Supporting Information for the manuscript:

Combine NMR-based metabolic profiling and genome mining for the accelerated discovery of archangiumide, an allenic macrolide from the myxobacterium *Archangium violaceum* SDU8

Jia-Qi Hu [1,a], Jing-Jing Wang [1,a], Yue-Lan Li [a], Li Zhuo [a], Ai Zhang [b], Haiyan Sui [a], Xiaoju Li [a], Tao Shen [c], Yizhen Yin [a], Zhi-Hong Wu [a], Wei Hu [a], Yue-Zhong Li [a,*], Changsheng Wu [a,*]

[a] State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Shandong University, Qingdao, P.R. China

[b] Fetal Medicine Center, Qingdao Women and Children's Hospital, Qingdao University, Qingdao, P.R. China

[c] Key Lab of Chemical Biology (MOE), School of Pharmaceutical Sciences, Shandong University, Jinan, PR China

* Corresponding authors:

wuchangsheng@sdu.edu.cn; Tel. (+86) 532-58631538; ORCID ID 0000-0003-1310-0089.

lilab@sdu.edu.cn; Tel. (+86) 532 58631539; Fax. (+86) 532 58631539; ORCID ID, 0000-0001-8336-6638.

[1] These authors contributed equally to this work

EXPERIMENTAL SECTION

General experimental procedures.

NMR spectra were recorded in CD_3OD or DMSO-d_6 on a Bruker AVNEO 600 MHz or 500 MHz calibrated on the residual of deuterated solvent. The HR-Q-ToF-ESI-MS analysis was performed on a rapid separation liquid chromatography system (Dionex, UltiMate3000, UHPLC) coupled with an ESI-Q-ToF mass spectrometer (Bruker Daltonics, Impact HD). HPLC analysis (210 nm, 254 nm, 280 nm and 365 nm) was performed with an Agilent 1260 series HPLC apparatus (Agilent technologies Inc., Santa Clara, CA, U.S.A.), using a 250 × 4.6 mm Luna 5 µm C_{18} (2) 100 Å column equipped with a guard column containing C_{18} 4 × 3 mm cartridges (Phenomenex Inc., Torrance, CA, U.S.A.). Semi-preparative HPLC separation was performed on a reversed-phase column (Phenomenex Luna 5 µm C_{18} (2) 100 Å column, 250 × 10 mm). HR-Q-ToF-ESI-MS instrument was equipped with a C_{18} column (Thermo Fisher Scientific, C_{18}, 250 × 4.6 mm, 5 µm). TLC analysis was developed by precoated silica gel GF$_{254}$ (Qingdao Haiyang Chemical Co., Ltd., Qingdao, China). The expansion of the compound was conducted by using a solvent system of $CHCl_2$/MeOH (18:1), and then visualized with iodine and anisaldehyde/sulfuric acid, respectively. According to the experiments, all organic solvents and chemicals were of analytical or HPLC grade.

Isolation and identification of the myxobacterium *Archangium violaceum* SDU8

The bacterial strain, designated SDU8, was isolated from an upper soil sample, which was collected from Qingdao campus garden of Shandong University, Shandong Province, P.R. China. To isolate myxobacteria, the soil sample was air-dried for one month before enrichment procedure to reduce the contamination of other bacteria and fungi. Air-dried soil was spread over cell mats of living *Escherichia coli* smeared on WCX basal medium (0.1% $CaCl_2 \cdot 2H_2O$, 25 µg/mL cycloheximide and 1.5% agar, w/v), which is a classical technique for the enrichment of myxobacteria.[1] The isolate was transferred from the colony edge onto fresh VY/2 medium agar plates (DSMZ medium 9) several times for purification. The purified strain SDU8 had a typical morphological feature of genus *Archangium*, for example, the swarm colonies exhibited branched

radial veins with reddish to violet pigmentation. The full-length 16S rRNA gene sequence (1536 bp; under the GenBank accession number MT860702) was retrieved from the genome of strain SDU8. Phylogenetic analysis based on the 16S rRNA gene sequences showed that SDU8 was closest to the strains *Archangium violaceum* DSM 14727^{T} and *Archangium gephyra* DSM 2260^{T} with similarity of 98.9% and 98.8% (Figure S14), respectively. Consequently, we concluded that SDU8 was affiliated with the species *Archangium violaceum.*

Culturing conditions and fermentation.

The myxobacterium *Archangium violaceum* SDU8 had a characteristic colony status when grown on the VY/2 medium containing (L^{-1}) 5 g baker's yeast, 1 g CaCl$_2$, 0.5 mg VB$_{12}$, 15 g agar, and 1.97 g 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), at pH 7.2. The strain was first activated on VY/2 agar plate for 3 days at 30 °C. Mycelium of the colony was inoculated into 100 mL of sterile liquid VY/2 medium in Erlenmeyer flasks (300 mL), which were continuously shaken at 200 rpm, 30 °C for 3 days. For fermentation, the 100 mL of seed culture was equally divided into three parts and then transferred into fresh liquid VY/2 medium in Erlenmeyer flasks (300 mL), and then continuously shook at 200 rpm, 30 °C for 7 days.

Metabolomics

To compare different extraction methods, six flasks of SDU8 were fermented for metabolomics research. After fermentation, all the cultures were centrifuged to remove the pellets. Three portions of the supernatants were adsorbed with 2 g of HP-20 resin (Shanghai Yuanye Biological Technology Co., Ltd., Shanghai, China) overnight. After filtration to remove the supernatant, the resin was rinsed with 5% methanol in water, and the adsorbed compounds were recovered with 100% methanol, followed by evaporation under vacuum at 40 °C. The recovered crude extracts were lyophilized and then re-dissolved in 0.5 mL of 50% CD$_3$OD in D$_2$O for NMR analysis. The other three portions of SDU8 were extracted with 100 mL of ethyl acetate (EtOAc) in funnel. The

following methods for compounds recovery and NMR analysis were same with those for HP-20 extraction.

NMR measurement and data analysis

Suppression methods were applied to remove the undesired signal caused by residual water. The ^{1}H NMR spectrum was acquired by Bruker AVNEO 600 MHz (Bruker technologies Inc., Karlsruhe, Germany) at 25 °C. Manual shimming and tuning generally took over 20 min on the first sample, using the corresponding deuterated reagent as internal lock. Each ^{1}H NMR spectrum consisted of 16 scans using the following parameters: 0.18 Hz/point, spectral width (SW) = 20 ppm, free induction decays resolution = 0.36 Hz, dwell time (DW) = 42 μs and pre-scan delay (DE) = 22 μs. ^{13}C NMR spectrum was measured with 256 scans using the following parameters: 0.54 Hz/point, spectral width (SW) = 237 ppm, free induction decays resolution = 1.09 Hz, dwell time (DW) = 14 μs and pre-scan delay (DE) = 18 μs. COSY required 2 scans, and obtained 16 dummy scans. The other parameters: spectral width (SW) = 20 ppm, free induction decays resolution = 12 Hz (F2) / 47 Hz (F1), dwell time (DW) = 42 μs, pre-scan delay (DE) = 22 μs and decimation rate of digital filter = 1680. HSQC achieved 24 scans. The other parameters: spectral width (SW) = 8 (F2) / 220 (F1) ppm, free induction decays resolution = 10 Hz (F2) / 130 Hz (F1), dwell time (DW) = 95 μs, pre-scan delay (DE) = 6.5 μs and decimation rate of digital filter = 3800. HMBC achieved 32 scans. The other parameters: spectral width (SW) = 20 (F2) / 240 (F1) ppm, free induction decays resolution = 6 Hz (F2) / 71 Hz (F1), dwell time (DW) = 42 μs, pre-scan delay (DE) = 6.5 μs and decimation rate of digital filter = 1680. NOESY required 32 scans, and obtained 4 dummy scans. The other parameters: spectral width (SW) = 20 ppm, free induction decays resolution = 0.38 Hz, dwell time (DW) = 40 μs, pre-scan delay (DE) = 6.5 μs and decimation rate of digital filter = 1600. The one-dimensional and two-dimensional NMR spectra were manually phased and baseline corrected by MestReNova (version 6.1.0, Mestrelab Research), and calibrated on the residual of the deuterated CD_3OD (3.31 ppm), or DMSO-d_6 (2.50 ppm).

Compound isolation

To purify the target compound revealed by metabolomics study, the ethyl acetate extraction of the 20 L of SDU8 fermentation was concentrated by rotary evaporation at 40 °C under vacuum. The resulting residue (1.42 g) dissolved in MeOH was applied to Sephadex LH-20 (GE Healthcare Bio-Sciences Inc., Uppsala, Sweden) with CH_2Cl_2−MeOH (2:1) to afford 11 fractions (FrA−FrK). All the obtained fractions were subjected to ^1H NMR and TLC analysis, the fraction H and I containing the signals of track were combined. The resultant mixture (0.977 g) was further separated by reverse phase preparative medium-pressure liquid chromatography (RP−MPLC) on an irregular C_{18} column connected to a Buchi labortechnik AG chromatography system (Buchi, Flawil, Japan) to generate 98 fractions (Fr1−Fr98), eluting with solvent system [solvent A: H_2O; solvent B: methanol; gradient: 10% B to 100% B]. Subsequently, four fractions (Fr31−Fr34) containing targeted compound based on TLC analysis were combined and further purified by semi-preparative RP-HPLC [column: Phenomenex Luna 5 μm C_{18} (2) 100 Å, 250 × 10 mm; solvent A: H_2O; solvent B: ACN; gradient: 40% B to 60% B in 30 min; flow rate 1.8 mL/min; 30 °C; UV detection at 210 nm], to afford **1** (20.5 mg).

Archangiumide (1): white powder, needle crystals (MeOH); $[\alpha]^{25}_D$ +37.0 (c 0.05, MeOH); $[\theta]^{25}$(nm) (MeOH) −66153 (201), 41177 (228), -32576 (252); IR v_{max} 3375, 2928, 1942, 1717, 1442, 1263, 1118, 1076, 1040, 1027, 962, 881 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (125 MHz, DMSO-d_6) data were showed in Table 1; HRMS (ESI) *m/z*: [M + H]$^+$ Calcd for $C_{19}H_{27}O_6$ 351.1802; Found 351.1797.

Bioassays

Antimicrobial activity: agar diffusion assay followed our previous method.[2] Single colony of Gram-positive bacteria *Staphylococcus aureus* and *Bacillus subtilis,* gram-negative bacterium *Escherichia coli,* and fungus *Candida albicans* were picked and inoculated into 8 mL of LB liquid medium for overnight growth at 37 °C, respectively. Next day, the cultures were then 100-fold diluted to 1×10^7 CFU/mL with fresh LB liquid medium, to grow at 37 °C until OD_{600} 0.4~0.6.

Archangiumide was dissolved in methanol (2 mg/mL), and 20 μL of the solution was applied on a paper disk. The disks were then placed onto an agar plate containing a soft agar overlay with the microorganisms. Kanamycin (for *S. aureus* and *B. subtilis*) and apramycin (for *C. albicans* and *E. coli*) at a concentration of 1 mg/mL were used as positive controls, and the solvent methanol as the negative control. After incubation at 37 °C for 18 h, archangiumide did not give any growth inhibition zones (in mm) against the tested strains.

