Sophopterocarpan A, a novel pterocarpine derivative with a benzotetrahydrofuran-fused bicyclo [3.3.1] nonane from *Sophora flavescens*

Hui Zhu,[§] Ya-Nan Yang,[§] Kuo Xu, Jing Xie, Zi-Ming Feng, Jian-Shuang Jiang, and Pei-Cheng Zhang*

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, People's Republic of China

Supporting information

The List of Contents

No	Content	Page
1	Experimental Section.	S1
2	Table S1. NMR Assignments for 1 in CD_3OD.	S2
3	Figure S1. Compound **1** inhibited MCF-7 cells growth.	S4
4	Table S2. Energies of the conformers with Boltzmann distribution over 1%.	S7
5	Figure S2. ^1H NMR spectrum of compound **1** (500 MHz, CD_3OD).	S8
6	Figure S3. The expansion for the ^1H NMR spectrum of compound **1** (δ: 3.50–4.30).	S8
7	Figure S4. The expansion for the ^1H NMR spectrum of compound **1** (δ: 4.80–5.45).	S9
8	Figure S5. The expansion for the ^1H NMR spectrum of compound **1** (δ: 6.18–6.82).	S9
9	Figure S6. The expansion for the ^1H NMR spectrum of compound **1** (δ: 3.01–3.18).	S10
10	Figure S7. ^{13}C NMR spectrum of compound **1** (125 MHz, CD_3OD).	S10
11	Figure S8. HSQC spectrum of compound **1**.	S11
12	Figure S9. HMBC spectrum of compound **1**.	S11
13	Figure S0. ^1H-^1H COSY spectrum of compound **1**.	S12
14	Figure S11. ROESY spectrum of compound **1**.	S12
15	Figure S12. HRESIMS spectrum of compound **1**.	S13
16	Figure S13. IR spectrum of compound **1**.	S13
17	Figure S14. UV spectrum of compound **1**.	S14
18	Figure S15. ECD spectrum of compound **1** (in MeOH).	S14

EXPERIMENTAL SECTION

General Experimental Procedures Optical rotation was measured by a JASCO P-2000 polarimeter (JASCO, Easton, MD, USA). UV spectrum was obtained on a JASCO V-650 spectrophotometer (JASCO). ECD spectra were measured on JASCO J-815 spectrometer (JASCO). IR spectrum was recorded on a Nicolet 5700 spectrometer with an FT-IR microscope transmission method (Thermo Scientific, Waltham, MA, USA). NMR spectra measurements were performed on Bruker 500 MHz spectrometer (Bruker-Biospin, Billerica, MA, USA) in CD_3OD. High-resolution electrospray ionization mass spectrometry (HRESIMS) was performed on an Agilent 6520 HPLC-Q-TOF (Agilent Technologies, Waldbronn, Germany). Column chromatography was performed with macroporous resin (Diaion HP-20, Mitsubishi Chemical Corp., Tokyo, Japan), Rp-18 (50 μm, YMC, Kyoto, Japan), and silica gel (100−200 mesh, Qingdao Marine Chemical Inc. Qingdao, People's Republic of China). Preparative HPLC was carried out on a Shimadzu LC-6AD instrument (Shimadzu Corp., Tokyo, Japan) with an SPD-20A detector (UV-detector), using a YMC-Pack ODS-A column (250 × 20 mm, 5 μm, Japan). HPLC-DAD analysis was performed using an Agilent 1260 series system (Agilent Technologies) with an Apollo C18 column (250 × 4.6 mm, 5 μm, Grace Davison).

Plant Material The roots of *S. flavescens* were purchased from Beijing PushengLin pharmaceutical company, grew in Hebei. The plant was authenticated by Professor Lin Ma, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College. A voucher specimen (ID-5-2438) was deposited at the Institute of Materia Medica, Chinese Academy of Medical Sciences, Beijing 100050, People's Republic of China.

Extraction and Isolation The dried roots of *S. flavescens* (39.0 kg) were extracted with 95% EtOH (200 L) under reflux (2 × 2h). After filtration, the residual was extracted one more time with 70% EtOH for 2 h. Then, the extracts were merged and concentrated in vacuo, to obtain the crude extracts (8.4 kg), which were suspended in water and then liquid-liquid extracted, to get four extractions, including petroleum ether extraction, ethyl acetate (EtOAc) extraction, *n*-Butyl alcohol extraction and water extraction. The EtOAc extraction was separated into seven fractions (Fr. A–G) by silica gel column chromatography eluting with PE, PE/EtOAc (1:1), EtOAc, EtOAc/acetone (1:1), acetone, acetone/methanol (1:1), methanol. Among above seven fractions, Fr. C was further separated by silica gel column chromatography eluting with a CH_2Cl_2/MeOH

(40:1-2:1) to afford 15 fractions (Fr. C-1–Fr. C-15). Furthermore, fraction Fr. C-5 (4.3 g) was chromatographed over a reverse-phase C_{18} silica gel columns eluting with MeOH/H_2O (40-100%, with 10% stepwise increase of MeOH) to get seven subfractions. Subfraction 2 (185 mg) was purified by preparative HPLC using MeOH/H_2O (40:60, *v*/*v*) at 5 mL/min to yield **1** (35.2 mg, t_R = 43.0 min).

Sophopterocarpan A: white powder; [α]20 D 33.9 (*c* 0.10 MeOH); UV (MeOH) λ_{max} 287 (3.17) nm; IR (KBr) ν_{max}: 3284, 2968, 1670, 1598, 1499, 1467, 1429, 1271, 1088, 925, 816, 663 cm^{-1}; (+)-HRESIMS *m/z* 305.1022 $[M + H]^+$ (calcd for $C_{16}H_{17}O_6$, 305.1020).

Formononetin: ^1H NMR (500 MHz, DMSO-d_6) δ: 8.30 (1H, s, H-2), 7.92 (1H, d, *J* = 8.5, Hz, H-5), 6.90 (1H, dd, *J* = 2.0, 8.5 Hz, H-6), 6.80 (1H, d, *J* = 2.0 Hz, H-8), 7.50 (1H, d, *J* = 8.0 Hz, H-2′, H-6′), 6.97 (1H, d, *J* = 8.0 Hz, H-3′, H-5′), 3.77 (3H, s, 4′-OCH_3); ^{13}C NMR (125 MHz, DMSO-d_6) δ: 153.0 (C-2), 124.3 (C-3), 174.6 (C-4), 127.7 (C-5), 115.5 (C-6), 158.9 (C-7), 102.1 (C-8), 157.5 (C-9), 116.1 (C-10), 123.1 (C-1′), 130.1 (C-2′, 6′), 113.6 (C-3′, 5′), 163.6 (C-4′), 55.2 (4′-OCH_3).

