
p10-2C	Amp^r, Kan^r,used for inactivation of *orf(-1), orf(-2)*	This study
pCWori-*atr27*	Vector for expression of *atr27*	This study

Table S9. *In vitro* biochemical reaction of Atr27 using L-Phe as substrate with surrogate redox partner proteins from *Synechococcus elongatus* PCC7942.

Components	Final concentration	Tris-HCl buffer			PBS buffer			HEPES buffer	
		pH 7.6	pH 8.0	pH 8.5	pH 6.8	pH 7.4	pH 8.0	pH 7.0	pH 7.5
Atr27	10 μM	√	√	√	√	√	√	√	√
L-Phe	500 μM	√	√	√	√	√	√	√	√
NADPH	1 mM	√	√	√	√	√	√	√	√
*se*Fdx	10 μM	√	√	√	√	√	√	√	√
*se*FdR	10 μM	√	√	√	√	√	√	√	√
total volume	100 μL	√	√	√	√	√	√	√	√

Note: The reaction mixtures were incubated at 30℃ for overnight. *se*Fdx and *se*fdR were the surrogate redox partner proteins from *Synechococcus elongatus* PCC7942.

√: item applied

Table S10. *In vitro* biochemical reaction of Atr27 using L-Phe as substrate with surrogate redox partner proteins from spinach.

Components	Final concentration	Tris-HCl buffer			PBS buffer			HEPES buffer	
		pH 7.6	pH 8.0	pH 8.5	pH 6.8	pH 7.4	pH 8.0	pH 7.0	pH 7.5
Atr27	10 μM	√	√	√	√	√	√	√	√
L-Phe	500 μM	√	√	√	√	√	√	√	√
NADPH	1 mM	√	√	√	√	√	√	√	√
*sp*Fdx	0.02 mg	√	√	√	√	√	√	√	√
*sp*FdR	0.2 U/mL	√	√	√	√	√	√	√	√
total volume	100 μL	√	√	√	√	√	√	√	√

Note: The reaction mixtures were incubated at 30℃ for overnight. *sp*Fdx and *sp*fdR were the surrogate redox partner proteins from spinach.

√: item applied

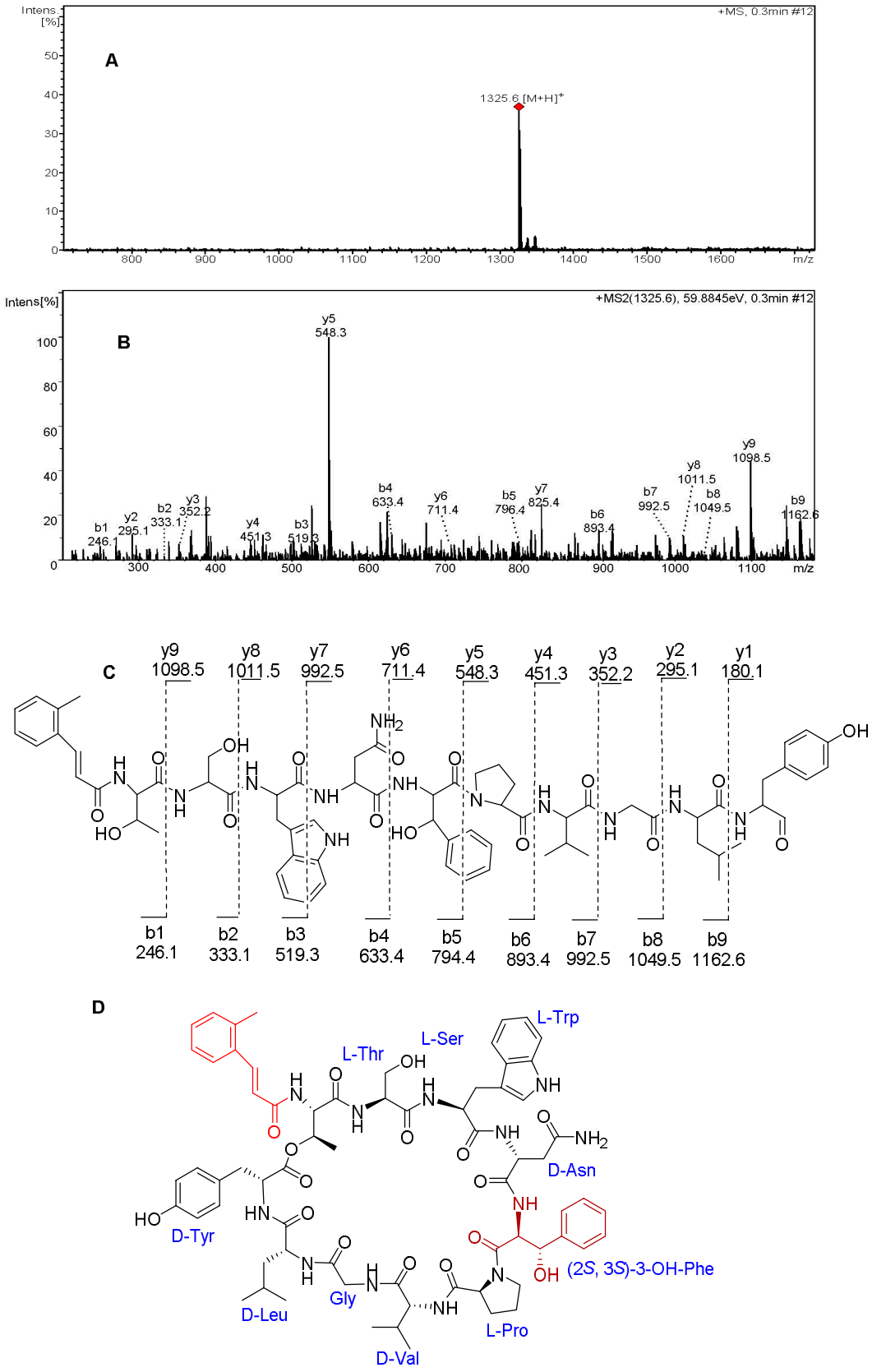

Figure S1. MS and MS^2 fragmentation of atratumycin (**1**). (A) The parent ion of the compound at *m*/*z* 1325.6201 $[M+H]^+$; (B) the MS^2 spectrum of the parent ion; (C) the designation of ions from the MS^2 spectrum; (D) atratumycin (**1**) from genetically engineered *Streptomyces atratus* SCSIO ZH16NS-80S.

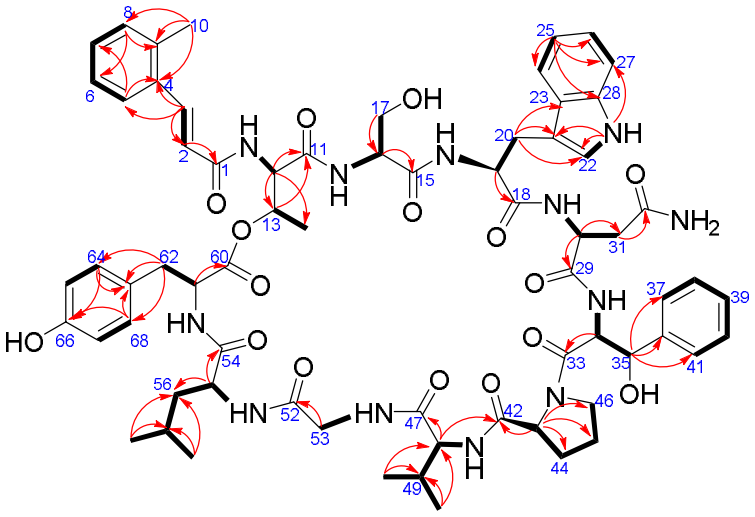

Figure S2. Selected COSY (bold) and HMBC (red arrows) correlations of atratumycin (**1**).

