Supplementary information

Discoipyrroles A-D: Isolation, Structure Determination and Synthesis of Potent Migration Inhibitors from Bacillus hunanensis

Youcai Hu[1], Malia B. Potts[2], Dominic Colosimo[1], Mireya L. Herrera-Herrera[2], Aaron G. Legako[1], Muhammed Yousufuddin[4], Michael A. White[2,3] and John B. MacMillan[1,*]

[1]Department of Biochemistry, [2]Department of Cell Biology, [3]Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390, USA. [4]Department of Chemistry, Center for Nanostructured Materials, University of Texas at Arlington, Arlington, TX 76019, USA.

e-mail: john.macmillan@utsouthwestern.edu

Table of Contents

1. Experimental Details

General Procedures. Optical rotations were recorded with an AUTOPOL AP IV-6W polarimeter equipped with a halogen lamp (589 nm). UV spectra were recorded on a Shimadzu UV-1601 UV–VIS spectrophotometer. ^{1}H and 2D NMR spectral data were recorded at 600 MHz in CD_3OD or DMSO-d_6 on a Varian System spectrometer, and chemical shifts were referenced to the corresponding residual solvent signal (CD_3OD: δ_H 3.31/ δ_C 49.15; DMSO-d_6: δ_H 2.50/ δ_C 39.51). ^{13}C NMR spectra were acquired at 100 MHz on a Varian System spectrometer. High resolution ESI-TOF mass spectra were provided by The Scripps Research Institute, La Jolla, CA. Low-resolution LC/ESI-MS data were measured using an Agilent 1200 series LC/MS system with a reversed- phase C18 column (Phenomenex Luna, 150 mm × 4.6 mm, 5 μm) at a flow rate of 0.7 mL/min. Preparative HPLC was performed on an Agilent 1200 series instrument with a DAD detector, using a Phenyl-Hexyl column (Phenomenex Luna, 250 ×10.0 mm, 5 μm) or Silica column (Phenomenex Luna, 250 ×10.0 mm, 10 μm). Sephadex LH-20 (GE Healthcare, Sweden) and ODS (50 mm, Merck) were used for column chromatography.

Collection and phylogenetic analysis of strain SNA-048 and SNE-038. The marine-derived bacterium, strain SNA-048, was isolated from a sediment sample collected from the brackish waters of Trinity Bay, Galveston Texas (N 29º 42.419', W 94º 49165' W). Bacterial spores were collected via stepwise centrifugation as follows: 2 g of sediment was dried over 24 h in an incubator at 35 ^{o}C and the resulting sediment added to 10 mL sea water (sH_2O) containing 0.05% Tween 20. After a vigorous vortex for 10 min, the sediment was centrifuged at 2500 rpm for 5 min (4 ^{o}C). The supernatant was removed and transferred into a new tube and centrifuged at 18,000 rpm for 25 min (4 ^{o}C) and the resulting spore pellet collected. The resuspended spore pellet (4 mL sH_2O) was plated on a humic acid media, giving rise to individual colonies of SNA-048 after two weeks. Genomic DNA of strain SNA-048 was isolated using standard methods[1] and was amplified using PCR with the Universal 16S rRNA primers FC27 and RC 1492 using the method of Mincer. The partial 16S rRNA sequence (1432 of 1492 bp) was compared to sequences in available databases using the Basic Local Alignment Search Tool and strain SNA-048 determined to be an *Bacillus hunanensis*. The 16S rRNA sequence of SNA-048 was deposited in the GenBank database with the accession # KC247801.
Strain SNE-038, was isolated from a sediment sample collected from an estuary on Kiawah Island, South Carolina and isolated on a seawater agar based media (10 g starch, 4 g yeast extract, 2 g peptone, 8 g agar, 1 L of 75% sH_2O). Phylogenetic analysis for SNE-038 was carried out as above and determined to be *Bacillus aquamaris*. The 16S rRNA sequence of SNE-038 was deposited in the GenBank database with the accession # KC493772.

Cultivation and extraction. Bacterium SNA-048 was cultured in 10 × 2.8 L Fernbach flasks each containing 1 L of a seawater based medium (10 g starch, 4 g yeast extract, 2 g peptone, 1 g $CaCO_3$, 40 mg $Fe_2(SO_4)_3 \cdot 4H_2O$, 100 mg KBr) and shaken at 200 rpm at 27 ^{o}C. After seven days of cultivation, sterilized XAD-7-HP resin (20 g/L) was added to adsorb the organic products, and the culture and resin were shaken at 200 rpm for 2 h. The resin was filtered through cheesecloth, washed with deionized water, and eluted with acetone. The acetone soluble fraction was dried in vacuo to yield 3.0 g of extract.

Isolation. The extract (3.0 g) was partitioned with *n*-hexane, CH_2Cl_2, and MeOH/H_2O. The MeOH/H_2O soluble layer (2.0 g) was submitted to HP-20 column eluting with H_2O to remove salts and eluting by MeOH to give 1.0 g salt free material. The MeOH eluted fraction was fractionated by flash column chromatography on ODS (50 μm, 50 g), eluting with a step gradient of MeOH and H_2O (100:0–30:70), and 25 fractions (Fr.M1–Fr.M25) were collected. Fractions 15 and 16 was combined and purified by reversed phase HPLC (Phenomenex Luna, Phenyl-Hexyl, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 210 nm) using a gradient solvent system from 30% to 90% CH_3CN (0.1% FA) over 30 min to afford discoipyrrole A (**1**, 2.0 mg, t_R = 21.6 min) and discoipyrrole B (**2,** 0.8 mg, t_R = 25.5 min). Fraction M17 was purified by Sephadex LH-20 eluted with MeOH to give 8 subfractions (Fr.M17-1–Fr.M17-8). Subfraction Fr.M17-5 was purified by reversed phase HPLC (same method as that using for **1** and **2** to afford discoipyrrole D (**4,** 0.4 mg, t_R = 18.5 min). Subfraction Fr.M17-8 was purified by the same reversed phase HPLC using a gradient solvent system from 20% to 80% CH_3CN (0.1% FA) over 30 min to afford discoipyrrole C (**3,** 1.0 mg, t_R = 14.9 min) and 4-hydroxysattabacin (**5,** 0.5 mg, t_R = 18.5 min). The DCM soluble layer (0.9 g) fractionated by flash column chromatography on ODS (50 μm, 50 g), eluting with a step gradient of MeOH and H_2O (100:0–20:80), and 9 fractions (Fr.D1–Fr.D9) were collected. Fr.D4 was prufied by flash column chromatography on silica gel (50 g), eluting with a step gradient of Hexane and EtOAC (60:40–0:100) to give **5** (500 mg). Fr.D5 was purified by Sephadex LH-20 eluted with MeOH to give 8 subfractions (Fr.D5-1–Fr.D5-8). Subfraction Fr.D5-4 was purified by reversed phase HPLC using a gradient solvent system from 30% to 80% CH_3CN (0.1% FA) over 30 min to afford 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**, 0.8 mg, t_R = 10.8 min) and *p*-hydroxybenzaldehyde (0.7 mg, t_R = 20.8 min).

To obtain enough material for crystallization of **1**, large scale fermentation (120 L broth) was carried out, and the above isolation process was repeated to offer 20 mg discoipyrrole A (**1**).

Discoipyrrole A (**1**) yellow solid; $[\alpha]_D$ 0^o (*c* 1.2 MeOH); UV (MeOH) λ_{max} (log ε) 214 (4.25), 263 (3.82), 324 (3.29), 396 (3.46). ^{1}H NMR (600 MHz, CD_3OD or DMSO-d_6) and ^{13}C NMR (100 MHz, CD_3OD or DMSO-d_6) see Table S1. ESI-MS *m*/*z* 464.2 $[M + Na]^+$, 440.2 $[M - H]^-$. HRESIMS *m*/*z* 442.1652 $[M + H]^+$ ($C_{27}H_{24}NO_5$, calcd 442.1655), 440.1550 $[M - H]^-$ ($C_{27}H_{22}NO_5$, calcd 440.1498).

Discoipyrrole B (**2**) yellow solid; $[\alpha]_D$ 0^o (*c* 0.6 MeOH); UV (MeOH) λ_{max} (log ε) 210 (4.02), 263 (3.74), 383 (3.43); 1D and 2D NMR data see Table S2. ESI-MS *m*/*z* 448.2 $[M + Na]^+$, 424.2 $[M - H]^-$. HRESIMS *m*/*z* 426.1697 $[M + H]^+$ ($C_{27}H_{24}NO_4$, calcd 426.1705).