Table S1 NMR Assignments for Sophopterocarpan A in CD_3OD[a]

No.	δ_H	δ_C	HMBC[b]	^1H-^1H COSY[c]	ROESY[d]
1		199.4			
2	5.40, dd (1.0, 10.0)	129.6	1,4,5,8	3	
3	6.66, d (10.0)	155.2	1,5,8,9	2	
4		75.4			
5	5.17, d (9.0)	87.2	3,4,6,7,9,1′,6′	6	6
6	4.02, dd (1.0, 9.0)	50.3	4,5,7,8,1′,2′,4′,5′,6′	5,7	5
7	4.20, br s	70.0	1,5,6,8,9,6′	8	5′
8	3.07, m	57.1	1,2,4,6,7,9	9	
9	4.08, d (4.0)	75.7	1,3,4,5,7,8	8	2,3
1′		161.5			
2′	6.26, d (2.0)	97.6	1′,3′,4′,6′		
3′		162.4			
4′	6.27, dd (2.0, 8.5)	106.8	6,2′,6′	5′	
5′	6.74, dd (1.0, 8.5)	123.8	6,1′,2′,3′,4′,6′	4′	7

6′		122.4	
3′-OCH_3	3.64, s	55.8	3′

[a]^{1}H NMR (500 MHz) (δ in ppm, J in Hz), ^{13}C NMR (125 MHz). [b]HMBC correlations are from the proton(s) to the indicated carbon(s). [c]^{1}H-^{1}H COSY correlations are from the proton(s) to the indicated proton(s). [d]ROESY correlations are from the proton(s) to the indicated proton(s).

X-ray crystallography analysis. Colorless plate crystals of compound **1** were obtained from the system of MeOH-acetone (1:1). A crystal (0.380 mm × 0.200 mm × 0.090 mm) was separated and mounted on a glass fiber which data was collected by Rigaku MicroMax 002+ CCD detector equipped with a graphite monochromator and *CuKa* radiation. Crystal data as follows: $C_{16}H_{16}O_8$, M = 304.29, space group orthorhombic, $P2_1$; unit cell dimensions, a = 7.0259 (3) Å, b = 6.8005 (3) Å, and c = 14.1172 (6) Å; V = 673 (5) Å^3, Z = 2, ρ_{calcd} = 1.502 mg/mm^3, F (000) = 320. The 4433 measurements yielded 2228 independent reflections after equivalent data were averaged and Lorentz and polarization corrections were applied. Using the SHELXS-97 program the structure was solved by direct methods and expanded using difference Fourier techniques, and refined by the program SHELXL-97 and full-matrix least-squares calculations. The final refinement gave R_1 = 0.0364, wR_2 = 0.0944. The flack parameters was 0.0(2). The crystallographic data have been deposited at The Cambridge Crystallographic Data Centre (deposition number CCDC 1496284). The data is available free of charge via www.ccdc.cam.ac.uk/products/csd/ request.

***In vitro* cytotoxic activity**. Sophopterocarpan A was screened for its antitumor activity against several human tumor cell lines including hepatoma (HepG2), breast cancer (MCF-7), and lung carcinoma (A549) cell lines using the standard MTT assay.[1] Each experiment was carried out in triplicates using curcumin as a positive control.

Autophagy Modulator Screening. Chinese hamster ovary (CHO) cells stably expressing green fluorescent protein-LC3 (GFP-LC3) was maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FBS) and 800 μg/mL geneticin. Stable GFP-LC3 cells were incubated with compound **1** or curcumin at a concentration of 20 μM (DMSO) for 6 h. Cells were harvested with trypsin, washed with phosphate buffer saline (PBS) and then PBS containing 0.05% saponin and finally suspended in 1 ml of PBS. GFP-LC3 fluorescence was measured using a flow cytometer (Partec, Muenster, Germany) with more than 10,000 events, and the data was analyzed using flow cytometry software (FCS) Express 5. Each experiment was carried out in

triplicates.

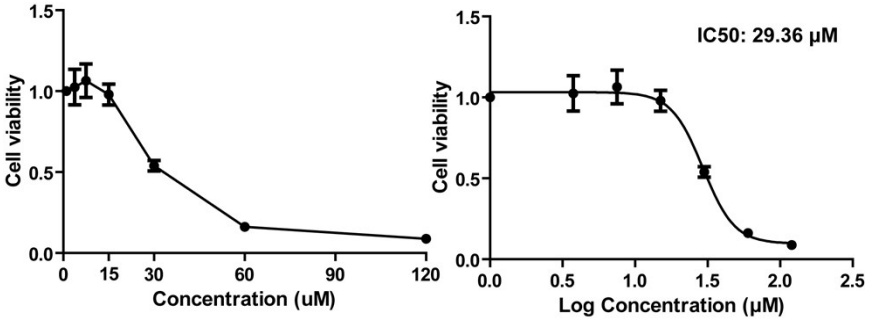


Figure S1. Compound **1** inhibited MCF-7 cells growth.

1 X. R. Lu, X. Pan, Y. Yang, M. Ji, X. G. Chen, Z. Y. Xiao, and Z. Z. Liu, *Eur. J. Med. Chem.*, 2017, **135**, 260–269.

Determination of the absolute configuration of Sophopterocarpan A (1) by measurement of the electronic circular dichroism (ECD) spectrum and by comparison with calculated ECD data. Compound **1** has only one pair of enantiomers (**1a**: 4*R*, 5*R*, 6*R*, 7*R*, 8*S*, 9*R* and **1b**: 4*S*, 5*S*, 6*S*, 7*S*, 8*R*, 9*S*). Conformational analysis of **1a** was performed by using the MMFF94 molecular mechanics force field. The molecule of **1a** showed two lowest energy conformers (Table S1). The conformers were further optimized at the B3LYP/6-311+G(d,p) level. ECD spectra of different conformers were simulated using a Gaussian function with a half-bandwidth of 0.4 eV. The overall theoretical ECD spectra were obtained according to the Boltzmann weighting of each conformers. The theoretically calculated ECD spectrum of **1a** (Figure S2) were in good agreement with the experimental ECD spectrum of **1**, which allowed the assignment of the absolute configuration of **1** as 4*R*, 5*R*, 6*R*, 7*R*, 8*S*, 9*R*.

Cartesian coordinate of low-energy optimized conformers of **1a-C1** optimized: no imaginary frequency

Zero-point correction=	0.302763 (Hartree/Particle)
Thermal correction to Energy=	0.320825
Thermal correction to Enthalpy=	0.321769
Thermal correction to Gibbs Free Energy=	0.257762
Sum of electronic and zero-point Energies=	-1070.327673
Sum of electronic and thermal Energies=	-1070.309611
Sum of electronic and thermal Enthalpies=	-1070.308666
Sum of electronic and thermal Free Energies=	-1070.372673

Standard orientation:

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z

1	6	0	0.795768	-0.590460	-0.871259
2	6	0	1.379466	0.652968	-0.638987
3	6	0	2.711168	0.823718	-0.279923
4	6	0	3.486416	-0.341710	-0.152394
5	6	0	2.923584	-1.609236	-0.381799
6	6	0	1.582079	-1.732591	-0.749856
7	6	0	-0.628296	-0.361074	-1.340371
8	6	0	-0.831900	1.128281	-0.936518
9	8	0	0.503194	1.699599	-0.808253
10	6	0	-1.733312	-1.356020	-0.891453
11	6	0	-2.432316	-0.969502	0.443728
12	6	0	-1.601805	1.374871	0.388812
13	6	0	-0.771706	1.064620	1.618496
14	6	0	-0.769555	-0.138610	2.213634
15	6	0	-1.538314	-1.264157	1.652590
16	8	0	-1.454163	-2.400108	2.104417
17	6	0	-2.863056	0.500896	0.429177
18	1	0	-0.643533	-0.417877	-2.436885
19	8	0	4.804727	-0.339242	0.198913
20	6	0	5.438129	0.912646	0.449546
21	1	0	-1.280820	-2.344534	-0.765111
22	1	0	-3.414826	0.746864	1.346422
23	8	0	-3.694579	0.735439	-0.708350
24	1	0	-1.342453	1.697082	-1.714831
25	1	0	-3.308540	-1.618582	0.534048
26	8	0	-2.709006	-1.519767	-1.916158
27	8	0	-2.059353	2.728644	0.434741
28	1	0	3.111113	1.814255	-0.104271
29	1	0	3.556137	-2.483930	-0.271542
30	1	0	1.166535	-2.719454	-0.935978
31	1	0	-0.172746	1.881032	2.013570
32	1	0	-0.173080	-0.343139	3.098130
33	1	0	6.472466	0.677354	0.704408
34	1	0	5.419092	1.555616	-0.438653
35	1	0	4.966863	1.438174	1.288943
36	1	0	-3.817772	1.698993	-0.772596
37	1	0	-3.231800	-0.695069	-1.919426
38	1	0	-1.311303	3.308863	0.210752

Cartesian coordinate of low-energy optimized conformers of **1a-C2** optimized: no imaginary frequency

Zero-point correction= 0.302767 (Hartree/Particle)

Thermal correction to Energy= 0.320823

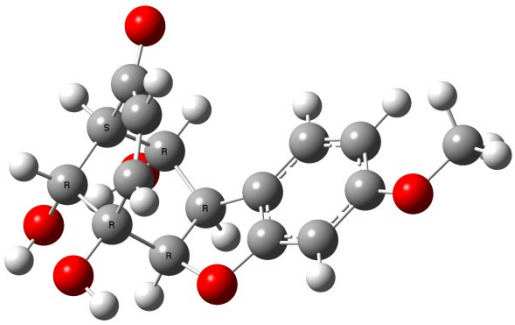
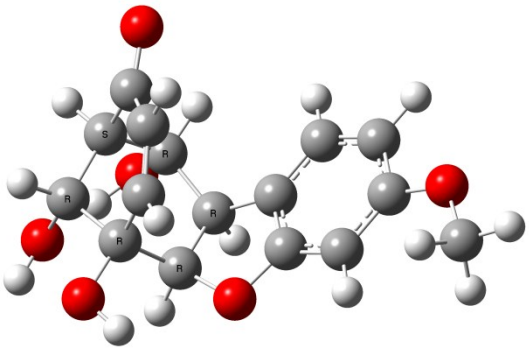
Thermal correction to Enthalpy= 0.321767
Thermal correction to Gibbs Free Energy= 0.257819
Sum of electronic and zero-point Energies= -1070.327770
Sum of electronic and thermal Energies= -1070.309714
Sum of electronic and thermal Enthalpies= -1070.308770
Sum of electronic and thermal Free Energies= -1070.372717

Standard orientation:

Center Number	Atomic Number	Atomic Type	Coordinates (Angstroms) X	Y	Z
1	6	0	0.840623	-0.267471	-0.897329
2	6	0	1.254934	1.018661	-0.534201
3	6	0	2.547788	1.314472	-0.143877
4	6	0	3.474455	0.253173	-0.116610
5	6	0	3.088441	-1.046510	-0.473299
6	6	0	1.765937	-1.298775	-0.874184
7	6	0	-0.600368	-0.180399	-1.362761
8	6	0	-1.001398	1.219845	-0.814751
9	8	0	0.244780	1.948483	-0.614454
10	6	0	-1.563486	-1.353813	-1.034792
11	6	0	-2.311458	-1.202036	0.321334
12	6	0	-1.796304	1.224059	0.518640
13	6	0	-0.931077	0.898815	1.719792
14	6	0	-0.771632	-0.348746	2.188918
15	6	0	-1.389073	-1.500878	1.507853
16	8	0	-1.160838	-2.656424	1.845887
17	6	0	-2.931347	0.193012	0.450970
18	1	0	-0.604013	-0.125241	-2.459326
19	8	0	4.732926	0.599709	0.281645
20	6	0	5.726140	-0.420512	0.350041
21	1	0	-0.983666	-2.281422	-1.001269
22	1	0	-3.511655	0.268484	1.380472
23	8	0	-3.785828	0.432248	-0.668505
24	1	0	-1.584898	1.790587	-1.538455
25	1	0	-3.094706	-1.965863	0.330422
26	8	0	-2.505905	-1.540784	-2.086289
27	8	0	-2.428657	2.493665	0.698080
28	1	0	2.852983	2.315557	0.139999
29	1	0	3.799325	-1.862959	-0.450440
30	1	0	1.484801	-2.308421	-1.162212
31	1	0	-0.442355	1.740940	2.202889
32	1	0	-0.152311	-0.563558	3.055077
33	1	0	5.903213	-0.873130	-0.632981

34	1	0	6.636960	0.073870	0.690950
35	1	0	5.446234	-1.203325	1.065187
36	1	0	-4.034818	1.372867	-0.636102
37	1	0	-3.135863	-0.798382	-2.014670
38	1	0	-1.764111	3.187441	0.545074

Table S2. Conformational Analysis of 1a.