Anti-cancer assay: archangiumide was assayed against the human breast cancer cell line MCF-7, liver carcinoma cell line HepG2 and hepatocyte cell line HL-7702 using the MTT colorimetric assay.[3] Fresh cell suspensions in the logarithmic growth phase were seeded in a 96-well microtiter plate. The newly attached cells were treated with varying concentrations of drugs by dilution from 100 μg/mL to 0.1 μg/mL (dissolved in DMSO). After 48 h of incubation with the test compounds, the cultures were treated with MTT (2 mg/mL) solution prepared in the growth medium. The cells were incubated for an additional 4 h at 37°C until a purple precipitate was visible. The purple formazan crystals were dissolved with DMSO, and the absorbance of each sample was measured at 492 nm using the microplate reader. For all the MTT assay experiments, triplicates were conducted, and taxol was used as the positive control. IC_{50} values were calculated using GraphPad Prism software (Version 6.01), and $IC_{50} > 100$ μg/mL was regarded as inactive.

DPPH radical-scavenging assay: the method for α, α-diphenyl-β-picrylhydrazyl (DPPH) free radical scavenging was performed according to the protocol of manufacturer (kit BC4750, Solarbio). In brief, 0.4 mg/mL of archangiumide was reacted with DPPH for 30 minutes at room termperature in the dark. The absorbance (A) of the resulting solution was measured at 515 nm with a UV spectrophotometer. Vitamin C (0.4 mg/mL) was used as the positive control. DPPH radical-scavenging rate was calculated by the formula $[D\% = (A_{control} - A_{sample})/A_{control}]$,[4] whereby an ignorable rate of < 5% was observed for archangiumide. The experiment was conducted in duplicate.

Anti-inflammatory activity：RAW 264.7 murine macrophages (8.0×10^4 cells/well) were seeded in a 96-well plate and treated with 1 μg/mL of LPS, in the absence or presence of the tested

compound with a concentration gradient ranging from 0.1~50 µM. After 24 h's incubation, 100 µL of supernatant was removed to a new 96-well plate and mixed with 100 µL of Griess reagent (0.1% naphthylethylenediamine and 1% sulfanilamide in 5% H_3PO_4 solution) for 15 min. Absorbance was measured at 570 nm on the Model 680 plate reader (Bio-rad). Nitrite concentration was calculated from a $NaNO_2$ standard curve. Didox (100 µM) was used as a positive control. The $IC_{50} > 50$ µM was regarded as non-active.

X-ray crystal structure determination of archangiumide (1).

The white needle crystals of archangiumide (**1**) were obtained from MeOH solvent by evaporating in a glass vial under ambient condition. The suitable crystals were selected under the microscope and were mounted at room temperature in vaseline to protect crystals. Single crystal X-ray diffraction data were collected on a Rigaku Oxford XtaLAB Synergy-DW diffractometer equipped with *mono*-chromated Cu Kα radiation at 100 K. The detector was Hypix 6000HE and crystal-to-detector distances were maintained as 50 mm during the measurement. The data were reduced and integrated with CrysAlisPro software. The structure of archangiumide was established by direct methods [5] and refined on F^2 through full-matrix least squares fitting using OLEX2, on the basis of anisotropic displacement parameters with the final least-squares refinement indexes ≥ 2σ. A SCALE program was used for numerical absorption corrections. Molecular graphics were carried out by using OLEX2 and Diamond software. The crystallographic data for archangiumide was summarized in Table S2. This structure has been deposited in CCDC database under the accession number 2058107.

Genome sequencing, assembly, and annotation.

Genomic DNA of *Archangium violaceum* SDU8 was extracted by the whole genome DNA sequencing kit (Oxford Nanopore Technologies Inc., Oxford, U.K.). Genome sequencing was conducted by Benagene (Wuhan, China), by joining the strength of Illumina and Nanopore technologies. Illumina sequencing produced a total of 2.64 G of raw data, and filtered to generate

2.63 G of clean data. After removing connector, short fragments and low-quality data, a total of 3,120,498,886 bp clean data for assembly was obtained. This sequencing was based on the pass reads (Q>7) with extracting the longest sequence of reads, and the total valid data was approximately 1 G. Unicycler (0.4.8) software (https://github.com/rrwick/Unicycler) was used to assemble the filtered reads. Genes were annotated based on the BLAST against databases COG (https://www.ncbi.nlm.nih.gov/COG/), KEGG (https://www.kegg.jp/kegg/), Refseq (https://www.ncbi.nlm.nih.gov/refseq/), Uniprot (https://www.uniprot.org/). The genome has been deposited at GenBank under the accession NO. CP069396.

HR-QToF-ESI-MS analysis.

The HR-Q-ToF-ESI-MS analysis was performed on an UHPLC system (Dionex, UltiMate3000, UHPLC) coupled with an ESI-Q-ToF mass spectrometer (Bruker Daltonics, Impact HD). The chromatographic separation was done using a Luna C_{18} HPLC 5 μm column 250 × 4.6 mm (Phenomenex) at a flow rate of 0.8 mL/min and a column temperature of 30 °C. 3 μL of sample was eluted with a gradient of solvent A (water) and B (methanol). The initial percentage of B was 5%, which was linearly increased to 100% in 30 min, followed by a 10 min isocratic period, and then re-equilibrated with original conditions in 6 min. ESI-positive mode was used for scanning. Nitrogen was used as drying and nebulizing gas. The gas flow was set at 6 L/min at 200 °C, and the nebulizer pressure was 0.0 bar. The MS data was acquired over mass range of *m/z* 50−1500. Water and methanol were LC-MS grade (Optima, Fisher Scientific, NJ, U.S.A.). The data was analyzed by DataAnalysis software (Bruker Daltonics), and Quadratic + HPC mode was internal calibration with sodium formate.

[13]C isotope labeling of archangiumide

0.4 g of [1-^{13}C]-sodium acetate or [2-^{13}C]-sodium acetate was added to one liter of SDU8 culture after 36 h of incubation. After an additional 5.5 days fermentation, the culture was harvested by extraction with ethyl acetate. The extract was dried under vacuum, and subjected to Sephadex LH-20 column eluted with methanol to resolve 10 fractions. The fraction containing archangiumide

was separated by semi-preparative RP-HPLC to afford 1.80 mg and 6.6 mg from [1-^{13}C]-sodium acetate and [2-^{13}C]-sodium acetate-labelled experiments, respectively. Finally, the labeled archangiumide was subjected to ^{13}C NMR analysis (methanol-d_4, 150 MHz).

Table S1. Annotations for archangiumide biosynthetic gene cluster *arc* (GeneBank Accession NO. MW488041).

ORF	Gene size (bp)	Gene name	Proposed function	NCBI similarity	ID%
1	1167	*arcL*	oxidation	N-acetyl-gamma-glutamyl-phosphate reductase [*Vitiosangium* sp. GDMCC 1.1324]; WP_108065195.1	91%
2	1107	*arcF*	loading ACP	2-oxo acid dehydrogenase subunit E2 [*Chthonomonas calidirosea*]; WP_075163492.1	40%
3	12819	*arcA*	type I PKS	polyketide synthase [*Pseudomonas batumici*]; AKQ22649.1	45%
4	19323	*arcB*	type I PKS	polyketide synthase [*Pelosinus fermentans*]; WP_052697319.1	46%
5	1143	*arcM*	oxidation	LLM class flavin-dependent oxidoreductase [*Okeania* sp. SIO3B3]; NEP84979.1	65%
6	7110	*arcC*	type I PKS	polyketide synthase [*Rhizobium* sp. BK315]; WP_165915351.1	43%
7	1263	*arcG*	beta-branching methyl biosynthesis	HMG-CoA synthase family protein [Bacillus wiedmannii]; WP_098202504.1	72%
8	258	*arcH*	beta-branching methyl biosynthesis	acyl carrier protein [*Gammaproteobacteria* bacterium]; RKZ52898.1	67%
9	1284	*arcI*	beta-branching methyl biosynthesis	beta-ketoacyl-ACP synthase [Gammaproteobacteria bacterium]; RKZ52897.1	58%
10	777	*arcJ*	beta-branching methyl biosynthesis	enoyl-CoA hydratase [Gammaproteobacteria bacterium]; RKZ52895.1	54%
11	747	*arcK*	beta-branching methyl biosynthesis	enoyl-CoA hydratase/isomerase family protein [*Methylocaldum* sp. 0917]; WP_133719094.1	59%
12	984	*arcD*	*trans*-AT	acyl hydrolase [*Thioploca* sp.]; HEC83863.1	42%
13	2277	*arcE*	*trans*-AT	[acyl-carrier-protein] S-malonyltransferase [*Beggiatoa* sp. 4572_84]; OQY57292.1	60%
14	777	*arcO*	Unknown function	type I methionyl aminopeptidase [*Corallococcus llansteffanensis*]; WP_120641935.1	82%
15	1221	*arcN*	oxidation	cytochrome P450 [*Gynuella sunshinyii*]; WP_044617477.1	43%

Table S2. X-ray crystallographic data for archangiumide (**1**)

Empirical formula	$C_{19}H_{26}O_6$
Formula weight	350.41
Temperature/K	100.0
Crystal system	monoclinic
Space group	$P2_1$
a/Å	9.07240(10)
b/Å	6.74730(10)
c/Å	16.2178(2)
α/°	90
β/°	91.2770(10)
γ/°	90
Volume/Å^3	992.51(2)
Z	2
ρ_{calc}g/cm^3	1.233
μ/mm^{-1}	0.777
F(000)	396.0
Crystal size/mm^3	0.13 × 0.12 × 0.11
Radiation	Cu Kα (λ = 1.54184)
2Θ range for data collection/°	5.45 to 152.976
Index ranges	-10 ≤ h ≤ 11, -8 ≤ k ≤ 8, -20 ≤ l ≤ 20
Reflections collected	18619
Independent reflections	3999 [R_{int} = 0.0458, R_{sigma} = 0.0290]
Data/restraints/parameters	3999/6/257
Goodness-of-fit on F^2	1.136
Final R indexes [I>=2σ (I)]	R_1 = 0.0304, wR_2 = 0.0845
Final R indexes [all data]	R_1 = 0.0330, wR_2 = 0.0879
Largest diff. peak/hole / e Å^{-3}	0.24/-0.22
Flack parameter	0.01(8)

Table S3. Prediction of stereochemistry of the reduced carbon in archangiumide (**1**). The presence of a diagnostic Asp residue within LDD motif correlates with B-type ketoreduction (3*R*), whereas the absence of the Asp residue correlates with A-type ketoreduction (3*S*). The KR-predicted stereochemistry at positions C-3, C-6, and C-17 of archangiumide matched with the crystallography-confirmed structure in Figure 2. The sequence alignment of KR domains was present in Figure S10, whereby the second KR domain in the module 8 of ArcB was thought to be inactive due to the defect in the NADP(H) binding motif GGxGxxG.

