Novel phenylpropanoids and isoflavone glycoside are isolated and identified from the carob pods (*Ceratonia siliqua* L.)

Zhi-Tian Peng[a], Ying-Jie Xia[b], Takuya Yashiro[c], Xiong Gao[b], Tina Ting-Xia Dong[a,b], Karl Wah-Keung Tsim[a,b] & Huai-You Wang[a,d,*]

[a]Shenzhen Key Laboratory of Edible and Medicinal Bioresources, HKUST Shenzhen Research Institute, Shenzhen, China; [b]Division of Life Science and Center for Chinese Medicine, The Hong Kong University of Science and Technology, Hong Kong, China; [c]Cosmetic Division, Nabocul Cosmetics Company Limited, Tokyo, Japan; [d]Institute of Pharmacy, School of Pharmacy, Henan University, Kaifeng, China.

* Corresponding author:

Huai-You Wang, Institute of Pharmacy, School of Pharmacy, Henan University, Kaifeng, China. E-mail: wanghy707@163.com

ABSTRACT:

Two new phenylpropanoids (**1** and **2**) and one new isoflavone glycoside (**3**), along with nine known compounds (**4**−**12**), were isolated from the pod of *Ceratonia siliqua* L.. Their chemical structures were elucidated based on extensive spectroscopic analyses (1D and 2D NMR, UV, IR, and HRESIMS) and compared with the literature data. In addition, all isolated compounds were evaluated *in vitro* for inhibitory activity against acetylcholinesterase (AChE). Compounds **4**, **5**, and **12** showed inhibitory

activity against acetylcholinesterase (AChE) with IC_{50} values ranging from 15.0 to 50.2 μM.

KEYWORDS: *Ceratonia siliqua* L.; phytochemical; phenylpropanoid; isoflavone; acetylcholinesterase inhibition

Contents

1. General experimental procedures

The NMR spectra were measured on an AVANCE III 600 MHZ Nuclear Magnetic Resonance Spectrometry operating at 600 MHz for ^{1}H and 150 MHz for ^{13}C. The HRESIMS data were recorded on maXis impact ESI-Q-TOF-MS (Bruker Daltonics, USA). IR spectra were run on a Thermo/Nicolet Nexus 470 FT-IR spectrophotometer (Waltham, MA, USA) with KBr pellets. UV spectra were acquired on a Shimadzu UV-2450 UV-visible spectrophotometer (Tokyo, Japan). Optical rotations were measured using a Rudolph Autopol IV automatic polarimeter (Hackettstown, NJ, USA). ODS (40-63 μm, Merck, Germany), Sephadex LH-20 (Pharmacia), Silica gel (200−300 mesh, Qingdao Marine Chemical Inc., Qingdao, People's Republic of China) and macroporous adsorption resin D101 (Cangzhou Bon Adsorber Technology Co., Ltd.) were used for open column chromatography. Semi-preparative HPLC was performed on a NP7005C pump system (Jiangsu Hanbon Science & Technology Co., Ltd, China), equipped with a NU3000C ultraviolet detector monitoring at 210 and 254 nm. Semi-preparative RP-HPLC column (YMC-Pack ODS-A, 250 × 20 mm, 5 μm) and semi-preparative RP-HPLC column (YMC-Pack ODS-A, 250 × 10 mm, 5 μm) were applied to the isolation of compounds. Fractions were monitored by TLC, and spots on precoated silica gel GF_{254} plates were visualized under a UV lamp at 254 nm and sprayed with reagent of 10% vitriol in EtOH. All purified compounds submitted for bioassay were at least 95% pure as judged by HPLC and supported by ^{1}H NMR analysis.

2. Extraction and isolation

The dried powder pod of *C siliqua* (20.0 kg) after removal of seeds was refluxed with 95% aqueous EtOH (140 L × 2, each for 1.5 h) and 70% aqueous EtOH (140 L, each for 1.5 h), successively. After filtration, the solution was evaporated under reduced pressure to give a residue (12.4 kg) that was then separated by macroporous adsorption resin D101 CC eluting successively with pure water and 95% aqueous EtOH to obtain the EtOH fraction (462.7 g). The EtOH fraction (460.0 g) was then dissolved in 50% aqueous MeOH and partitioned with petroleum ether and EtOAc. The dried EtOAc extract (135.0 g) was subjected to silica gel CC eluted successively with a gradient of petroleum ether–EtOAc (1:1 to 0:1, v/v) and then CH_2Cl_2–MeOH (7:1 to 1:1, v/v) to yield ten fractions (A–J). Fraction B (5.1 g) was isolated by silica gel CC using gradient of petroleum ether–EtOAc (4:1 to 0:1, v/v) as mobile phase to obtain four sub-fractions (B1–B4). Sub-fraction B2 (2.2 g) was repeatedly treated to semi-preparative HPLC (isocratic 24% aqueous MeOH) to produce **12** (6.3 mg, t_R 20.39 min). Fraction C (8.7 g) was fractionated by silica gel CC with mobile phase of gradient of petroleum ether–EtOAc (3:1 to 0:1, v/v) to afford five sub-fractions (C1–C5). Sub-fraction C3 (2.4 g) was subjected to Sephadex LH-20 column with CH_2Cl_2–MeOH (1:1, v/v), followed by semi-preparative HPLC (isocratic 18% aqueous MeOH) to gain **10** (2.9 mg, t_R 16.59 min). Fraction D (10.4 g) was separated by silica gel CC eluted with gradient of petroleum ether–EtOAc (3:1 to 0:1, v/v) to provide three sub-fractions (D1–D3). Sub-fraction D2 (4.5 g) was applied to Sephadex LH-20 column with CH_2Cl_2–MeOH (1:1, v/v), followed by

semi-preparative HPLC (isocratic 37% aqueous MeOH) to obtain sub-fractions (D2a–D2c). Compounds **11** (19.1 mg, t_R 31.4 min) was purified from sub-fractions D2b by semi-preparative HPLC with mobile phase of isocratic 32% aqueous MeOH. Sub-fraction D3 (3.2 g) was sujected to Sephadex LH-20 column using CH_2Cl_2–MeOH (1:1, v/v) as mobile phase, followed by semi-preparative HPLC with mobile phase of isocratic 48% aqueous MeOH to yield **5** (71.6 mg, t_R 22.39 min) and **6** (3.7 mg, t_R 29.61 min). Fraction F (18.1 g) was subjected to ODS CC eluted with a gradient of aqueous MeOH from 15% to 100% to afford five sub-fractions (F1–F5). Sub-fraction F4 (4.2 g) was isolated by Sephadex LH-20 column with mobile phase of CH_2Cl_2–MeOH (1:1, v/v), followed by semi-preparative HPLC with mobile phase of isocratic 36% aqueous MeOH to obtain **01** (23.8 mg, t_R 37.14 min) and **9** (125.5 mg, t_R 30.06 min). Sub-fraction F5 (3.9 g) was treated to Sephadex LH-20 column eluted with mobile phase of CH_2Cl_2–MeOH (1:1, v/v) and further separation was followed by semi-preparative HPLC (isocratic 45% aqueous MeOH) to produce **4** (17.7 mg, t_R26.62 min). Fraction I (20.0 g) was subjected to ODS CC using a stepwise gradient of aqueous MeOH from 15% to 100% to give five sub-fractions (I1–I5). Separation of sub-fraction I2 (3.0 g) by Sephadex LH-20 column using CH_2Cl_2–MeOH (1:1, v/v) as mobile phase was performed and further purification was followed by semi-preparative HPLC (isocratic 21% aqueous MeOH) to yield **7** (13.6 mg, t_R 50.10 min). Sub-fraction I5 (2.7 g) was sujected to Sephadex LH-20 column with mobile phase of CH_2Cl_2–MeOH (1:1, v/v) to give three sub-fractions (I5a–I5c). Compounds **02** (48.3 mg, t_R 27.27 min) and **03** (3.3 mg, t_R 56.42 min) were purified from

sub-fractions I5b, and I5c by semi-preparative HPLC eluting with isocratic 55% and 49% aqueous MeOH, respectively. Fraction J (5.0 g) was separated by Sephadex LH-20 column using mobile phase of CH_2Cl_2–MeOH (1:1, v/v) and further purification was followed by semi-preparative HPLC eluting with isocratic 30% aqueous MeOH to yield **8** (42.6 mg, t_R 21.95 min).

Table S1. ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR Data of **1** in CD_3OD (δ in ppm, J in Hz)

No.	δ_H (J in Hz)	δ_C	No.	δ_H (J in Hz)	δ_C
1	-	127.7	3″,5″	-	146.5
2	7.18, s	111.6	4″	-	140.4
3	-	149.3	7″	-	166.9
4	-	150.6	1′	5.68, d (5.4)	95.9
5	6.80, d (7.8)	116.4	2′	3.51, m	74.1
6	7.06, d (7.8)	124.2	3′	3.51, m	78.0
7	7.62, d (15.6)	147.2	4′	3.46, m	71.3
8	6.39, d (15.6)	115.2	5′	3.70, m	76.3
9	-	169.1	6′	4.31, m	64.4
1″	-	120.6		4.52, d (11.4)	
2″,6″	7.14, s	110.6	OCH_3	3.88, s	56.4

Table S2. ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR Data of **2** in CD_3OD (δ in ppm, J in Hz)

No.	δ_H (J in Hz)	δ_C	No.	δ_H (J in Hz)	δ_C
1	-	134.5	3′,4′	1.18, d (7.2)	19.3
2,6	6.76, s	105.6	1″	4.89, d (7.8)	105.2
3,5	-	154.3	2″	3.48, m	75.7
4	-	136.2	3″	3.43, m	77.8
7	6.61, d (16.2)	134.8	4″	3.42, m	71.3
8	6.30, dt (15.6, 6.0)	124.4	5″	3.22, m	78.3
9	4.71, d, (6.0)	66.0	6"	3.67, dd (12.0, 5.4)	62.5
1′	-	178.5		3.78, dd (12.0, 2.4)	
2′	2.60, m	35.2	OCH_3-3,5	3.86, s	57.0