No.	Conformer	Population
1a-C1		0.52
1a-C2		0.48

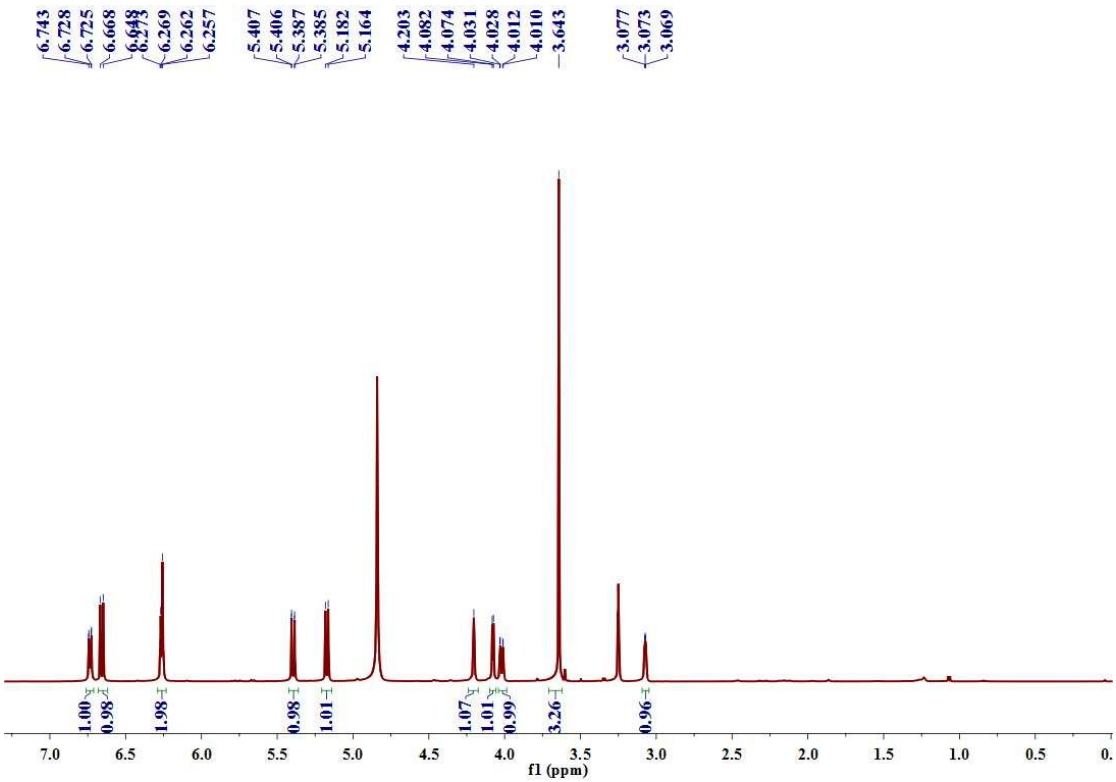


Figure S2. ^{1}H NMR spectrum of compound **1** (500 MHz, CD_3OD).

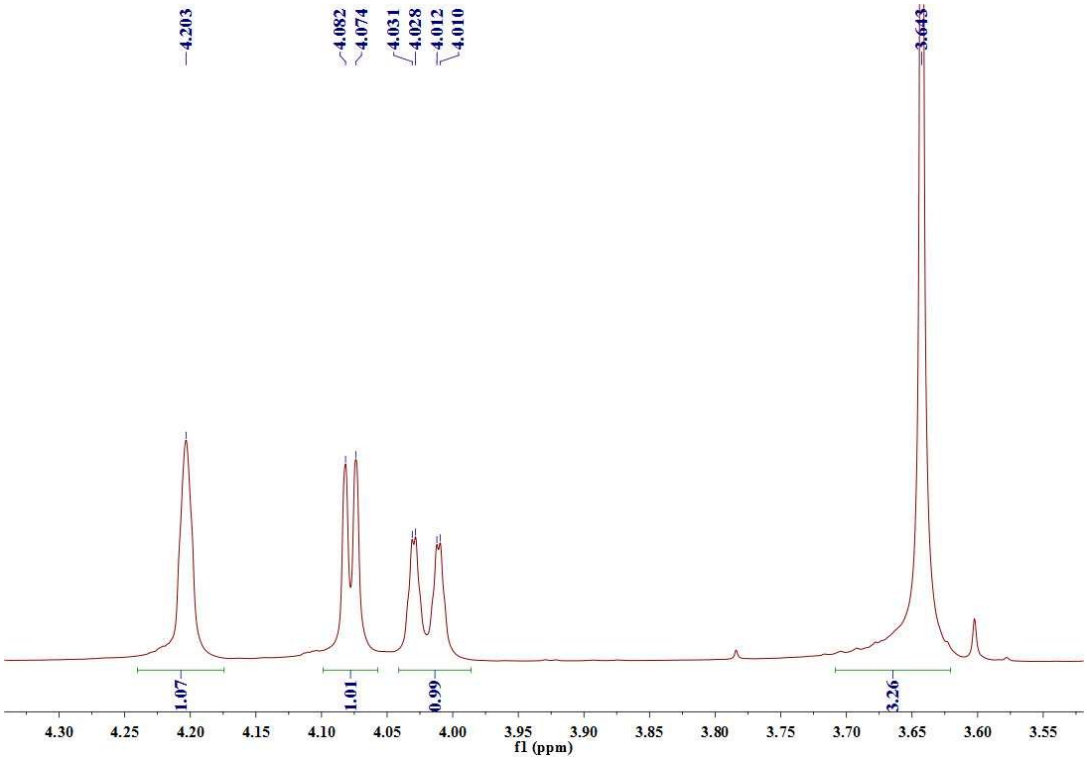


Figure S3. The expansion for the ^{1}H NMR spectrum of compound **1** (δ: 3.50–4.30).

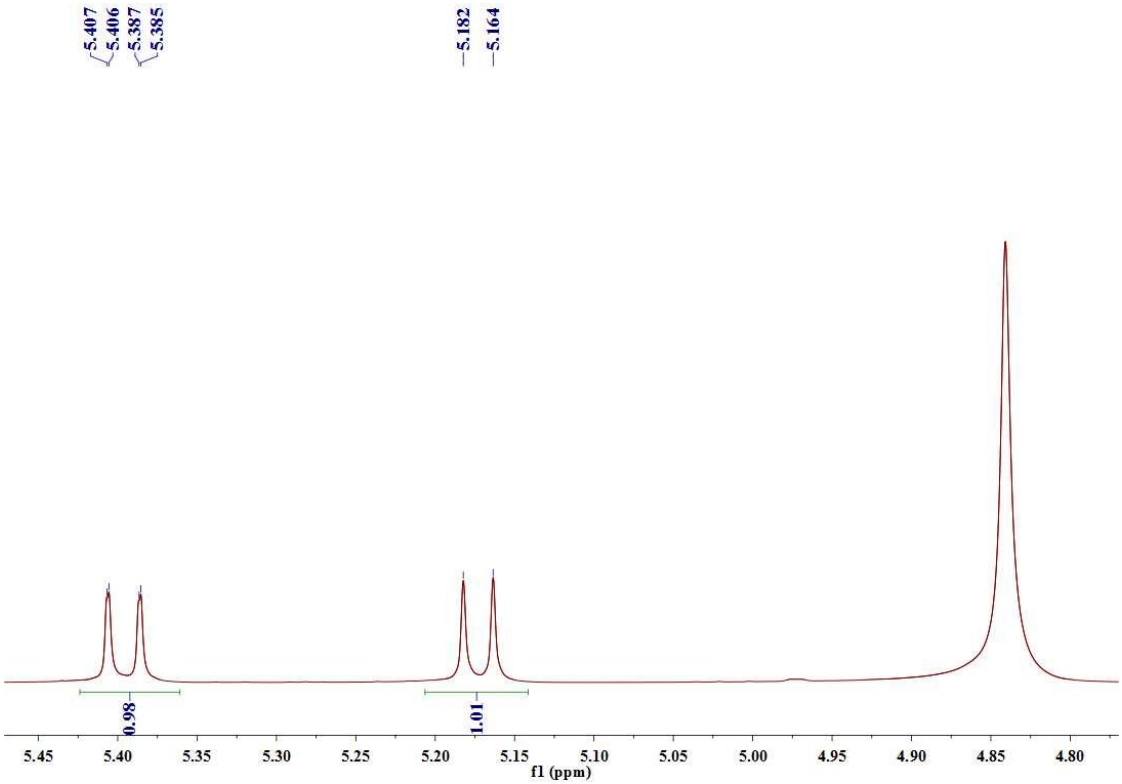


Figure S4. The expansion for the ^{1}H NMR spectrum of compound **1** (δ: 4.80–5.45).

