Supporting information

Perforalactones D and E, two new C-20 quassinoids with potential activity in inducing lysosomal biogenesis from twigs of *Harrisonia perforata* (Blanco) Merr.

Shuai Liu,[a,c,g] Cui-Shan Zhang,[a,c,g] Xiao-Qian Ran,[b,c,g] Xiao-Han Tang,[a] Ya-Rong Guo,[d] Ying Yan,[e] Yong-Gang Yao,[b,c,f] Xiao-Jiang Hao,[a] Rong-Can Luo,*[b,c] and Ying-Tong Di*[a,c]

[a.] *State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China*

[b.] *Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650204, China*

[c.] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China;

[d.] *School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, China*

[e.] *Guizhou Chemical Drug Research and Development Engineering Technical Center, Guizhou Medicinal University, Guiyang 550004, China*

[f.] *CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China*

[g.] These authors contribute equally to this work.

[h.] *E-mail:* luorongcan@mail.kiz.ac.cn (R.-C. Luo) and diyt@mail.kib.ac.cn (Y.-T. Di)

General experimental procedures

CD spectra were determined on the Applied Photophysics circular dichroism spectrometer (Applied Photophysics, Leatherhead, Surrey, UK). Optical rotation measurements were conducted with a Jasco P-1020 automatic polarimeter. IR spectra were surveyed on a Bio-Rad FTS-135 as KBr pellets. ^{1}H, ^{13}C, ^{1}H-^{1}H COSY, HSQC, HMBC and ROESY spectra were collected on Bruker DRX-500 instruments (Bruker, Bremerhaven, Germany). Semi-preparative HPLC separations were performed on an Agilent 1260 liquid chromatograph (Agilent Technologies, USA) with a Waters XSelect CSH C-18 column (5 μm, 10×250 mm). Analytical TLC systems were carried out on silica gel 60 F254 plates (Qingdao Marine Chemical Inc., Qingdao, China). Column chromatography (CC) was performed using silica gel (200-300 mesh and 60-80 mesh, Qingdao Marine Chemical, Inc., Qingdao, China) and Sephadex LH-20 (40-70 μm, Amersham Pharmacia Biotech AB, Uppsala, Sweden). Lichroprep RP-18 gel (40-63 μm; Merck, Darmstadt, Germany). And spots were visualized by heating silica gel plates sprayed with 5% H_2SO_4 in EtOH.

Plant material

The branches of *Harrisonia perforata* (Blanco) Merr. were collected from Hainan Province, China, in January 2018. The plant samples were identified by Prof. Sheng-Zhuo Huang from Institute of Tropical Biotechnology, Chinese Academy of Tropical Agricultural Science. A voucher specimen (NO. 20180104) was deposited at the State Key Laboratory of Phytochemistry and Plant Resource in West China, Kunming Institute of Botany, Chinese Academy of Science (CAS).

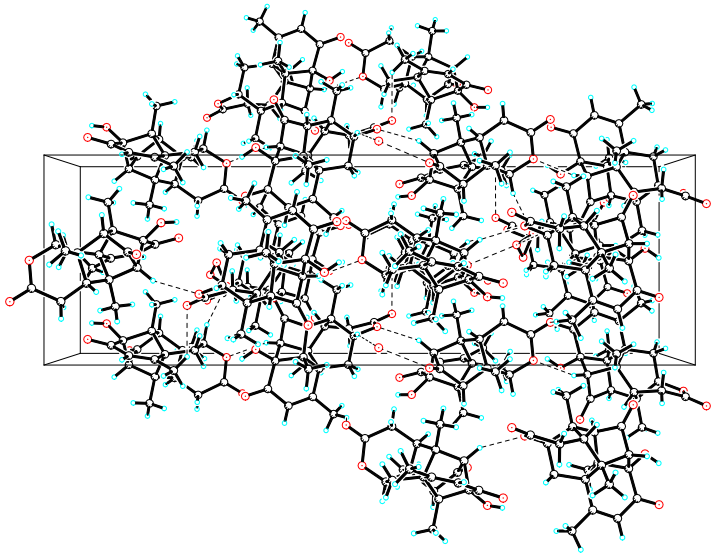
Extraction and Isolation.

The air-dried powder of the plant material (100 kg) was extracted with 95% EtOH under reflux thrice. The crude extract was obtained by reflux. After suspension in water,

the combined extract was successively partitioned with petroleum ether, and ethyl acetate. The ethyl acetate extract (980 g) was then subjected to MCI gel column eluted with MeOH-H_2O (3:7 to 10:0) to give five major fractions (Fr1–Fr5). Fr3 (45.6g) was then chromatographed on a silica gel column eluted with PE-EtOAc (from 1:0 to 1:1), to give five subfractions (Fr3-1- Fr3-5). The fraction Fr3-2 was subjected to a C18 silica gel column (MeOH/H_2O 3:7 to 10: 0) and further purified by Sephadex LH-20 (MeOH) and semipreparative HPLC (CH_3CN/H_2O, 32:68) to obtain compounds **1** (6.1 mg) and **2** (15.6 mg).

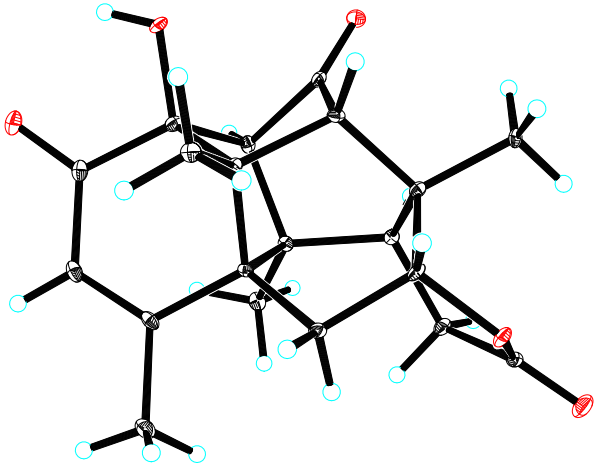
X-ray crystallographic data

Single crystal culture and confirmation: First, compound **2** was added to a bottle and dissolved by the addition of MeOH/H_2O (10:1). Then, the bottle was sealed with parafilm, which only reserves two tiny holes on it, then remained at room temperature for 3 days. Some crystals appeared, and for single crystal parsing, crystals were selected with sizes of 0.460m x 0.370m x 0.360m. All diffraction data were obtained on a Bruker Smart Apex CCD diffractometer equipped with graphite-monochromated Cu Kα radiation. CCDC-2110146 (**2**), contain the supplementary crystallographic data. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre (http://www.ccdc.cam.ac.uk/). Thermal ellipsoids are shown at the 30% level.



View of the pack drawing of **2**.

Hydrogen-bonds are shown as dashed lines.



View of a molecule of **2** with the atom-labelling scheme.

Displacement ellipsoids are drawn at the 30% probability level.

Crystal data for **2** ($C_{20}H_{22}O_5$)•3(H_2O), M = 738.80, a = 10.2285(3) Å, b = 10.2285(3) Å, c = 31.7311(10) Å, α = 90°, β = 90°, γ = 90°, V = 3319.8(2) Å^3, T = 101.(2) K, space group $P41212$, Z = 4, μ(Cu Kα) = 0.913 mm^{-1}, 29917 reflections measured, 3278

independent reflections (R_{int} = 0.0226). The final R_1 values were 0.0575 ($I > 2\sigma(I)$). The final $wR(F^2)$ values were 0.1702 ($I > 2\sigma(I)$). The final R_1 values were 0.0575 (all data). The final $wR(F^2)$ values were 0.1702 (all data). The goodness of fit on F^2 was 1.120. Flack parameter = 0.10(3).