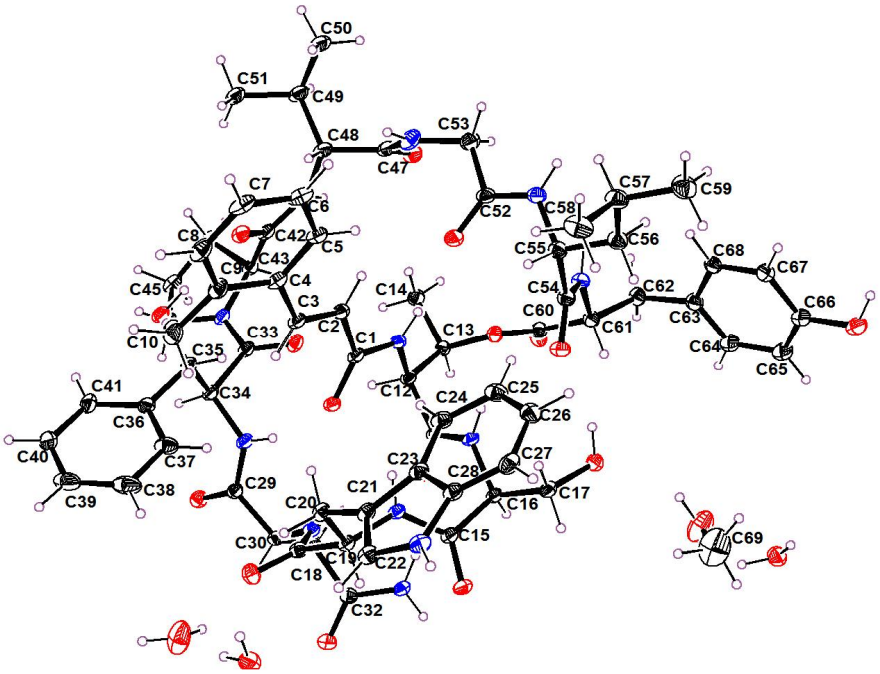

Figure S3. The X-ray crystallographic structure of atratumycin (**1**).

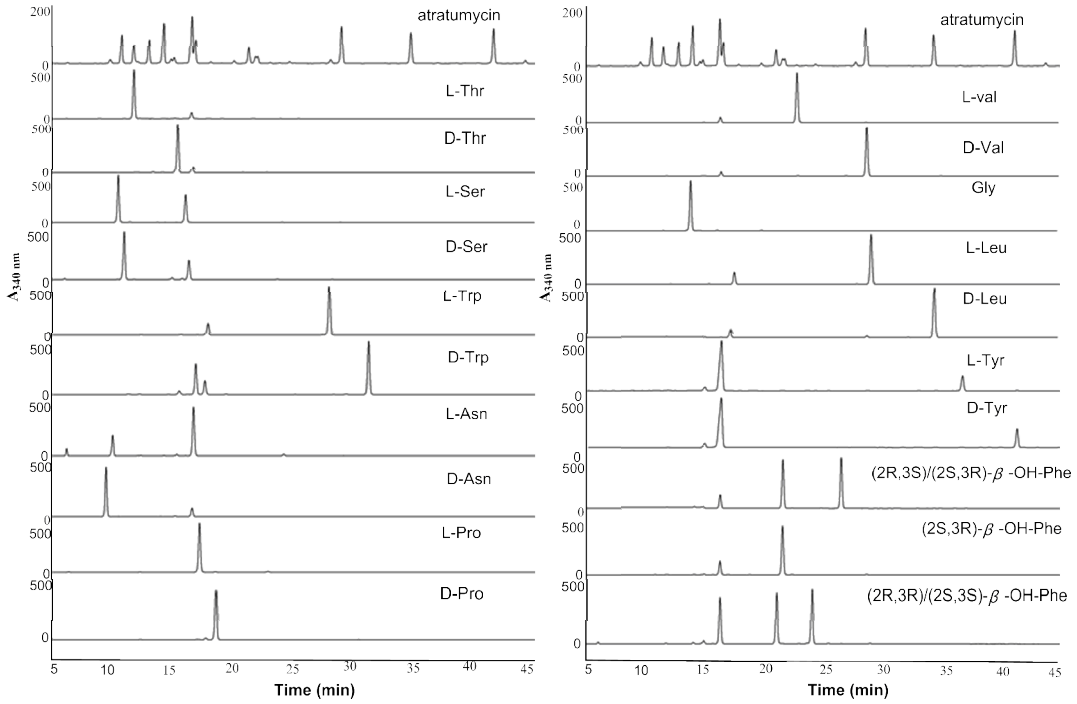

Figure S4. The HPLC analyses of Marfey's reaction products of atratumycin (**1**) and standard amino acids.

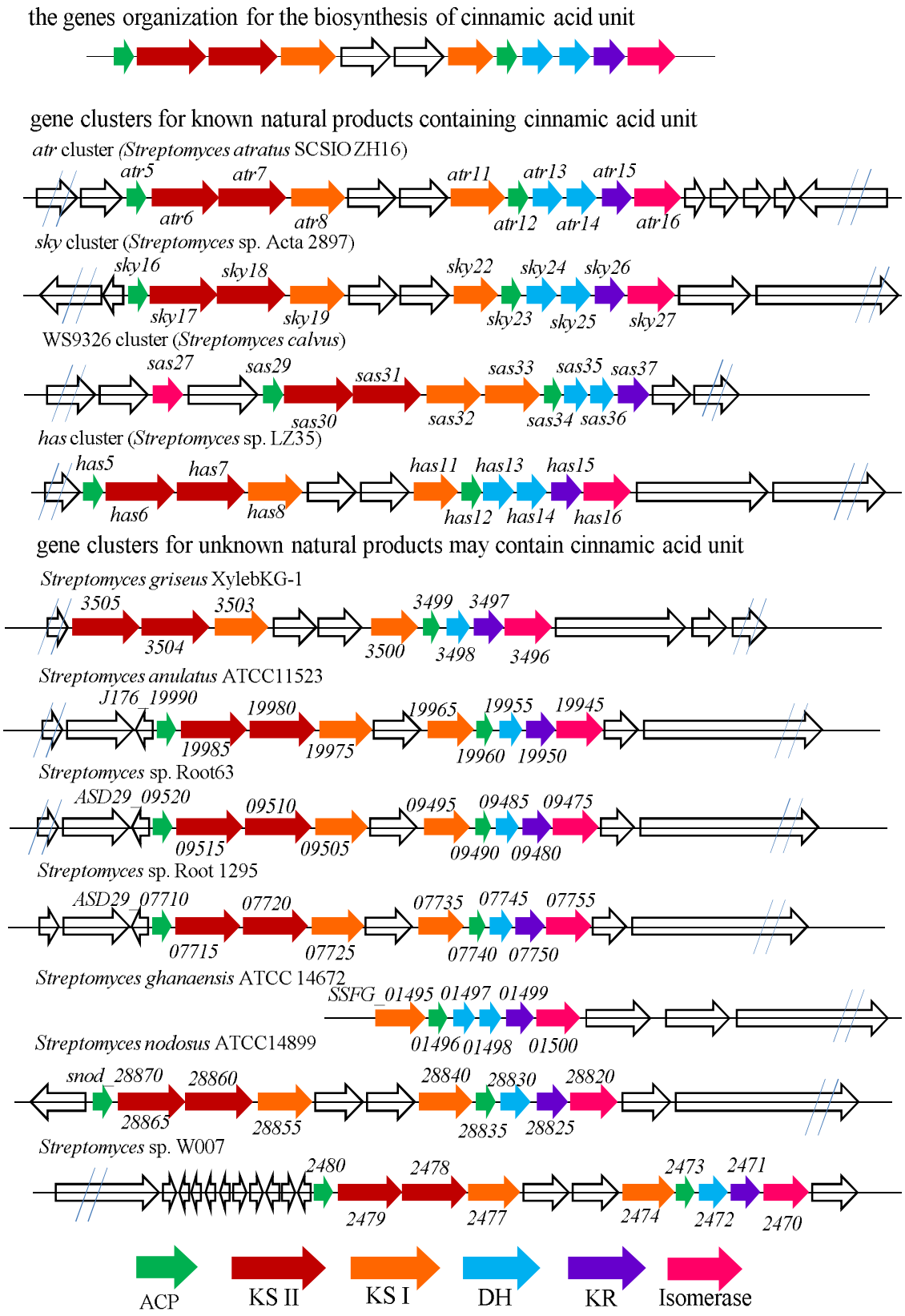

Figure S5. Organization of gene clusters containing cinnamic acid biosynthesis gene cassette. Gene encoding homologues of Atr5 (acyl carrier protein, ACP), Atr6 (β-ketosynthase II, KS II), Atr8 (β-ketosynthase I, KS I), Atr13 (dehydratase, DH), Atr15 (Ketoreductase, KR) and Atr16(isomerase) are colored in green, crimson, orange, navy, purple and magenta, respectively.