KR domain	Conserved Asp motif	KR type (position in acetate unit)	Corresponding position of -OH	Predicted stereochemistry/geometry in archangiumide	Observed stereochemistry/geometry in archangiumide
KR1	IRD	B-type (3*R*)	17	*R*	*R*
KR3+(DH)	LRD	B-type (3*R*)	12	12*E*-double bond	Allene group
KR6+(DH)	IQD	B-type (3*R*)	6	6*E*-double bond	6*E*-double bond
KR7+(DH)	IQD	B-type (3*R*)	4	4*E*-double bond	epoxy
KR8a	IQD	B-type (3*R*)	3	*S*	*S*
KR8b	Inactive				

Figure S1. NMR (methanol-d_4, 600 MHz) metabolic comparison of different extraction methods by using HP-20 and EtOAc.
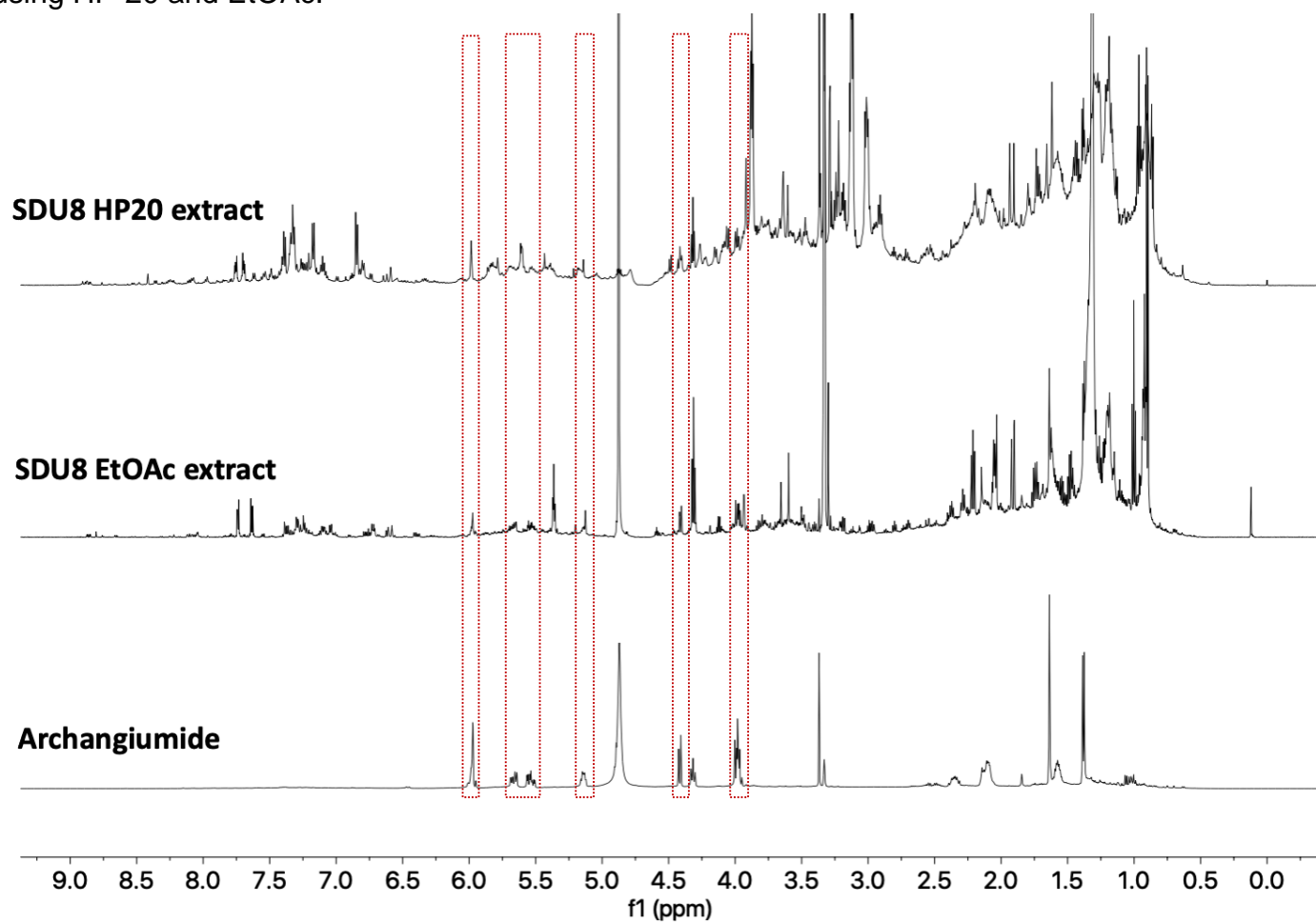
Figure S2. ^{1}H-^{1}H COSY (methanol-d_4, 600 MHz) metabolic profiling of SDU8 EtOAc metabolome.
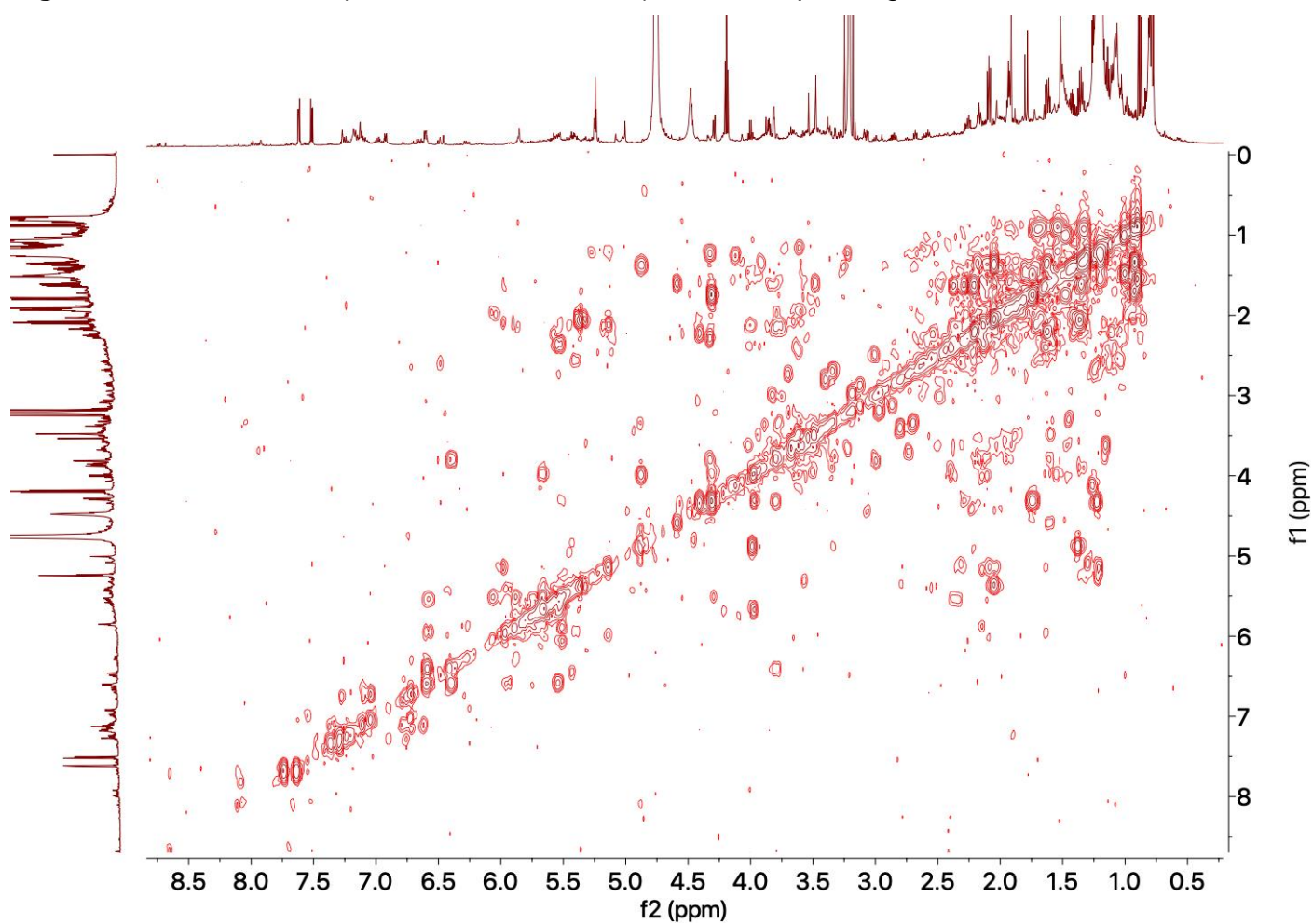
Figure S3. ^{1}H-^{13}C HSQC (methanol-d_4, 600 MHz) metabolic profiling of SDU8 EtOAc metabolome.
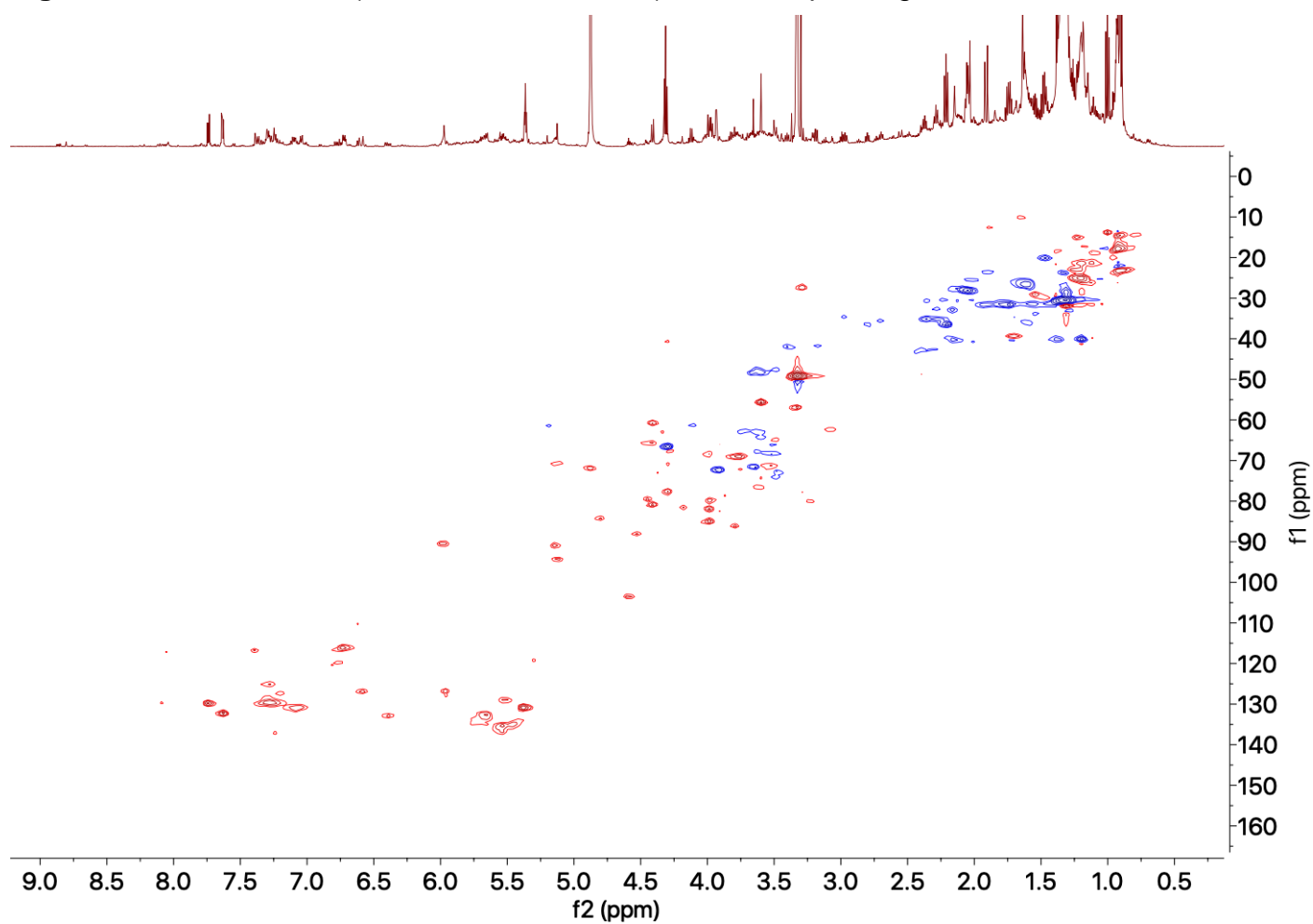
Figure S4. ^{1}H-^{13}C HMBC (methanol-d_4, 600 MHz) metabolic profiling of SDU8 EtOAc metabolome.
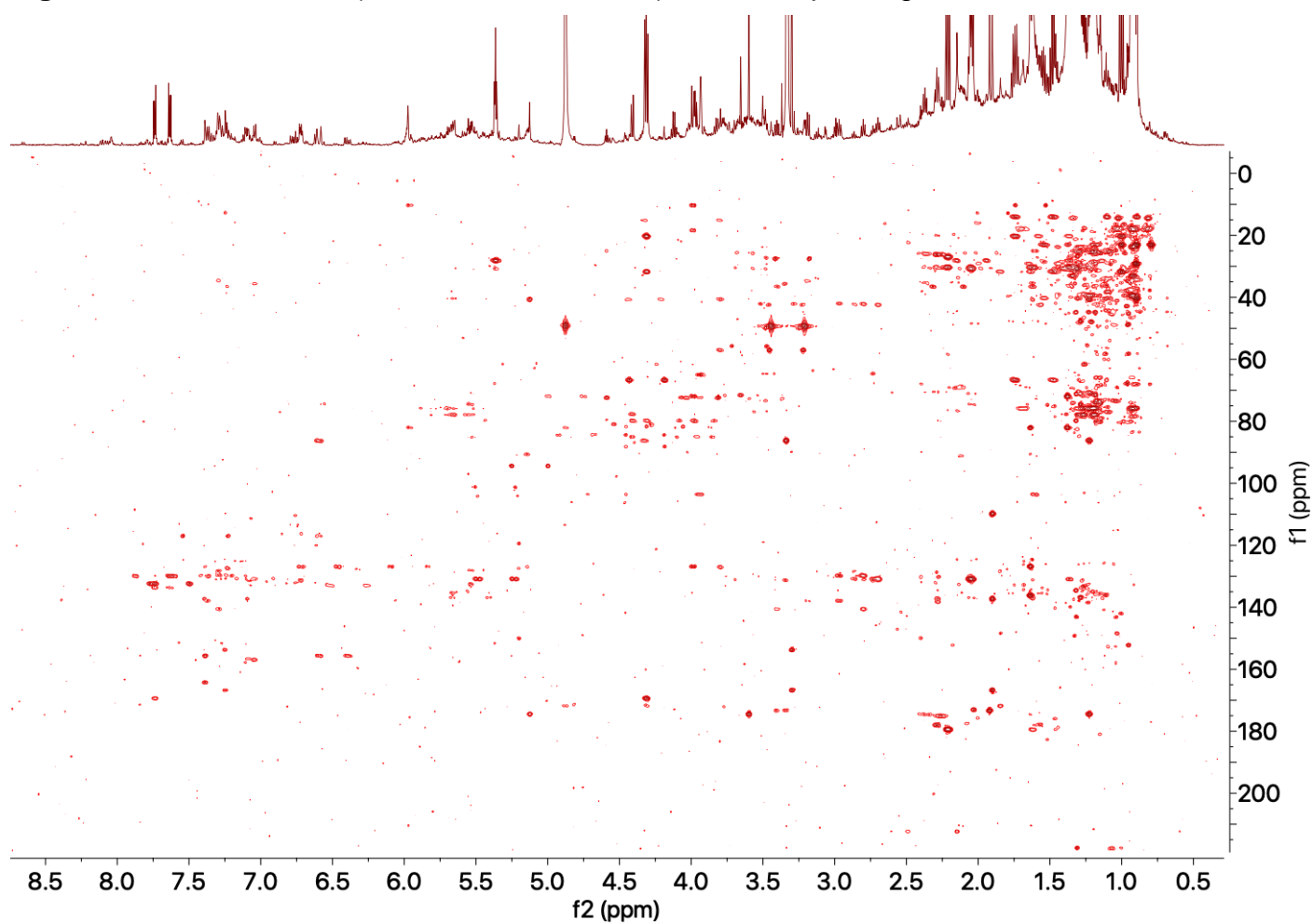
Figure S5. SMART 2.0 analysis of HSQC spectrum of SDU8 EtOAc crude extract. Top 12 structures based on cosine similarity score in the parentheses suggested that polyketidic macrolides were potentially contained in the mixture.[6]