Table S3. ^{1}H (600 MHz) and ^{13}C (150 MHz) NMR Data of **3** in CD_3OD (δ in ppm, J in Hz)

No.	δ_H (J in Hz)	δ_C	No.	δ_H (J in Hz)	δ_C
2	8.24, s	155.5	3′,5′	7.00, d (8.4)	114.9

3	-	124.7	4′	-	161.4
4	-	182.5	1″	5.09, d, (7.8)	102.0
5	6.69, s	100.2	2″	3.55, m	74.8
6	-	157.7	3″	3.44, m	78.0
7	-	158.7	4″	3.43, m	71.1
8	-	130.7	5″	3.50, m	78.4
9	-	151.4	6"	3.91, dd, (12.0, 2.4)	62.3
10	-	107.7		3.73, dd, (12.0, 5.4)	
1′	-	124.2	OCH_3-4′	3.83, s	55.8
2′,6′	7.50, d (8.4)	131.4	OCH_3-8	3.91, s	62.4

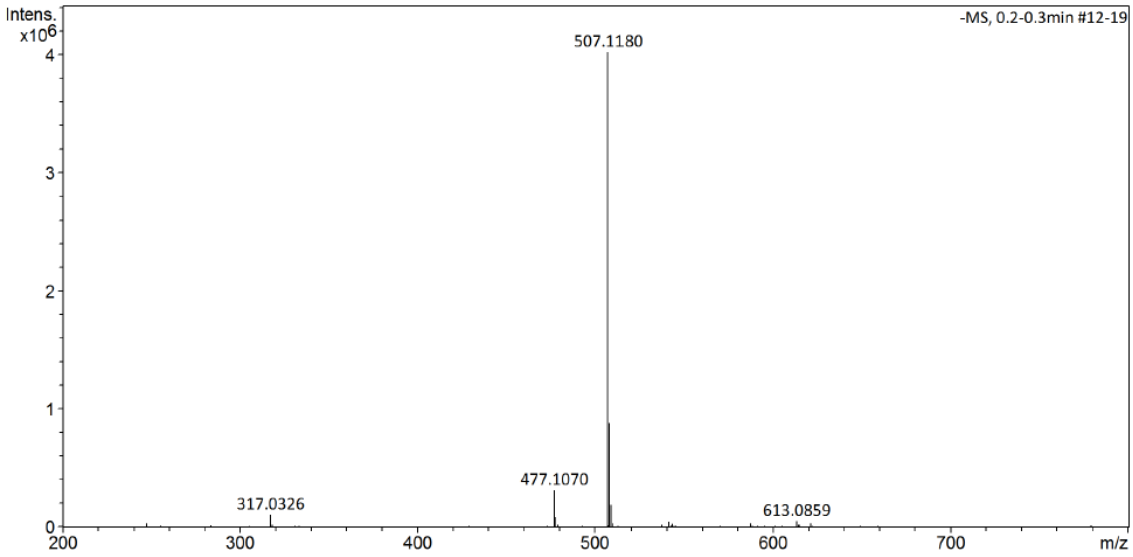

Figure S1 HRESIMS spectrum of compound **1**

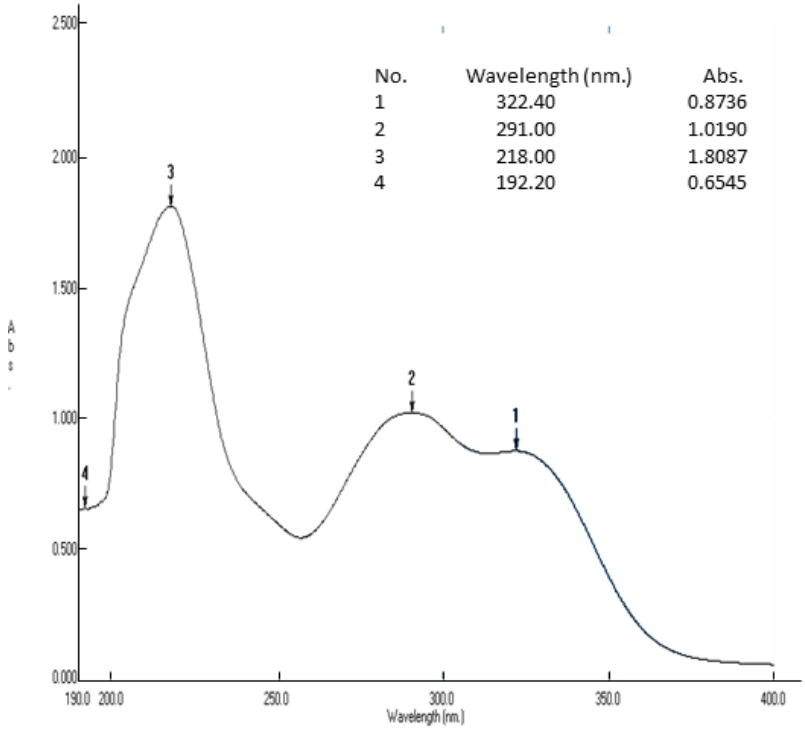

Figure S2 UV spectrum of compound **1**

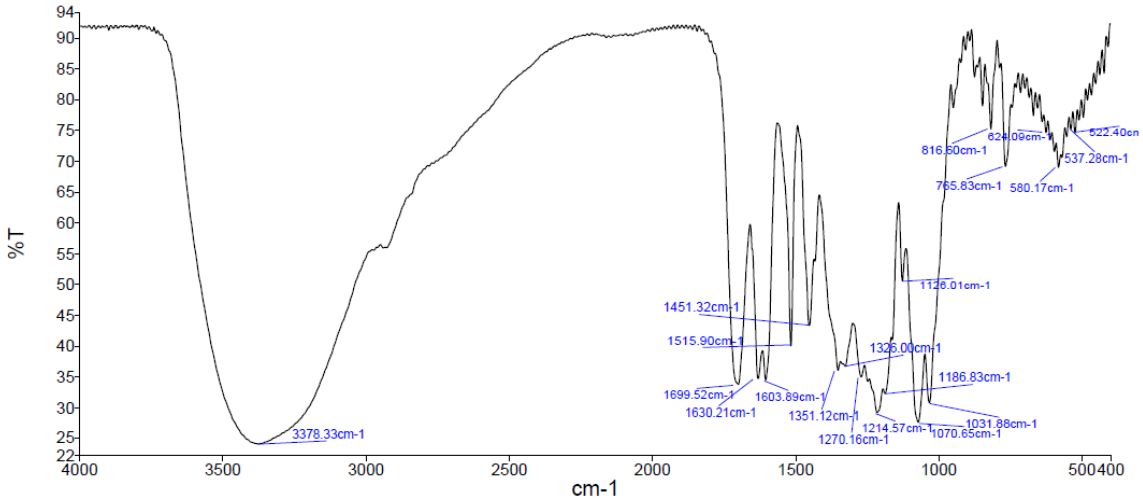

Figure S3 IR spectrum of compound **1**

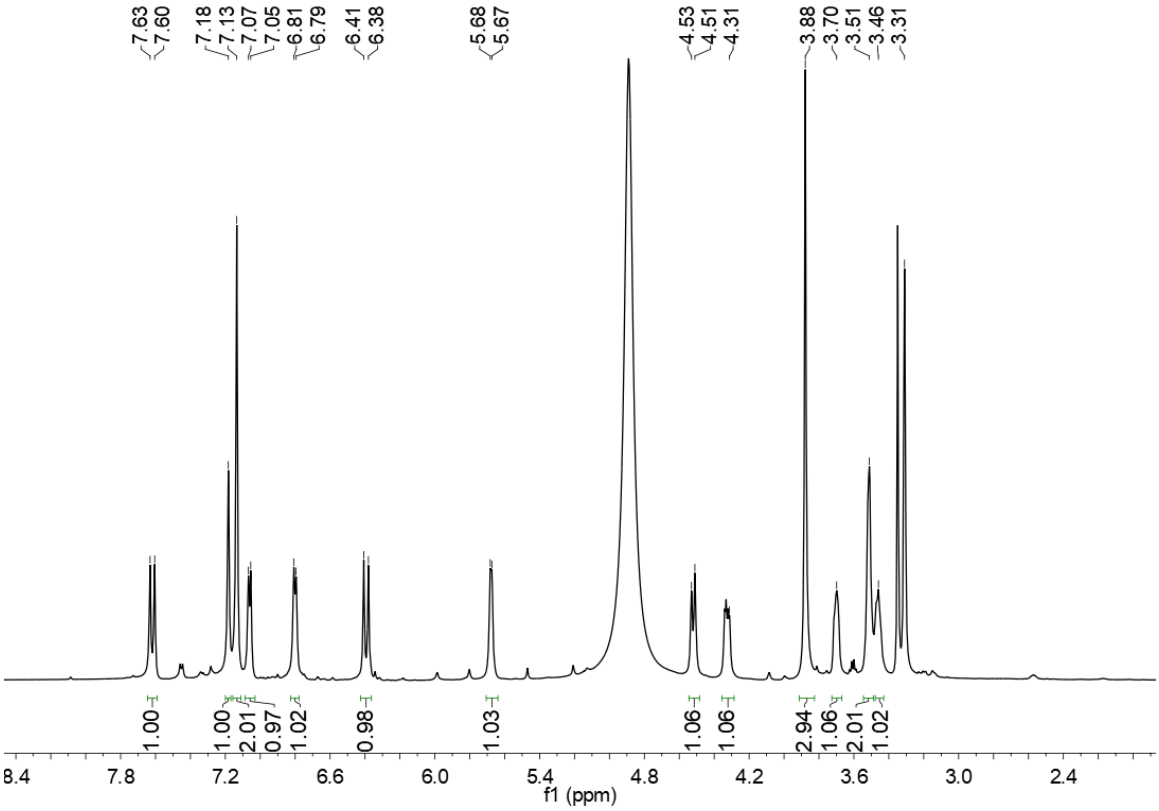

Figure S4 ^{1}H NMR spectrum of compound **1** in CD_3OD (600 MHz）