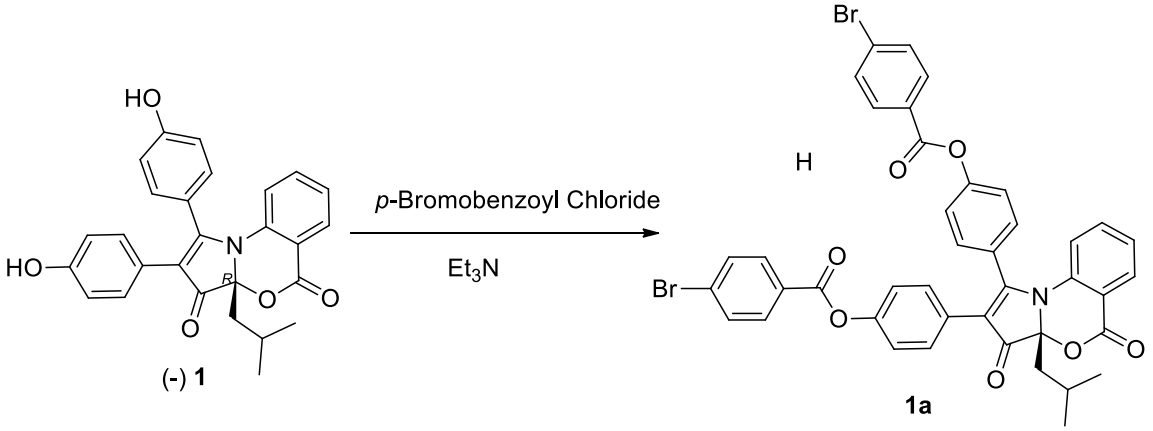
Discoipyrrole C (**3**) yellow solid; $[\alpha]_D$ 0^o (*c* 1.0 MeOH); UV (MeOH) λ_{max} (log ε) 212 (3.91), 264 (3.52), 385 (3.14); 1D and 2D NMR data see Table S3. ESI-MS *m*/*z* 362.2 $[M + Na]^+$, 338.2 $[M - H]^-$. HRESIMS *m*/*z* 340.1554 $[M + H]^+$ ($C_{20}H_{22}NO_4$, calcd 340.1549).

Discoipyrrole D (**4,** mixture of diastereomers at C-1) yellow solid; UV (MeOH) λ_{max} (log ε) 210 (4.28), 272 (3.81), 325 (3.32), 400 (3.47); 1D and 2D NMR data see Table S4. ESI-MS *m*/*z* 653.2 $[M + Na]^+$, 629.2 $[M - H]^-$. HRESIMS *m*/*z* 631.2444 $[M + H]^+$ ($C_{38}H_{35}N_2O_7$, calcd 631.2444).

2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) yellow solid; UV (MeOH) λ_{max} (log ε) 237 (3.87), 307 (3.94), 352 (3.91) ; ^{1}H NMR (600 MHz, CD_3OD) and ^{13}C NMR (100 MHz) see Table S5. ESI-MS *m*/*z* 221.1 $[M + H]^+$, 219.1 $[M - H]^-$.

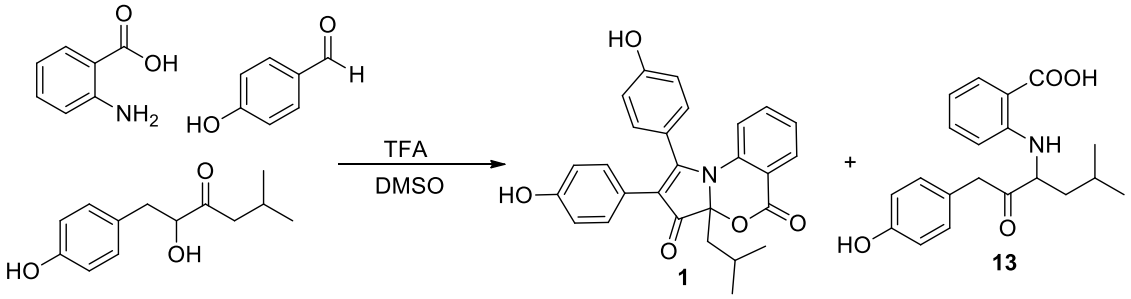
Separation of the two enantiomers of discoipyrrole A (1). **1** (9.0 mg) was separated by AD-H Chiral column (Daicel chemical IND. LTD, 250 × 4.6, 5 μm, 2.5 mL/min, UV = 340 nm) eluted by 15% ethanol in hexane to afford (–)-**1** (4.5 mg, t_R = 16.2 min) and (+)-**1** (4.4 mg, t_R = 19.7 min) as showed in Figure S4.
(–) Discoipyrrole A (4.5 mg) yellow solid; $[\alpha]_D$ –78.6 (*c* 0.28 MeOH).
(+) Discoipyrrole A (4.4 mg) yellow solid; $[\alpha]_D$ +76.1 (*c* 0.28 MeOH).

Synthesis of 8, 12-di-*p*-bromobenzoate derivative of (–) discoipyrrole A (1a).



A solution of (–) discoipyrrole A (4.0 mg) in dry CH_2Cl_2 (2.0 mL) was treated with triethylamine (4.0 μL) under nitrogen. A solution of 4-bromobenzoyl chloride (6.0 mg) in dry CH_2Cl_2 (300 μL) was added at 0 ^oC. After completion of the reaction (about 2 h) as monitored by TLC, the solvent was removed under N_2. The residue was purified by normal phase HPLC (Phenomenex Luna, siliga, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 340 nm) using a gradient solvent system from 10% to 40% EtOAC in Hexane over 20 min to afford a yellow powder (-) **1a** (4.7 mg, t_R = 15.8 min) which was crystallized from MeOH/$CHCl_3$ (10:1). ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.09 (dd, *J* = 7.9, 1.1 Hz, 1H), 8.04 (d, *J* = 8.5 Hz, 2H), 8.00 (d, *J* = 8.5 Hz, 2H), 7.67 (d, *J* = 8.5 Hz, 2H), 7.62 (d, *J* = 8.5 Hz, 2H), 7.35 (dd, *J* = 7.5, 7.5 Hz, 1H), 7.34-7.31 (m, 4H), 7.23 (d, *J* = 8.8 Hz, 2H), 7.22 (dd, *J* = 8.0, 8.0 Hz, H), 7.07 (d, *J* = 8.8 Hz, 2H), 6.33 (d, *J* = 8.2 Hz, 1H), 2.37 (dd, *J* = 14.2, 6.0 Hz, 1H), 2.00 (dd, *J* = 14.2, 7.1 Hz, 1H), 1.77 (m, *J* = 6.5 Hz, 1H), 0.95 (d, *J* = 6.7 Hz, 3H), 0.84 (d, *J* = 6.7 Hz, 3H). ESI-MS *m*/*z* 806.1 $[M + H]^+$, 804.1 $[M - H]^-$. HRESIMS *m*/*z* 806.0381 $[M + H]^+$ ($C_{41}H_{29}BrNO_7$, calcd 806.0389).

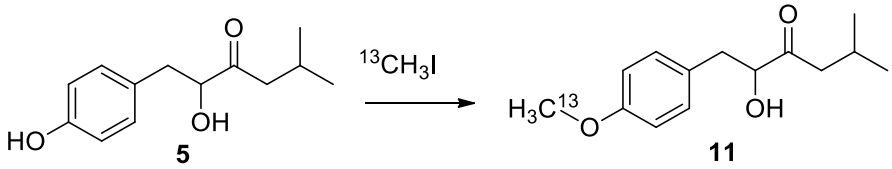
Synthesis of discoipyrrole A (1)



To a solution of *p*-hydroxybenzaldehyde (1.2 mg) and anthranilic acid (1.9 mg) in DMSO (400 μL) was added a solution of 4-hydroxysattabacin (2.2 mg) in DMSO (200 μL) and TFA (6 μL), and the mixture was stirred at 45°C for 30 hours before added 10 mL H_2O and extracted with EtOAC. The organic layer was dried and purified by reversed phase HPLC (Phenomenex Luna, Phenyl-Hexyl, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 210 nm) using a gradient solvent system from 40% to 90% CH_3CN (0.1% FA) over 30 min to give discoipyrrole A (**1**, t_R = 16.9 min, 1.1 mg) and **13** (t_R = 16.0 min, 2.5 mg).
Discoipyrrole A (**1**)**:** ^{1}H NMR (600 MHz, CD_3OD) see Table S1 and attached spectrum comparing with natural discoipyrrole A. ESI-MS *m*/*z* 464.2 $[M + Na]^+$, 440.2 $[M - H]^-$.

Compound **13:** ^{1}H NMR (600 MHz, CD_3OD): δ 7.91 (d, *J* = 7.9 Hz, 1H), 7.16 (ddd, *J* = 7.8, 7.8, 1.6 Hz, 1H), 6.92 (d, *J* = 8.5 Hz, 2H), 6.67 (d, *J* = 8.5 Hz, 2H), 6.58 (dd, *J* = 7.8, 7.8 Hz, 1H), 6.27 (d, *J* = 8.0 Hz, 1H), 4.05 (dd, *J* = 9.2, 5.3 Hz, 1H), 3.70 (d, *J*= 15.8 Hz, 1H), 3.57 (d, *J*= 15.8 Hz, 1H), 1.79 (m, 1H), 1.63 (m, 2H), 0.96 (d, *J*= 6.7 Hz, 3H), 0.89 (d, *J*= 6.7 Hz, 3H). ^{13}C NMR (100 MHz, CD_3OD): δ 214.1, 173.7, 157.1, 150.3, 133.9, 133.2, 125.6 (×2), 131.5 (×2), 116.2, 116.1, 116.0, 111.7, 62.0, 44.2, 42.2, 26.0, 23.2, 21.8, ESI-MS *m*/*z* 342.1 $[M + H]^+$, 340.1 $[M - H]^-$.