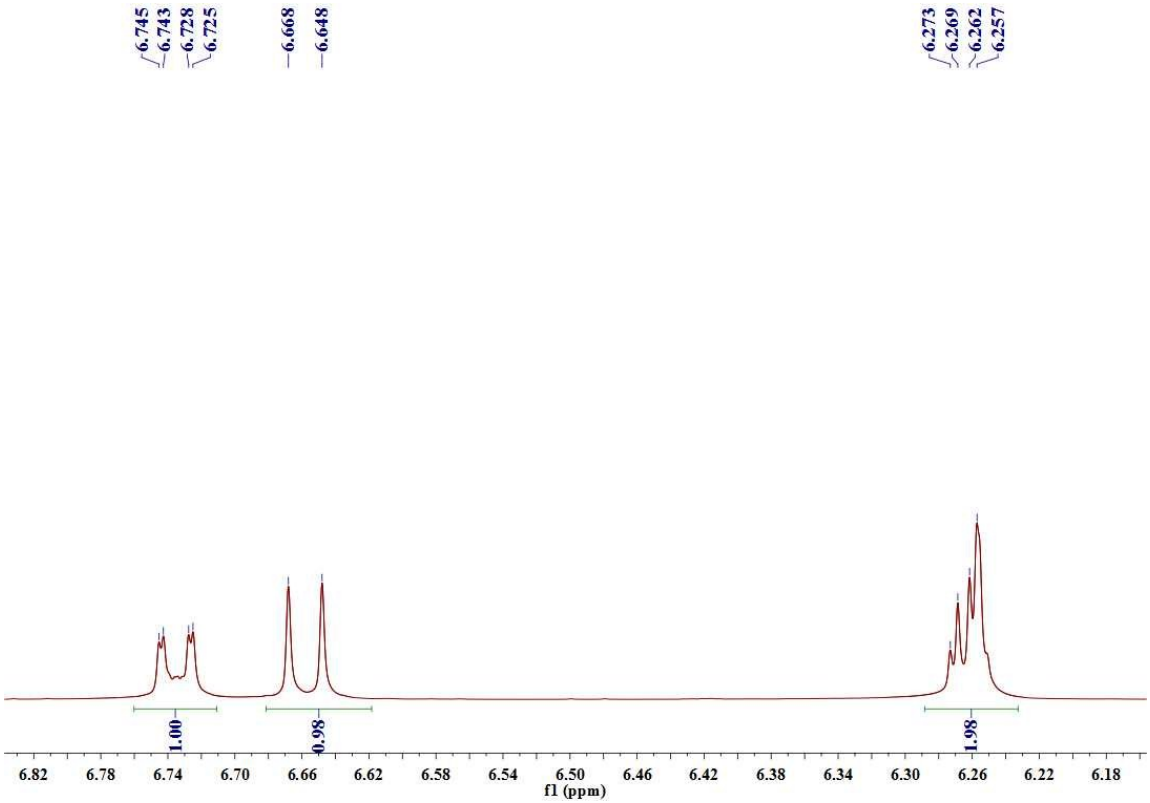


Figure S5. The expansion for the ^{1}H NMR spectrum of compound **1** (δ: 6.18–6.82).

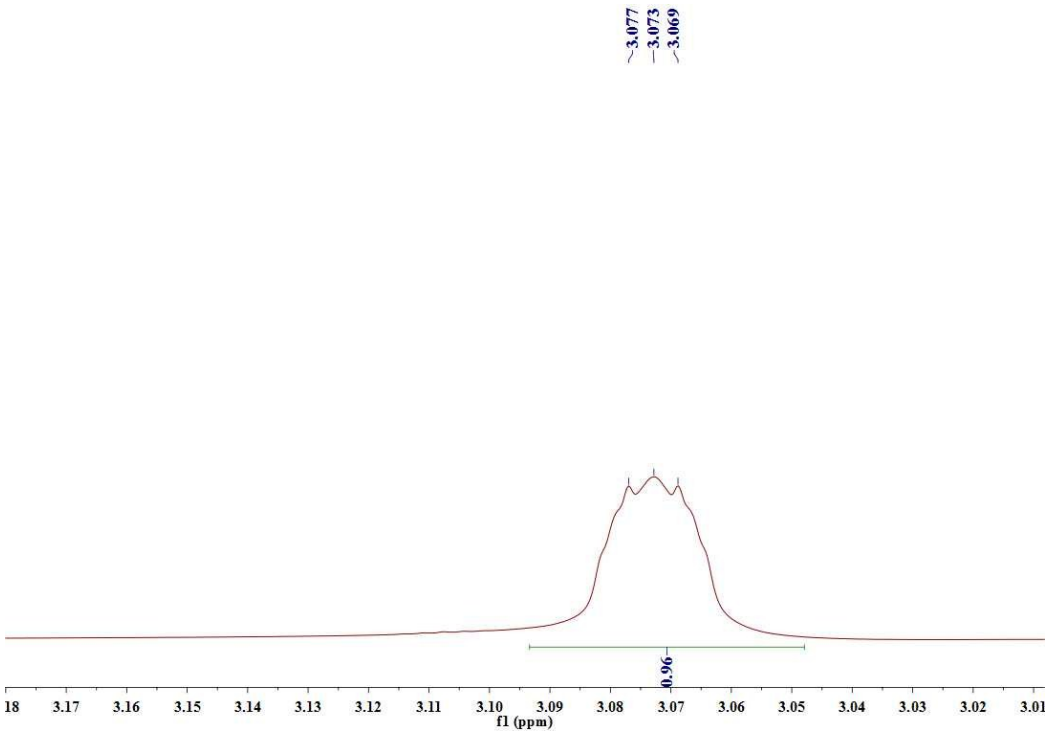


Figure S6. The expansion for the ^{1}H NMR spectrum of compound **1** (δ: 3.01–3.18).