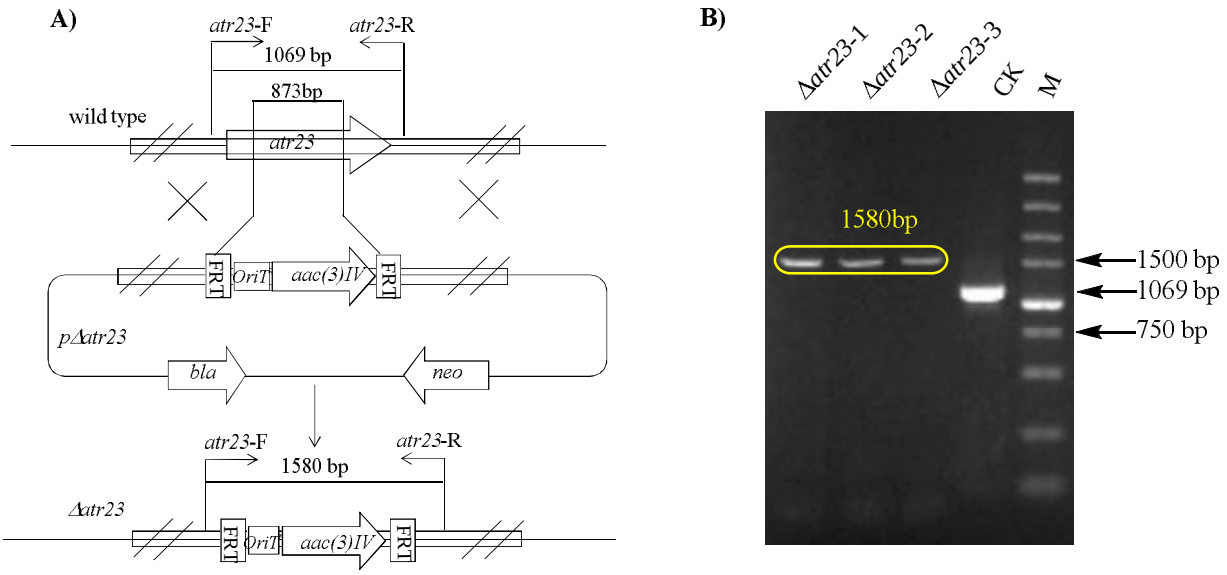

Figure S6. Disruption of *atr23* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *atr23*. B) PCR analysis of the control strain (*S. atratus* SCSIO ZH16NS) and the Δ*atr23* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*atr23(1-3)*: using the genomic DNA of three different colonies of Δ*atr23* mutant as template.

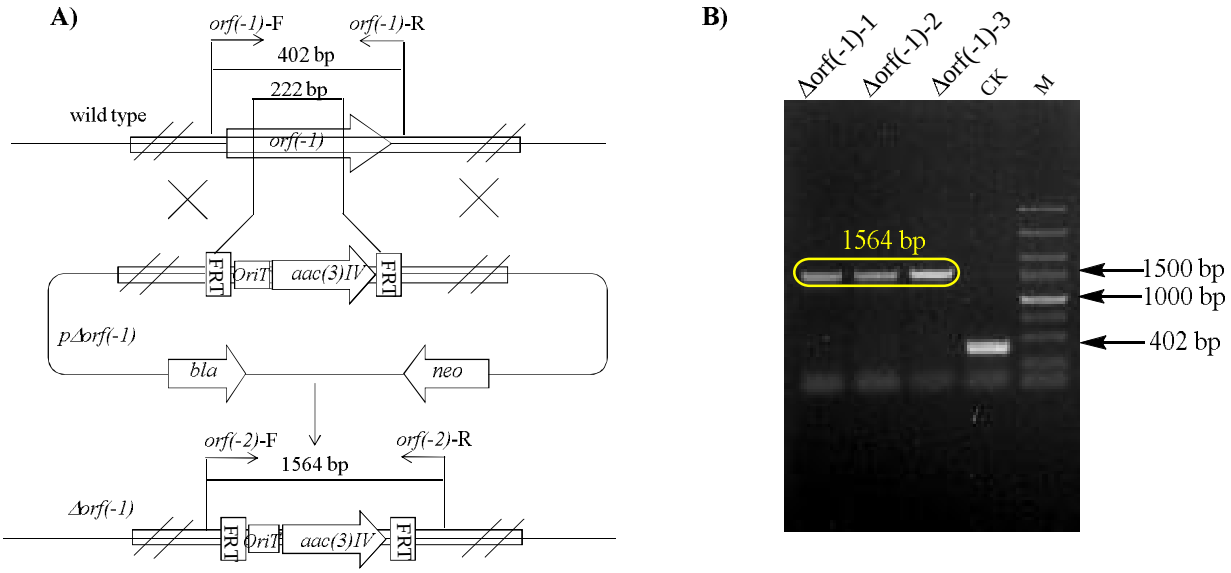

Figure S7. Disruption of *orf(-1)* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *orf(-1)*. B) PCR analysis of the control strain (*S.atrtus* SCSIO ZH16NS) and the Δ*orf(-1)* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*orf(-1) (1-3)*: using the genomic DNA of three different colonies of Δ *orf(-1)* mutant as template.

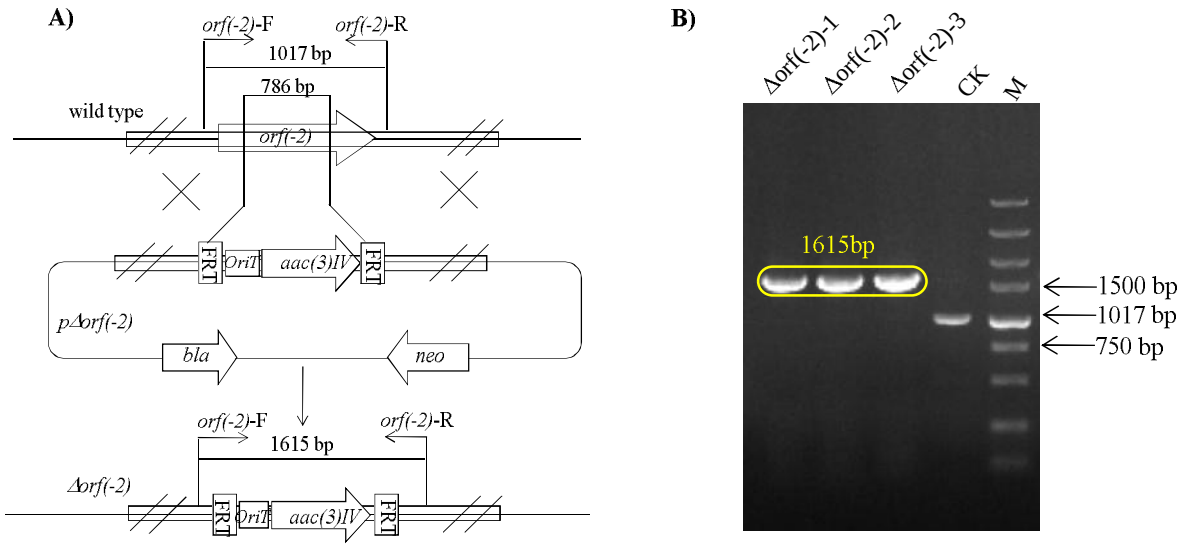

Figure S8. Disruption of *orf(-2)* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *orf(-2)*. B) PCR analysis of the control strain (*S. atratus* SCSIO ZH16NS) and the Δ*orf(-2)* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*orf(-2) (1-3)*: using the genomic DNA of three different colonies of Δ*orf(-3)* mutant as template.