Synthesis of *p*-[^{13}C]-methoxysattabacin (11)



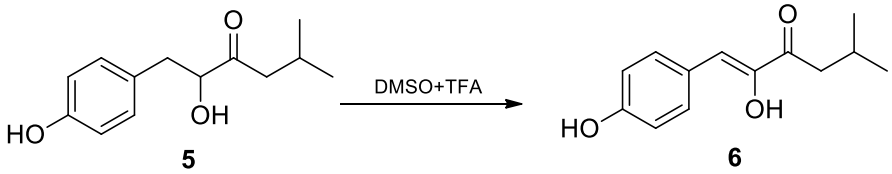
To a solution of 222 mg of hydroxysattabacin (**5**) in 4 mL DMF was added 68 μL $^{13}CH_3I$ and 148 mg K_2CO_3. The mixture was stirred at room temperature overnight. The product was extracted with EtOAc, washed with brine, and dried. The crude product was chromatographed on silica gel using ethyl-acetate-hexane (3:7) to give 200 mg (85%) of *p*-[^{13}C]-methoxysattabacin (**11**). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.06 (d, *J* = 8.6 Hz, 2H, H-2 and H-6), 6.74 (d, *J* = 8.6 Hz, 2H, H-3 and H-5), 4.20 (dd, *J* = 7.5, 4.5 Hz, 1H, H-8), 3.67 (d, $^1J_{CH}$ = 144 Hz, 3H, -OCH_3), 2.96 (dd, *J* = 14.2, 4.5 Hz, 1H, Hα-7), 2.68 (dd, *J* = 14.2, 7.5 Hz, 1H, Hβ-7), 2.29 (d, *J* = 6.9 Hz, 1H, H-10), 2.06 (m, 1H, H-11), 0.83 (d, *J* = 6.9 Hz, 6H, H-12 and H-13).). ^{13}C NMR (100 MHz, $CDCl_3$): δ 211.7 (C-9), 158.4 (C-4), 130.3 (C-2 and C-6), 128.7 (C-1), 113.8 (C-3 and C-5), 77.6 (C-8), 55.1 (-OCH_3), 47.3

(C-10), 40.5 (C-7), 24.3 (C-11), 22.6 (C-12), 22.5 (C-13). ESI-MS *m*/*z* 238.1 $[M + H]^+$, 236.1 $[M - H]^-$.

Synthesis of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (14)

To a solution of *p*-hydroxybenzaldehyde-1-^{13}C (1.3 mg) and anthranilic acid (2.0 mg) in DMSO (400 μL) was added a solution of *p*-[$^{13}CH_3$]-methoxysattabacin (**11**, 2.5 mg) in DMSO (200 μL) and TFA (6 μL), and the mixture was stirred at 45°C for 30 hours before added 10 mL H_2O and extracted with EtOAC. The organic layer was dried and purified by reversed phase HPLC (Phenomenex Luna, Phenyl-Hexyl, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 210 nm) using 60% CH_3CN (0.1% FA) to give 4-[^{13}C]-12-*O*-[^{13}C]-methyl-discoipyrrole A (**14,** t_R = 19.5 min). 1D and 2D NMR data see Table S7. ESI-MS *m*/*z* 456.1 $[M - H]^-$, 458.1 $[M + H]^+$, 480.2 $[M + Na]^+$.

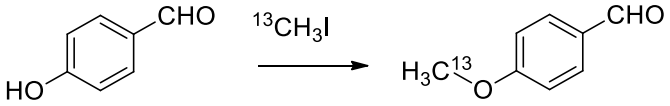
Oxidation of 4-hydroxysattabacin



5 **6**

A solution of 4-hydroxysattabacin (**5**, 2.2 mg) in dry DMSO-d_6 (300 μL) was added 20 μL TFA and stirred at 45°C for 2 days. LC-MS analysis detected the presence of enol **6** and *p*-hydroxybenzaldehyde (see Figure S7).

Synthesis of *p*-[^{13}C]-methoxybenzaldehyde



To a solution of 29 mg of benzaldehyde in 500 μL DMF was added 15 μL $^{13}CH_3I$ and 35 mg K_2CO_3. The mixture was stirred at room temperature for 5 hours. The product was extracted with EtOAc, washed with brine, and dried to give 30 mg (92%) of *p*-[^{13}C]-methoxybenzaldehyde. ^{1}H NMR (600 MHz, $CDCl_3$): δ 9.87 (s, 1H), 7.83 (d, J = 8.7 Hz, 2H), 7.00 (d, J = 8.7 Hz, 2H), 3.87 (*d,* $^1J_{CH}$ = 145 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 191.1, 164.8, 132.2 (×2), 130.1, 114.5 (×2), 55.8 (C-8). ESI-MS *m*/*z* 138.1 $[M + H]^+$, 136.1 $[M - H]^-$.

Procedure to generate 4-[^{13}C]-discoipyrrole A (10) using precursor-driven biosynthesis.
25 mg of *p*-hydroxybenzaldehyde-1-^{13}C was added to autoclaved A1 media containing 8 g/L of agar. The resulting agar plate (24.5 cm × 24.5 cm) were inoculated with *B. hunanensis* and allowed to grow for 7 days. After 7 days, the agar was soaked in MeOH (500 mL) and subsequently concentrated. The resulting crude extract was analyzed by LC-MS to reveal **10**. Further purification by reversed phase HPLC (Phenomenex Luna, Phenyl-Hexyl, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 210 nm) using a gradient solvent system from 40% to 90% CH_3CN (0.1% FA) over 30 min to give 4-[^{13}C]-discoipyrrole A (**10**, t_R = 16.9 min). ^{1}H NMR (600 MHz, CD_3OD) see attached spectrum. ESI-MS: *m*/*z* 465.1 $[M + H]^+$, 441.1 $[M - H]^-$.

Procedure to generate 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (12) using precursor-driven biosynthesis.
30 mg of *p*-[^{13}C]-methoxysattabacin (**11**, 62 mg) was added to autoclaved A1 media containing 8 g/L of agar. The resulting agar plate (24.5 cm × 24.5 cm) were inoculated with *B. hunanensis* and allowed to grow for 7 days. After 7 days, the agar was soaked in MeOH (500 mL) and subsequently concentrated. The resulting crude extract was purified by reversed phase HPLC (Phenomenex Luna, Phenyl-Hexyl, 250 × 10.0 mm, 2.5 mL/min, 5 mm, UV = 210 nm) using 60% CH_3CN (0.1% FA) to give 12-*O*-[^{13}C]-methyl-discoipyrrole A (**12,** t_R = 19.5 min). 1D and 2D NMR data see Table S6. ESI-MS: *m*/*z* 479.2 $[M + Na]^+$, 455.2 $[M - H]^-$.

General Procedure for Generation of Cell-free Spent Media.
Bacterium SNE-038 was cultured in a 2.8 L Fernbach flasks containing 1 L of a seawater based A1 medium (0.5 g starch, 0.2 g yeast extract, 0.1 g peptone) and shaken at 200 rpm at 27 ^oC. After seven days, the culture was transferred to sterile centrifuge tubes and centrifuged at 5200 RPM for 15 min. The supernatant was filtered using 0.22 μm polyethersulfone filter system (nonpyrogenic sterile polystyrene). The resulting supernatant was heated to 90 ºC for five minutes followed by filtration through a 5KD filter. The pH of the media was measured as 9.0.

Procedure for generation of 1 from Biosynthetic Precursors in "Spent" Media
10 mg/L of **5**, 8 mg/L of *p*-hydroxybenzaldehyde and 8 mg/L of anthranilic acid were added to the cell free spent A1 media at 37 ºC and allowed to shake for 5 days. The reaction was monitored at 2, 4, and 5 days by removing a 10 mL aliquot and extracted with ethyl acetate (3 x 10 mL). After 5 days the remaining culture was extracted with ethyl acetate (3 x 1L). The organic phase was conc. *in vacuo* and the products analyzed by LC-MS.

Cytotoxicity assays: Cell lines were cultured in 10 cm dishes (Corning, Inc.) in NSCLC cell-culture medium: RPMI/L-glutamine medium (Invitrogen, Inc.), 1000 U/mL penicillin (Invitrogen, Inc.), 1 mg/mL streptomycin (Invitrogen, Inc.), and 5% fetal bovine serum (Atlanta Biologicals, Inc.). Cell lines were grown in a humidified environment in the presence of 5% CO_2 at 37 ^oC. For cell viability assays, HCC366, A549 cells (60 μL) were plated individually at a density of 750 and 500 cells/well, respectively, in 384 well microtiter assay plates (Bio-one; Greiner, Inc.). After incubating

the assay plates overnight under the growth conditions described above, purified compounds were dissolved and diluted in DMSO and subsequently added to each plate with final compound concentrations ranging from 40 μM to 1 nM and a final DMSO concentration of 0.5%. After an incubation of 96 h under growth conditions, Cell Titer Glo reagent (Promega, Inc.) was added to each well and mixed. Plates were incubated for 10 min at room temperature and luminescence was determined for each well using an Envision multi-modal plate reader (Perkin-Elmer, Inc.). Relative luminescence units were normalized to the untreated control wells.

Migration through collagen. BR5 fibroblasts were suspended in Micro-matrix gel (rat tail collagen, 4 mg/mL) at a concentration of 30×10^6 cells per ml. 2 μL were placed in the center of each well in a 96-well plate and incubated for 10 minutes to allow collagen polymerization. The remainder of the well was then coated with Micro-matrix gel without cells for 10 minutes, followed by two washes with DMEM and addition of growth medium containing PDGF to stimulate migration, with or without DMSO or the indicated natural products in triplicate wells per condition. After overnight incubation cells were fixed, stained for actin and DNA, and imaged. Migration was quantitated as the number of nuclei outside the rim of the initial collagen plug. Triplicates were averaged, and values were normalized to treatment with DMSO alone to obtain % migration vs. control.