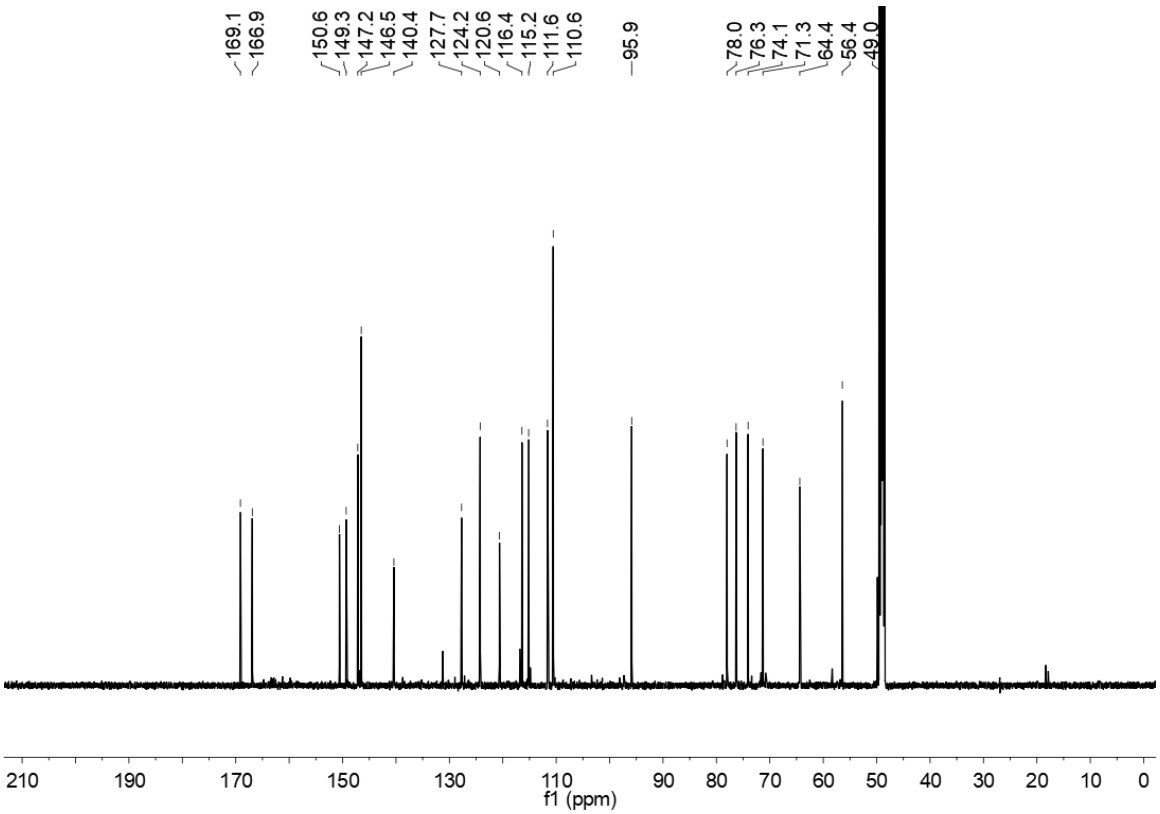

Figure S5 ^{13}C NMR spectrum of compound **1** in CD_3OD (150 MHz）

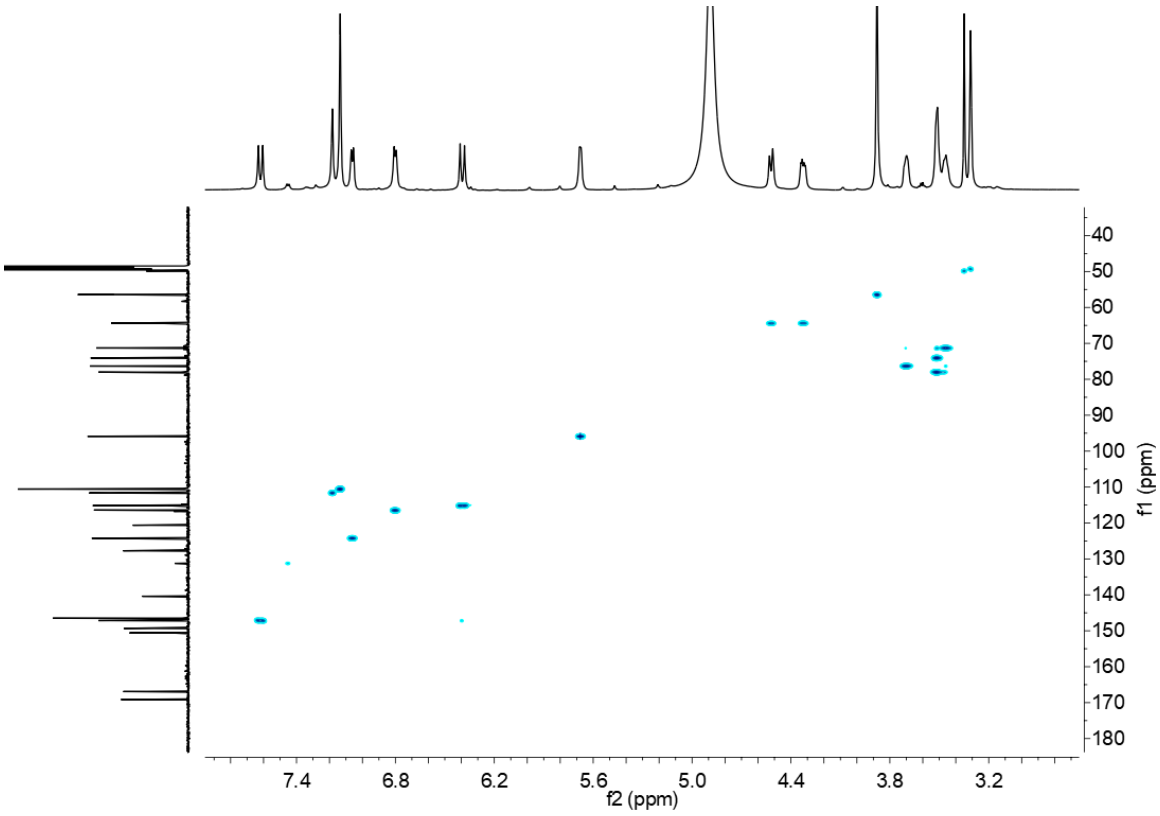

Figure S6 HSQC spectrum of compound **1** in CD_3OD

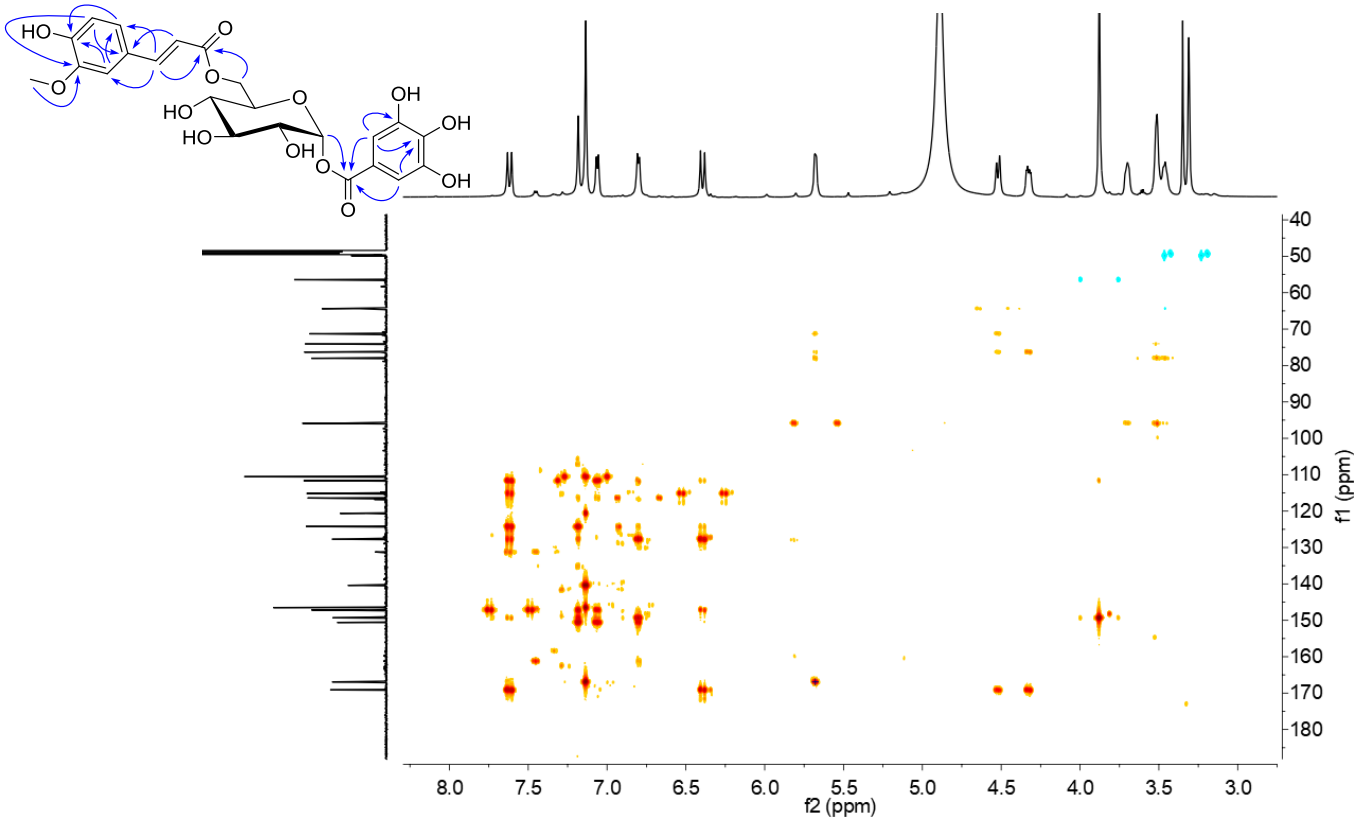

Figure S7 HMBC spectrum of compound **1** in CD_3OD

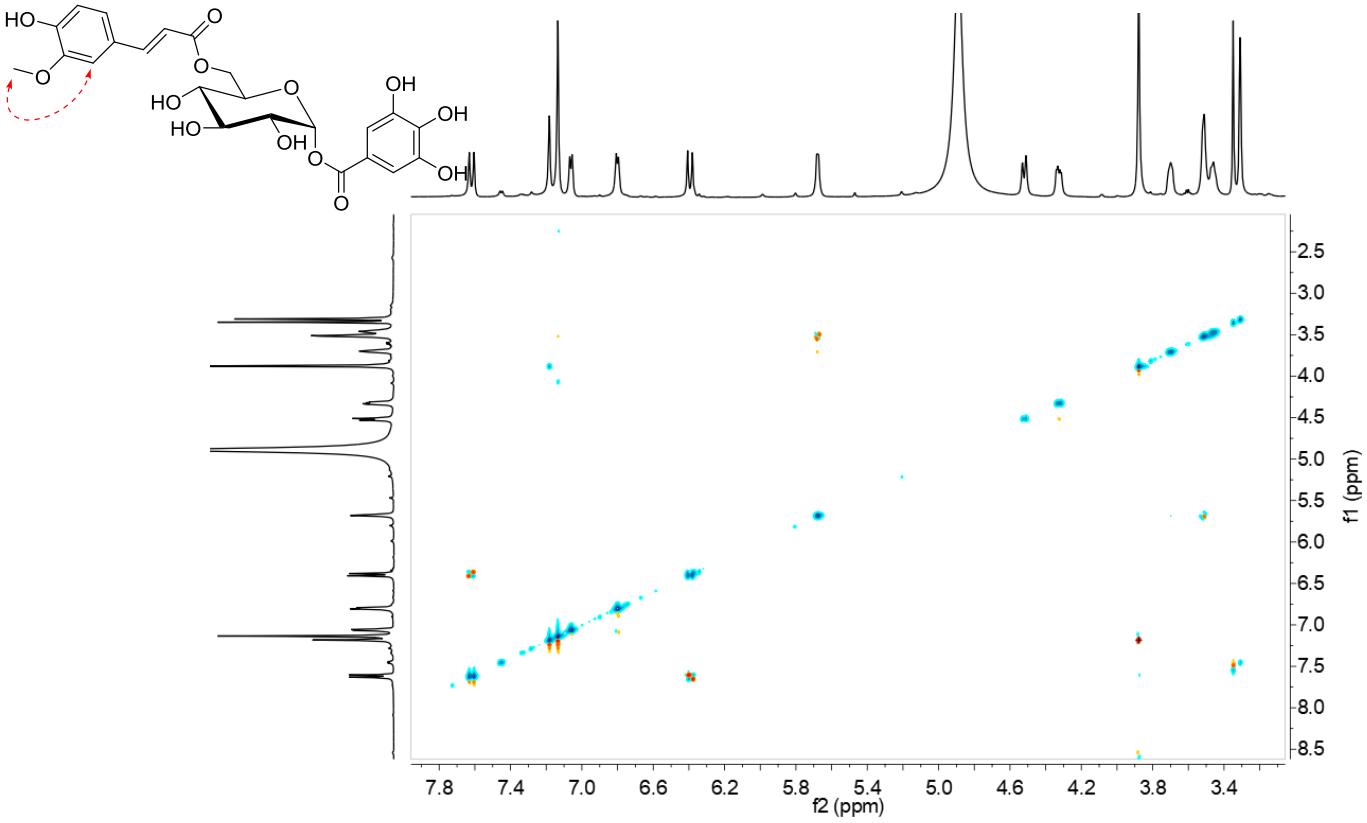

Figure S8 NOESY spectrum of compound **1** in CD_3OD

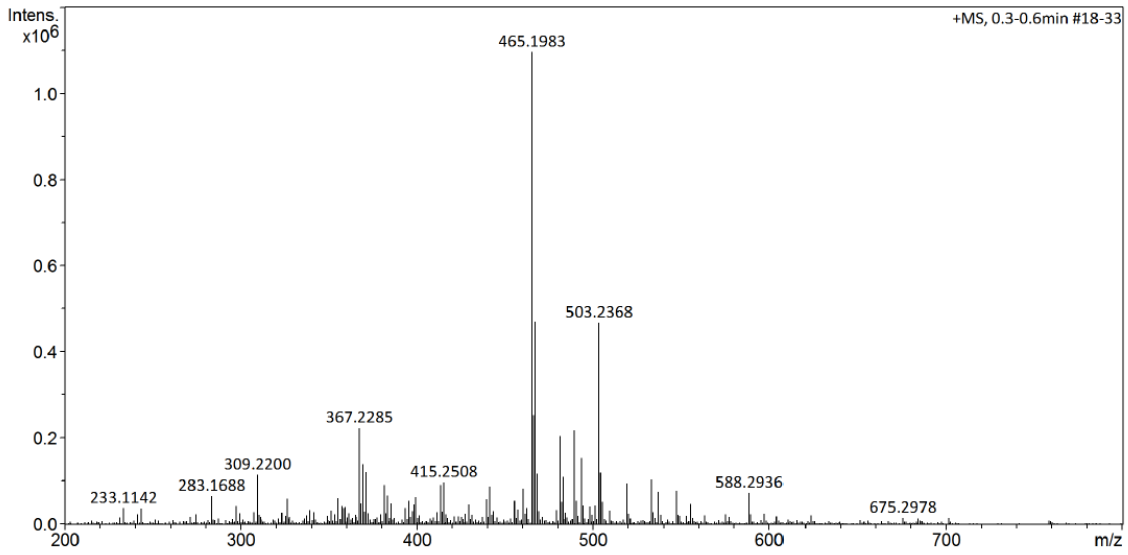