Table S1. Crystal data and structure refinement for 2

Identification code	global	
Empirical formula	$C_{40}H_{50}O_{13}$	
Formula weight	738.80	
Temperature	101(2) K	
Wavelength	1.54178 Å	
Crystal system	Tetragonal	
Space group	$P4_12_12$	
Unit cell dimensions	a = 10.2285(3) Å	α= 90°.
	b = 10.2285(3) Å	β= 90°.
	c = 31.7311(10) Å	γ= 90°.
Volume	3319.8(2) $Å^3$	
Z	4	
Density (calculated)	1.478 Mg/m^3	
Absorption coefficient	0.913 mm^{-1}	
F(000)	1576	
Crystal size	0.460 x 0.370 x 0.360 mm^3	
Theta range for data collection	4.54 to 72.34°.	
Index ranges	-12<=h<=12, -12<=k<=12, -39<=l<=17	
Reflections collected	29917	
Independent reflections	3278 [R(int) = 0.0226]	
Completeness to theta = 72.34°	99.6 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.73 and 0.67	

Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	3278 / 0 / 245
Goodness-of-fit on F^2	1.120
Final R indices [I>2sigma(I)]	R1 = 0.0575, wR2 = 0.1702
R indices (all data)	R1 = 0.0575, wR2 = 0.1702
Absolute structure parameter	0.10(3)
Largest diff. peak and hole	1.357 and -0.930 e.$Å^{-3}$

Biological Assay

Lysosome-Tracker Red staining

The U251 cells cultured in Lab-Tek II Chamber Slide were treated with compound **1** or **2** and then stained by lysosome-tracker red (500 nM) for 1 hour. For better live-cell imaging, glass-bottom dishes were also used in cell growing through Olympus FV1000 confocal microscope. Images were analyzed with FV10-ASW 2.1 Viewer.

Quantitative real-time PCR

Total RNA was isolated from the U251 cells treated with or without compounds **1** and **2** using TRIZOL (Invitrogen, 15596-018). Around 1.5 μg total RNA was used to synthesize single-strand cDNA using the M-MLV Reverse Transcriptase (Promega, M170A) in a final volume of 25 μL according to the manufacturer's instructions. The relative mRNA levels of *LAMP1*, *LAMP2* and, *CTSB*, *ATP6V0E1*, *ARSB*, and *CTSD* were quantified by using quantitative real-time PCR (qRT-PCR), with normalization to the *GAPDH* gene. The qRT-PCR was performed in a total volume of 20 μL containing 2 μL of diluted products, 10 μL of SYBR Master Mix (Takara), 0.2 uL 10 μM each primer (**Table S2**), on an BIO-RAD Real-time PCR detection system. The qRT-PCR thermal cycling conditions were composed of a denaturation cycle at 95°C for 5 min, followed by 40 cycles of 95 °C for 10 s and 58 °C for 30 sec.

Statistics and reproducibility

Data analysis was carried out using GraphPad Prism 8. The Student's *t*-test was used to detect the mRNA expression difference between groups. Values were expressed as mean ± standard error. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, not significant.

Table S2. Primer pairs for measuring mRNA levels of the targeted genes in U251 cells

Primer	Sequence (5'-3')	Product length (bp)
LAMP1 Forward	TCTCAGTGAACTACGACACCA	151
LAMP1 Reverse	AGTGTATGTCCTCTTCCAAAAGC	
LAMP2 Forward	GAAAATGCCACTTGCCTTTATGC	184
LAMP2 Reverse	AGGAAAAGCCAGGTCCGAAC	
CTSB Forward	ACAACGTGGACATGAGCTACT	85
CTSB Reverse	TCGGTAAACATAACTCTCTGGGG	
ATP6V0E1 Forward	GTCCTAACCGGGGAGTTATCA	101
ATP6V0E1 Reverse	AAAGAGAGGGTTGAGTTGGGC	
ARSB Forward	TCTTGCTGGCAGACGACCTA	121
ARSB Reverse	GGCTGCGTGTAGTAGTTGTCC	
CTSD Forward	CACCACAAGTACAACAGCGAC	77
CTSD Reverse	CCCGAGCCATAGTGGATGT	
GAPDH Forward	CTGGGCTACACTGAGCACC	101
GAPDH Reverse	AAGTGGTCGTTGAGGGCAATG	

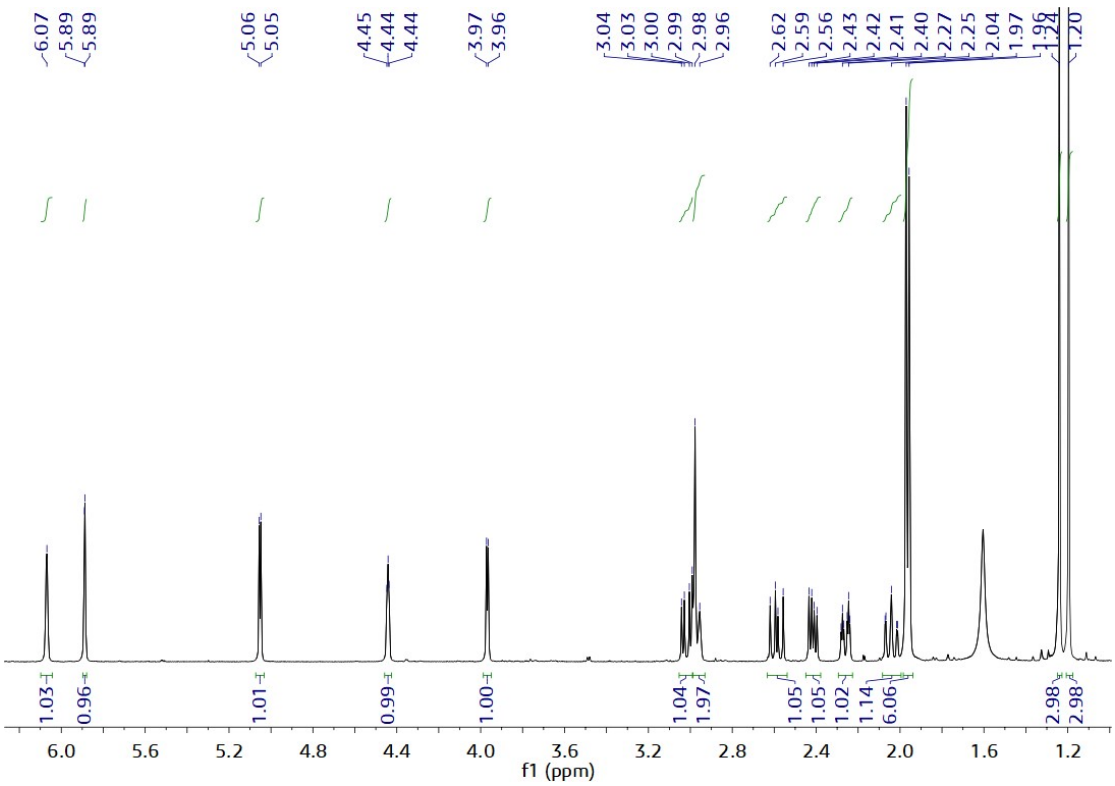


Figure S1 ^{1}H NMR spectrum of perforalactone D (**1**) in $CDCl_3$**.**

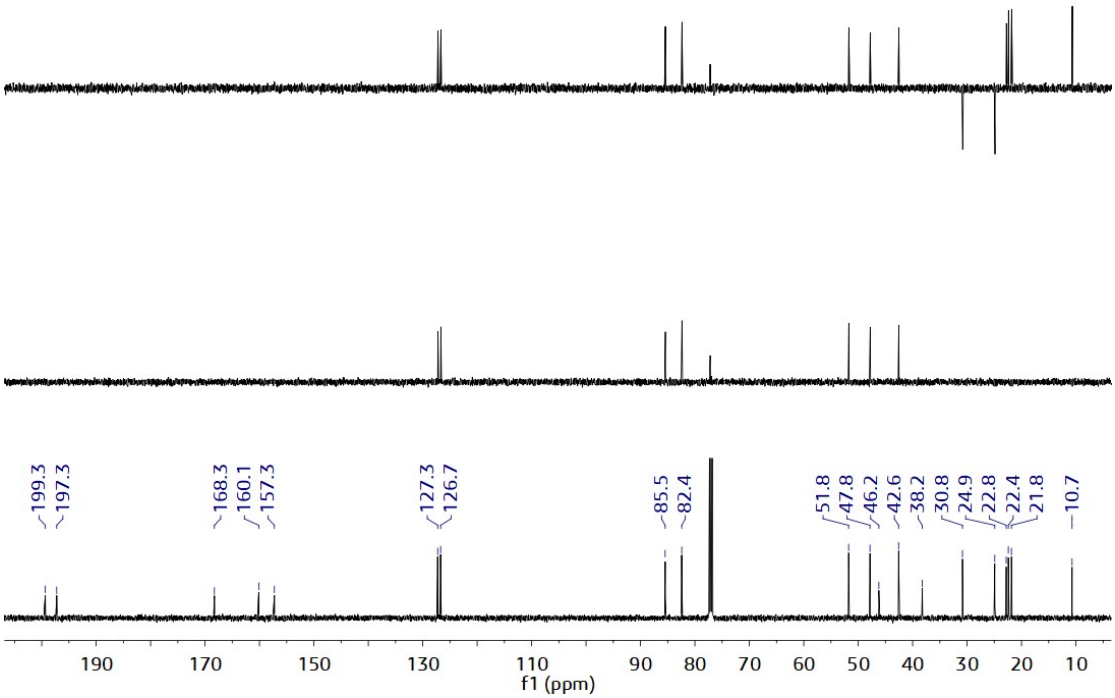


Figure S2 ^{13}C NMR spectrum of perforalactone D (**1**) in $CDCl_3$.