Cell migration assays and western blots. BR5 fibroblast cells are hTERT-immortalized, early passage human foreskin fibroblasts. BR5 were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum at 37°C in a 5% CO_2 humidified incubator. BR5 fibroblasts were suspended in Micro-matrix gel (rat tail collagen, 4 mg/ml) at a concentration of 30×10^6 cells per ml. 2 ul were placed in the center of each well in a 96-well plate and incubated for 10 minutes to allow collagen polymerization. The remainder of the well was then coated with Micro-matrix gel without cells for 10 minutes, followed by two washes with DMEM and addition of growth medium containing PDGF to stimulate migration, with or without DMSO or discoipyrroles in triplicate wells per condition. After overnight incubation cells were fixed, stained for DNA with propidium iodide, and imaged. Migration was quantitated as the number of nuclei outside the rim of the initial collagen plug. Triplicates were averaged, and values were normalized to treatment with DMSO alone to obtain % migration vs. control.

BR5 fibroblast cells in culture were treated with the indicated fractions, compounds, or DMSO alone for 18 hours. Whole cell lysate was collected, separated by electrophoresis through an SDS-PAGE gel, transferred to a PVDF membrane, and detected by anti-DDR2 antibody (R&D AF2538). 5% DMSO was used as a control for these experiments.

Functional Signature of Ontology Screen (FUSION)

All gene expression-based functional signature screens were performed in triplicate in the human colon tumor cell line HCT116 in 384-well microtiter plate format, with each perturbation (miRNA, siRNA, natural product fraction) being added individually to each well at a concentration described below. Cells were propagated in DMEM (Cellgro) supplemented with 10% FBS (serum Lot M0017 from Atlanta Biologicals), L-Glutamine,

and NEAA. Transfection of Thermo Fisher Scientific's (Dharmacon) human si*ARRAY*® library SMARTpools and miRNA mimics was based on the Wet Reverse Transfection Version 2.0 protocol with modifications developed for use with equipment and layout of the Eppley Cancer Center HTS Screening Facility. siRNA smartpools and miRNA mimics (final concentrations of 50 nM) were complexed with Dharmafect 4 (0.05 μl/well) reagent prior to plating with HCT116 cells (2,500 cells/well) in a final volume of 30 µl propagation medium. After 72 hrs, cells in each individual well were lysed in 20 μl of QuantiGene Plex 2.0 assay (Panomics) lysis buffer containing 10 μL of Proteinase K per mL lysis buffer and stored at -80°C until further assayed for gene expression signatures. Natural product library screens were performed in HCT116 cells (2,500 cells/well) plated in 50 μL of propagation medium (described above). 48 hours post plating, cells were treated with 60 μg/mL of natural product extract (0.3 µL) Cell lysates were generated and frozen as described above. Endogenous expression levels of ALDOC, NDRG1, BNIP3, BNIP3L, ACSL5, LOXL2, HPRT, and PPIB mRNA in each well were measured using QuantiGene Plex 2.0 assays (Panomics) with a Luminex 200 machine (Luminex) following manufacturer's protocols. The protocol is based on branched DNA signal amplification using eight fluorescent magnetic bead panels with each bead panel comprised of eight different fluorescent magnetic beads (8x8-plex with a total number of individual fluorescent bead classifications=64). Each cell lysate (40 μl) was added directly to a distinct bead/reagent mix and hybridized overnight at 54°C in a Cytomat 2 shaking incubator (1500 hz). After hybridization, branched signal amplification steps with preamplifier, amplifier, biotin label, and streptavidin-phycoerythrin reagents were performed according to manufacturer's protocol and measured by Luminex 200.

The miRIDIAN miRNA mimic library from Dharmacon used in this study contained 426 mimics inclusive of miRBase version 8.0 (with 20% internal redundancy). The siRNA library (also from Dharmacon) contained pools of 4 oligos targeting each of 780 kinases, phosphatases and kinase signaling accessory proteins.

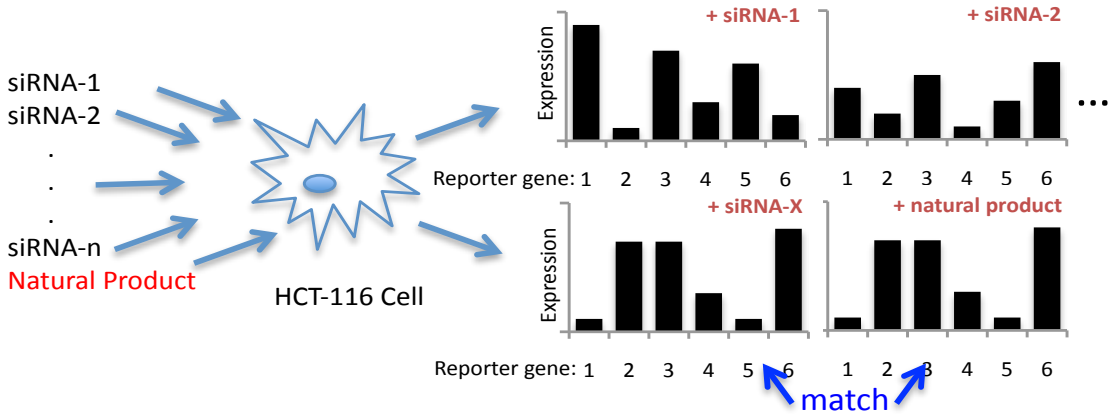


Figure S1. Generic representation of gene expression data obtained from FUSION. In this case the gene expression pattern of siRNA-X and natural product are identical and would suggest the natural product has a mechanism of action involving siRNA-X.

Artificial Seawater Recipe (10 L):

Solution **1** (dissolve into 8L):		Solution **2** (dissolve ino 1.93 L)	
NaCl	211.90 g	$MgCl_2 \cdot 6H_2O$	95.92 g
Na_2SO_4	35.5 g	$CaCl_2 \cdot 6H_2O$	13.44 g
KCl	5.99 g	$SrCl_2 \cdot 6H_2O$	0.218 g
$NaHCO_3$	1.74 g		
KBr	0.863 g		
Boric acid	0.230 g		
NaF	0.028 g		

Nutrient 1 (add 10 mL of the stock solution):
$NaNO_3$ 46.70 g in 1L deionized water (dH_2O)
Nutrient 2 (add 10 mL of the stock solution):
$NaH_2PO_4 \cdot H_2O$ 3.09 g in 1L dH_2O
Metals Stock 1- iron (add 10 mL of the stock solution):
$FeCl_3 \cdot 6H_2O$ 1.77 g in 1L d H_2O
$Na_2EDTA \cdot 2H_2O$ 3.09 g
Metals Stock 2- trace metals (add 10 mL of the stock solution):
$ZnSO_4 \cdot 7H_2O$ 0.073 mg in 1L dH_2O
$CoSO_4 \cdot 7H_2O$ 0.016 g
$MnSO_4 \cdot 2H_2O$ 0.54 g
$Na_2MoO_4 \cdot 2H_2O$ 1.48 mg
$Na2SeO_3$ 0.0173 mg
$NiCl_2 \cdot 6H_2O$ 1.49 mg
$Na_2EDTA \cdot 2H_2O$ 2.44 g
Vitamin Stock (add 10 mL of the stock solution):
Thiamine –HCl 0.1 g
Biotin 2.0 mg
Vitamin B12 1.0 mg

Complete reference 5

Hammerman, P. S.; Sos, M. L.; Ramos, A. H.; Xu, C.; Dutt, A.; Zhou, W.; Brace, L. E.; Woods, B. A.; Lin, W.; Zhang, J.; Deng, X.; Lim, S. M.; Heynck, S.; Peifer, M.; Simard, J. R.; Lawrence, M. S.; Onofrio, R. C.; Salvesen, H. B.; Seidel, D.; Zander, T.; Heuckmann, J. M.; Soltermann, A.; Moch, H.; Koker, M.; Leenders, F.; Gabler, F.; Querings, S.; Ansen, S.; Brambilla, E.; Brambilla, C.; Lorimier, P.; Brustugun, O. T.; Helland, A.; Petersen, I.; Clement, J. H.; Groen, H.; Timens, W.; Sietsma, H.; Stoelben, E.; Wolf, J.; Beer, D. G.; Tsao, M. S.; Hanna, M.; Hatton, C.; Eck, M. J.; Janne, P. A.; Johnson, B. E.; Winckler, W.; Greulich, H.; Bass, A. J.; Cho, J.; Rauh, D.; Gray, N. S.; Wong, K. K.; Haura, E. B.; Thomas, R. K.; Meyerson, M., Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer discovery* **2011,** *1* (1), 78-89.

Detailed Structural elucidation of discoipyrroles A-D

Discoipyrroles A (1)

Discoipyrrole A (**1**) was obtained as a yellow solid and was determined to have a molecular formula of $C_{27}H_{23}NO_5$, based on HR-ESIMS $[M + H]^+$ *m/z* 442.1652 and interpretation of NMR data. The UV spectrum of **1** exhibited absorption bands at 396, 324, 263 and 214 nm, indicative of a highly conjugated system. The ^{1}H NMR spectrum (Table S1) revealed a number of signals indicative of di-substituted aromatic rings (δ_H 6.3 – 8.0 ppm) and a small ketide fragment (δ_H 0.7 – 2.1 ppm), while the ^{13}C NMR spectrum contained multiple sp^2 carbons and two downfield signals indicative of carbonyls (δ_C 193.9 and 161.3 ppm). Detailed analysis of 1D and 2D NMR data (in both DMSO-d_6 and CD_3OD) allowed the assignment of the simple substructures ***I***–***IV*** (Figure 3) for **1**.