Eprinomectin (0.858)

Avermectin A1a (0.853)

Abamectin Component B1A (0.851)

mandelalide G (0.850)

Polycavernoside D (0.841)

mandelalide E (0.839)

microsporanate F (0.827)

avermectin B1c (0.824)

microsporanate E (0.819)

microsporanate D (0.817)

tolytoxin 23-acetate (0.808)

bryostatin 6 (0.808)
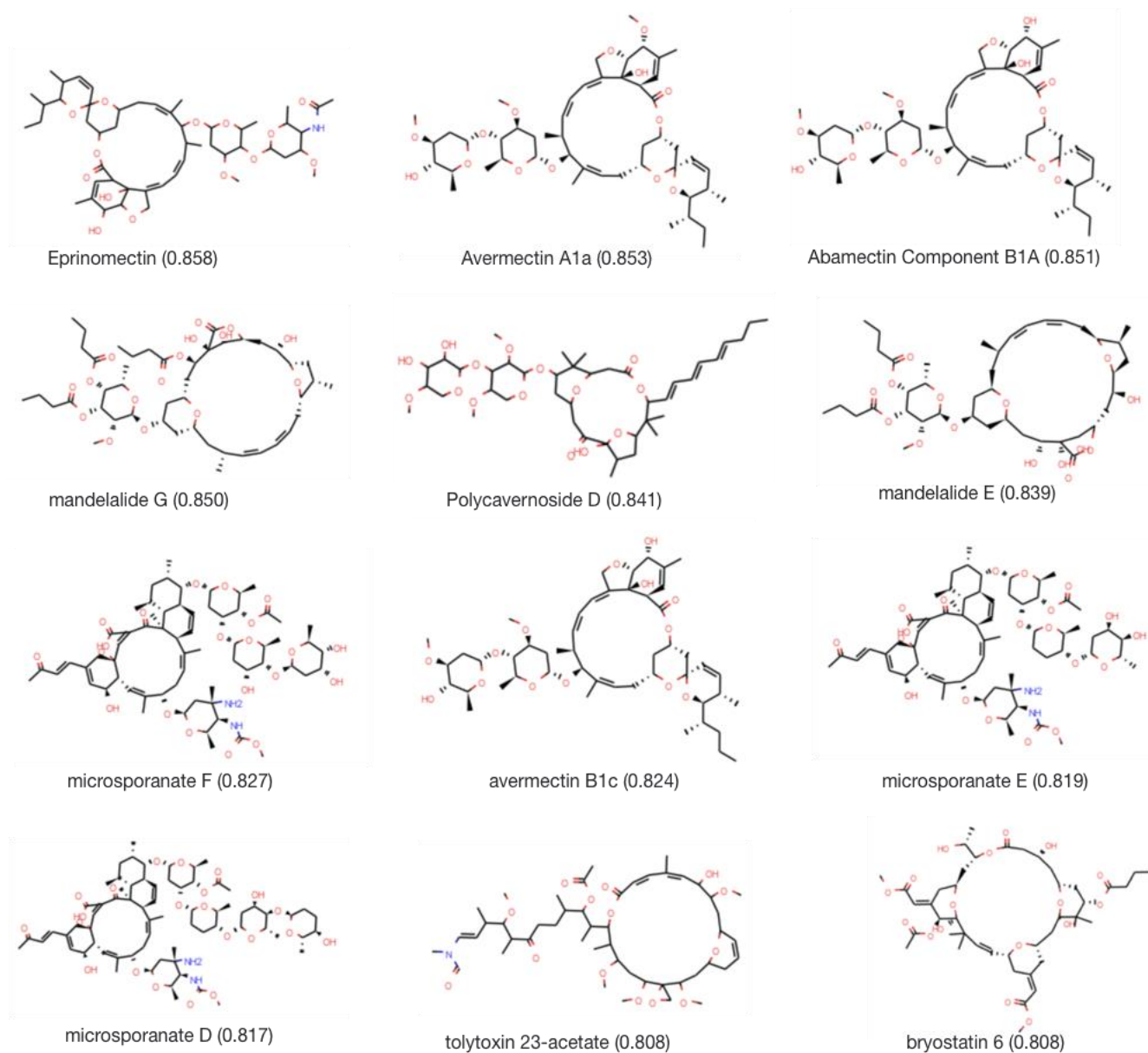
Figure S6. HMBC (⌣) and COSY (—) correlations for archangiumide.
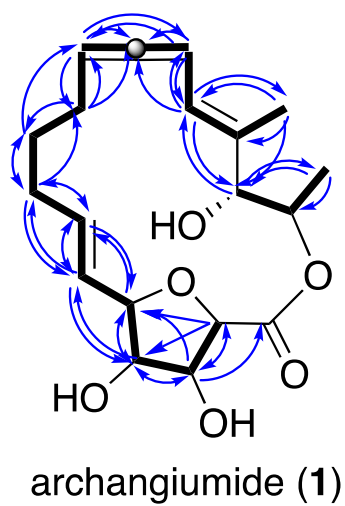
archangiumide (**1**)

Figure S7. Key NOESY correlations (H↔H) of archangiumide.
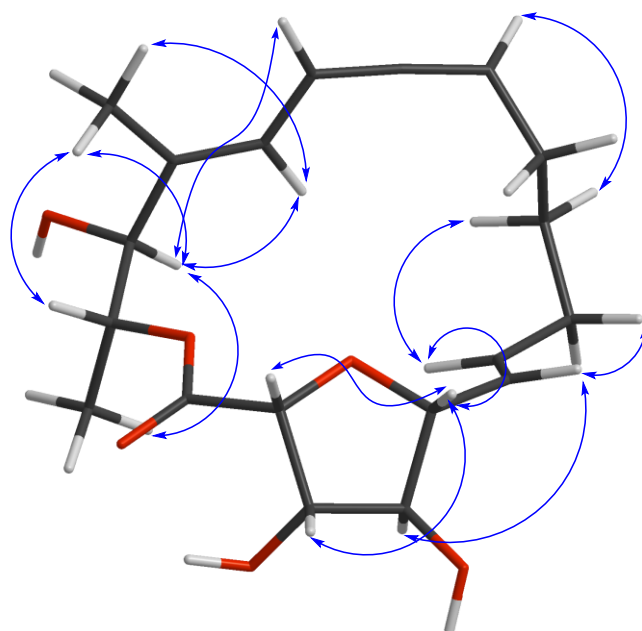
Figure S8. Phylogenetic analysis identified ArcE as malonyl-transferring AT, whereas ArcD is a acyl hydrolase (AH).[7]

Figure S9. Sequence alignment of KS domains of *arc* PKSs. The conserved HGTGT motif essential for decarboxylative condensation was indicated in box. The KS$_9$ and KS$_{10}$ in ArcC were supposed to be inactive, since they lacked the requisite His residue in the first position of this catalytic pocket.