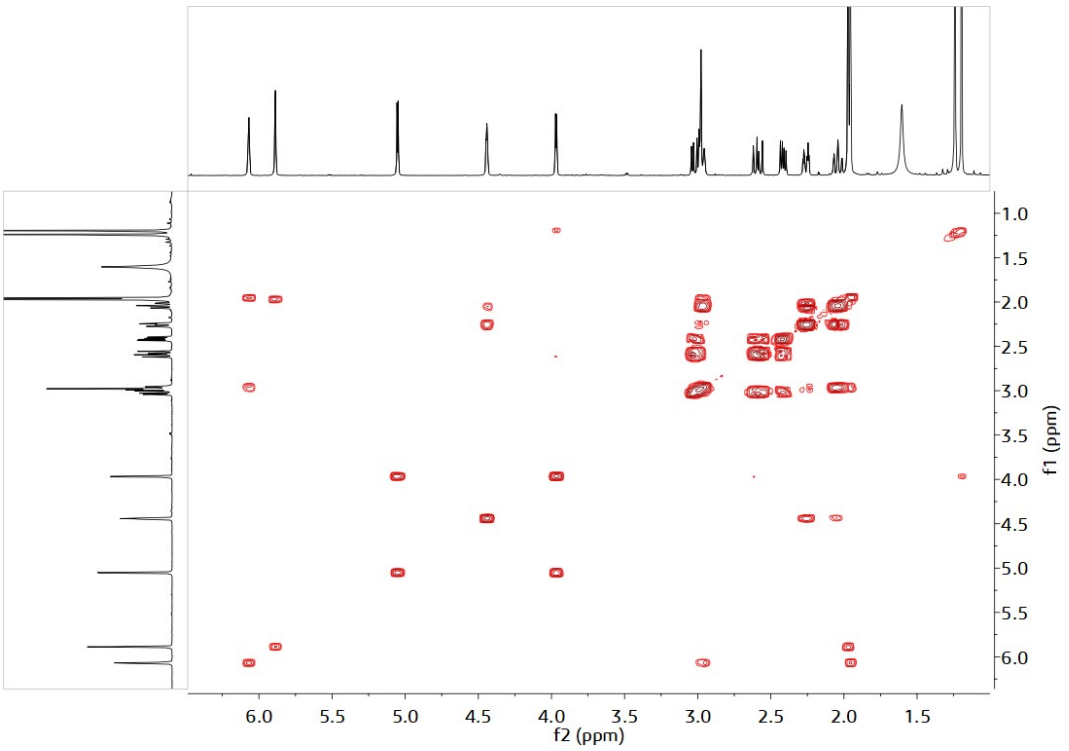


Figure S3 ^{1}H-^{1}H COSY spectrum of perforalactone D (1) in $CDCl_3$.

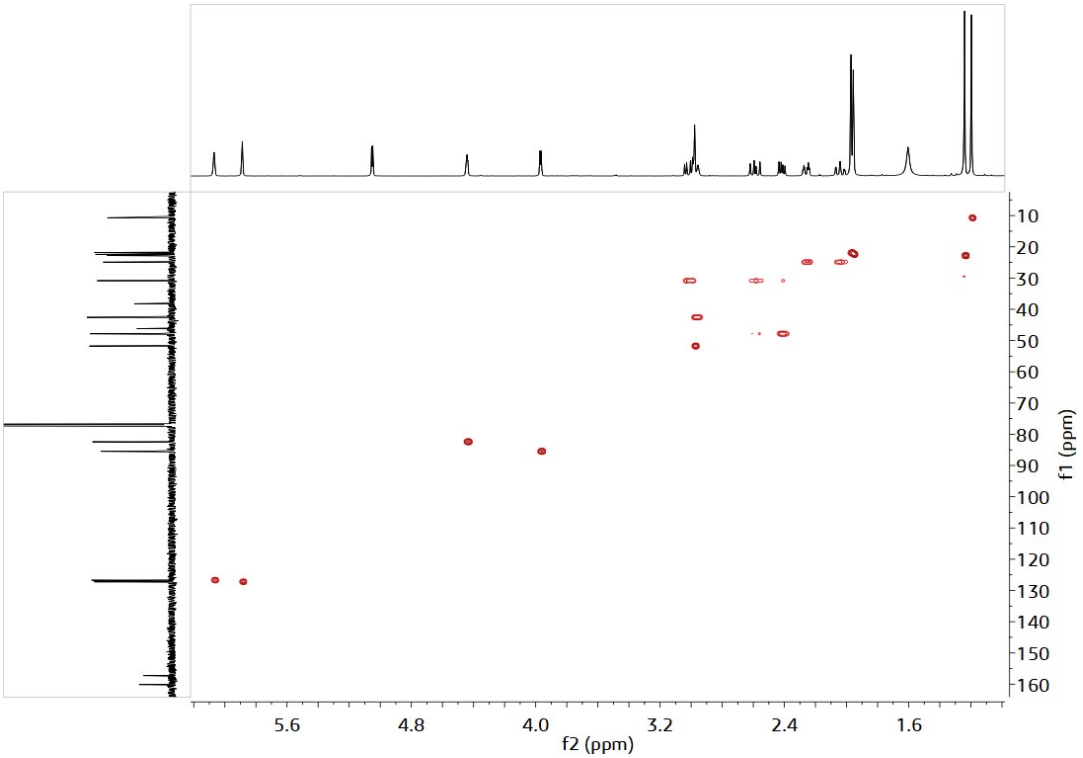


Figure S4 HSQC spectrum of perforalactones D (1) in $CDCl_3$.

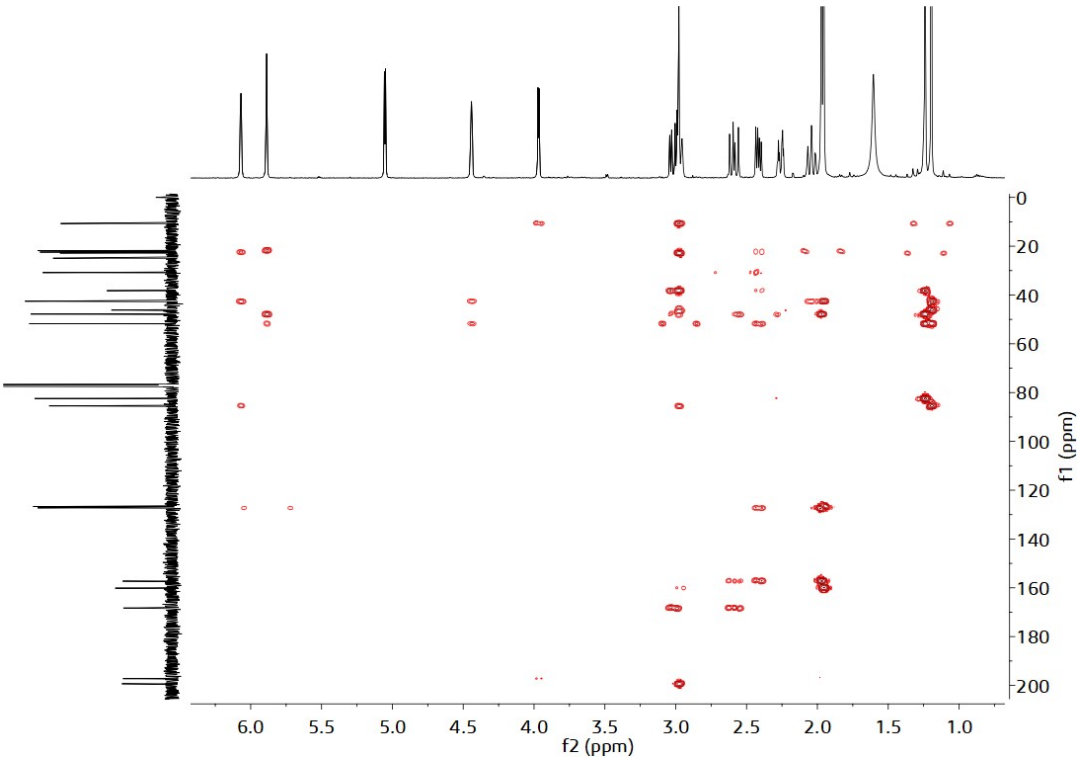


Figure S5 HMBC spectrum of perforalactone D (1) in $CDCl_3$.