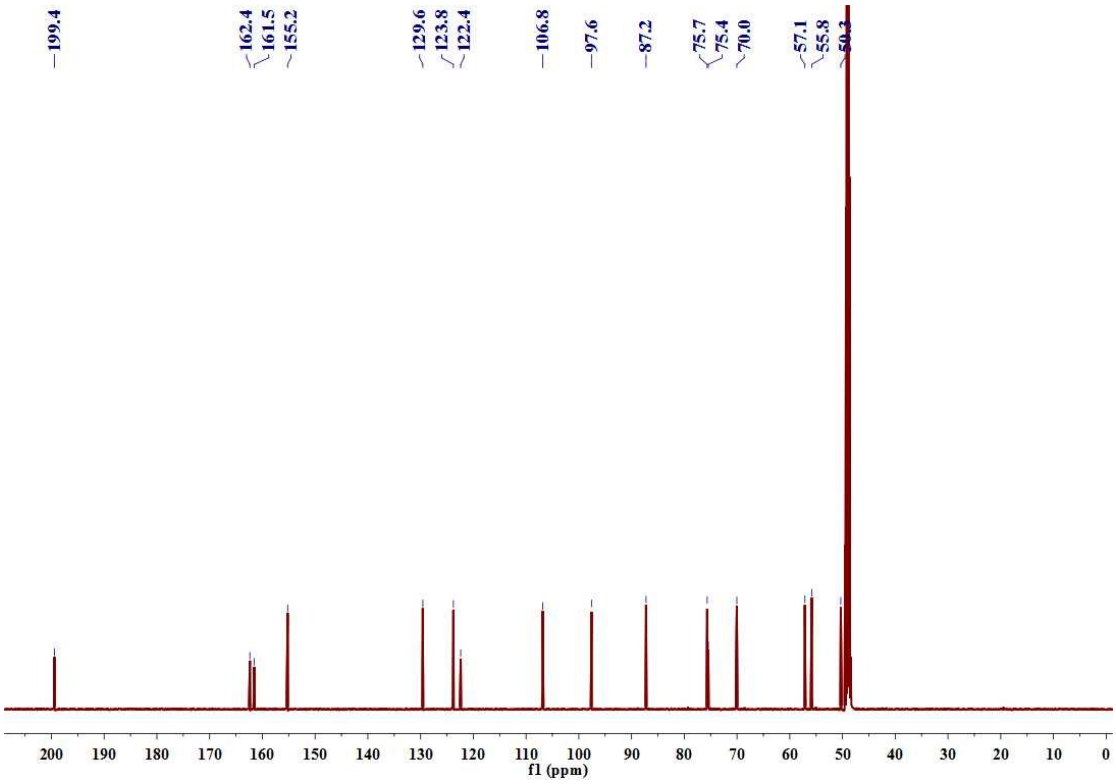


Figure S7. ^{13}C NMR spectrum of compound **1** (125 MHz, CD_3OD).

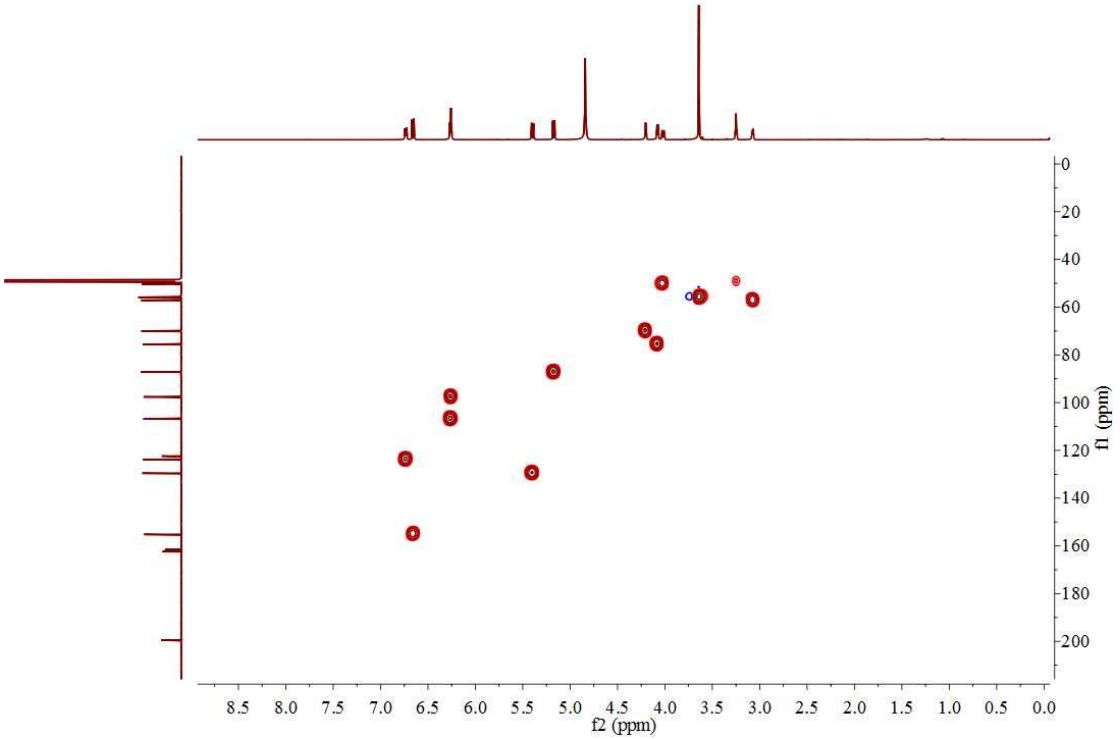


Figure S8. HSQC spectrum of compound **1**.

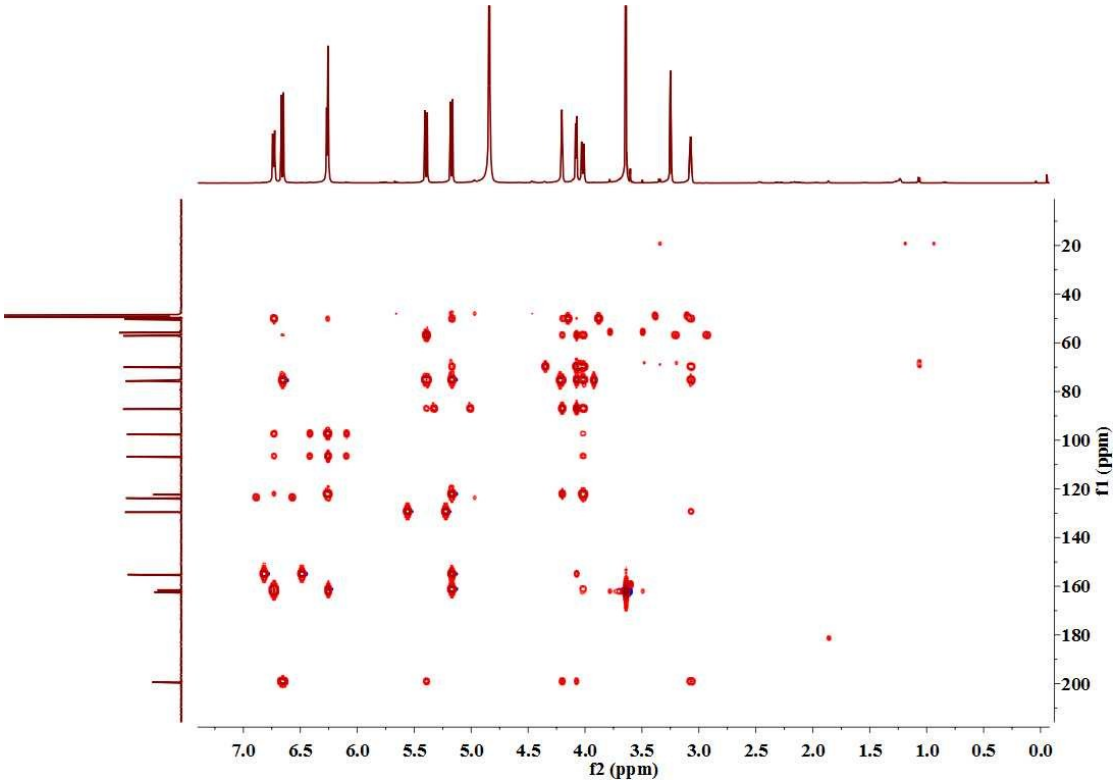


Figure S9. HMBC spectrum of compound **1**.