Figure S10. Sequence alignment of KR domains of *arc* PKSs. The NADP(H) binding motif GGxGxxG was indicated in blue box. Analysis of the diagnostic Asp presence within a LDD signature motif represented in red box allowed the classification of KR types.

```
                       GGxGxxG
KR6    GVYVVTGGTNGIGLEVAKLLARRGASKLVLMGVTPLPPKERWTHLVHDSSTPEYVRRKLE    60
KR2    -VYWITGGLGGLGTIFVRHLCATEGVTVVLSGRSPLTGDRQKALE---------------    44
KR1    GVYWITGGLGGLGLIFARHLGRVGSAKLVLTGRSPLSDEKRVQLD---------------    45
KR5    GVYWITGGMGGLGLHFARHLGQTKGVKVVLSGRSALDETKTAALE---------------    45
KR4    GVYWITGGMGGLGLHFARHLGQTKGVKVVLSGRSALDETKTAALE---------------    45
KR3    GVYWITGGMGGLGLHFARHLGQTKGVKVVLSGRSALDETKTAALE---------------    45
        ** :*** .*:*   ..: *      . .:** * : *      .

                                                     LDD
KR6    AFLELEGRGVSVTLYTGPLDDRARLGAFFGDIRRELGAIRGVVHSAGAMQAAKGDEFAFV   120
KR2    -ELRQEG--FQVEYVQVDVGEKEQVRRAFSEIKREHGRIDGIVHSAGILRDS-----LLL    96
KR1    -ALRASG--ATVEYLQGDIAVREDVQRIVAEIKRRHGRLNGVIHSAGLIRDA-----FII    97
KR5    -ELRRQQPHATFTYVAVDLGDRGAVAGAVEKLRAEYGGVKGVIHSAGVIQDA-----YVV    99
KR4    -ELRRQQPHATFTYVAVDLGDRGAVAGAVEKLRAEYGGVKGVIHSAGVIQDA-----YVV    99
KR3    -ELRRQQPHATFTYVAVDLGDRGAVAGAVEKLRAEYGGVKGVIHSAGVIQDA-----YVV    99
         *. .     .       : : :   . .:: . * : *::**** :: :       .:

KR6    RKSLKRMEEVFAPKIAGLDTLSSLLSADDLDFFVSFSSTAGQLPGFMRGMSDYGIANAYV   180
KR2    NKSGADAAQVFGSKVDGALNVDEVTKNEALDFFVLFASFAGVFGN--VGQGDYSGANAFL   154
KR1    KKTEAEIATVMAPKVGGVVNVDLATREEPLDFFVVFSSMTSLLGN--VGQADYSAANAFL   155
KR5    KKEEGQIEAVFAPKVKGLLNLEEATREEELDFVVLFSSLAGAMGS--AGQADYAGANAFL   157
KR4    KKEEGQIEAVFAPKVKGLLNLEEATREEELDFVVLFSSLAGAMGS--AGQADYAGANAFL   157
KR3    KKEEGQIEAVFAPKVKGLLNLEEATREEELDFVVLFSSLAGAMGS--AGQADYAGANAFL   157
        .*        *:. *: *  .:.      : ***.* *:* :. : .   * .**. ***::

KR6    DSFTEFQAHGG-----KPCFRSMAWVGWADT            206
KR2    DAFARARNGQVARGERRGRTVSIDWPLWRDG            185
KR1    DGYAEYRQRLVQEGERSGKTLSVSWPLWKEG            186
KR5    DAYAEYRQGQVREGKRKGRTVSVGWPLWREG            188
KR4    DAYAEYRQGQVREGKRKGRTVSVGWPLWREG            188
KR3    DAYAEYRQGQVKEGKRKGRTVSVGWPLWKEG            188
        *.::. :               *: *   * :
```

Figure S11. Sequence alignment of the ACP domains from *arc* PKSs. The serine residue that serves as the attachment site for the 4'-PP cofactor was highlighted in red box, the signature tryptophan residue for β-branching was marked in blue box.