Partial structures ***I*** (C-4 through C-8) and ***II*** (C-3, C-9 – C-12) were assigned as 1,4-disubstituted phenyl rings with an additional sp^2 substituent in the benzylic positions (C-4, C-3). Key NMR correlations for the structural assignment of ***I*** included HMBC correlation from H-6 (δ_H 7.15) to C-8 (δ_C 159.3) and quaternary carbon C-4 (δ_C 169.8), whereas key correlations for ***II*** include HMBC correlations from H-10 (δ_H 6.91) to C-12 (δ_C 156.2) and C-3 (δ_C 114.1). The exact nature of the substituent attached at the benzylic positions was not readily apparent from the initial investigation of the NMR data. However, the difference in chemical shifts between H-6 (H-6') and H-7 (H-7') of ~ 0.3 ppm, provided evidence that C-4 was part of a vinyl group rather than a carbonyl group. This is predicated on considerable literature NMR data for *p*-hydroxybenzaldehydes (as well as esters, acids, amides) where in all cases the protons equivalent to H-6 and H-7 have a $\Delta\delta$ of ~1.0 ppm (Figure S1). Alternatively, in *p*-hydroxycinnamates where there is heteroatom substitution in the a or b position, the $\Delta\delta$ value is always ~0.3 ppm (Figure S2). A similar $\Delta\delta$ of ~ 0.3 between H-10 (H-10') and H-11 (H-11') and the ^{13}C chemical shift of C-3 supports the structure of ***II***.[11]

Partial structure ***III*** (C-17 through C-22) was determined to be an anthranilate moiety, with sequential COSY correlations from H-20 through H-23 clearly defining the 1,2 disubstituted benzene ring. The upfield chemical shift of H-20 (δ_H 6.31) was highly suggestive of an electron-donating atom at C-19, while the ^{13}C chemical shift of C-19 (δ_C 136.9) was more compatible with an aniline over a phenol. Moreover, H-23 (δ_H 7.94) showed HMBC correlations with the expected aryl carbons C-21 (δ_C 135.0) and C-19 as well as to the C-17 carbonyl (δ_C 161.3).

The aliphatic partial structure ***IV*** (C-13 through C-16) was established by COSY correlations from the diastereotopic methylene H-13 (δ_H 2.07, 1.92) to H-14 (δ_H 1.66) and from H-14 to methyl singlets H-15 (δ_H 0.85) and H-16 (δ_H 0.76), indicating the presence of an isobutyl group. HMBC correlations from H-14 to a carbon at δ_C 90.3 (C-1) and from H-13 to C-1 and to the C-2 carbonyl (δ_C 193.9) established ***IV***.

An additional structural feature we could deduce based on chemical shifts was an α,β -unsaturated ketone, with the C-2 carbonyl carbon connected to the benzylic carbons C-3 and C-4. The four partial structures and α,β-unsaturated ketone account for all of the carbon and hydrogens in the molecular formula and 15 of the 17 double bond equivalents, suggesting **1** contained two additional rings. The distinct chemical shift of the C-1 quaternary carbon (δ_C 90.3) suggested the presence of a hemiaminal functionality, which allowed us to propose two possible structures (Figure S3). The characteristic chemical shifts of C-17, C-18 and C-19 allowed us to assign the structure as depicted for **1**. A zero $[\alpha]_D$ was measured for **1**, suggesting that the C1 quaternary stereocenter is racemic.

Structural elucidation of discoipyrrole B:

Discoipyrrole B (**2**), determined to have a molecular formula of $C_{27}H_{23}NO_4$ on the basis of a HRESIMS $[M + H]^+$ of *m/z* 426.1697, differs from **1** by only one oxygen Based on the molecular formula and the similarity of the UV, ^1H, and ^{13}C NMR of **1** and **2** (Figure S5, Table S2), we could deduce that **2** lacked the hydroxyl group at C-12. This was fully supported by COSY correlations of a five-proton aromatic spin-system. This structure was further supported by additional COSY and HMBC correlations (Figure S5). Like **1**, discoipyrrole B (**2**) was isolated as a racemic mixture ($[\alpha]_D = 0$).

Structural elucidation of discoipyrrole C:

Discoipyrrole C (**3**) is the simplest of the analogs, as it lacks the C-17 to C-23 anthranilate moiety, which is replaced by ammonia. The molecular formula of **3** was determined as $C_{20}H_{21}NO_4$ based on the HRESIMS $[M + H]^+$ of *m/z* 340.1554, indicating eleven degrees of unsaturation, six fewer than **1**. Detailed evaluation of the ^1H and ^{13}C NMR revealed **3** contained all of the signals for partial structures ***I, II*** and ***IV,*** but lacked the characteristic signals of anthranilic acid (Table S3). Acquisition of the ^1H NMR in DMSO-d_6 identified exchangeable protons for the -NH (δ_H 7.96) and -OH (δ_H 6.01) of the C-1 hemiaminal. HMBC correlations from the -NH proton to C-1 (δ_C 86.9), C-2 (δ_C 199.9), C-3 (δ_C 105.9), C-4 (δ_C 170.1), C-5 (δ_C 121.5), and C-13 (δ_C 46.2), allowed assignment of a pyrrol-3-one core of **3** (Table S3 and Figure S5). As with **1** and **2**, compound **3** is racemic ($[\alpha]_D = 0^o$).

Structural elucidation of discoipyrrole D:

Discoipyrrole D (**4**) is the most complicated member of the family, with a molecular formula of of $C_{38}H_{34}N_2O_7$ on the basis of a HRESIMS $[M + H]^+$ of *m/z* 631.2444, which is considerably higher in molecular weight than **1**. Based on the molecular formula and the similarity of the UV, ^1H, and ^{13}C NMR spectra of **1** and **4** (Table S4), we could deduce that the partial structures ***I-IV*** were mostly intact, however, there were additional signals in the aromatic region as well as the oxymethine region of the ^1H NMR spectrum.

COSY and HMBC correlations allowed us to assign the additional molecular mass to a 3-(1*H*-indol-3-yl) propane-1,2-diol unit (Figure S5) connected by a C-C bond to C-22 of the anthranilate group. Sequential COSY correlations extending from H-27 through H-30 mapped a 1,2-disubstituted benzene ring. COSY correlations extending from H-32 through H-34 together with HSQC data suggested a 3,3-disubstituted propane-1,2-diol. Key HMBC correlations from H-24 to C-31, C-25 and C-26 established the indole ring, while correlations from H-32 to C-24, C-25, C-26, C-33, and C34 established a propane-1,2-diol substituent at the 3-position of the indole. Additional HMBC correlations from H-32 to C-21, C-22, and C-23 confirmed the C-C bond between C-32 and C-22.

The additional stereocenters in **4** (C-32/C-33), along with C-1 epimers give rise to a 1:1 diastereomeric mixture. We are able to fully assign the ^{1}H and ^{13}C NMR signals for each of the diastereomers and, after repetitive C18 HPLC, obtained materials enriched in each of the diastereomers for ^{1}H NMR analysis (Table S4). The relative configuration of C-32/C-33 in **4** was assigned as *anti* based on the 1D and 2D-ROESY data and comparison to data for conomicidines A and B, which contain a similar 3-(1H-indolyl) propane-1,2-diol unit.[12] Restricted rotation of the C-32/C-33 bond due to the presence of bulky substituents on C-32, allowed us to determine the relative configuration of **3**. The large $^{3}J_{HH}$ between C-32/C-33 (7.8 Hz), confined the possibilities as depicted by Newman projections of the *anti* and *syn* diastereomers (sse below). Analysis of the ROESY data showed strong NOE correlations between H-33 and H-34α (δ_H 3.55) and between H-33 and H-24 of the indole ring. H-32 showed NOE correlations with H-27, H-23 and H-34β (δ_H 3.44). Importantly, H-34 shows a NOE correlation with H-24 of the indole ring instead of H-21 or H-23 of the anthranilic acid ring. This data is consistent with the *anti* diastereomer.[13] The absolute configuration of C-33 has yet to be determined.