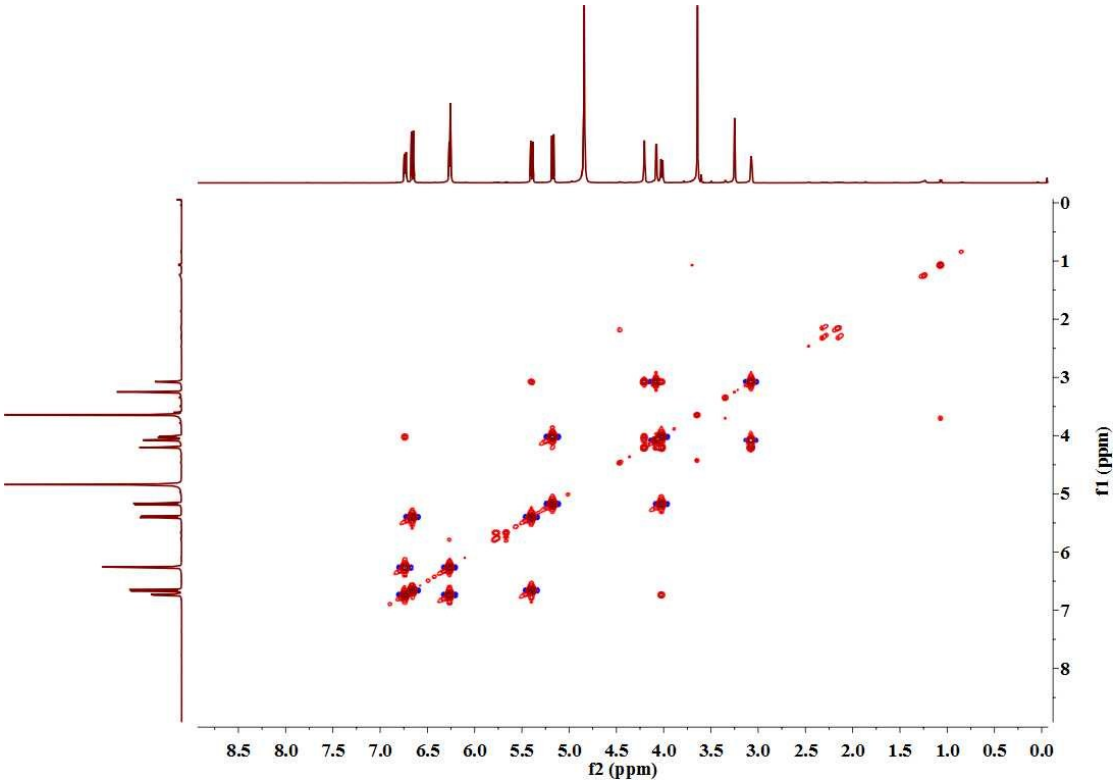


Figure S10. ^{1}H-^{1}H COSY spectrum of compound **1**.

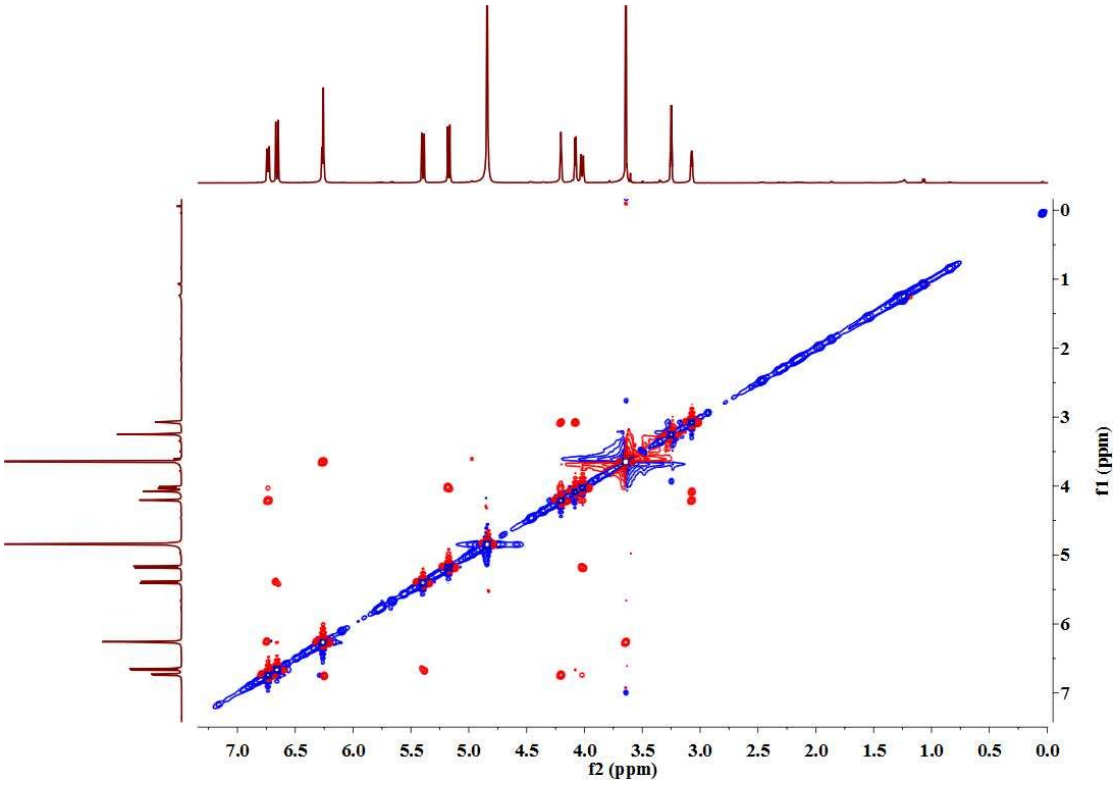
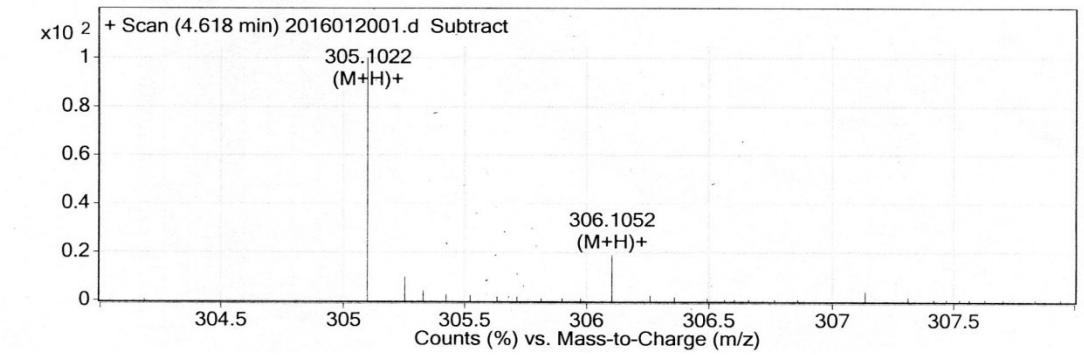


Figure S11. ROESY spectrum of compound **1**.