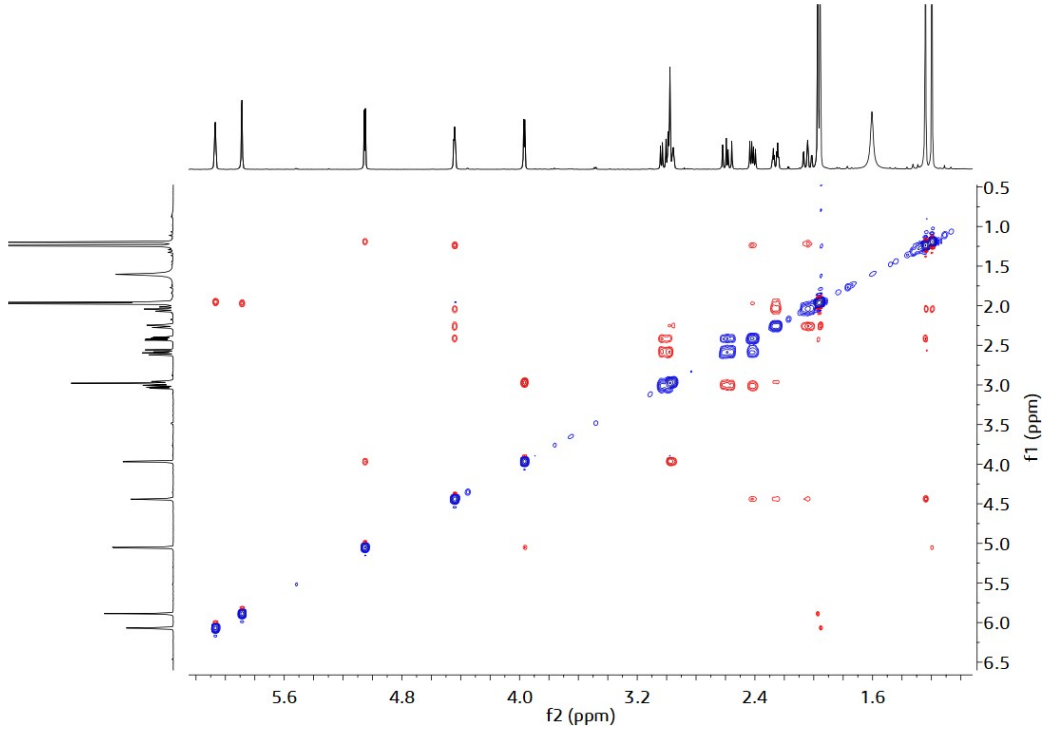


Figure S6 ROESY spectrum of perforalactone D (1) in $CDCl_3$.

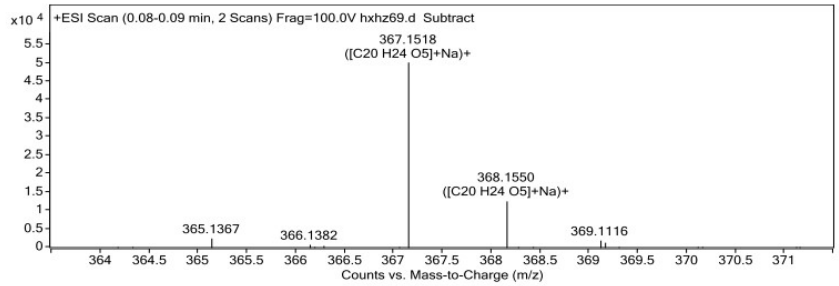
Qualitative Analysis Report

Data Filename	hxhz69.d	**Sample Name**	hxhz69
Sample Type	Sample	**Position**	P1-A3
Instrument Name	Instrument 1	**User Name**	
Acq Method	s.m	**Acquired Time**	4/16/2020 10:35:52 AM
IRM Calibration Status	Success	**DA Method**	Default.m
Comment			

Sample Group	**Info.**
Acquisition SW Version	6200 series TOF/6500 series Q-TOF B.05.01 (B5125.2)

User Spectra

Fragmentor Voltage	Collision Energy	Ionization Mode
100	0	ESI



Peak List

m/z	z	Abund	Formula	Ion
274.2739	1	16259.46		
318.3003	1	23972.25		
345.1698	1	14859.37		
353.1361	1	11999.54		
362.3262	1	8611.1		
367.1518	1	50153.58	C20 H24 O5	(M+Na)+
368.155	1	12744.48	C20 H24 O5	(M+Na)+
697.2986	1	12283.43		
711.3138	1	28421.73		
712.3167	1	12732.39		

Formula Calculator Element Limits

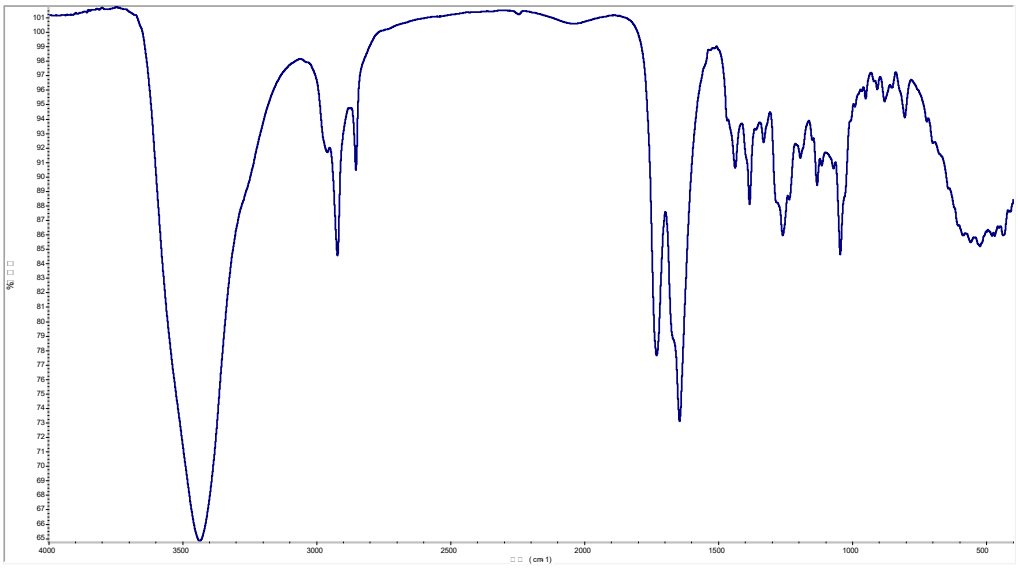
Element	Min	Max
C	3	60
H	0	120
O	0	30

Formula Calculator Results

Formula	CalculatedMass	CalculatedMz	Mz	Diff. (mDa)	Diff. (ppm)	DBE
C20 H24 O5	344.1624	367.1516	367.1518	-0.20	-0.54	9.0000