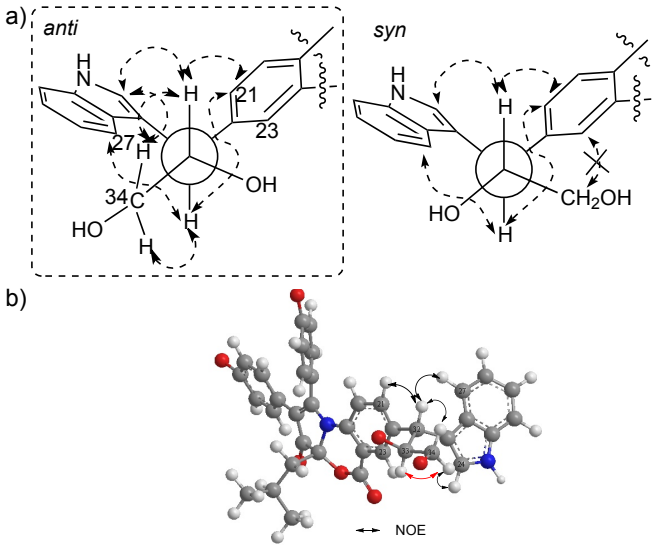
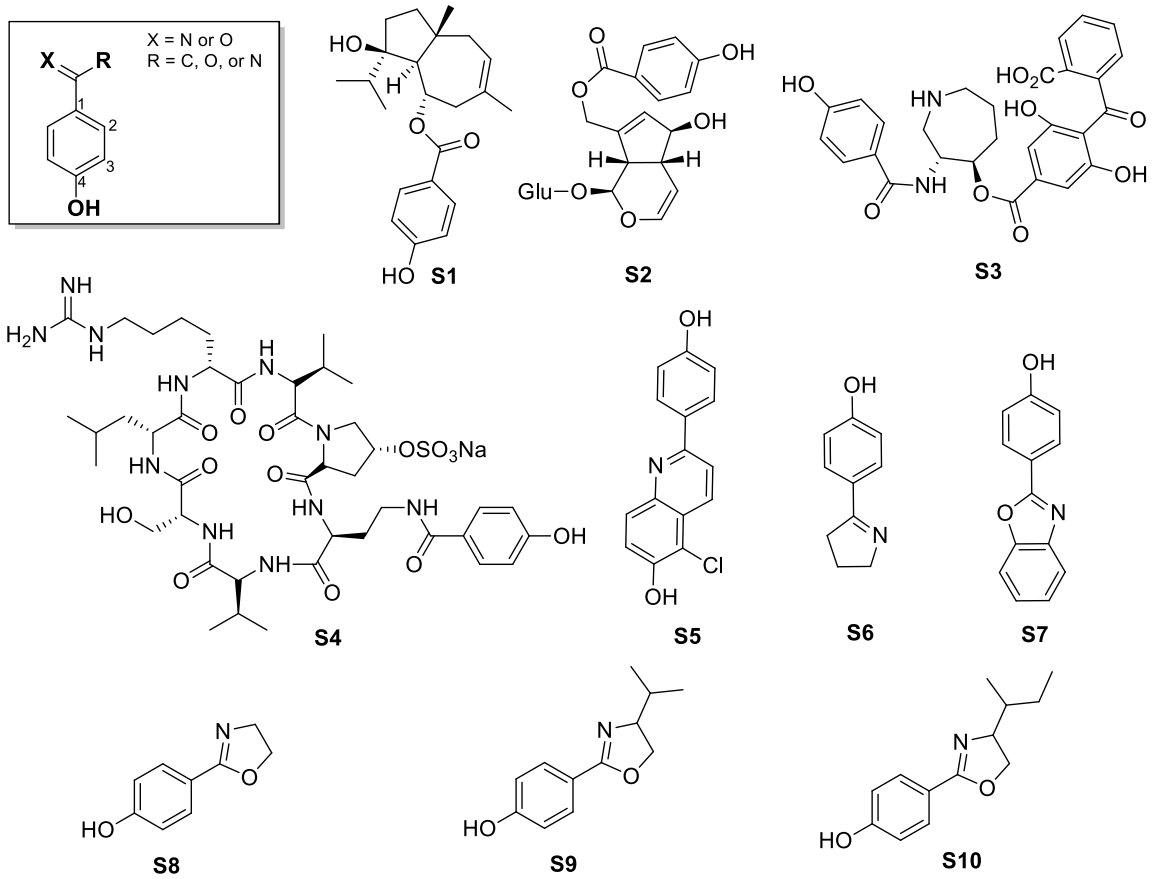


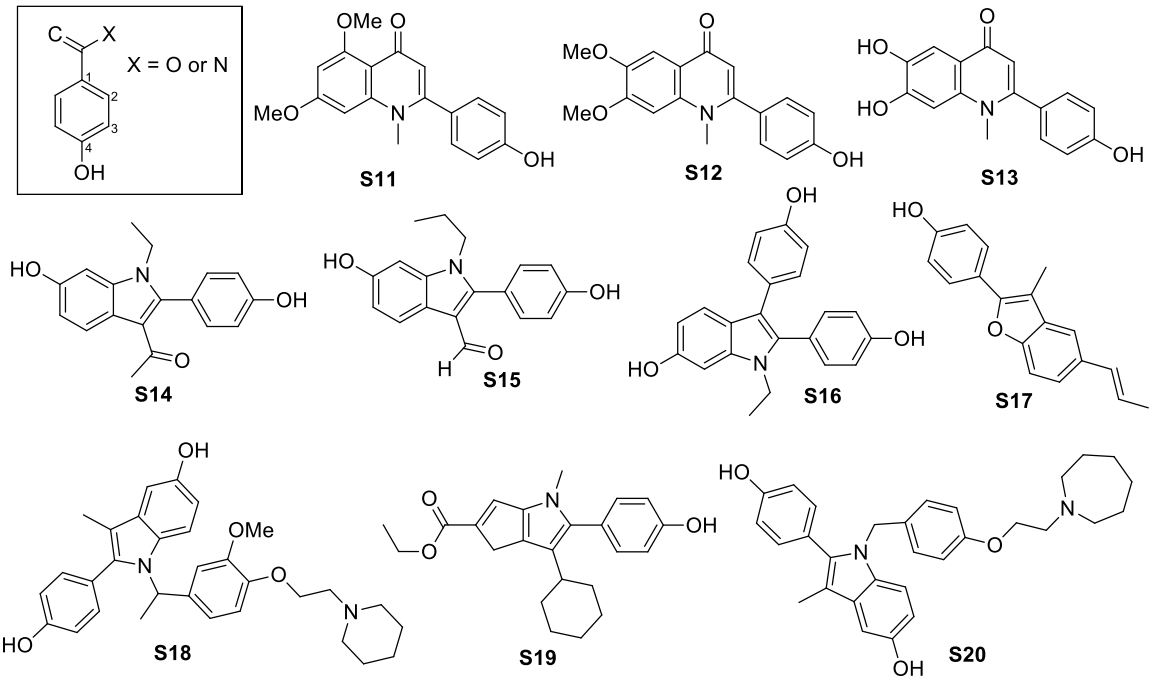
Figure S2: Examples of *p*-hydroxyl-benzoic acid derivatives with key ^{1}H NMR chemical shifts.



Comp.	$\delta_{H\text{-}2}$	$\delta_{H\text{-}3}$	$\Delta\delta = \delta_{H\text{-}2} - \delta_{H\text{-}3}$	References
S1	7.80	6.80	**1.00**	*Phytochemistry* **1983**, *22*, 2231-2233.
S2	7.80	6.75	**1.05**	WO2007/29263 A1, **2007**
S3	7.62	6.74	**0.88**	*J. Am. Chem. Soc.,* **1993**, *115*, 6452-6453
S4	7.72	6.77	**0.95**	*J. Org. Chem*., **1997**, *62*, 7766-7767
S5	8.03	6.91	**1.12**	WO2004/103973 A1, **2004**
S6	7.69	6.81	**0.88**	*Tetrahedron Lett.,* **2005**, *46*, 4213-4217
S7	8.13	7.06	**1.07**	*Org. Lett*., **2011**, *13*, 1804-1807
S8	7.71	6.77	**0.94**	*Euro. J. Med. Chem*., **2009**, *44*, 2994-3008
S9	7.70	6.71	**0.99**	*J. Am. Chem. Soc.,* **1995**, *117*, 8312-8321
S10	7.703	6.73	**0.99**	*J. Am. Chem. Soc.,* **1995**, *117*, 8312-8321

The ^{1}H NMR chemical shift differences ($\Delta\delta$) between the two aromatic protons (H-2 and H-3) are always ~ 1.0 ppm. This large difference is due to not only the significant upfield shift effect of OH on the H-3 protons, but also the downfield shift effect of the ester/amide on the H-2 protons.

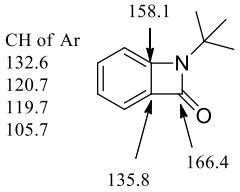
Figure S3: Examples of *p*-hydroxylbenzene derivatives with key ^{1}H NMR chemical shifts.

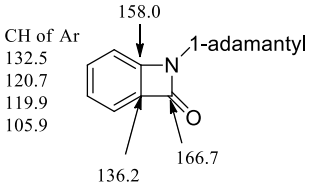


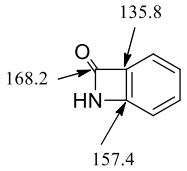
Comp.	δ_{H-2}	δ_{H-3}	$\Delta\delta = \delta_{H-2} - \delta_{H-3}$	References
S11	7.20	6.75	**0.45**	*Bioorg. Med. Chem.Lett.,* **2009***, 19,* 714-717.
S12	7.30	6.85	**0.45**	*J. Med. Chem.,* **2011**, *54*, 5722-5736.
S13	7.25	6.80	**0.45**	*J. Med. Chem.,* **2011**, *54*, 5722-5736.
S14	7.35	6.95	**0.40**	*Archiv der Pharmazie*., **1994,** *327,*481-492.
S15	7.35	7.00	**0.35**	*Archiv der Pharmazie*., **1994,** *327,*481-492.
S16	7.15	6.85	**0.30**	*Eur. J. Org. Chem,* **2001**, 1723-1729.
	7.10	6.76	**0.34**	
S17	7.37	6.98	**0.39**	*Phytochemistry,* **1984***, 23,* 2643-2645.
S18	7.17	6.87	**0.30**	*EP 1076558 B1,* **2003**
S18	7.19	6.92	**0.27**	*EP 1688420,* **2006**
S20	7.16	6.86	**0.30**	*J. Med. Chem.,* **2001**, *44*, 1654-1657.