MS Formula Results: + Scan (4.618 min) Sub (2016012001.d)

m/z	Ion	Formula	Abundance
305.1022	(M+H)+	C16 H17 O6	397488.8

Best	Formula (M)	Ion Formula	Score	Cross Sco	Mass	Calc Mass	Calc m/z	Diff (ppm)	Abs Diff (ppm)	Mass Match	Abund Match	Spacing Match	DBE
✓	C16 H16 O6	C16 H17 O6	87.63		304.0949	304.0947	305.102	-0.76	0.76	99.98	98.43	49.96	9
	C17 H12 N4 O2	C17 H13 N4 O2	87.62		304.0949	304.096	305.1033	3.63	3.63	99.59	99.03	50	14
	C17 H20 O S2	C17 H21 O S2	83.35		304.0949	304.0956	305.1028	2.09	2.09	99.86	83.61	50	8
	C8 H20 N2 O8 S	C8 H21 N2 O8 S	82.95		304.0949	304.094	305.1013	-2.91	2.91	99.74	82.46	49.98	0

Figure S12. HRESIMS spectrum of compound **1**.

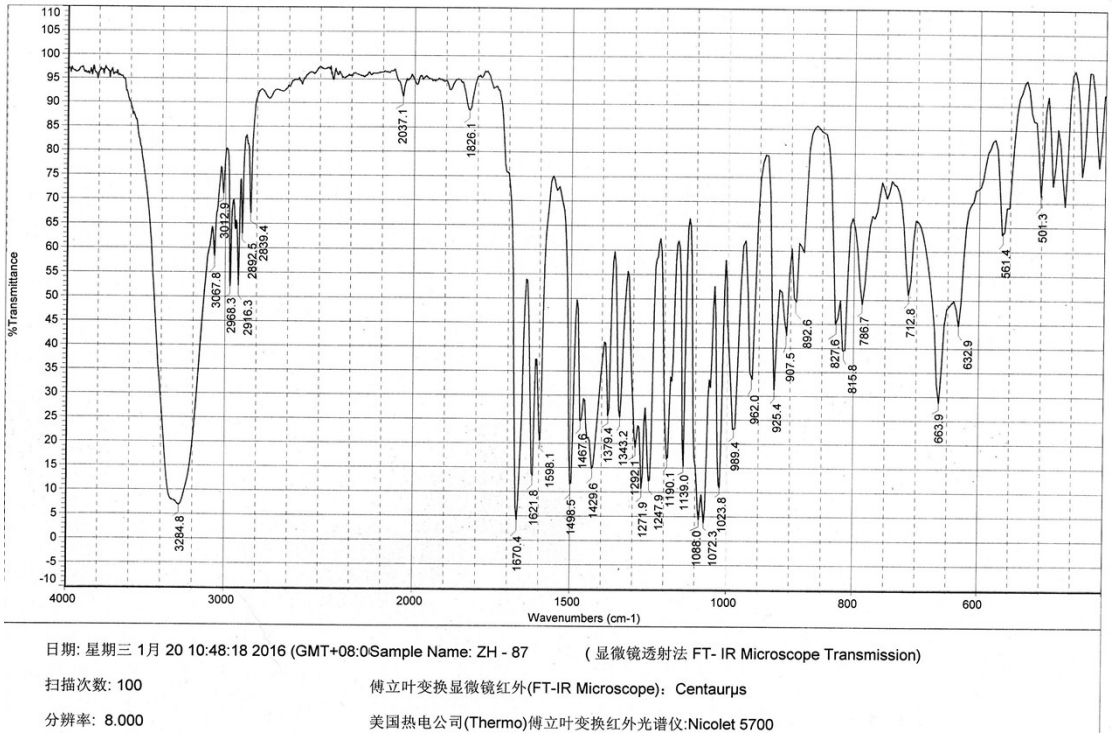


Figure S13. IR spectrum of compound **1**.

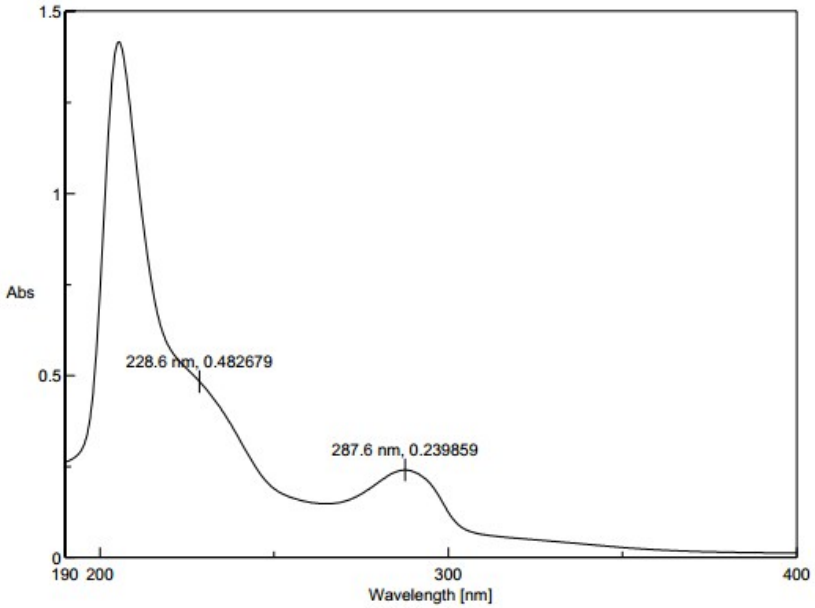


Figure S14. UV spectrum of compound **1**.

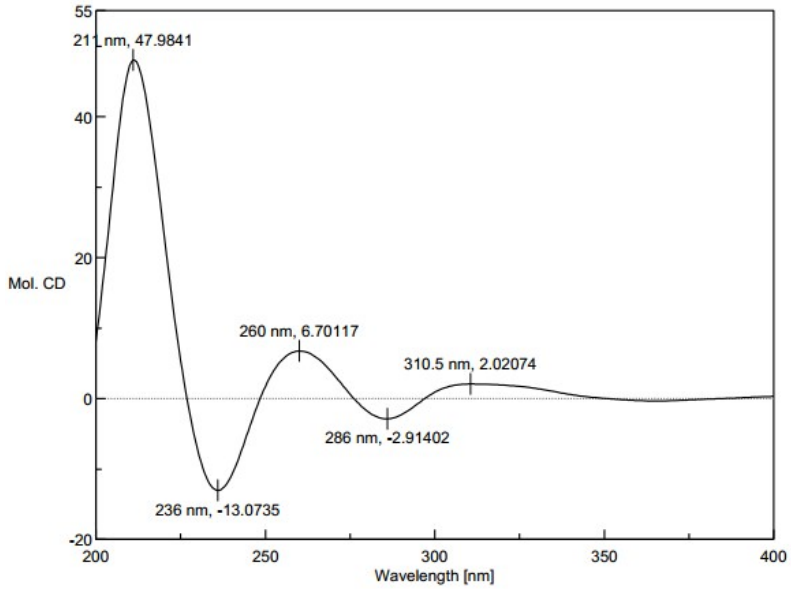


Figure S15. ECD spectrum of compound **1** (in MeOH).