Figure S9 HRESIMS spectrum of compound **2**

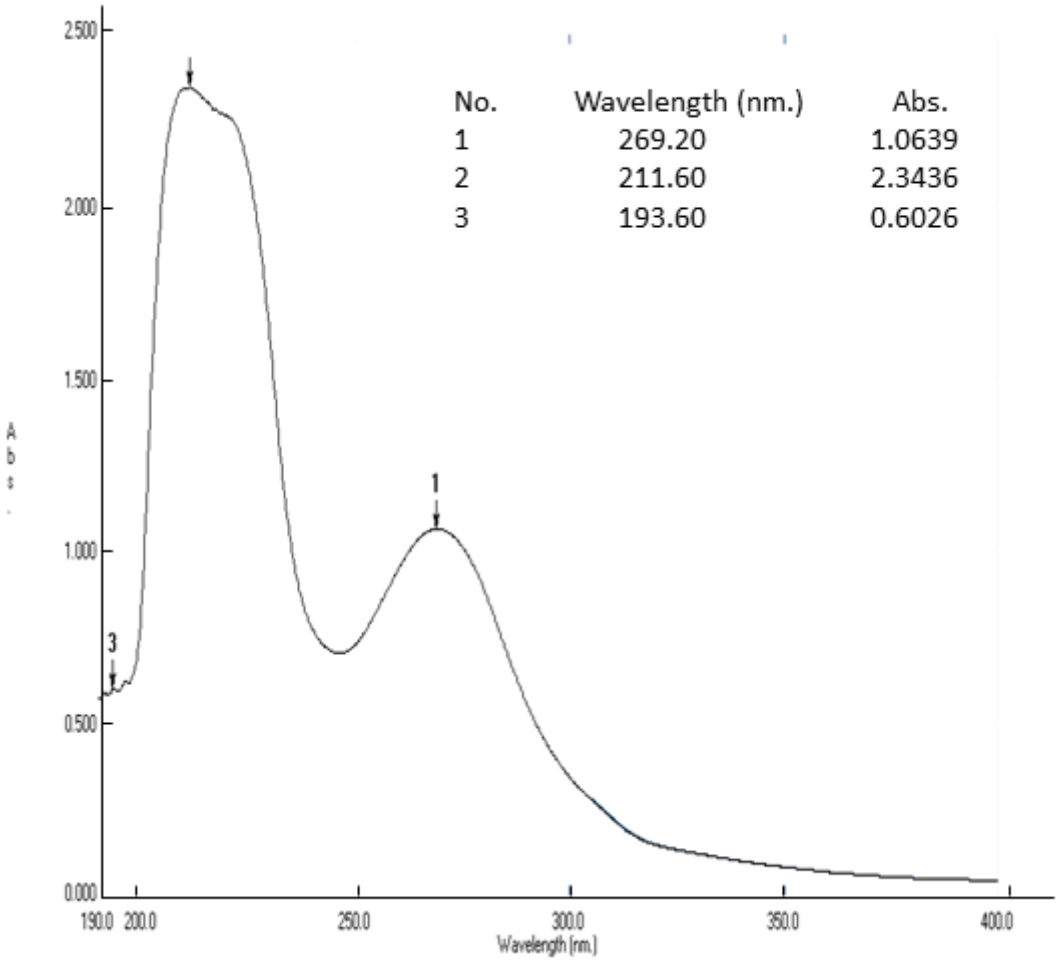

Figure S10 UV spectrum of compound **2**

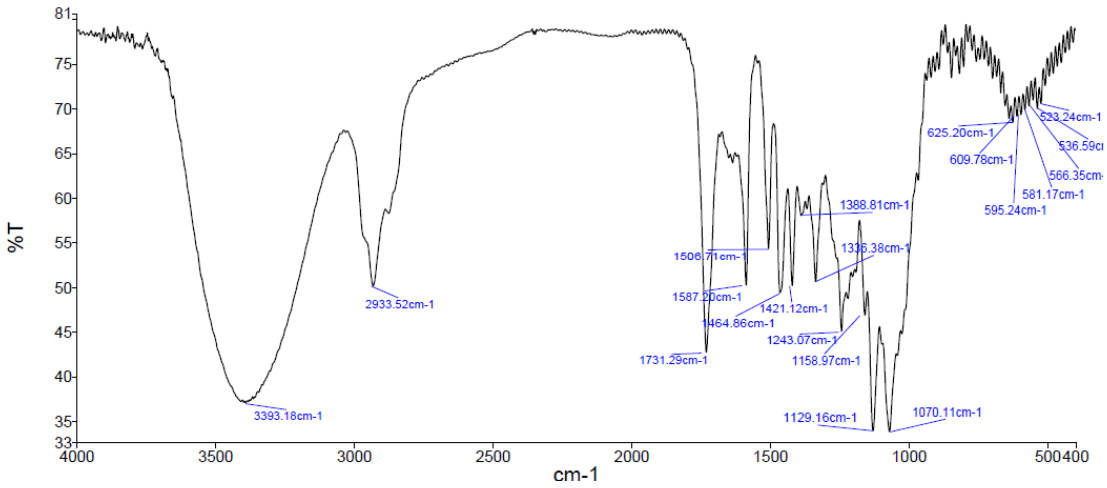

Figure S11 IR spectrum of compound **2**

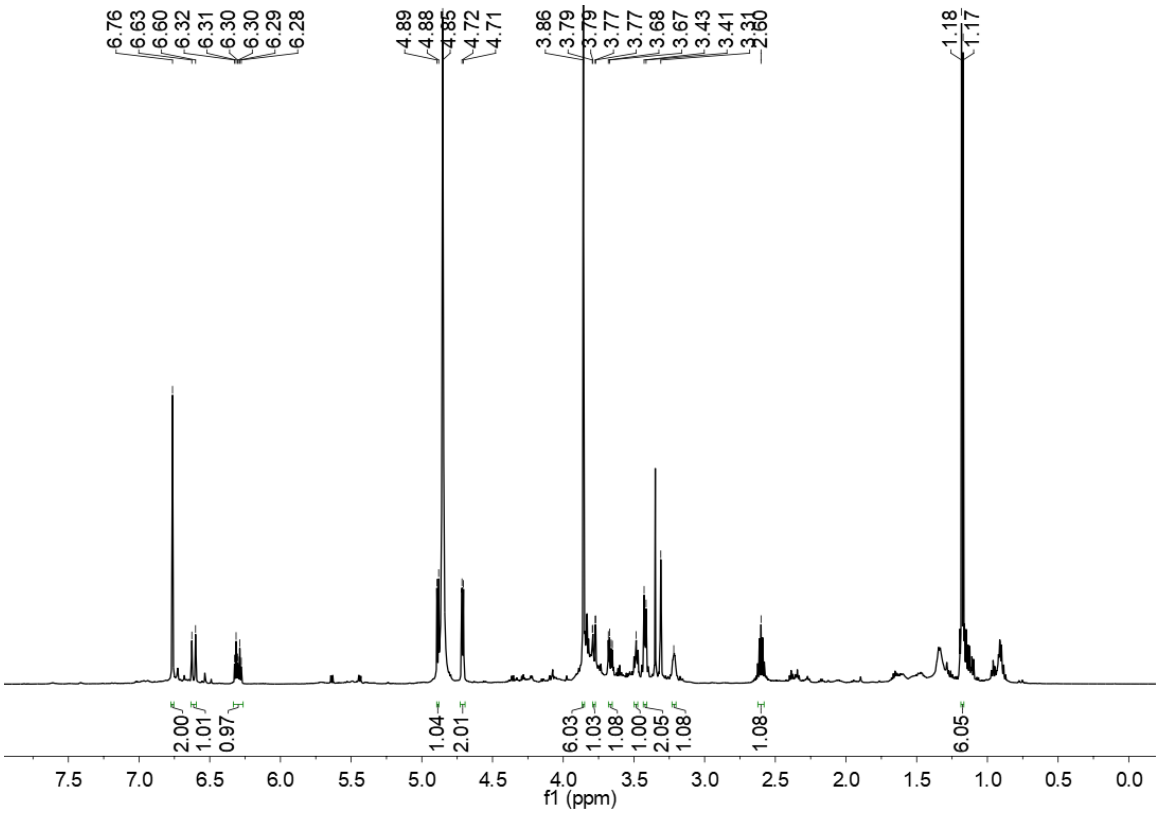

Figure S12 ^{1}H NMR spectrum of compound **2** in CD_3OD (600 MHz）

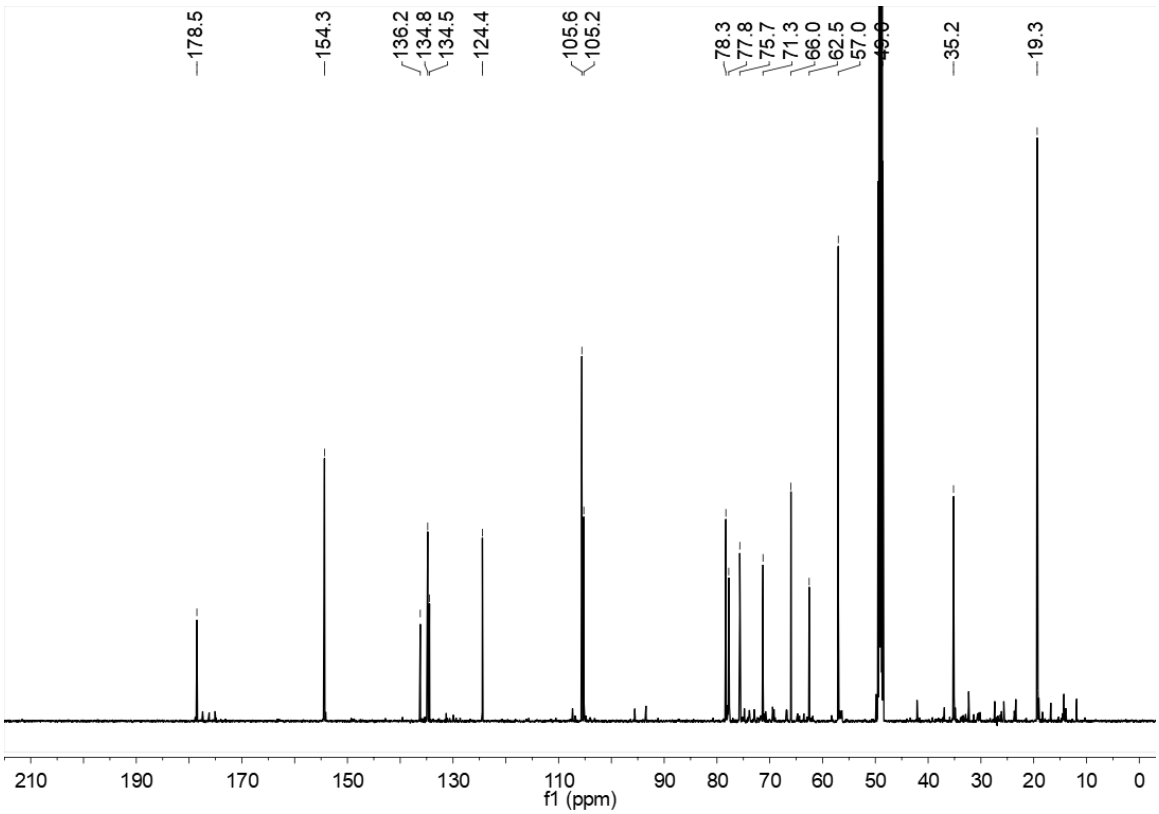

Figure S13 ^{13}C NMR spectrum of compound **2** in CD_3OD (600 MHz）