Figure S12. ^{13}C NMR spectra (CD$_3$OD, 150 MHz) of 1-^{13}C-labeled archangiumide (A), natural abundance archangiumide (B), and 2-^{13}C-labeled archangiumide (C). The ^{13}C enriched positions were labelled in the isotope-labelling groups.
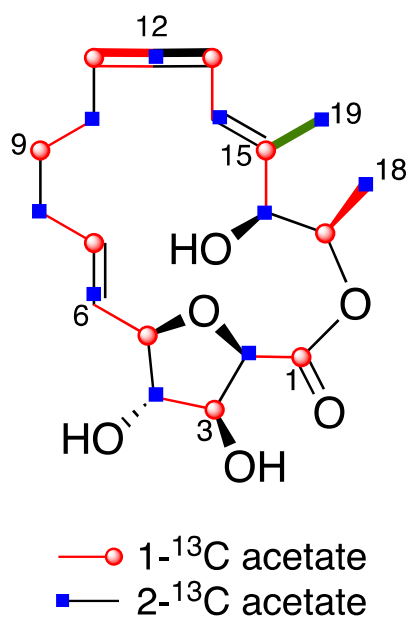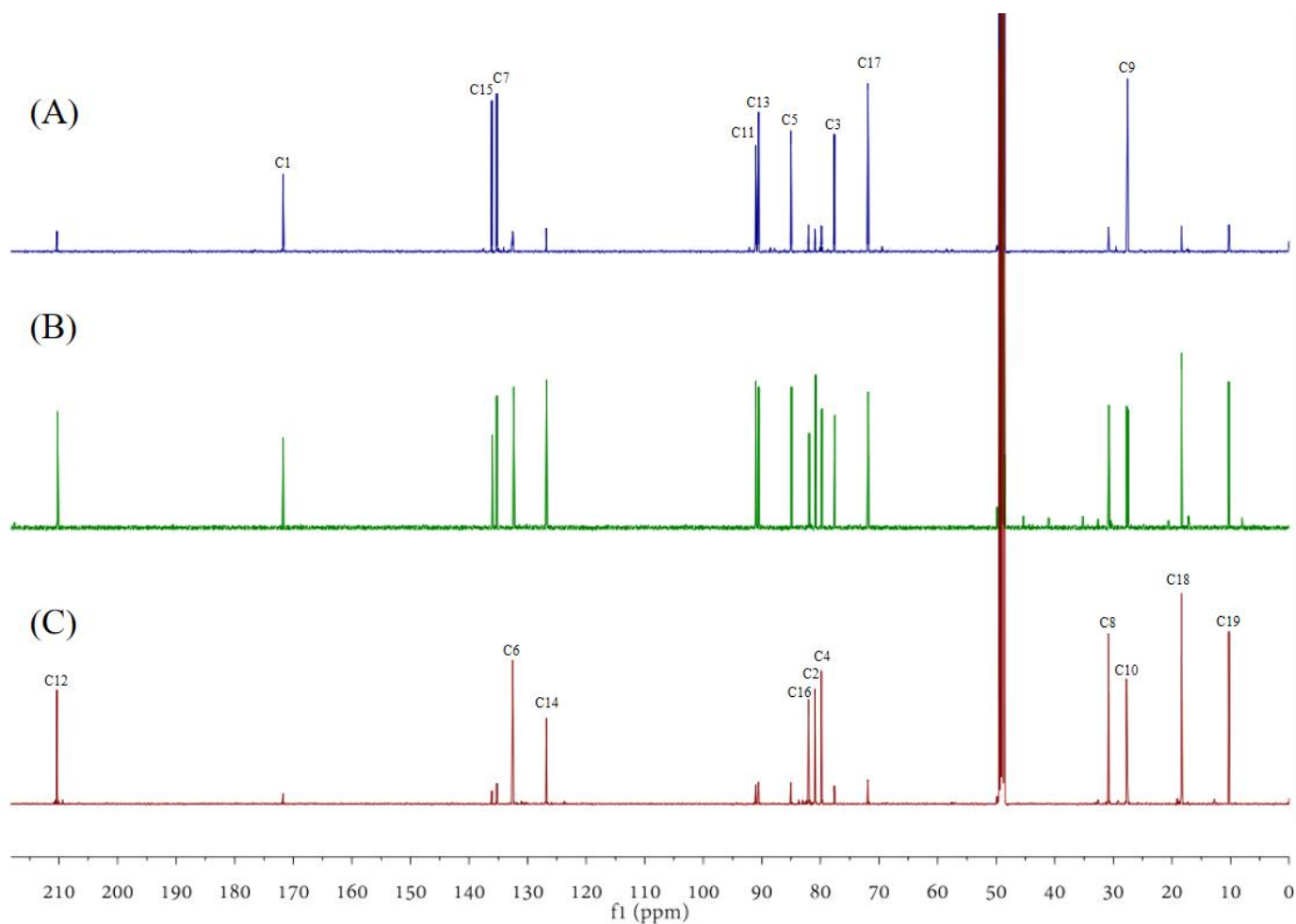
Figure S13. LC-MS analysis of SDU8 EtOAc crude extract. The same sample of SDU8 EtOAc crude extract for NMR analysis in Figure S1 was used for LC-MS analysis. Metabolites detection was done on the basis of UV (254 nm) and MS (total ion chromatography, TIC), simultaneously. The extracted ion chromatography (EIC) of *m/z* 351 presented the peak for archangiumide.
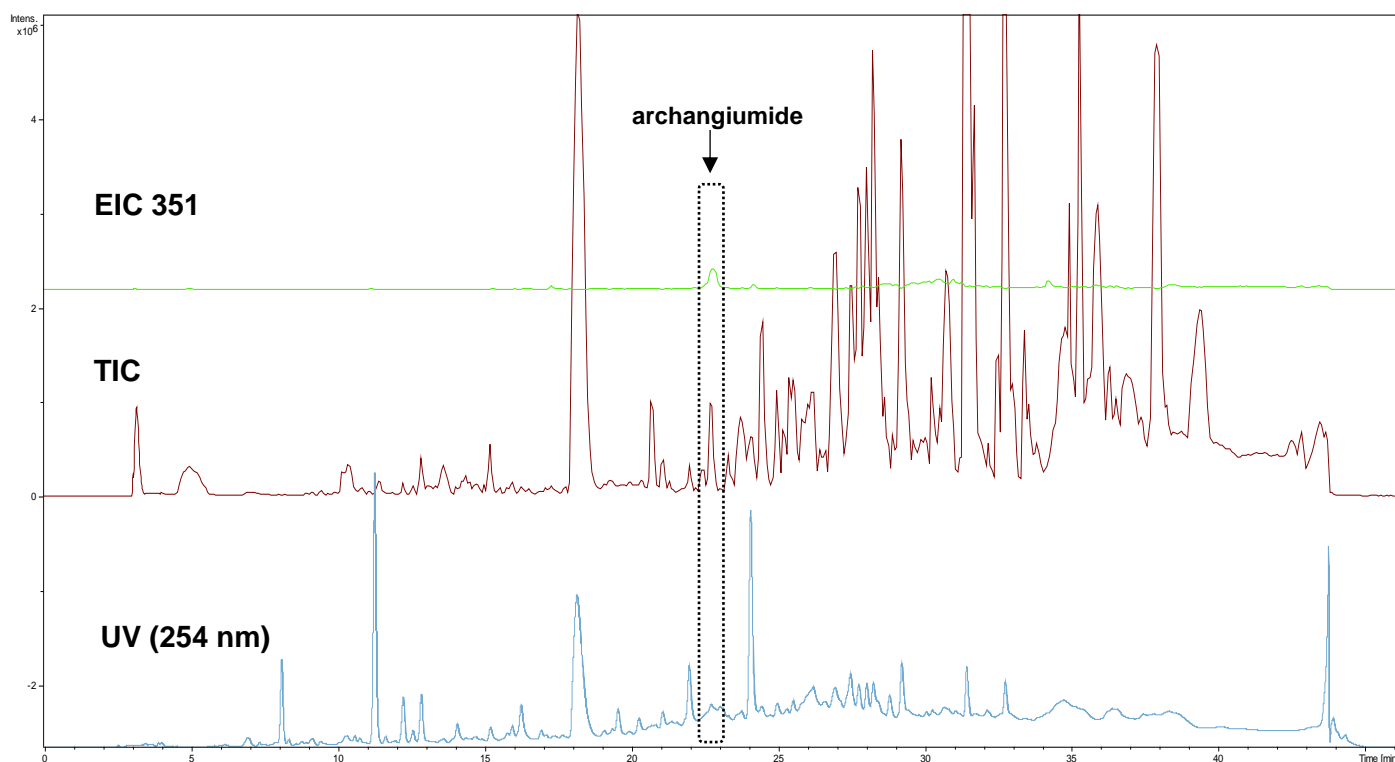