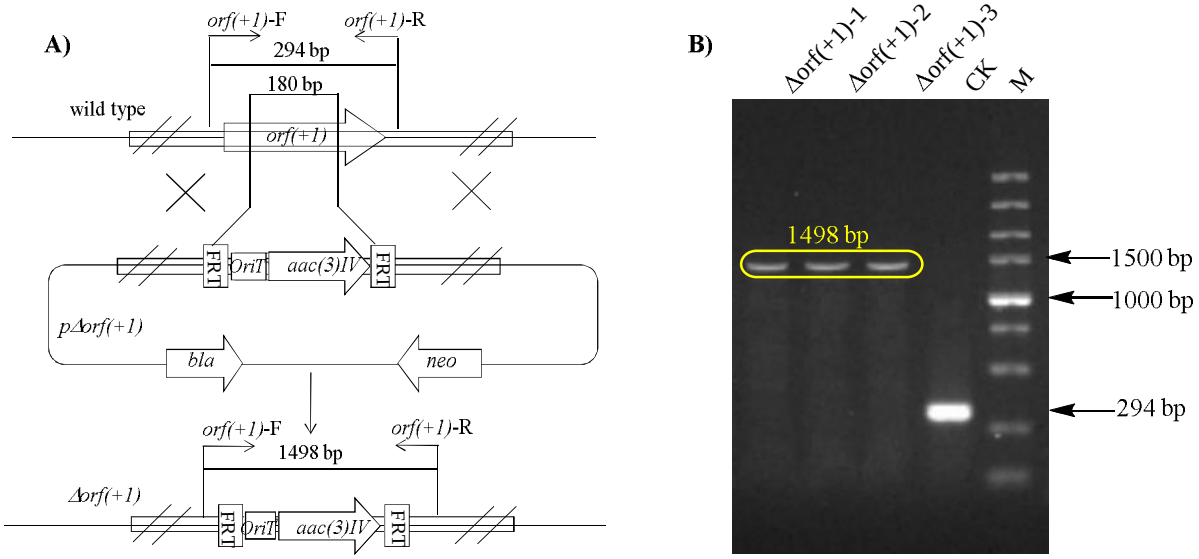

Figure S9. Disruption of *orf(+1)* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *orf(+1)*. B) PCR analysis of the control strain (*S.atrtus* SCSIO ZH16NS) and the Δ*orf(+1)* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*orf(+1) (1-3)*: using the genomic DNA of three different colonies of Δ*orf(+1)* mutant as template.

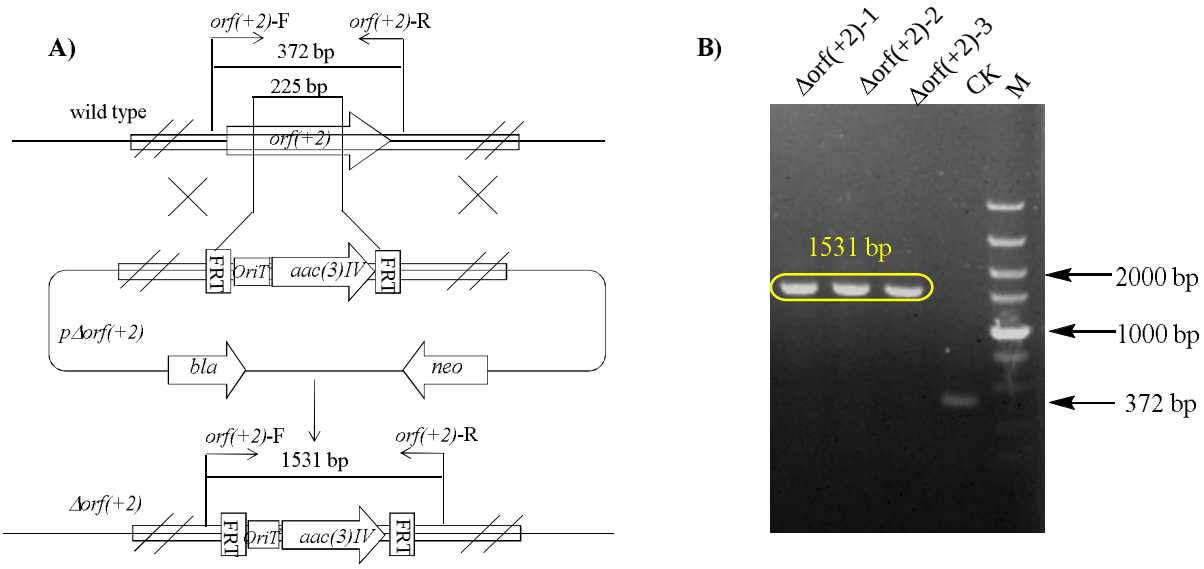

Figure S10. Disruption of *orf(+2)* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *orf(+2)*. B) PCR analysis of the control strain (*S. atratus* SCSIO ZH16NS) and the Δ*orf(+2)* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*orf(+2) (1-3)*:using the genomic DNA of three different colonies of Δ*orf(+2)* mutant as template.

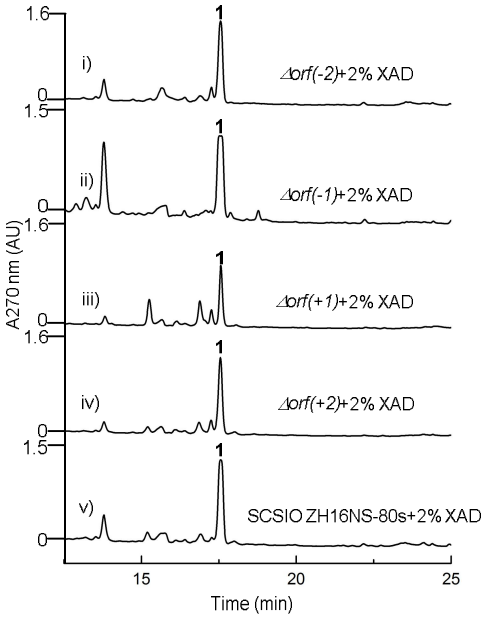

Figure S11. HPLC analyses of fermentation broth of mutants used for the determination of the *atr* cluster boundaries. i) Δ*orf(-2)*+2% XAD, ii) Δ*orf(-1)*+2% XAD, iii) Δ*orf(+1)*+2% XAD, iv) Δ*orf(+2)*+2% XAD, v) SCSIO ZH16NS-80S+2% XAD.

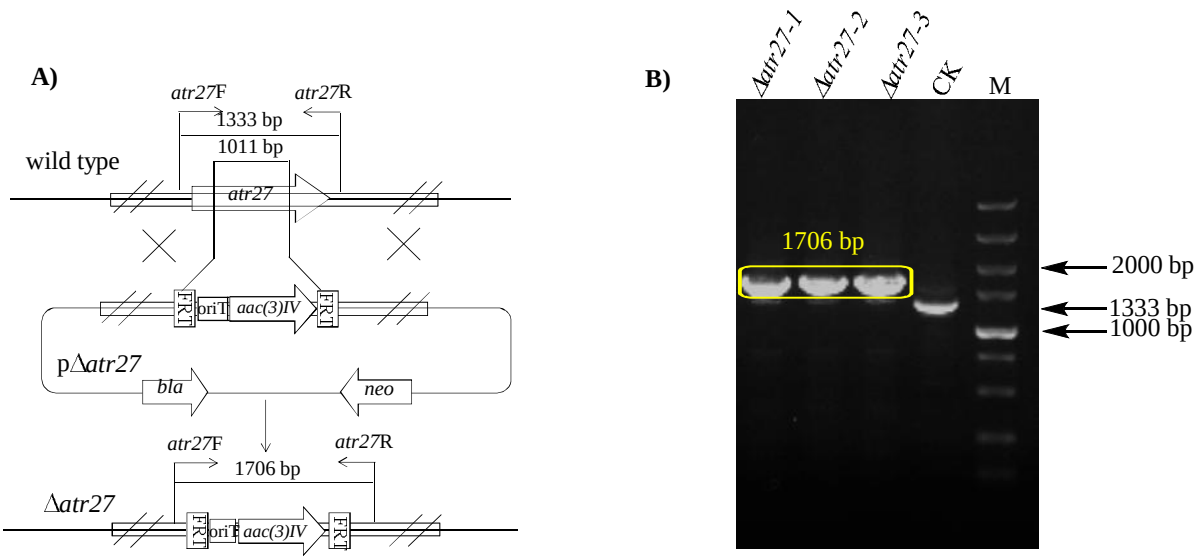

Figure S12. Disruption of *atr27* in *S. atratus* SCSIO ZH16NS-80S via PCR-targeting. A) Schematic representation for disruption of *atr27*. B) PCR analysis of the control strain (*S. atratus* SCSIO ZH16NS) and the Δ*atr27* mutants carried out with the primers listed in **Table S7**. M: DNA molecular ladder; CK: using the genomic DNA of *S. atratus* SCSIO ZH16NS-80S as template, Δ*atr27 (1-3)*: using the genomic DNA of three different colonies of Δ*atr27* mutant as template.