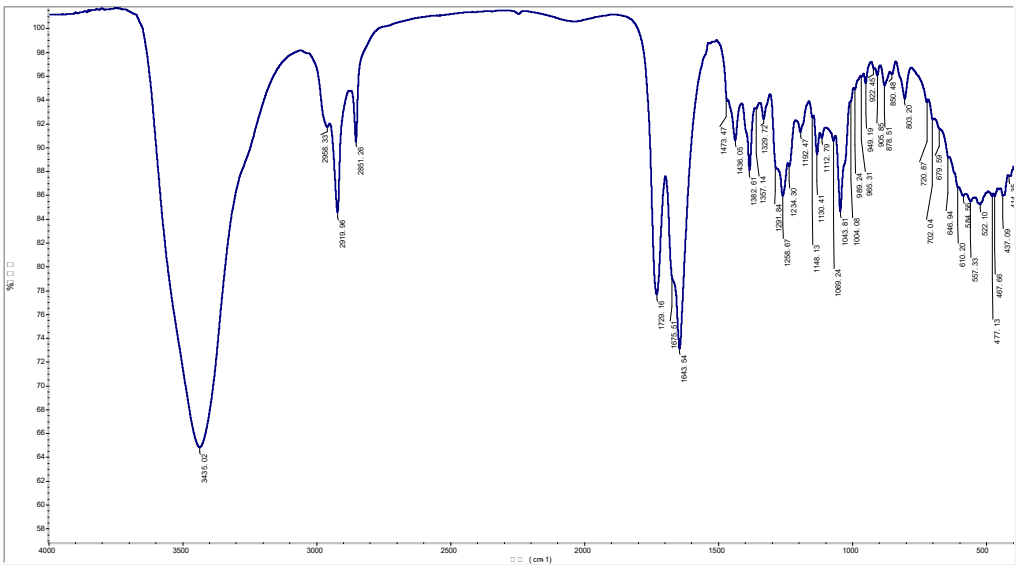
--- End Of Report ---

Figure S7 HR-ESI-MS spectrum of perforalactone D (**1**)



Sample Name: HXHZ69
KBr□□
□□□□□: □□□ 11□ 18 15:19:14 2019 (GMT+08:00)
□□□□□NICOLET iS10
Software version: OMNIC 9.8.372

□□□□□□□: 16
□□□□□□□□: 16
□□□□: 4.000
□□□□□□: 1.0
□□□□□□: 0.4747
□□: 80.00



Sample Name: HXHZ69
KBr□□
□□□□□: □□□ 11□ 18 15:19:14 2019 (GMT+08:00)
□□□□□NICOLET iS10
Software version: OMNIC 9.8.372

□□□□□□□: 16
□□□□□□□□: 16
□□□□: 4.000
□□□□□□: 1.0
□□□□□□: 0.4747
□□: 80.00

Figure S8 IR (KBr disk) spectrum of perforalactone D (1)

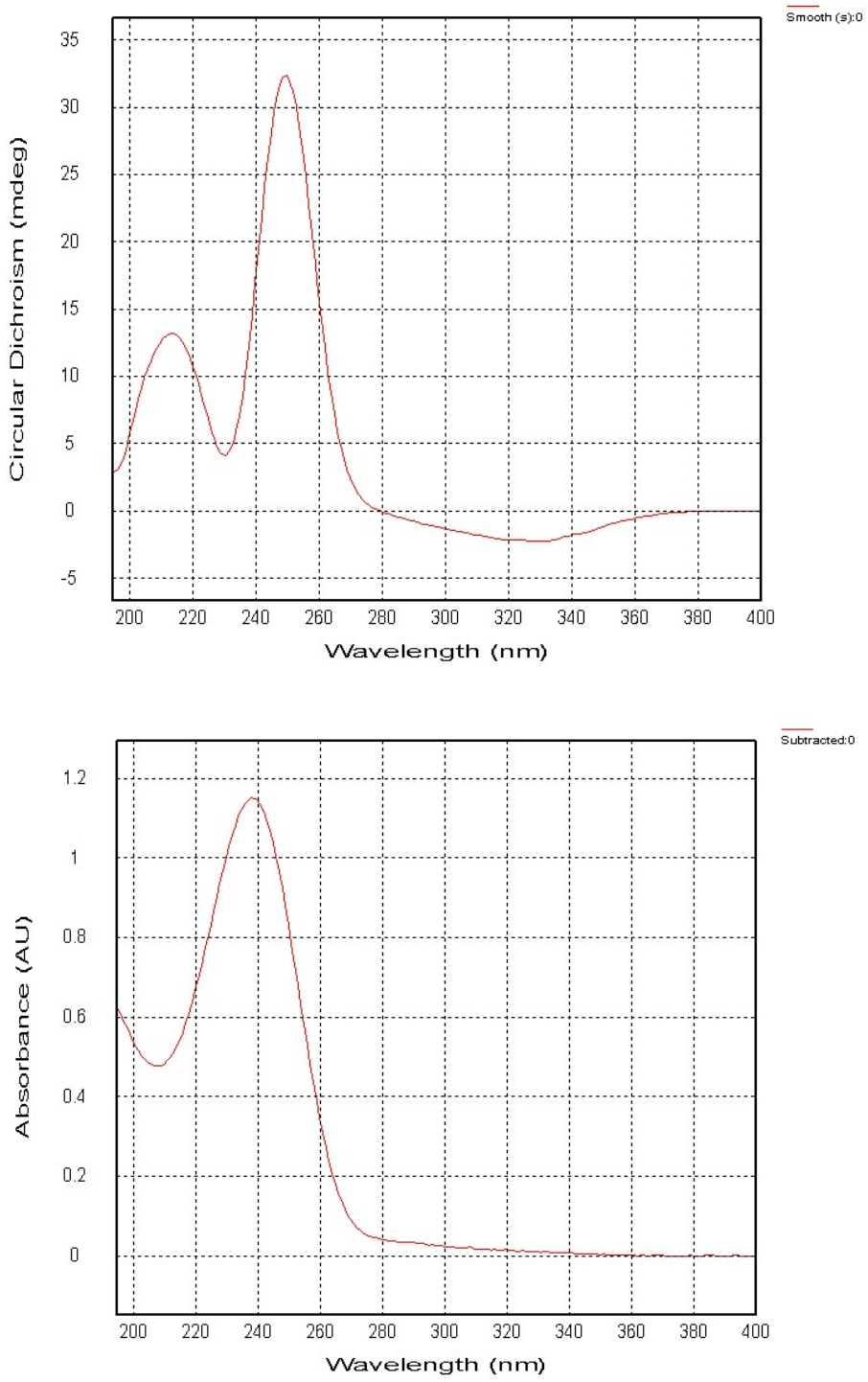


Figure S9 ECD spectrum of perforalactone D (**1**) **in methanol.**

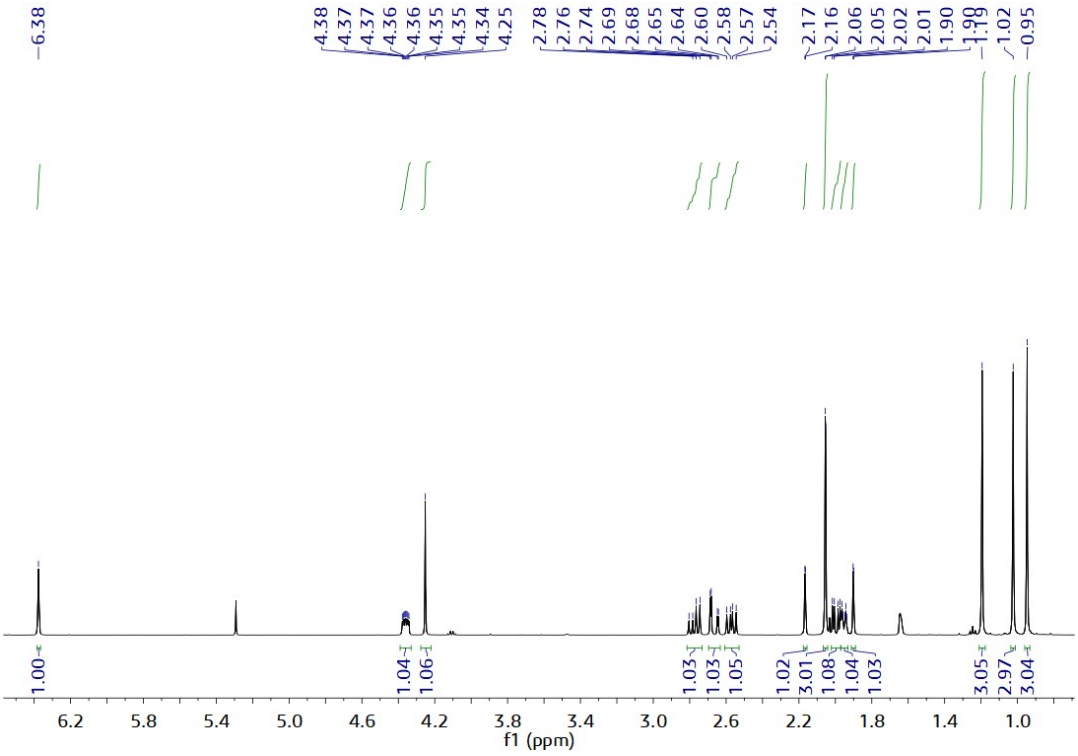


Figure S10 ^{1}H NMR spectrum of perforalactone E (2) in $CDCl_3$.