Figure S14. The phylogenetic tree constructed based on 16S rRNA gene sequences of myxobacteria. The neighbor-joining method was used. The position of strain SDU8 among related taxa shows it falls into the genus *Archangium*. Bootstrap values based on 1000 replications were listed as percentages at the branching points. The tree was rooted with *Anaeromyxobacter dehalogenans* DSM 21875 as the outgroup, and the NCBI accession number of each sequence was added in the parentheses.
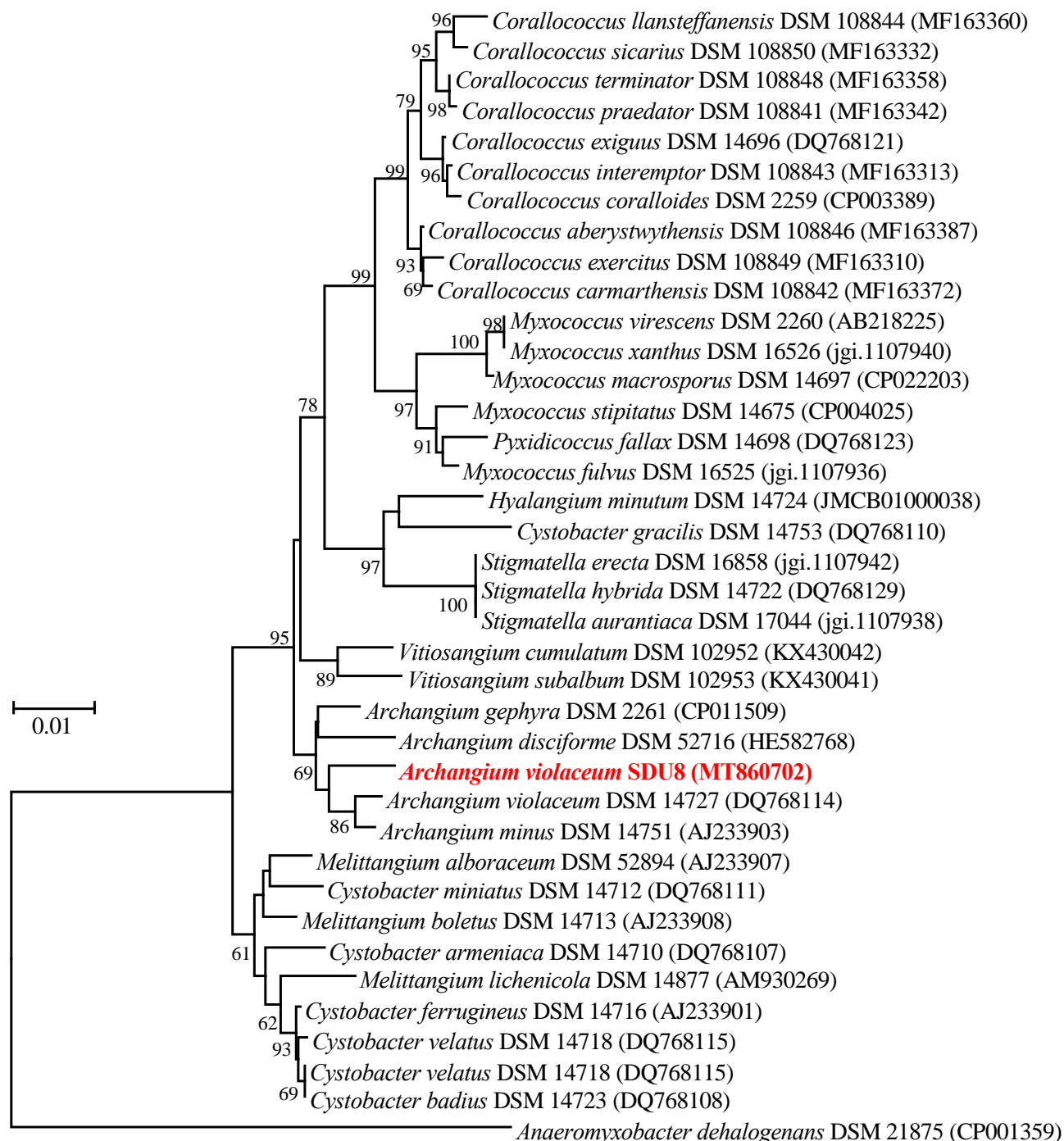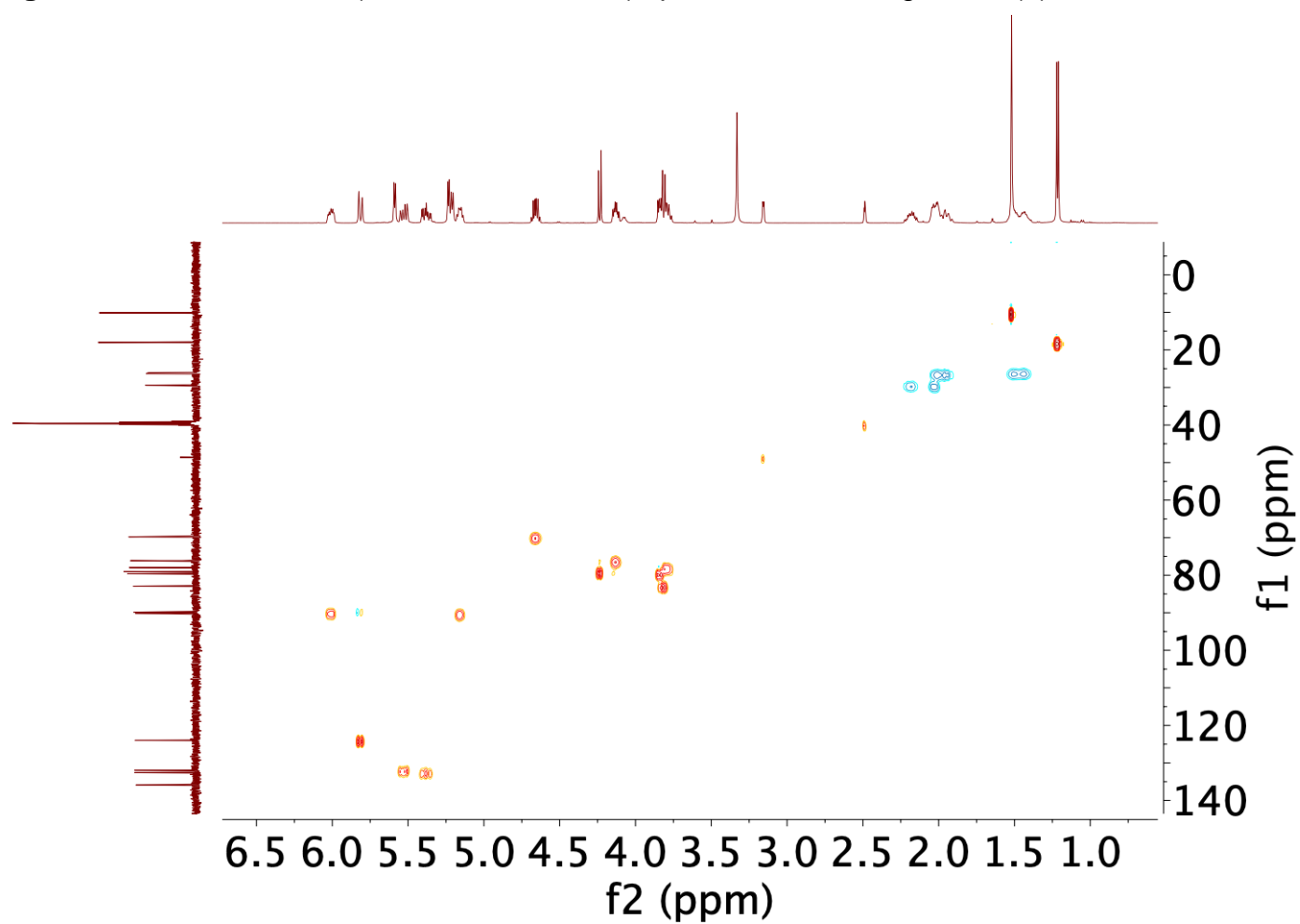
Figure S15. ^{1}H NMR (500 MHz, DMSO-d_6) of archangiumide (**1**).
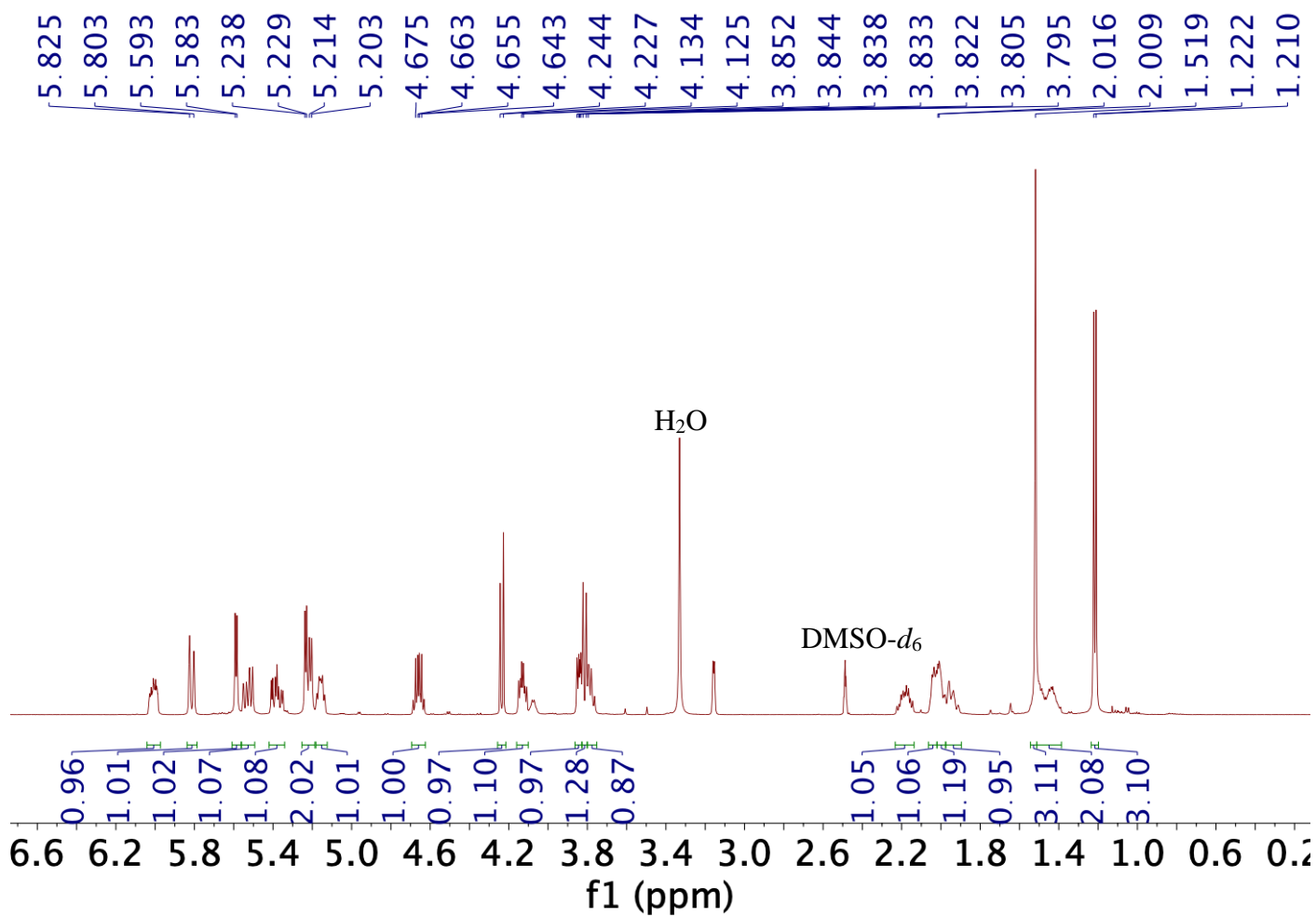
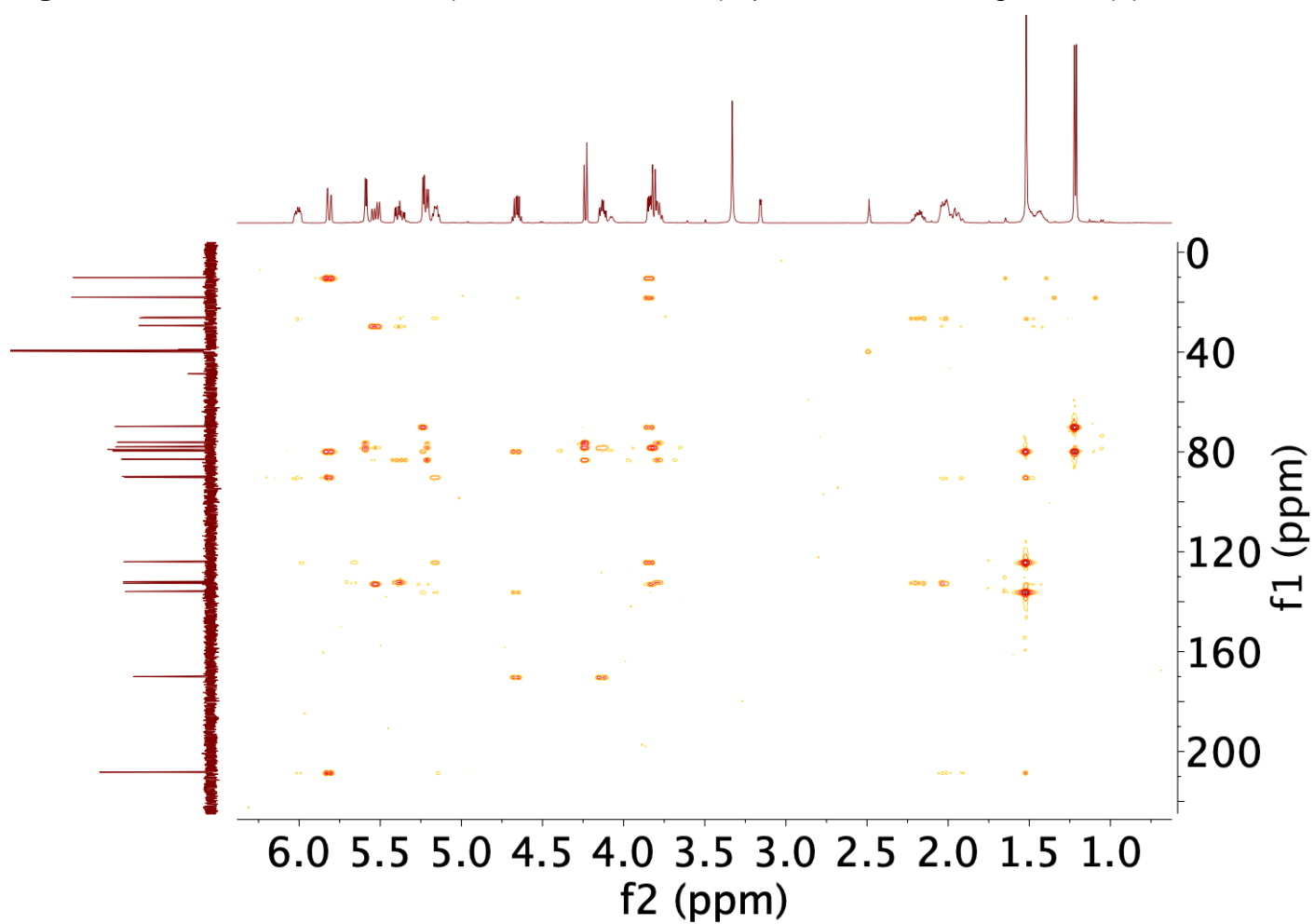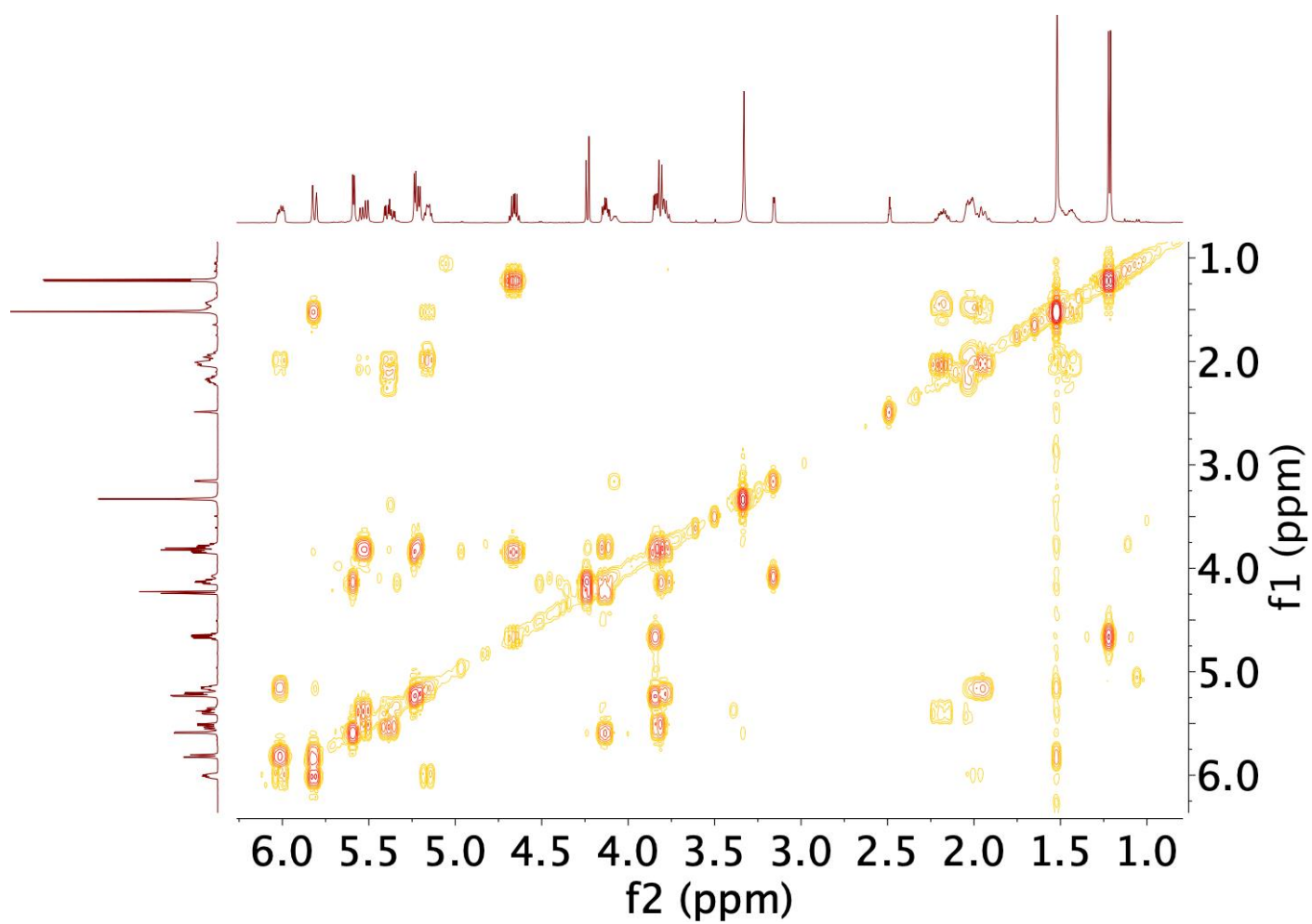**Figure S16.** ^{13}C NMR (125 MHz, DMSO-d_6) spectrum of archangiumide (**1**).
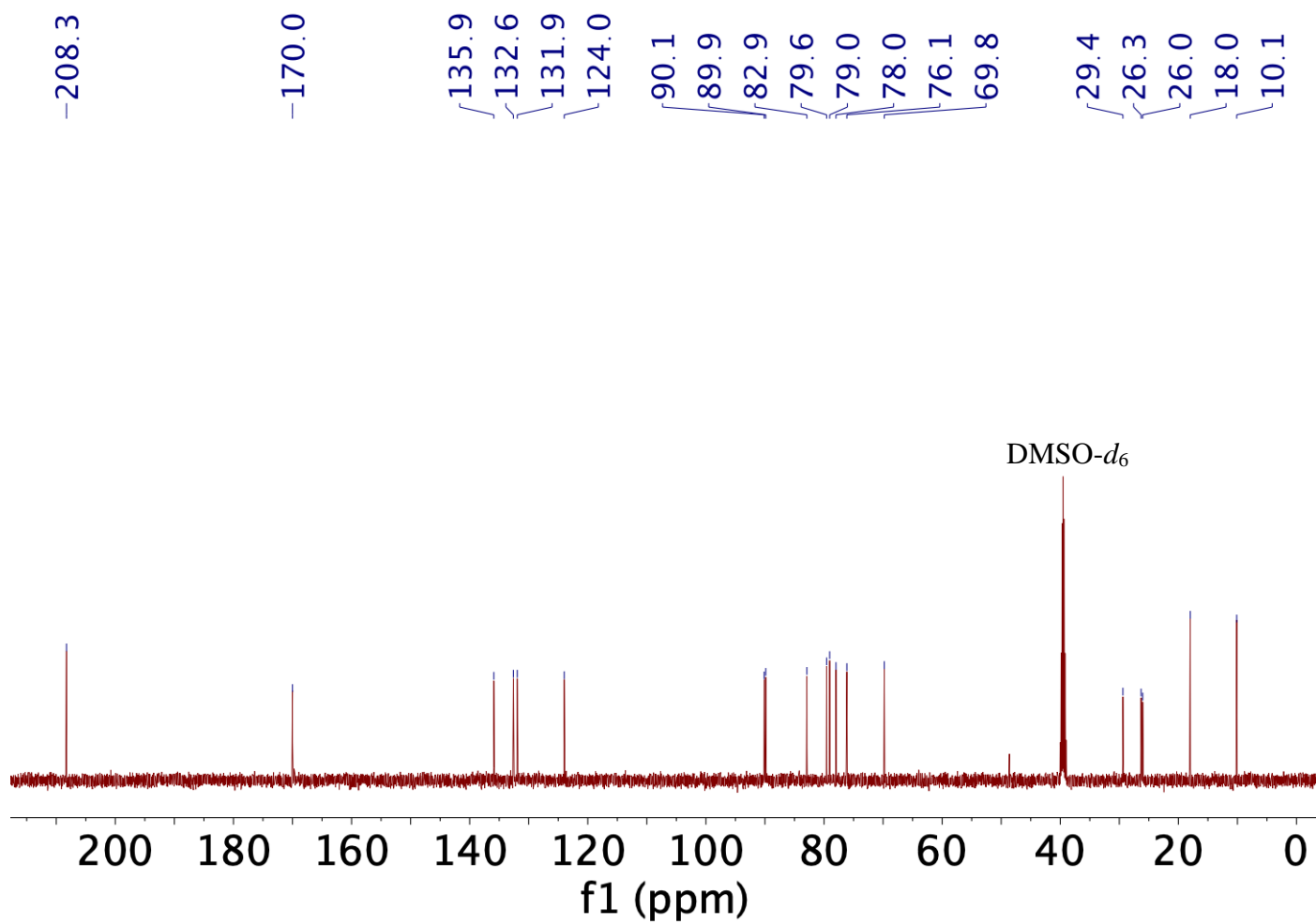
Figure S17. ^1H-^{13}C HSQC (500 MHz, DMSO-d_6) spectrum of archangiumide (**1**).