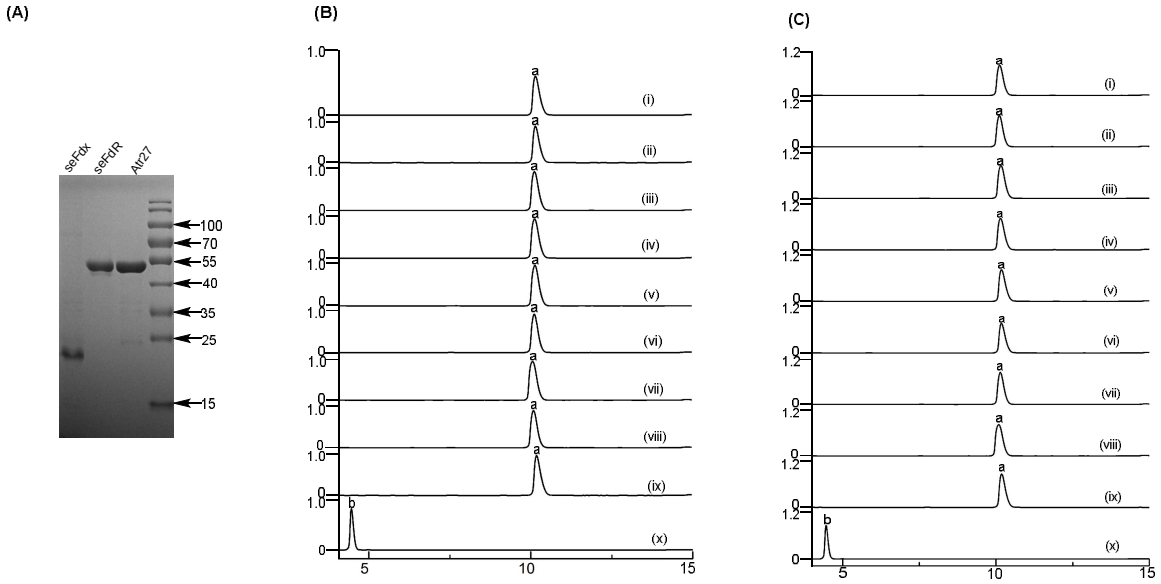

Figure S13. SDS-PAGE analyses of purified proteins and HPLC analyses of Atr27 reaction mixtures. (A) SDS-PAGE analyses of purified proteins, *se*Fdx, *se*FdR and Atr27. (B) HPLC analysis of Atr27 reaction mixtures using *se*Fdx and *se*FdR from Cyanobacterium *Synechococcus elongatus* PCC7942 as surrogate electron transfer system in different buffers with L-Phe as a substrate. (C) HPLC analysis of Atr27 reaction mixture using *sp*Fdx and *sp*FdR from spinach as surrogate electron transfer system in different buffers with L-Phe as a substrate (i) the reaction buffer of Atr27 is 50 mM Tris-HCl (pH 7.6), 100 mM NaCl, 1mM DTT; (ii) the reaction buffer of Atr27 is 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1mM DTT; (iii) the reaction buffer of Atr27 is 50 mM Tris-HCl (pH 8.5), 100 mM NaCl, 1mM DTT; (iv) the reaction buffer of Atr27 is 50 mM PBS (pH 6.8), 100 mM NaCl, 1mM DTT; (v) the reaction buffer of Atr27 is 50 mM PBS (pH 7.4), 100 mM NaCl, 1mM DTT; (vi) the reaction buffer of Atr27 is 50 mM PBS (pH 8.0), 100 mM NaCl, 1mM DTT; (vii) the reaction buffer of Atr27 is 50 mM HEPEs (pH 7.0), 100 mM NaCl, 1mM DTT; (viii) the reaction buffer of Atr27 is 50 mM HEPEs (pH 7.5), 100 mM NaCl, 1 mM DTT; (ix) the standard L-Phe (a); (x) the standard (*2S*, *3S*)-3-OH-Phe (b).

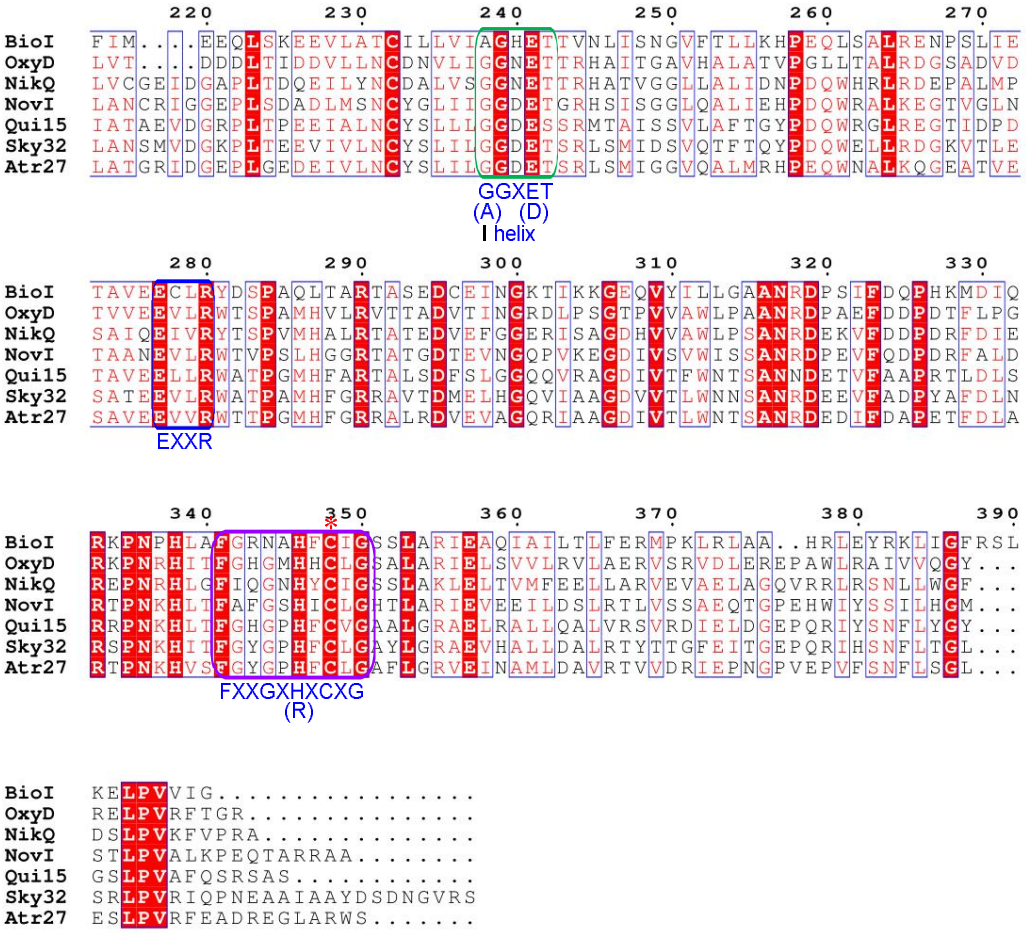

Figure S14. The sequence alignment of Atr27 with other cytochrome P450 involved in the β-hydroxylation of the amino acid in the biosynthesis of secondary metabolites. The I helix motif, the ExxR motif and the heme binding motif were marked in green, blue and purple square, respectively. And the well conserved cysteine residue was marked with red star. The following proteins (with GenBank ID) were used for sequences alignment: BioI (ABS74187.1) from *Bacillus velezensis* FZB42 involved in the biosynthesis of biotin; OxyD (3MGX_B) from *Amycolatopsis balhimycina* involved in the biosynthesis of vancomycin; NikQ(CAC11139.1) from *Streptomyces tendae* involved in the biosynthesis of nikkomycin; NovI (RED69302.1) from *Streptomyces* sp. 67 involved in the biosynthesis of novobiocin; Qui15 (AET98913.1) from *Streptomyces griseovariabilis* subsp. *bandungensis* involved in the biosynthesis of echinomycin; Sky32 (AEA30275.1) from *Streptomyces* sp. Acta 2897 involved in the biosynthesis of skyllamycin; Atr27 from *Streptomyces atratus* SCSIO ZH16 involved in the biosynthesis of atratumycin.