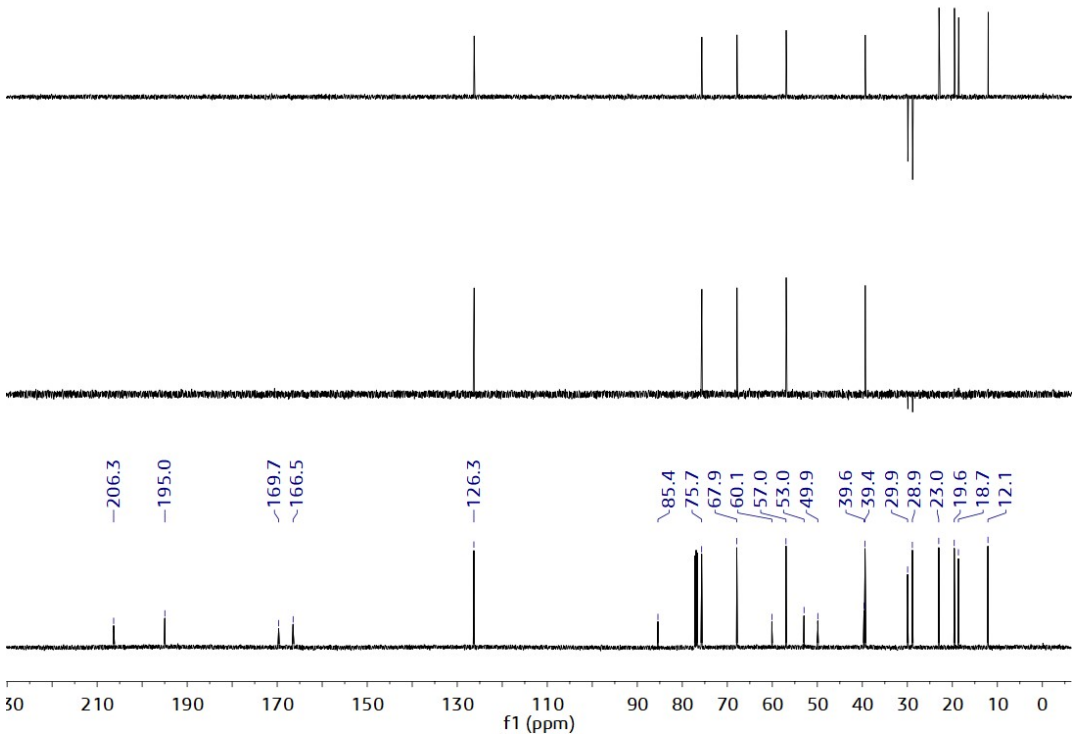


Figure S11 ^{13}C NMR spectrum of perforalactone E (2) in $CDCl_3$.

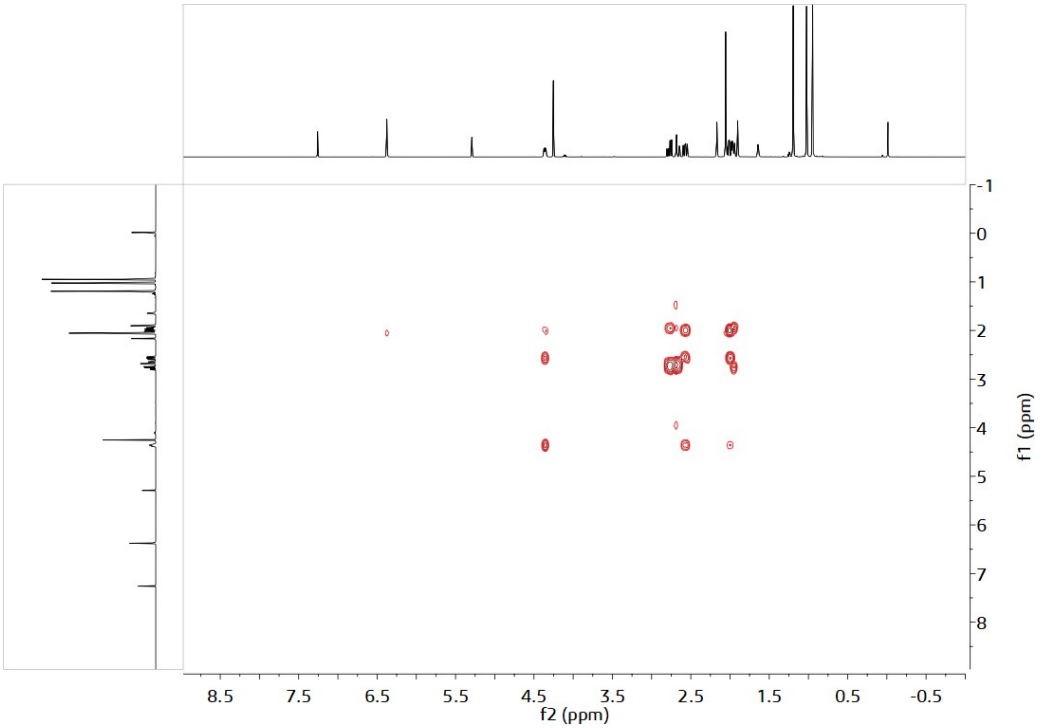


Figure S12 ^{1}H-^{1}H COSY spectrum of perforalactone E (2) in $CDCl_3$.

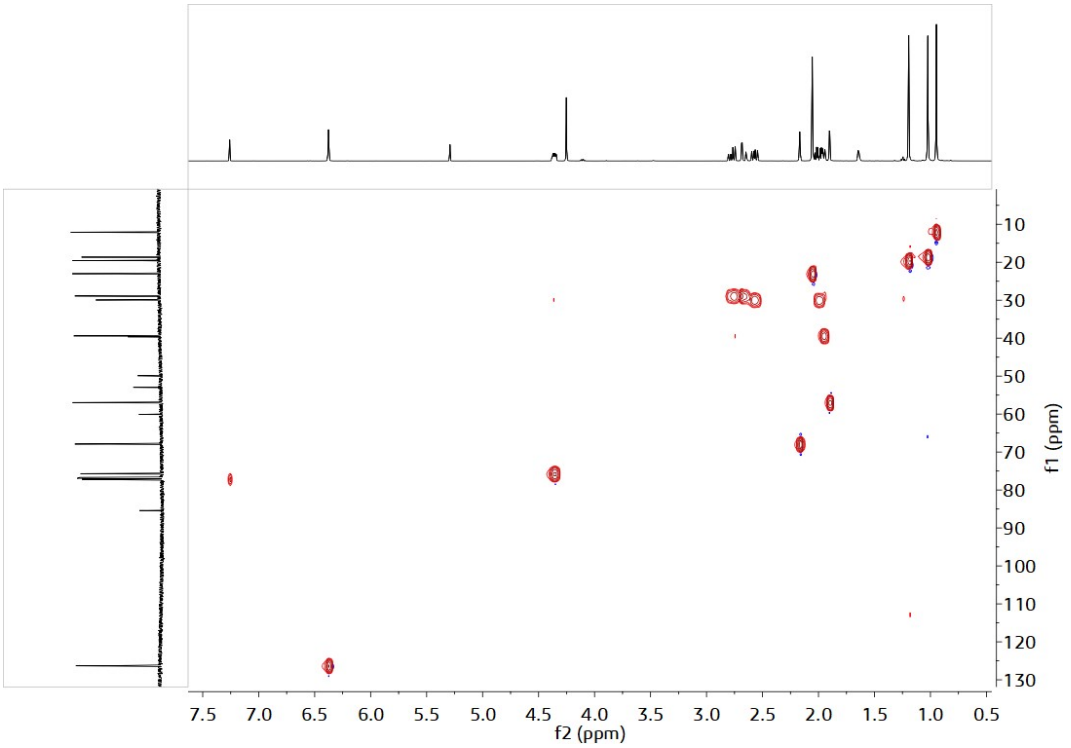


Figure S13 HSQC spectrum of perforalactone E (2) in $CDCl_3$.