There is no downfield chemical shift effect on the H-3 protons, so the ^{1}H NMR chemical shift differences ($\Delta\delta$) between the two aromatic protons H-2 and H-3 are consistent between 0.25-0.45 ppm**.**

Figure S4. Examples of synthetic pyrrolobenzoxazinone and N-alkylbenzoazetinones with key ^{13}C NMR chemical shifts

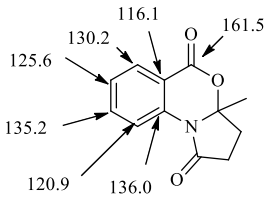


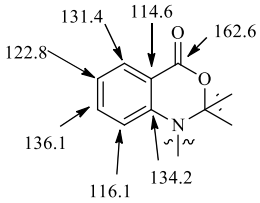


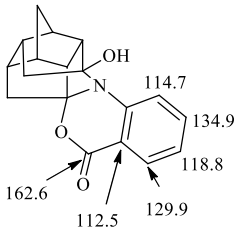


J. Org. Chem., **1984**, *49*, 3367-3372

Tetrahedron Lett,**1999**, *40*, 9271



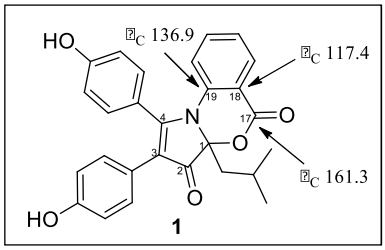


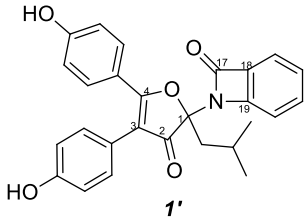


J. Org. Chem. **2010**, *75*, 3274-3282

Tetrahedron, **1996**, *52*, 12475-12482

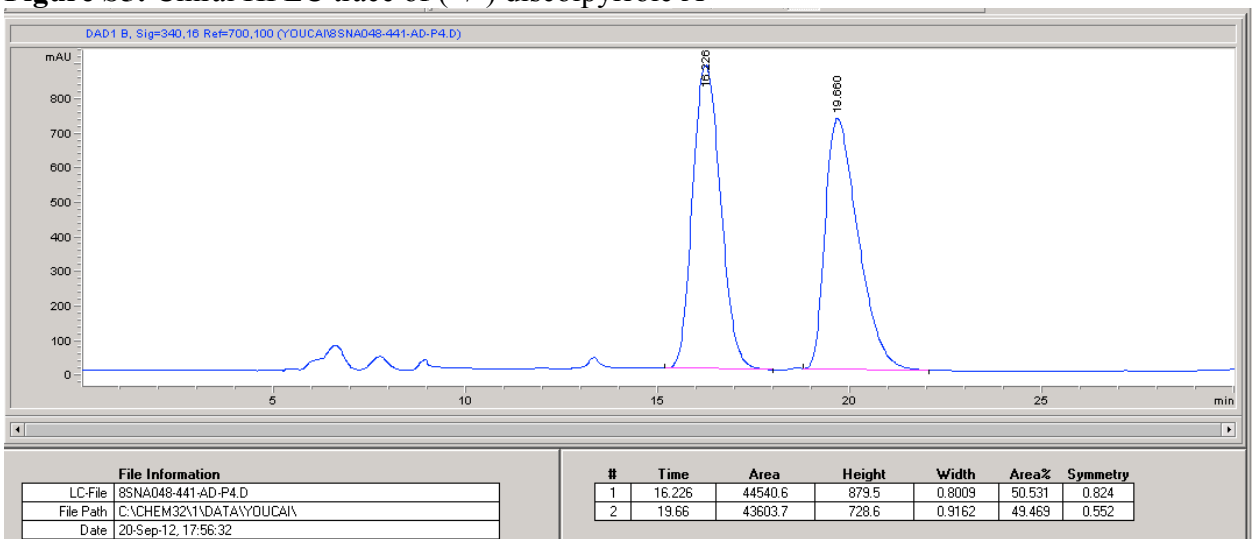
Heterocycles, **2000**, *54*, 2221-2224





Typical ^{13}C NMR chemical shift of synthetic pyrrolobenzoxazinone and N-alkylbenzoazetinones not only ruled out the alternative ***1'***, but supported the assignment of **1.**

Figure S5: Chiral HPLC trace of (+/-) discoipyrrole A



File Information	
LC-File	8SNA048-441-AD-P4.D
File Path	C:\CHEM32\1\DATA\YOUCAI\
Date	20-Sep-12, 17:56:32

#	Time	Area	Height	Width	Area%	Symmetry
1	16.226	44540.6	879.5	0.8009	50.531	0.824
2	19.66	43603.7	728.6	0.9162	49.469	0.552

Column: AD-H Chiral column (4.6 × 250mm; 5mm), Daicel chemical IND. LTD)
Solvents: 15% ethanol in hexane
Flow: 2.5 mL/min
Detector: DAD detector (210, 254, 280, 340 nm)

(–) Discoipyrrole A, t_R = 16.2 min
(+) Discoipyrrole A, t_R = 19.7 min

Figure S6: Key COSY and HMBC correlations of **2, 3** and **4**

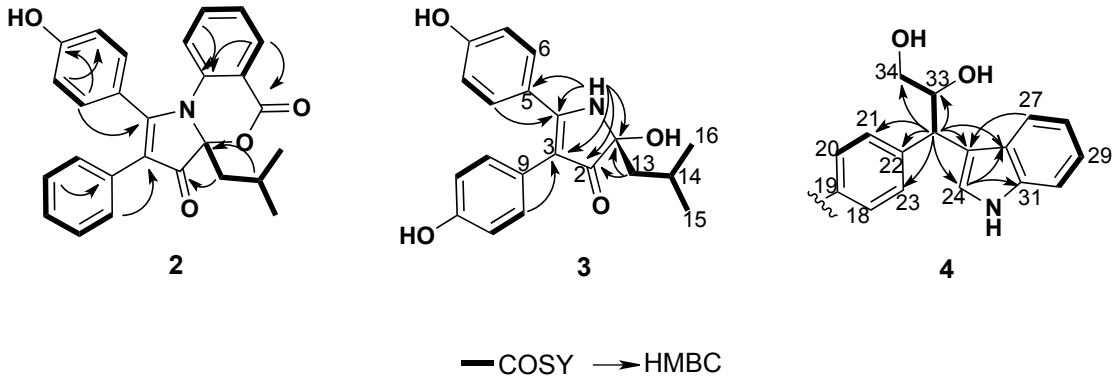
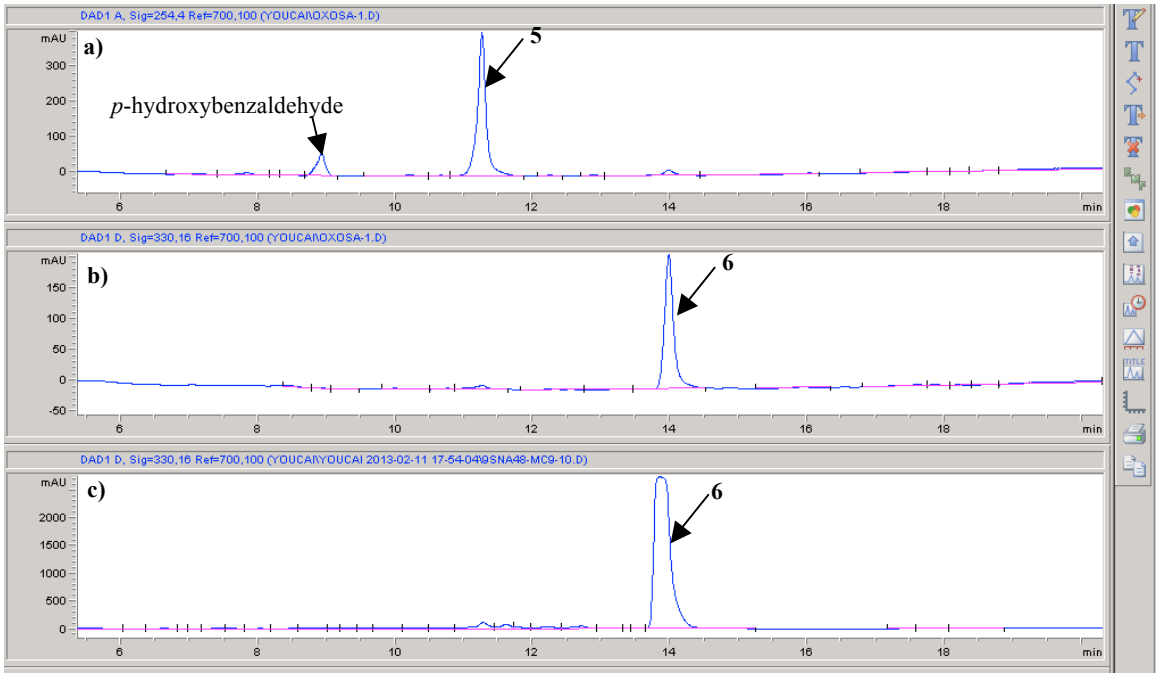


Figure S7: LC-MS trace of oxidative products of 4-hydroxysattabacin (**5**)



a) **5** treated with DMSO/TFA, observed product of *p*-hydroxybenzaldehyde detected by LC-MS at 254 nm;
b) **5** treated with DMSO/TFA, observed product of **6** detected by LC-MS at 330 nm;
c) Pure **6** detected by LC-MS at 330 nm.