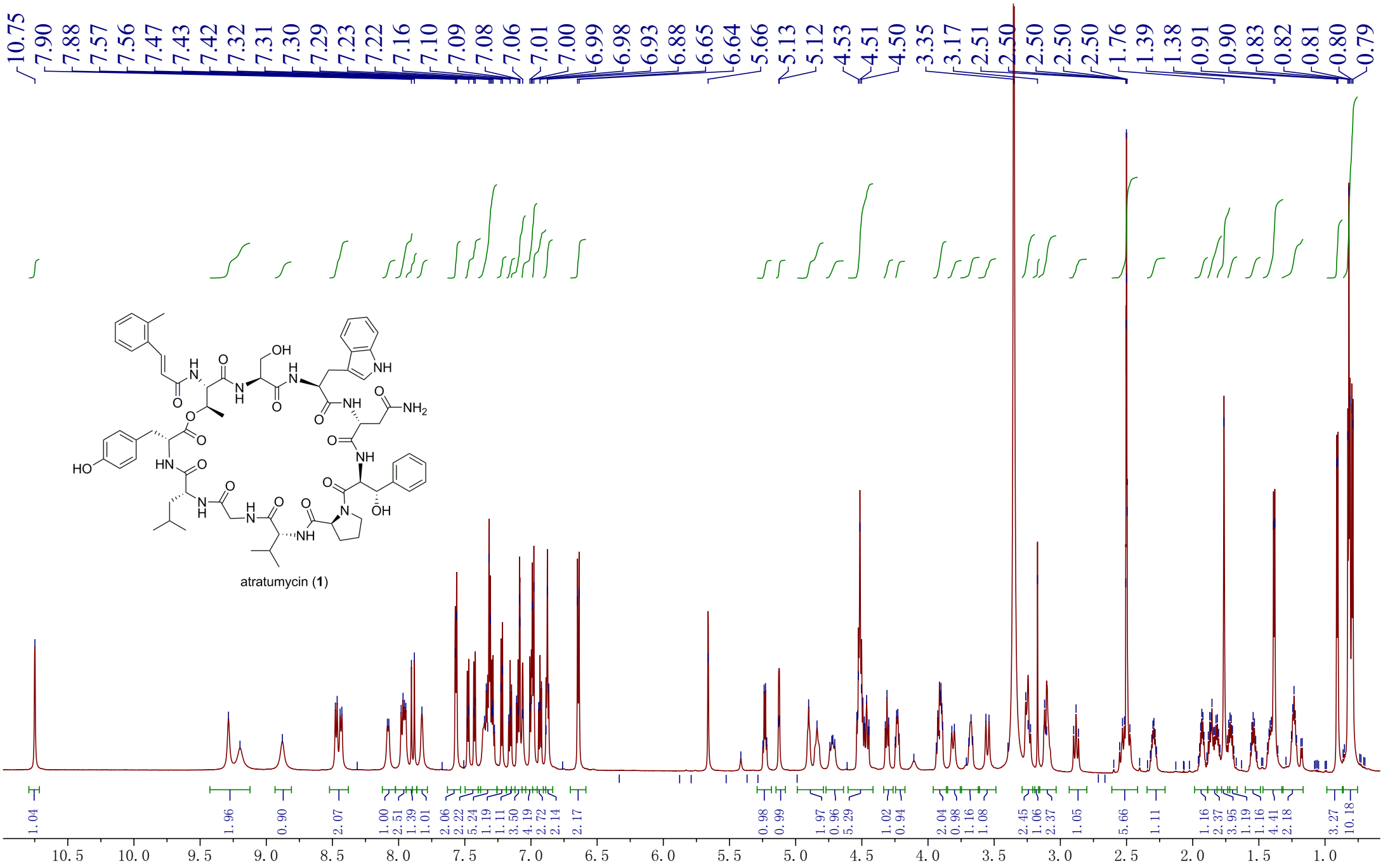

Figure S15. ^{1}H spectrum of atratumycin (**1**) in DMSO-d_6.

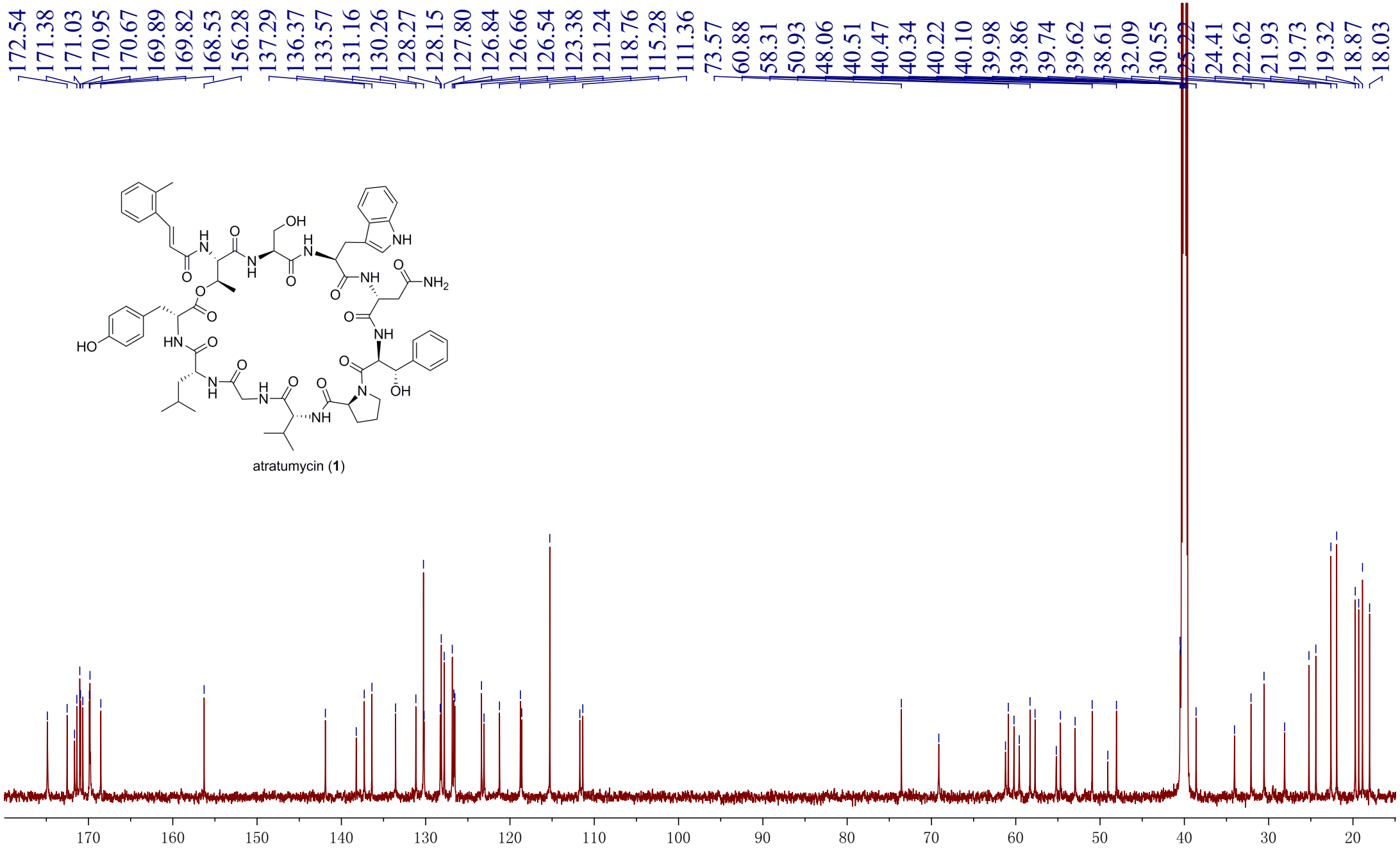

Figure S16. ^{13}C spectrum of atratumycin (**1**) in DMSO-d_6.