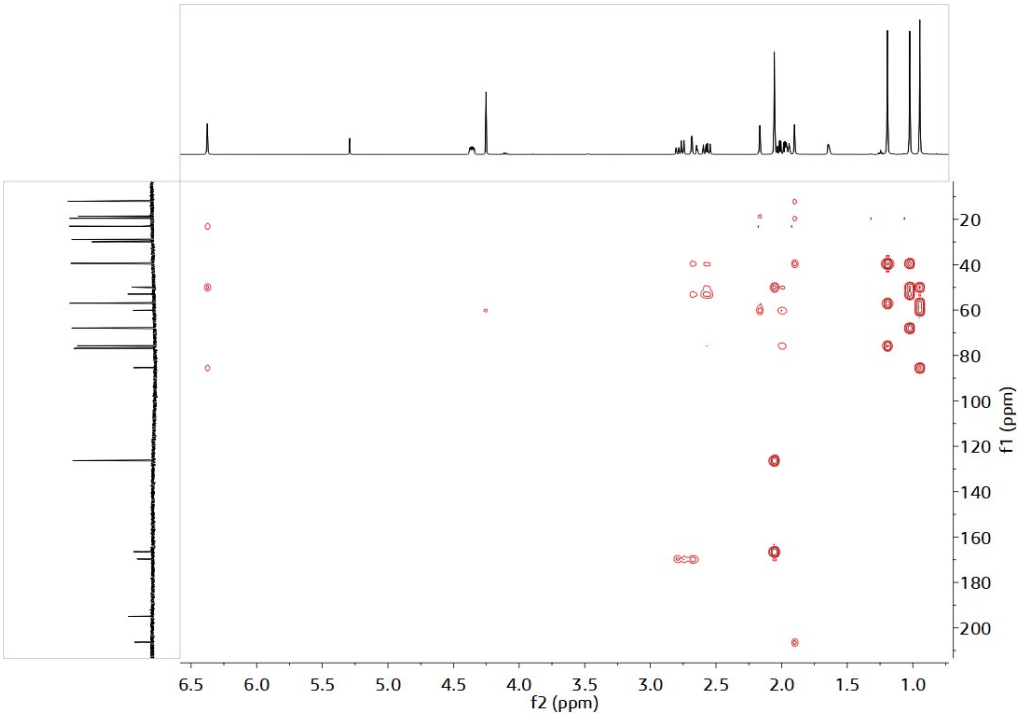


Figure S14 HMBC spectrum of perforalactone E (2) in $CDCl_3$.

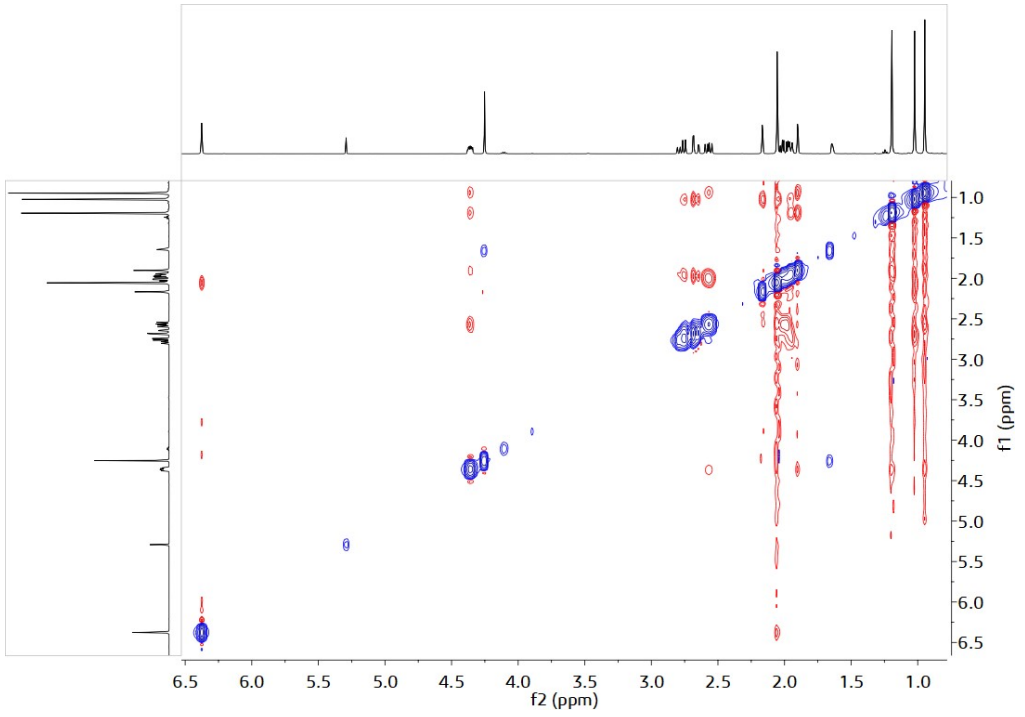


Figure S15 ROESY spectrum of perforalactone E (2) in $CDCl_3$.

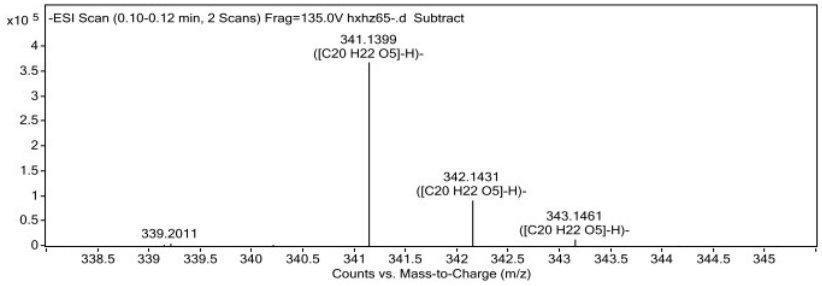
Qualitative Analysis Report

Data Filename	hxhz65-.d	**Sample Name**	hxhz65
Sample Type	Sample	**Position**	P1-A2
Instrument Name	Instrument 1	**User Name**	
Acq Method	s-.m	**Acquired Time**	4/16/2020 10:43:44 AM
IRM Calibration Status	Success	**DA Method**	Default.m
Comment			

Sample Group	**Info.**
Acquisition SW Version	6200 series TOF/6500 series Q-TOF B.05.01 (B5125.2)

User Spectra

Fragmentor Voltage	Collision Energy	Ionization Mode
135	0	ESI



Peak List

m/z	z	Abund	Formula	Ion
96.9603		382631.03		
124.9915	1	284327.59		
255.2334	1	279182.66		
283.2645	1	136074.89		
341.1399	1	368479.13	C20 H22 O5	(M-H)-
342.1431	1	92961.52	C20 H22 O5	(M-H)-
377.1168	1	117501.91		
393.2781	1	138949.59		
403.3068	1	101302.96		
1045.9938	1	113466.62		