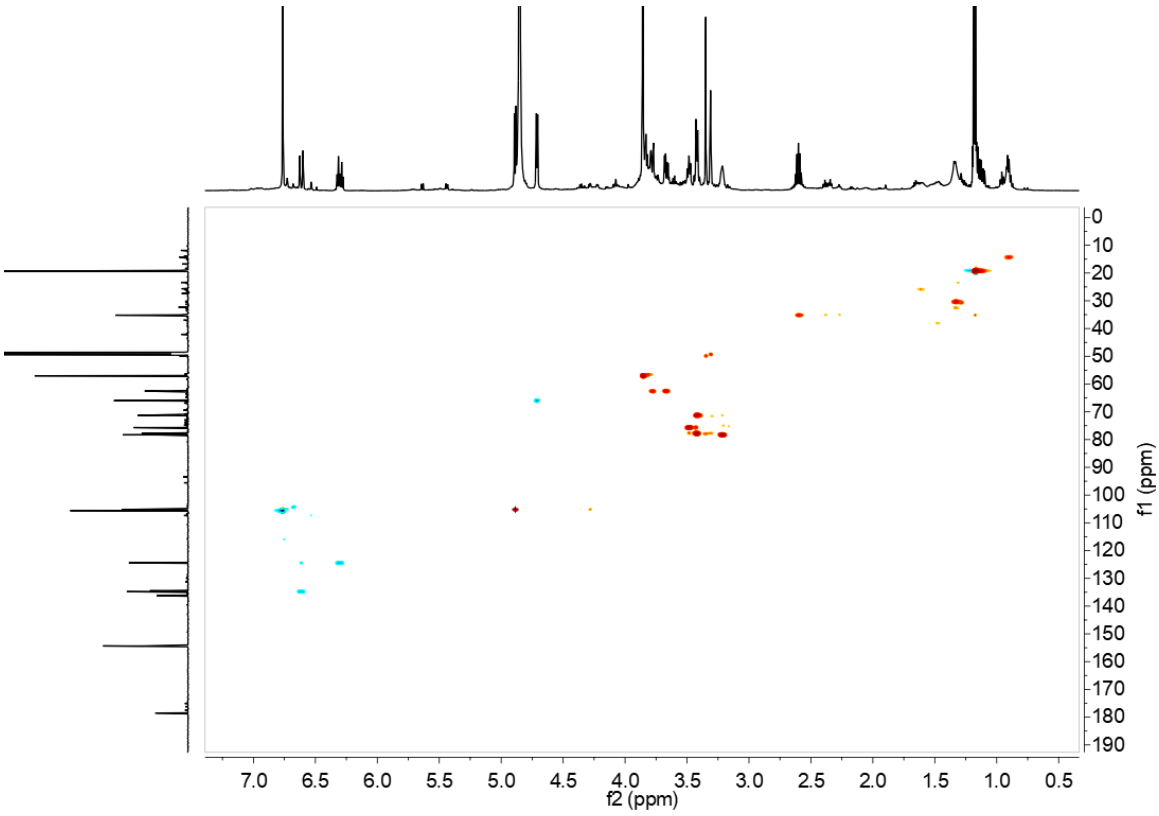

Figure S14 HSQC spectrum of compound **2** in CD_3OD

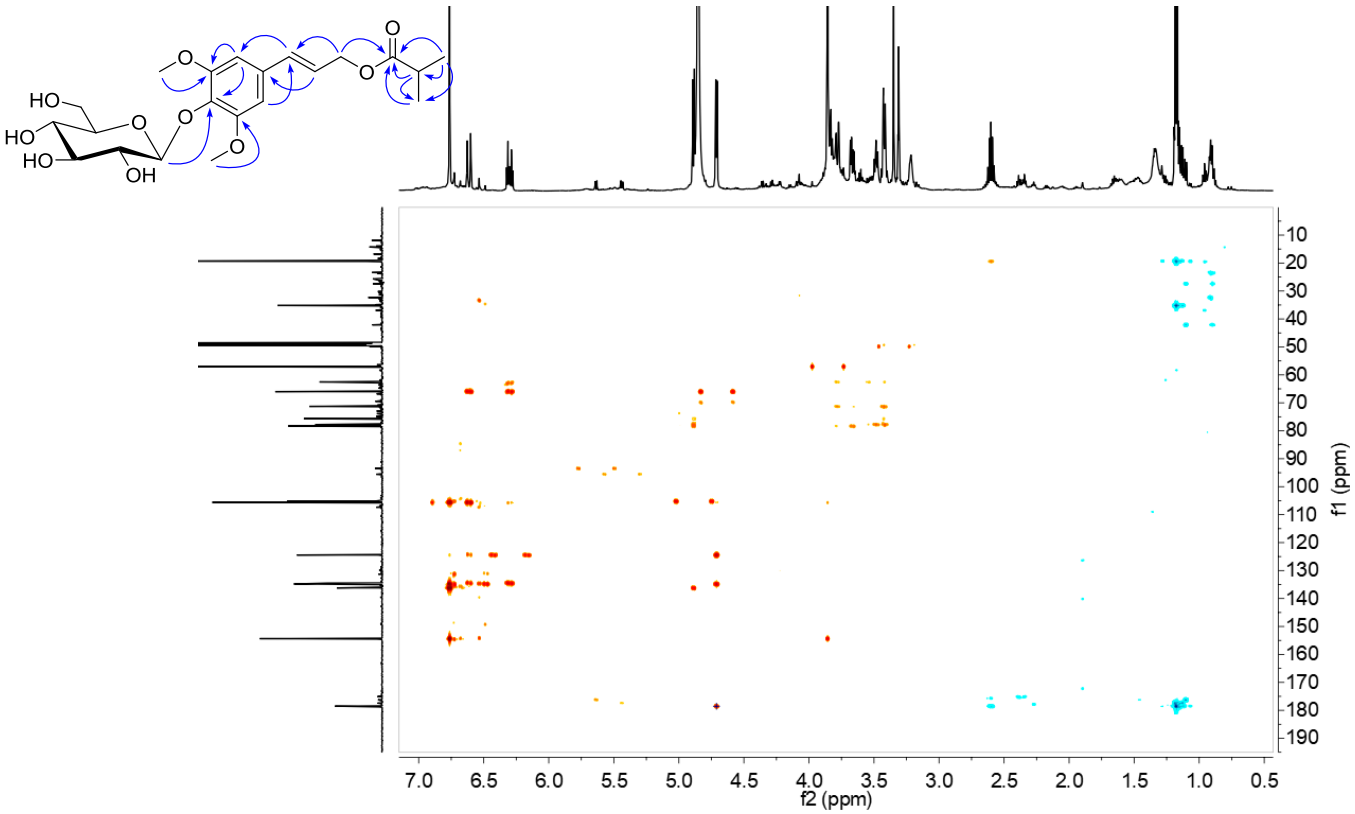

Figure S15 HMBC spectrum of compound **2** in CD_3OD

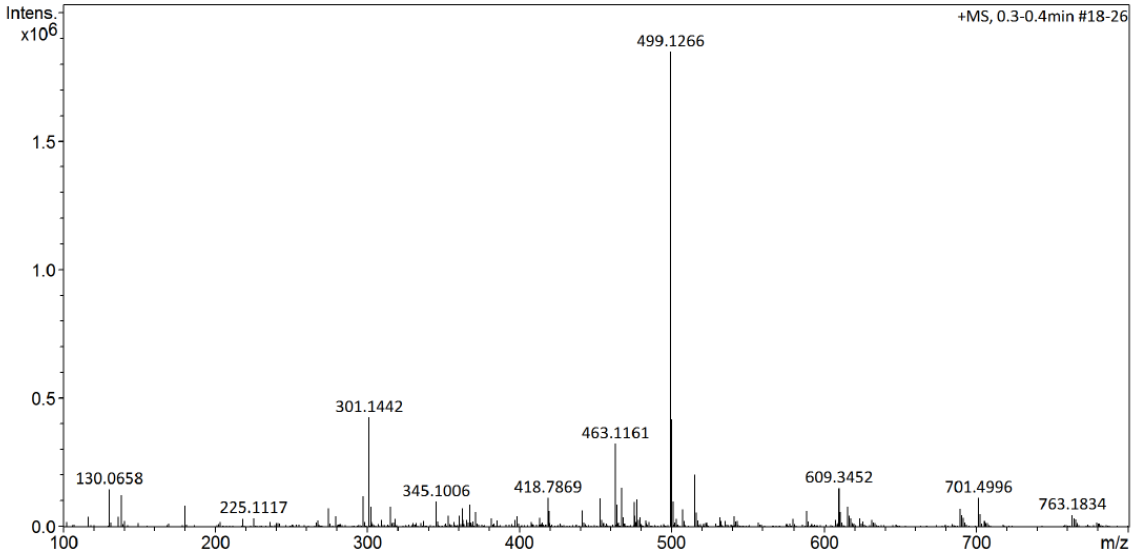

Figure S16 HRESIMS spectrum of compound **3**

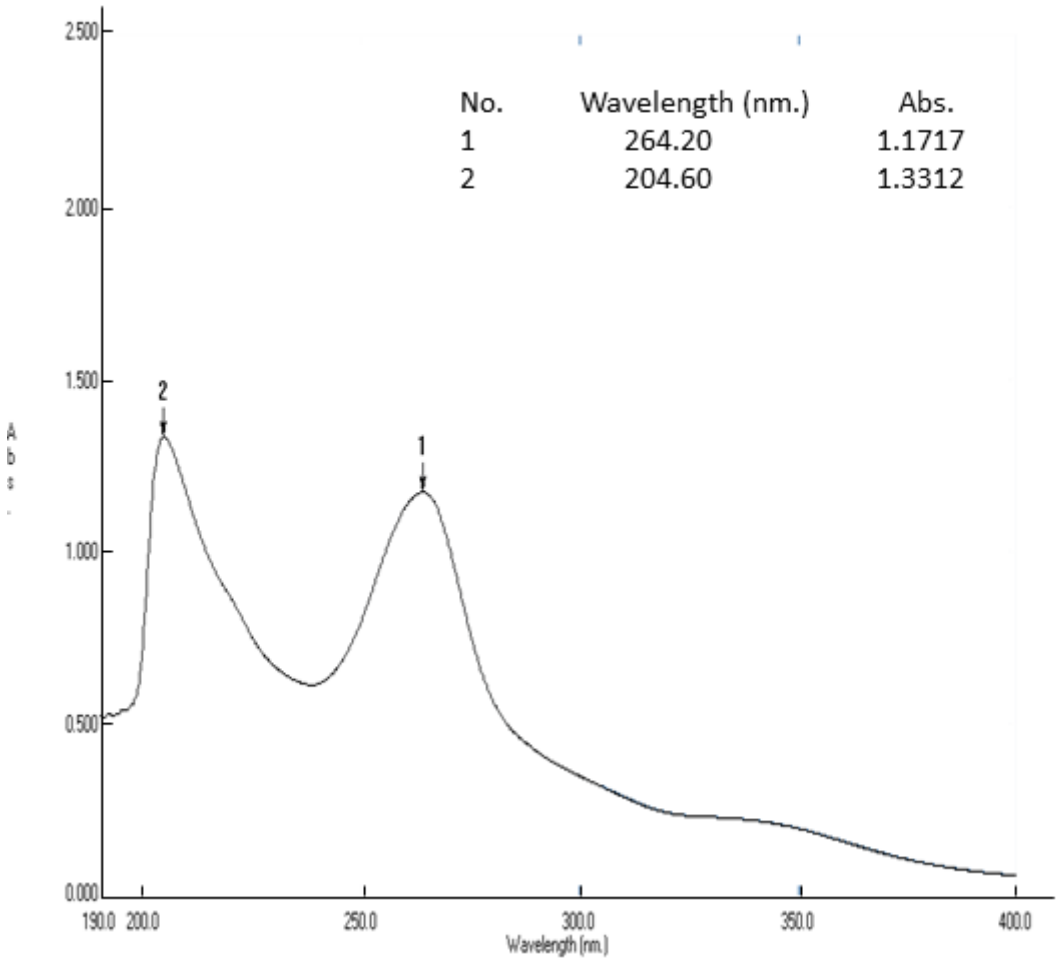

Figure S17 UV spectrum of compound **3**

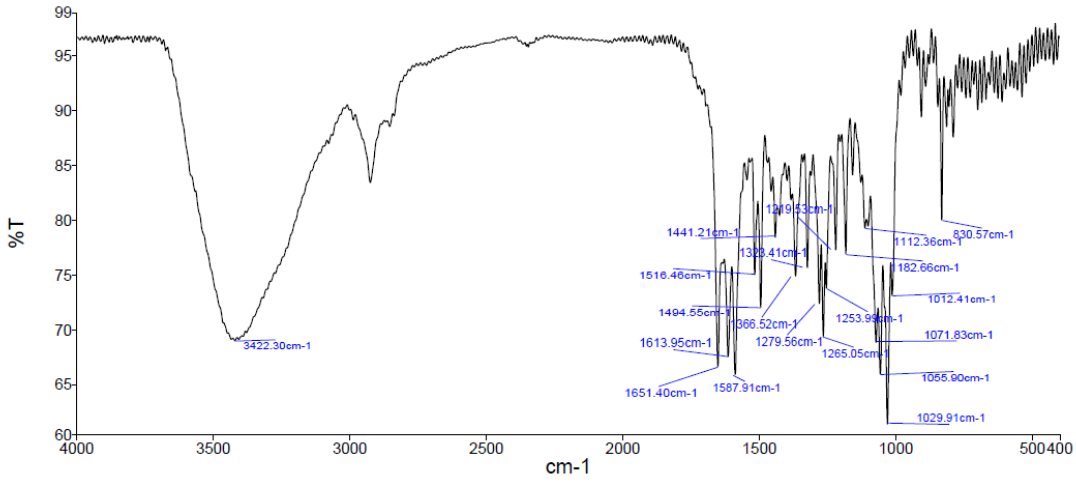

Figure S18 IR spectrum of compound **3**