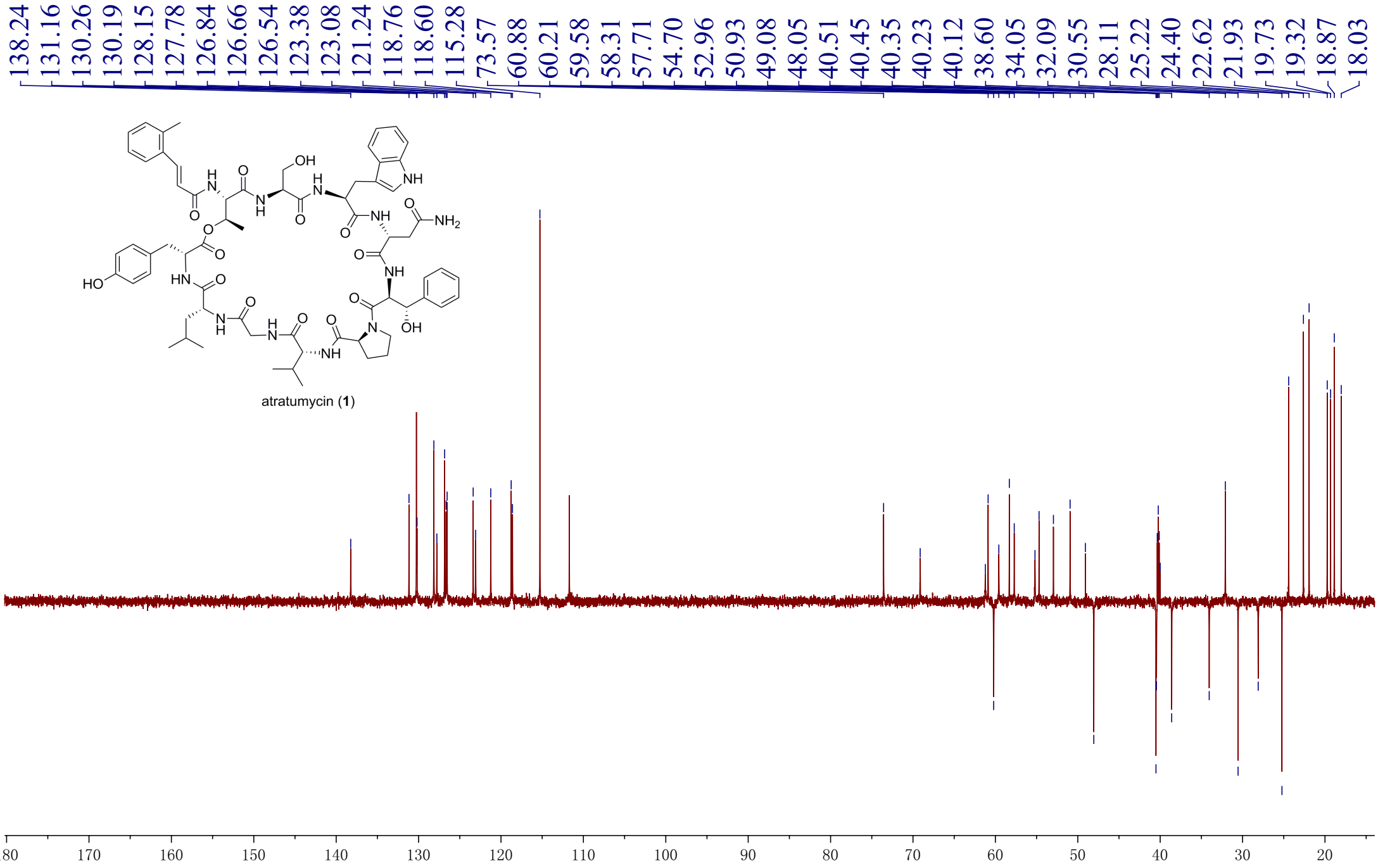

Figure S17. DEPT NMR spectrum of atratumycin (**1**) in DMSO-d_6.

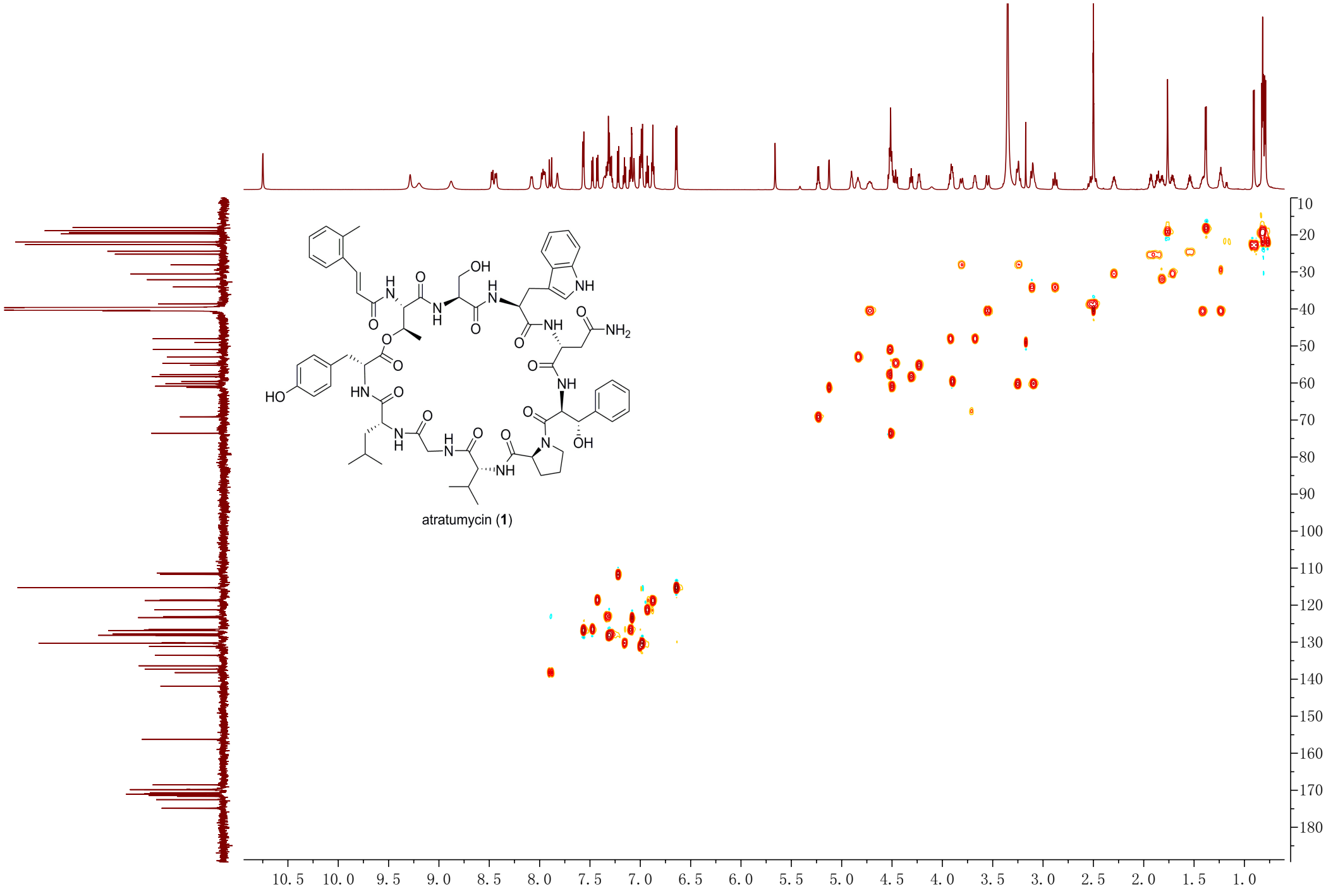

Figure S18. HSQC NMR spectrum of atratumycin (**1**) in DMSO-d_6.