Figure S18. ^{1}H-^{13}C HMBC NMR (500 MHz, DMSO-d_6) spectrum of archangiumide (**1**).

Figure S19. ^{1}H-^{1}H COSY NMR (500 MHz, DMSO-d_6) spectrum of archangiumide (**1**).

Figure S20. ^{1}H-^{1}H NOESY NMR (600 MHz, CD$_3$OD) spectrum of archangiumide (**1**).
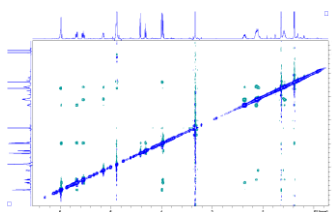
Figure S21. High-resolution mass spectrum of archangiumide
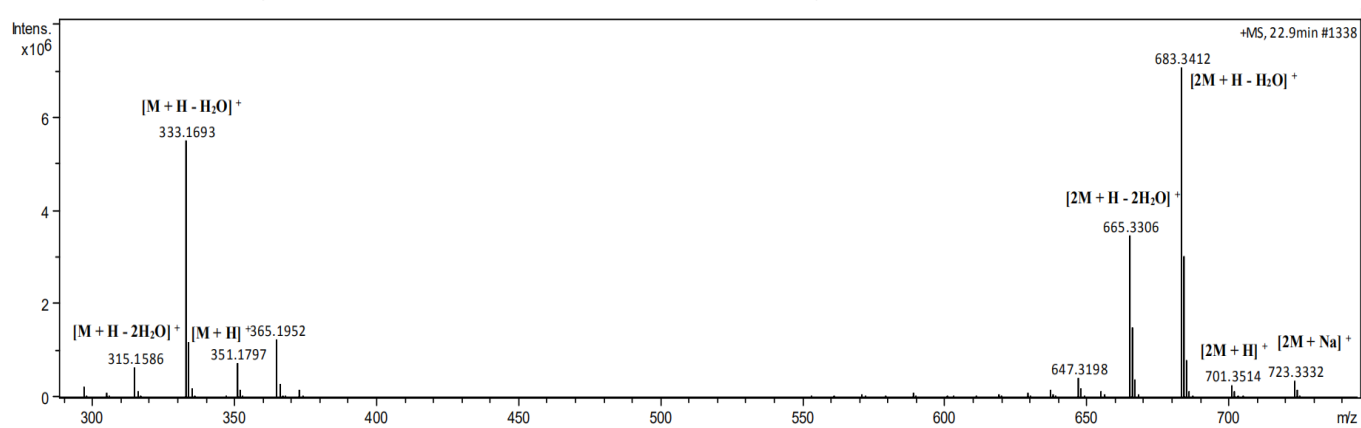
Figure S22. IR spectrum of archangiumide

Figure S23. UV spectrum of archangiumide

Figure S24. ORTEP structure generated from the single crystal diffraction of archangiumide (**1**) showing 50% probability ellipsoids. This structure has been deposited in CCDC database under the accession number 2058107.
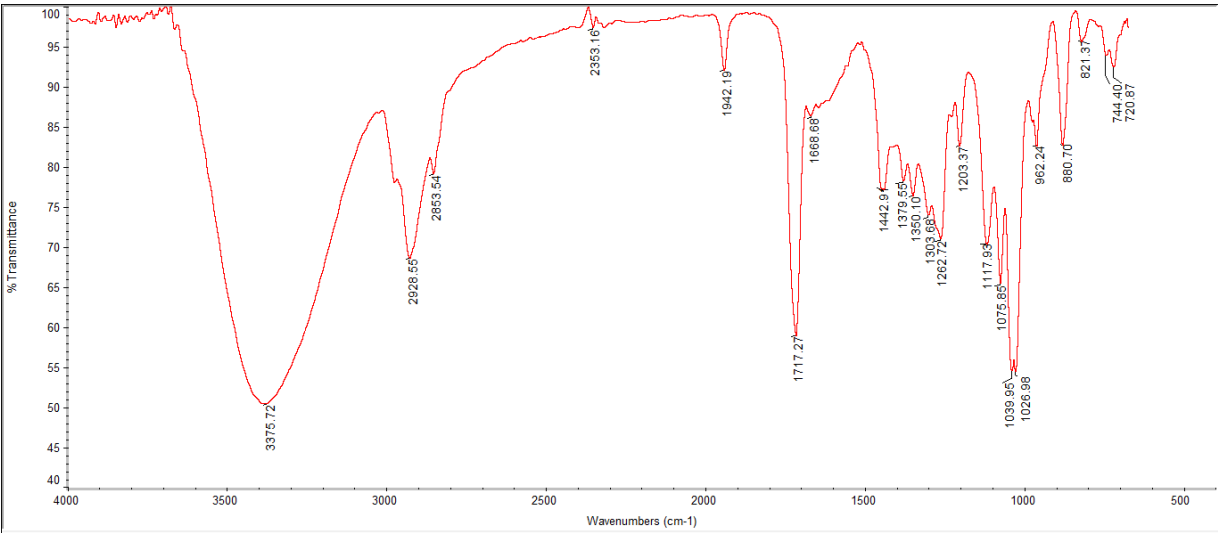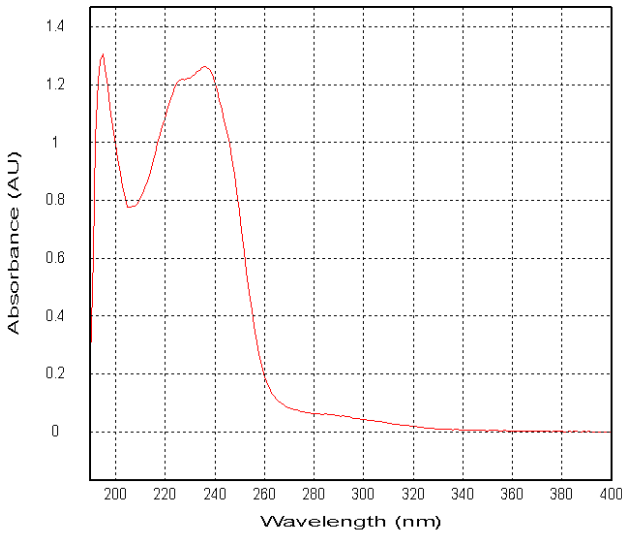

archangiumide (1)

References

(1) Reichenbach, H.; Dworkin, M. *The Prokaryotes*, 2nd ed.; Balows, A., Trüper, H. G., Dworkin, M., Harder, W., Schleifer, K.-H., Eds.; Springer-Verlag New York: New York, **1992**. https://doi.org/10.1007/978-1-4757-2191-1.

(2) Wu, C.; Van Wezel, G. P.; Hae Choi, Y. Identification of Novel Endophenaside Antibiotics Produced by Kitasatospora Sp. MBT66. *J. Antibiot. (Tokyo).* **2015**, *68* (7), 445–452. https://doi.org/10.1038/ja.2015.14.

(3) DeGraff, W. G.; Mitchell, J. B. Evaluation of a Tetrazolium-Based Semiautomated Colorimetric Assay: Assessment of Chemosensitivity Testing. *Cancer Res.* **1987**, *47* (4), 936–942.

(4) Hichri, F.; Omri Hichri, A.; Maha, M.; Saad Mana Hossan, A.; Flamini, G.; Ben Jannet, H. Chemical Composition, Antibacterial, Antioxidant and in Vitro Antidiabetic Activities of Essential Oils from Eruca Vesicaria. *Chem. Biodivers.* **2019**, *16* (8). https://doi.org/10.1002/cbdv.201900183.

(5) Schneider, T. R.; Sheldrick, G. M. Substructure Solution with SHELXD. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58* (10), 1772–1779. https://doi.org/10.1107/S0907444902011678.

(6) Reher, R.; Kim, H. W.; Zhang, C.; Mao, H. H.; Wang, M.; Nothias, L. F.; Caraballo-Rodriguez, A. M.; Glukhov, E.; Teke, B.; Leao, T.; Alexander, K. L.; Duggan, B. M.; Van Everbroeck, E. L.; Dorrestein, P. C.; Cottrell, G. W.; Gerwick, W. H. A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products. *J. Am. Chem. Soc.* **2020**, *142* (9), 4114–4120. https://doi.org/10.1021/jacs.9b13786.

(7) Moon, K.; Cui, J.; Kim, E.; Riandi, E. S.; Park, S. H.; Byun, W. S.; Kal, Y.; Park, J. Y.; Hwang, S.; Shin, D.; Sun, J.; Oh, K. B.; Cha, S.; Shin, J.; Lee, S. K.; Yoon, Y. J.; Oh, D. C. Structures and Biosynthetic Pathway of Pulvomycins B-D: 22-Membered Macrolides from an Estuarine Streptomyces Sp. *Org. Lett.* **2020**, *22* (14), 5358–5362. https://doi.org/10.1021/acs.orglett.0c01249.