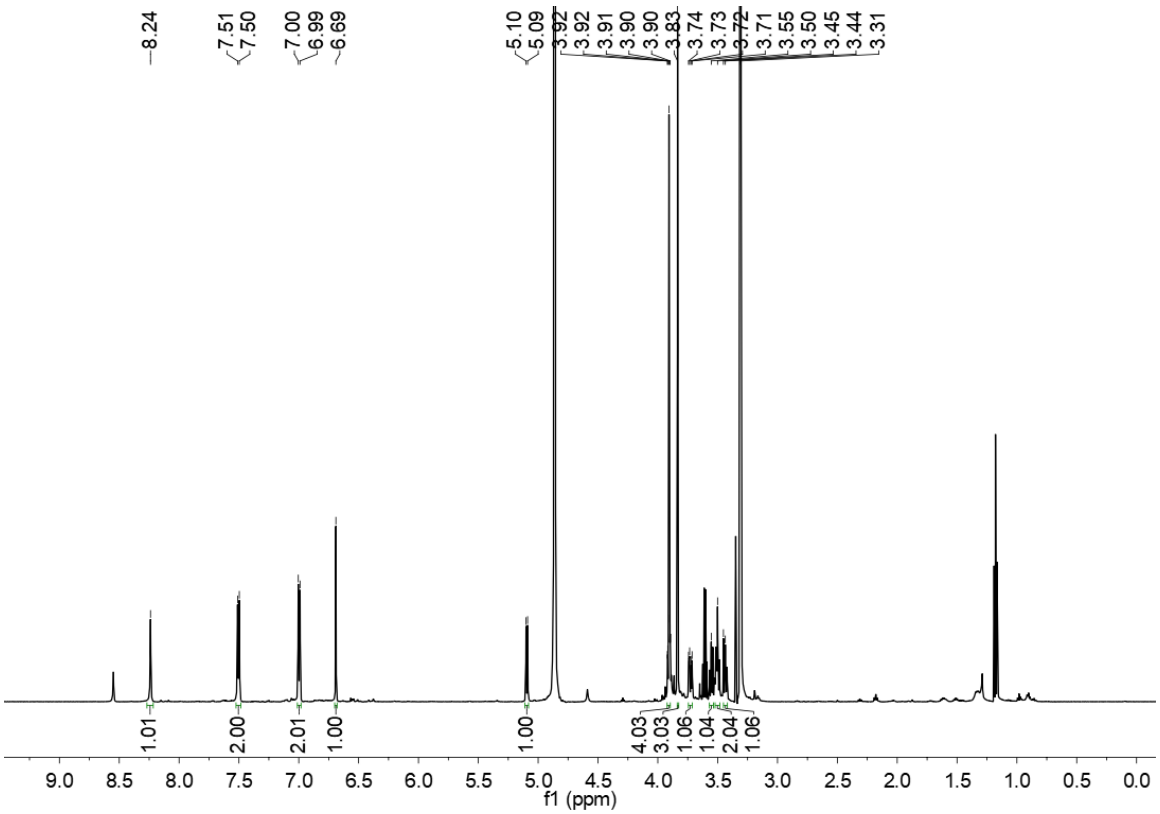

Figure S19 ^{1}H NMR spectrum of compound **3** in CD_3OD (600 MHz）

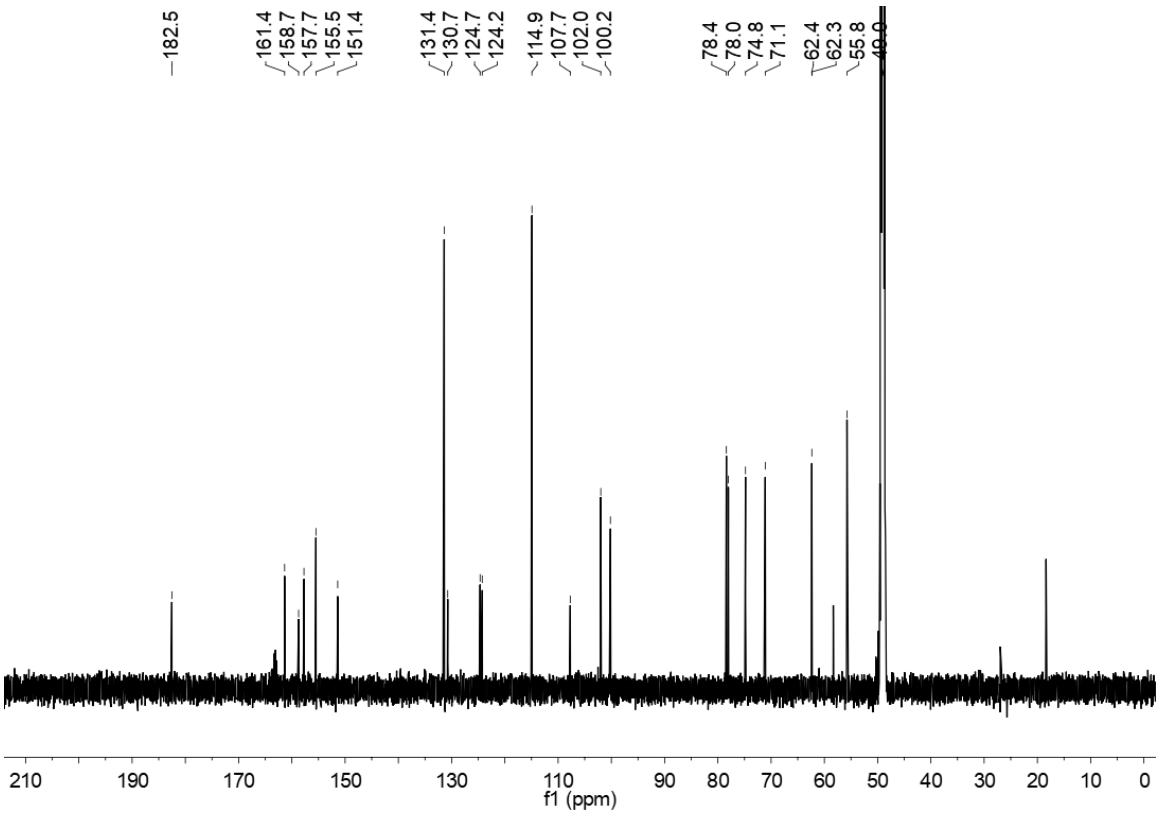

Figure S20 ^{13}C NMR spectrum of compound **3** in CD_3OD (150 MHz）

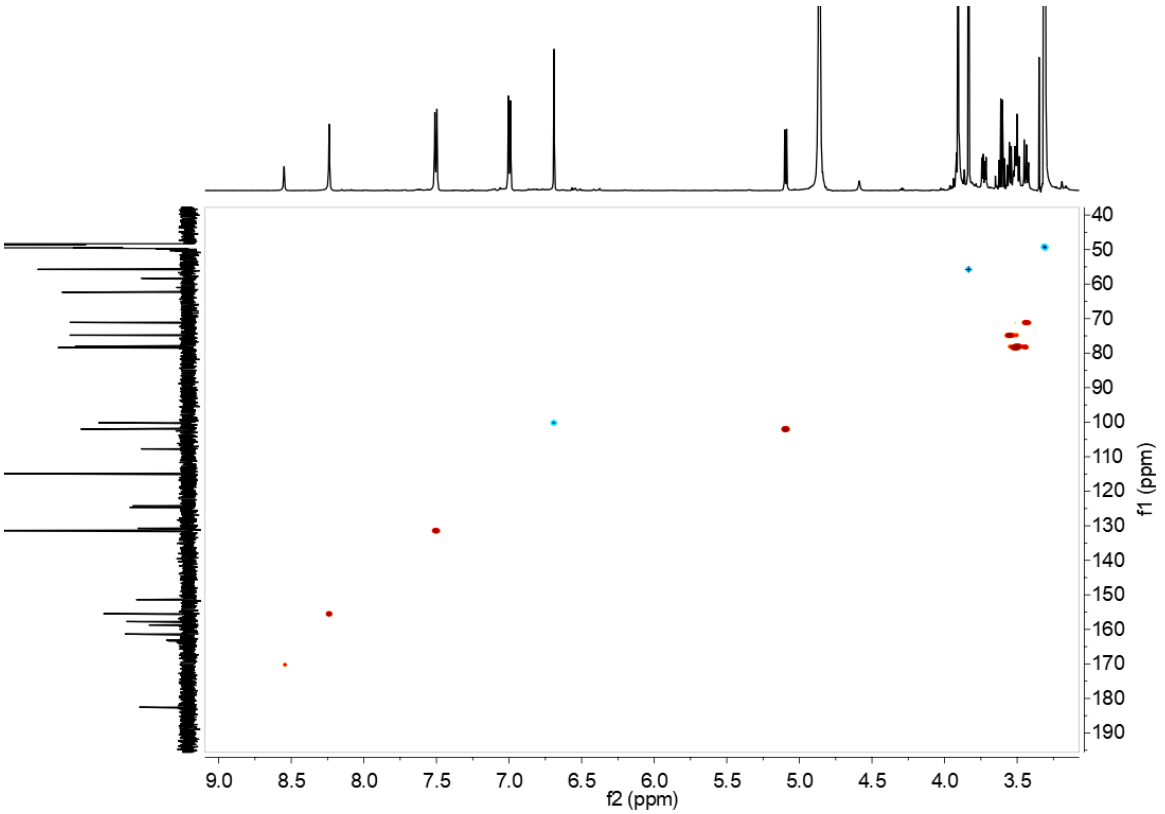

Figure S21 HSQC spectrum of compound **3** in CD_3OD

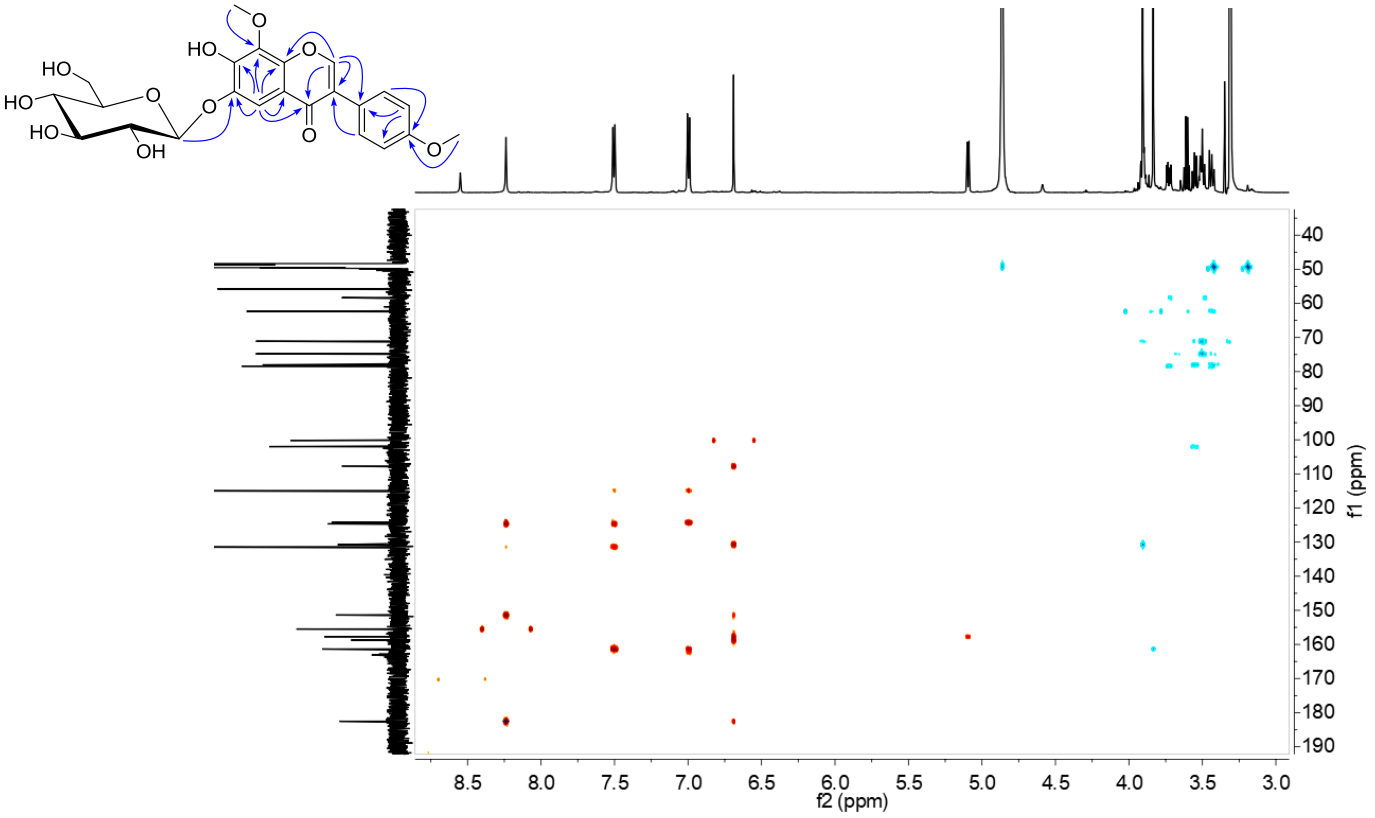

Figure S22 HMBC spectrum of compound **3** in CD_3OD