Formula Calculator Element Limits

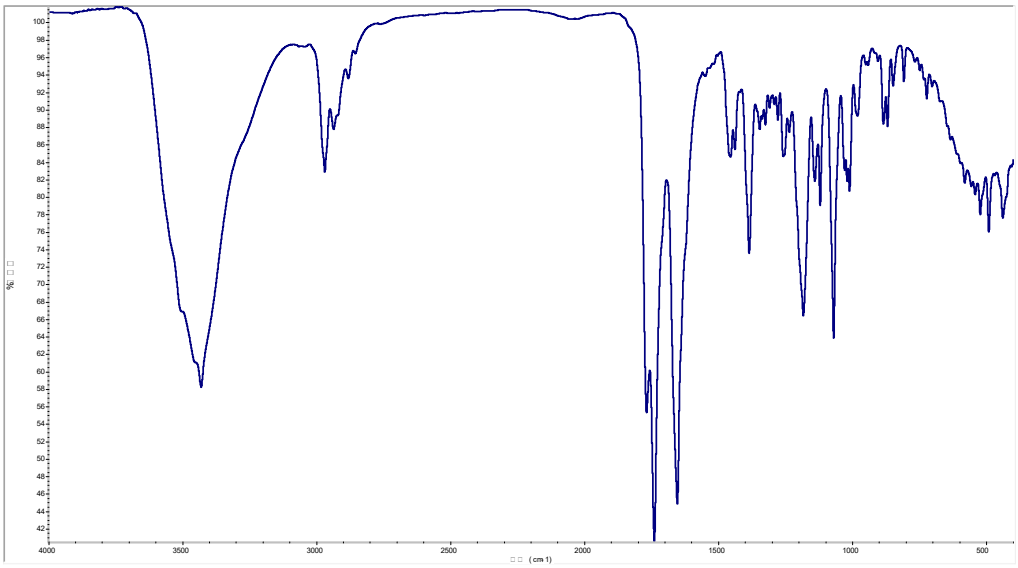
Element	Min	Max
C	3	60
H	0	120
O	0	30

Formula Calculator Results

Formula	CalculatedMass	CalculatedMz	Mz	Diff. (mDa)	Diff. (ppm)	DBE
C20 H22 O5	342.1467	341.1394	341.1399	-0.50	-1.47	10.0000

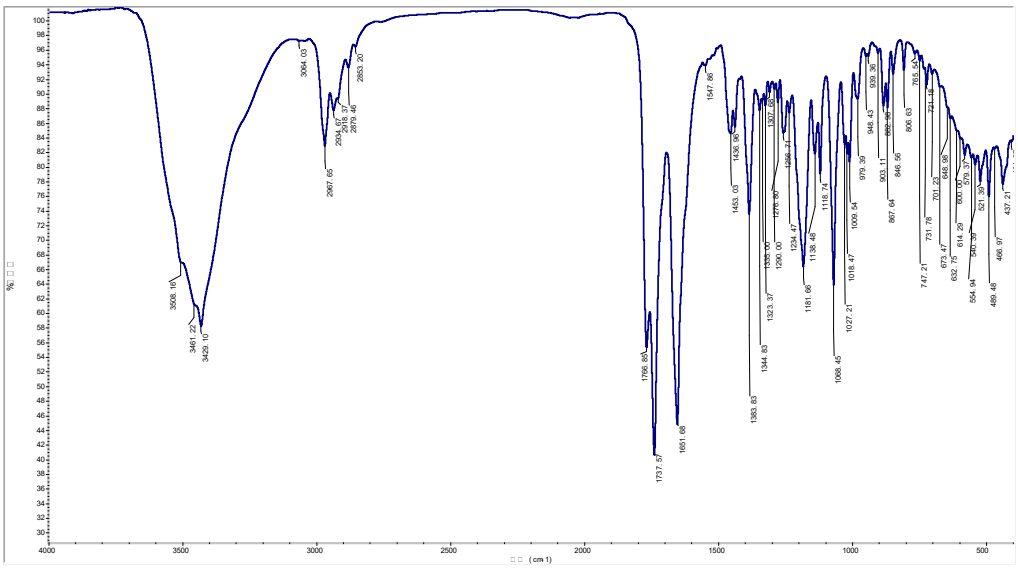
--- End Of Report ---

Figure S16 HR-ESI-MS spectrum of perforalactone E (**2**)



Sample Name: HXHZ65
KBr□□
□□□□□: □□□ 11□ 18 13:55:55 2019 (GMT+08:00)
□□□□□NICOLET iS10
Software version: OMNIC 9.8.372

□□□□□□□: 16
□□□□□□□□: 16
□□□□: 4.000
□□□□□□: 1.0
□□□□□□: 0.4747
□□: 80.00



Sample Name: HXHZ65
KBr□□
□□□□□: □□□ 11□ 18 13:55:55 2019 (GMT+08:00)
□□□□□NICOLET iS10
Software version: OMNIC 9.8.372

□□□□□□□: 16
□□□□□□□□: 16
□□□□: 4.000
□□□□□□: 1.0
□□□□□□: 0.4747
□□: 80.00

Figure S17 IR (KBr disk) spectrum of perforalactone E (**2**)

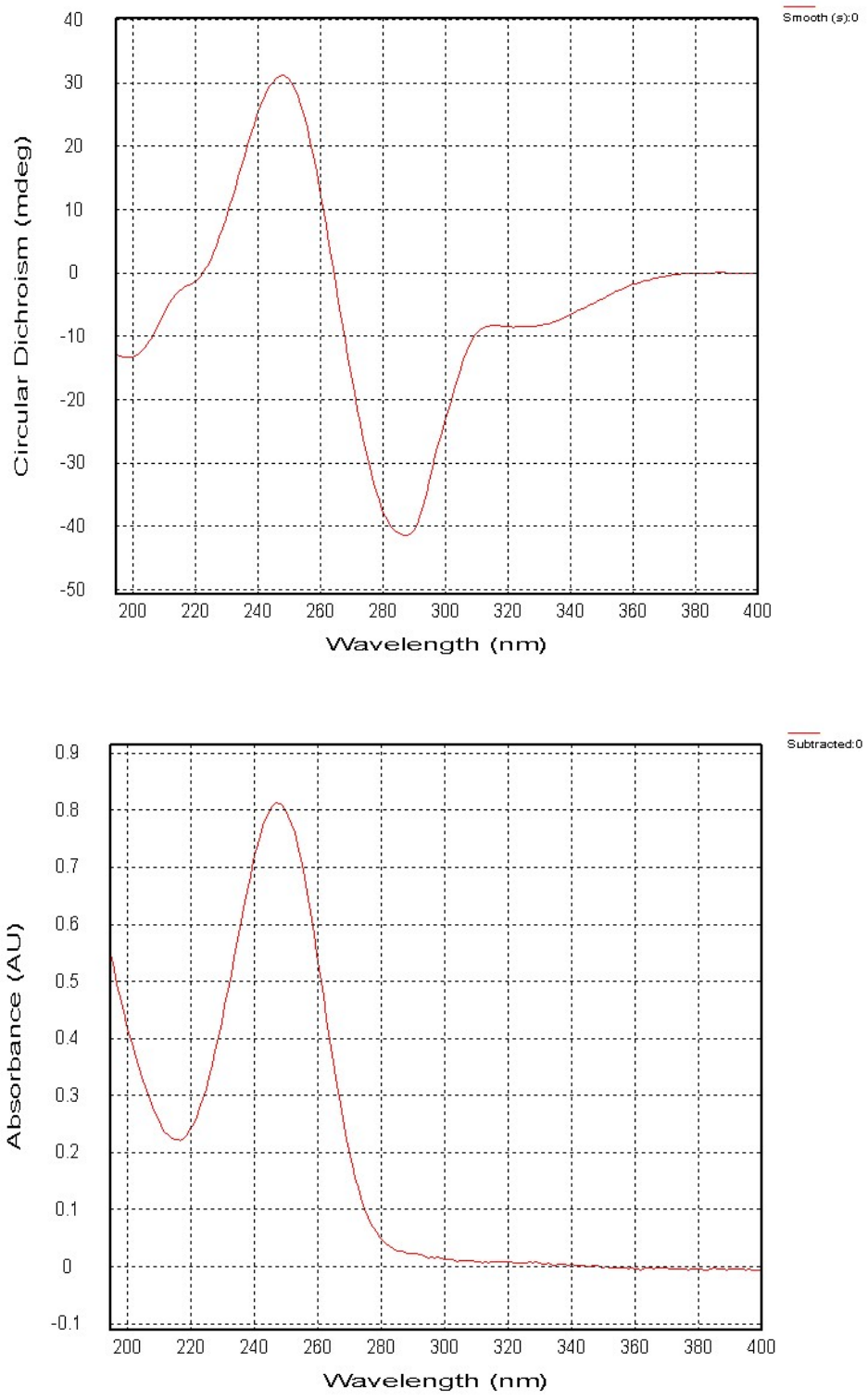


Figure S18 ECD spectrum of perforalactone E (**2**) **in methanol.**