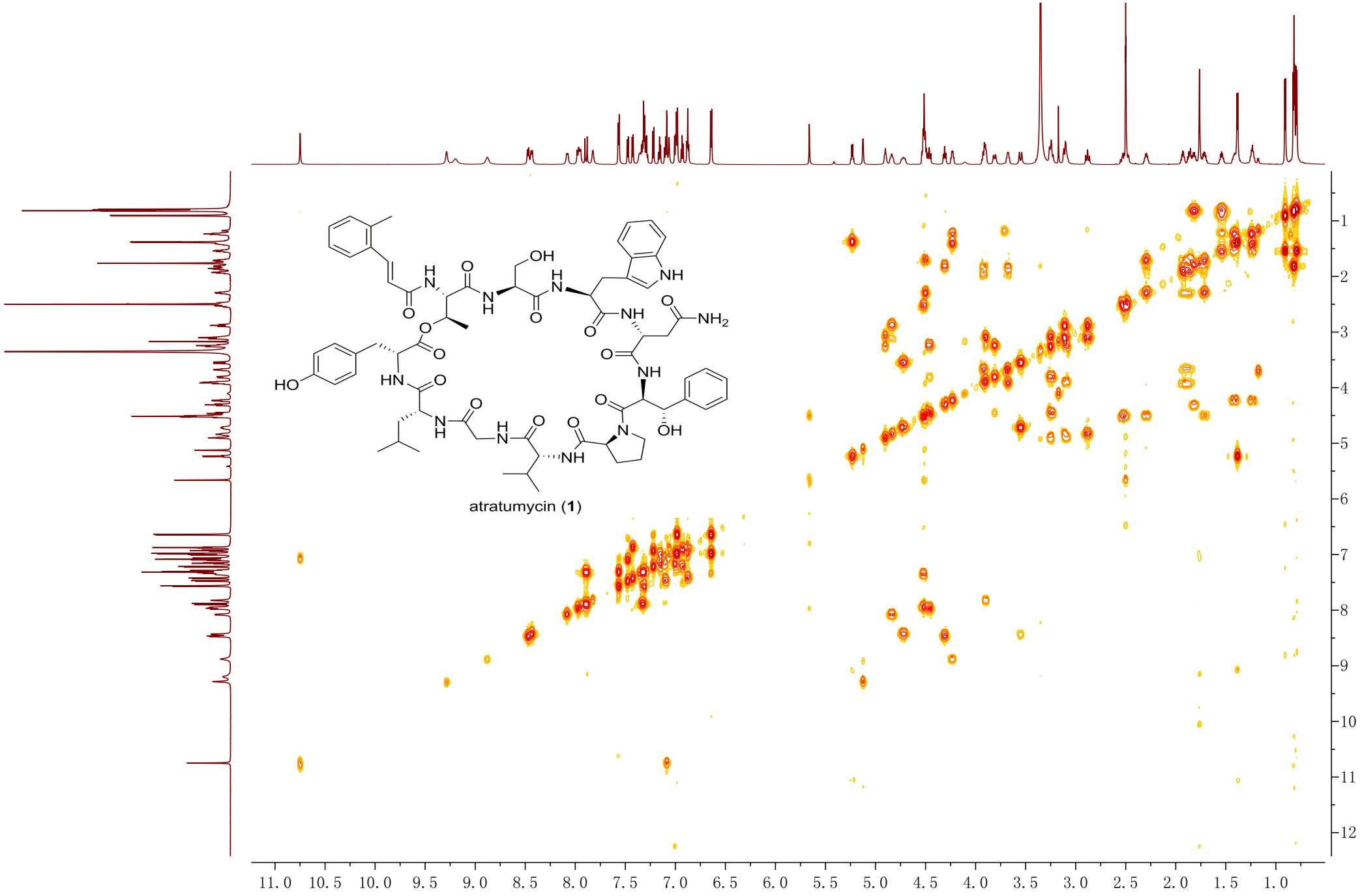

Figure S19. COSY NMR spectrum of atratumycin (**1**) in DMSO-d_6.

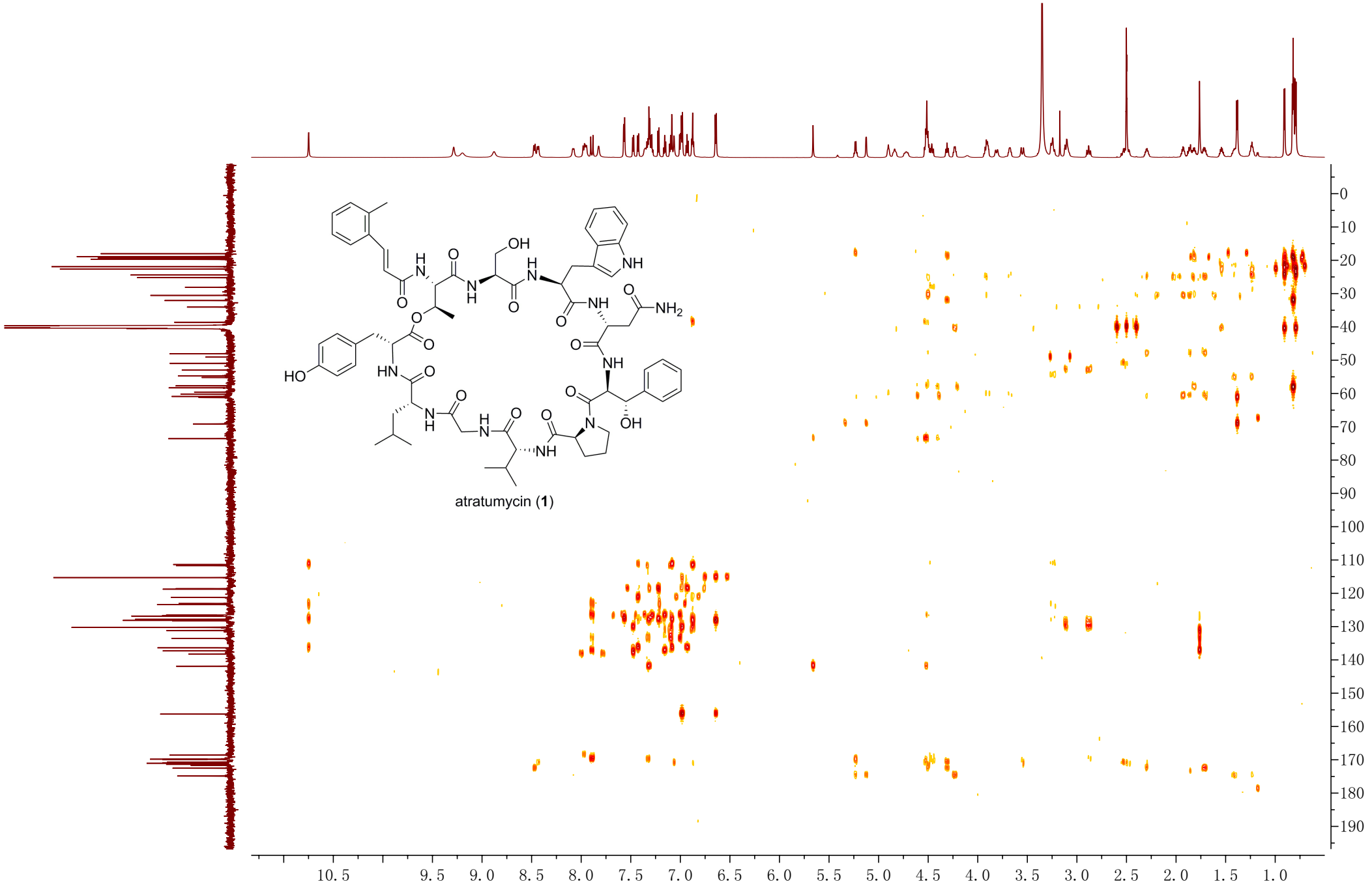

Figure S20. HMBC NMR spectrum of atratumycin (**1**) in DMSO-d_6.

References:

[1] Li, Y.; Zhang C.; Liu C.; Ju J.; Ma J. *Front. Microbiol.* **2018**, *9*, 1269-1277.

[2] Zhang, X.; Zhang, X. F.; Li, H. P.; Wang, L.Y.; Zhang, C.; Xing, X. H.; Bao, C. Y. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 5387-5396.

[3] Ma, J.; Huang, H.; Xie, Y.; Liu, Z.; Zhao, J.; Zhang, C.; Jia. Y.; Zhang. Y.; Zhang, H.; Zhang, T.; Ju, J. *Nat. Commun.* **2017**, *8*, 391.

[4] Lessard J. C. *Methods Enzymol*. **2013**, *533*, 181-189.

[5] Bradford M. M. *Anal. Biochem.* **1976**, *72*, 248-254.

[6] MacNeil, D. J.; Gewain, K. M.; Ruby, C. L.; Dezeny, G.; Gibbons, P. H.; MacNeil, T. *Gene* **1992**, *111*, 61-68.

[7] Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6640-6645.

[8] Gust, B.; Chandra, G.; Jakimowicz, D.; Yuqing, T.; Bruton, C. J.; Chater, K. F. *Adv. Appl. Microbiol*. **2004**, *54*, 107-128.

[9] Paget, M. S. B.; Chamberlin, L.; Atrih, A.; Foster, S. J.; Buttner, M. J. *J. Bacteriol.* **1999**, *181*, 204-211.