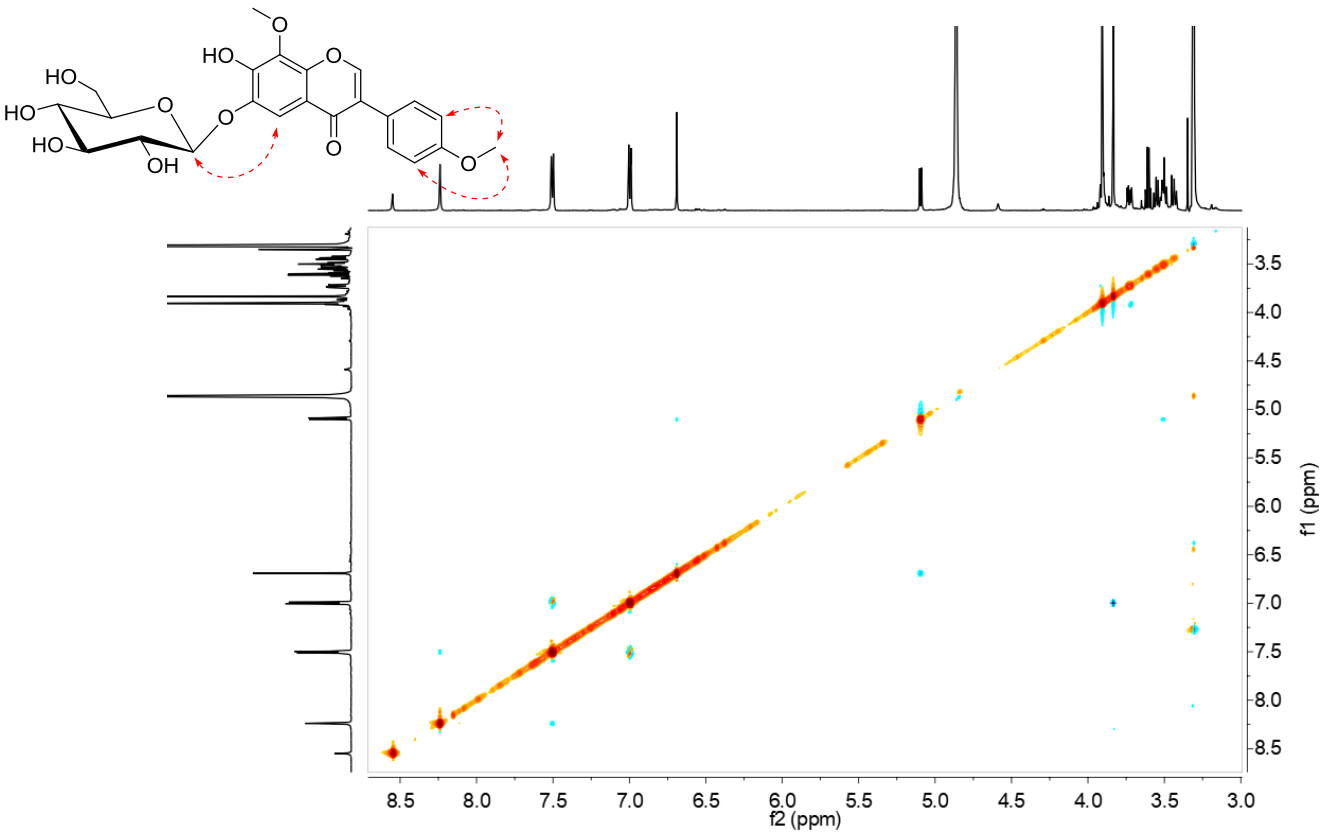

Figure S23 NOESY spectrum of compound **3** in CD_3OD

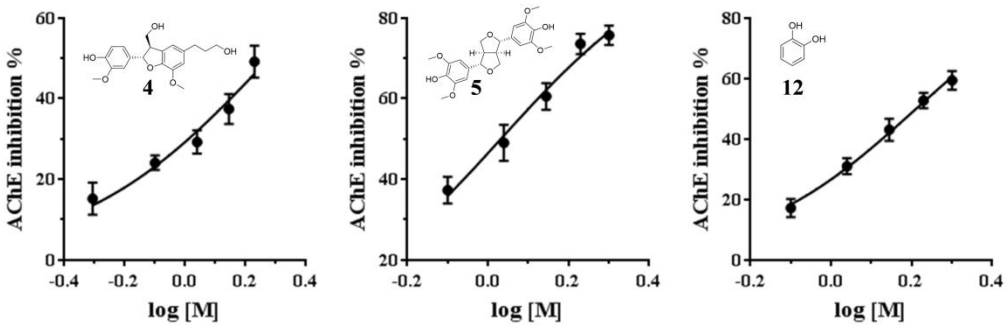

Figure S24 Chemical structure and AChE inhibitory effects of compound **4**, **5**, and **12**. The data was calculated by inhibitory percentage of control, and expressed as Mean ± SEM, *n* = 3, each with triplicate samples. BW284C51 (10 μM) was used as a positive control, and AChE inhibition rate was ~61.8%.