Column: C18 column (4.6 × 250mm; 5µm, Phenomenex Luna)
Solvents: 0', 10% CH_3CN (0.1% FA) in H_2O
17', 99% CH_3CN (0.1% FA) in H_2O
25', 99% CH_3CN (0.1% FA) in H_2O
Flow: 0.7 mL/min
Detector: DAD detector (254 and 330 nm)

Figure S8: Generation of *p*-hydroxybenzaldehyde from **5** under oxidation condition

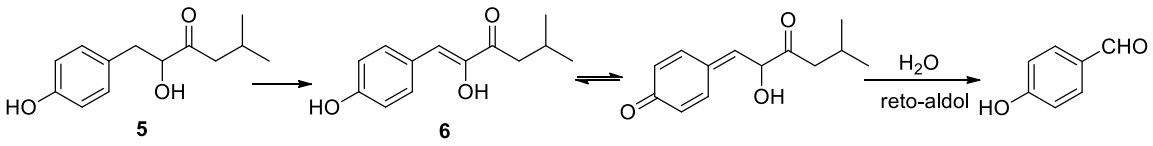


Figure S9. BR5 fibroblast migration assay for **1**–**4**. a, Discoipyyrole A at 1 μg/mL. b, Discoipyyrole B at 1 μg/mL. c, Discoipyyrole C at 1 μg/mL. d, Discoipyyrole D at 1 μg/mL.

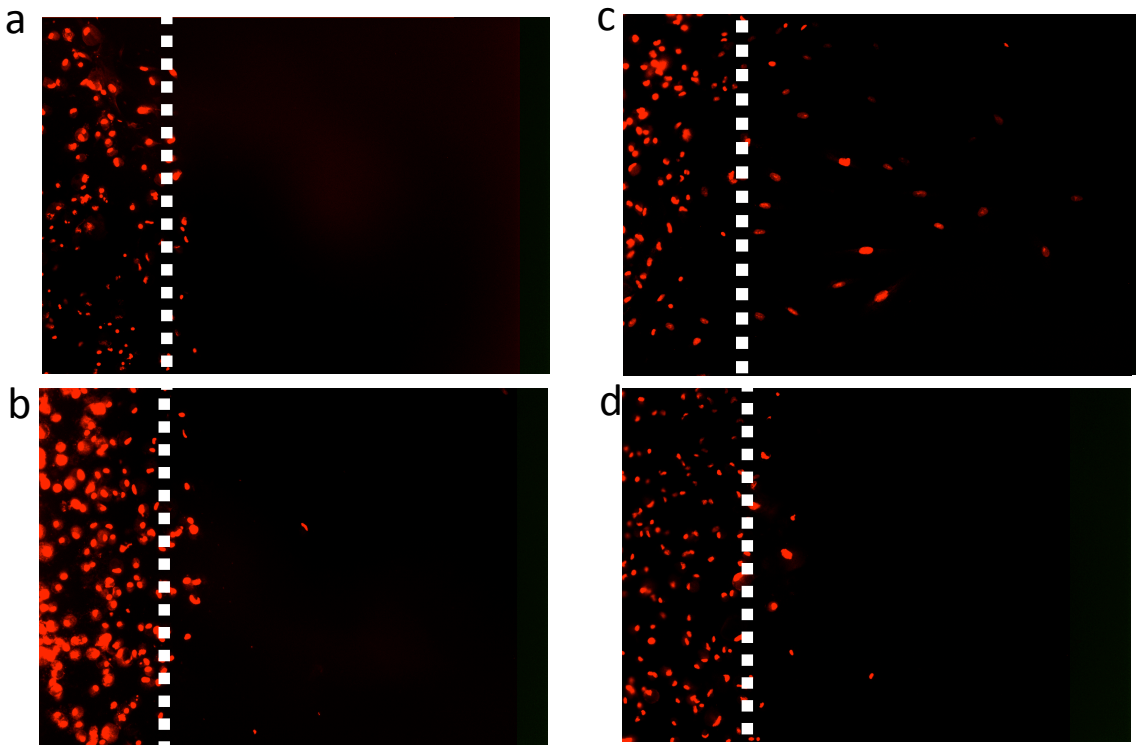
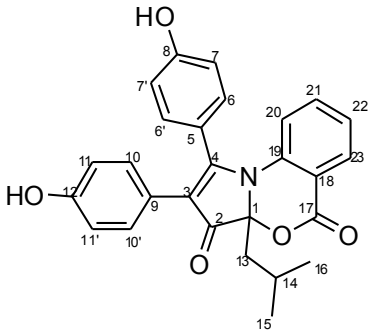


Table S1. 1D and 2D NMR data of discoipyrrole A (**1**) in CD_3OD and DMSO-d_6

no.	δ_H, mult. (*J* in Hz) [a]	δ_C [a]	COSY [a]	HMBC [a]	δ_H, mult. (*J* in Hz) [b]	δ_C [b]	HMBC [b]
1		92.7				90.3	
2		196.6				193.9	
3		116.5				114.1	
4		172.3				169.8	
5		121.1				119.1	
6, 6'	7.07, d (8.0)	131.7	7/7'	6, 8, 4	7.15, d (8.0)	130.2	4, 6, 8
7, 7'	6.85, d (8.0)	117.0	6/6'	5, 7, 8	6.84, d (8.0)	115.9	5, 7, 8
8		161.3				159.3	
9		121.8				119.9	
10, 10'	6.96, d (8.7)	131.6	11/11'	10, 12, 3	6.91, d (8.6)	129.9	10, 12, 3
11, 11'	6.63, d (8.7)	116.1	10/10'	9, 11, 12	6.61, d (8.6)	114.9	9, 11, 12
12		157.8				156.2	
13α	2.21, dd (14.1, 6.1)	43.1	13β, 14	14, 15, 16, 1, 2	2.07, dd (6.1, 14.0)	41.6	14, 15, 16, 1, 2
13β	1.99, dd (14.1, 6.5)		13α, 14	14, 15, 16, 1, 2	1.92, dd (6.4, 14.0)		14, 15, 16, 1, 2
14	1.75, m	25.3	13α, 15, 16	1, 13, 15, 16	1.66, m	23.4	1, 13, 15, 16
15	0.95, d (6.7)	24.4	14	13, 14, 16	0.85, d (6.7)	23.5	13, 16
16	0.84, d (6.7)	23.6	14	13, 14, 15	0.76, d (6.7)	22.7	13, 15
17		163.5				161.3	
18		119.3				117.4	
19		138.7				136.9	
20	6.47, d (7.8)	123.6	21	22, 18, 17	6.31, d (8.2)	121.5	22, 18
21	7.43 ddd (7.8, 7.8, 1.4)	136.2	22, 24	19, 23, 20, 22	7.47, dd (7.8, 8.2)	135.0	19, 23
22	7.28, ddd (7.8, 7.8, 1.4)	126.4	21, 23	20, 18, 21, 23	7.27, d (7.8, 7.8)	124.8	20, 18, 21, 23
23	8.03, dd (7.8, 1.4)	131.8	22	19, 21, 17	7.94, d (7.8)	130.2	19, 21, 17
OH-8					10.1, s		7/7', 8
OH-12					9.4, s		12, 11/11'

[a] in CD_3OD; [b] in DMSO-d_6.

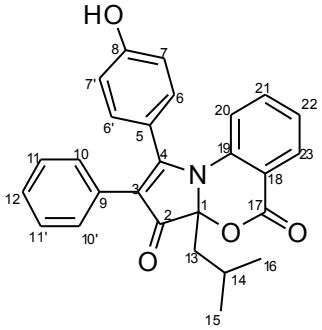


Discoipyrrole A (**1**)

Table S2. 1D and 2D NMR data of discoipyrrole B (**2**) in CD_3OD

no.	δ_H, mult. (*J* in Hz)	δ_C [a]	COSY	HMBC
1		92.6		
2		196.1		
3		116.2		
4		173.1		
5		120.6		
6/6'	7.09, d (8.1)	131.3	7/7'	4, 6/6', 8
7/7'	6.85, d (8.1)	116.8	6/6'	5, 7/7'
8		161.3		
9		130.7		
10/10'	7.13, dd (8.5, 1.6)	130.0	11/11', 12	3, 10', 12
11/11'	7.21, dd(7.5, 7.5)	128.8	10/10', 12	9, 11'
12	7.17, dd (7.2, 7.2)	127.8	10/10'	10
13α	2.22 dd (14.1, 6.1)	42.8	13b, 14	14, 15, 16, 1, 2
13β	1.99, dd (14.1, 6.1)		13a, 14	14, 15, 16, 1, 2
14	1.78, m	24.9	13a, 15, 16	15, 16
15	0.97, d (6.7)	24.0	14	13, 14, 16
16	0.85, d (6.7)	23.1	14	13, 14, 15
17		163.3		
18		119.3		
19		138.4		
20	6.50, d (7.8),	123.4	21	18, 22
21	7.45 ddd (8.2, 8.2, 1.4)	135.9	20, 22	19, 22, 23
22	7.30, ddd (7.8, 7.8, 1.4)	126.2	21, 23	18, 20, 21
23	8.05, dd (7.8, 1.4)	131.5	22	17, 19, 21

[a] derived from HSQC and HMBC experiments

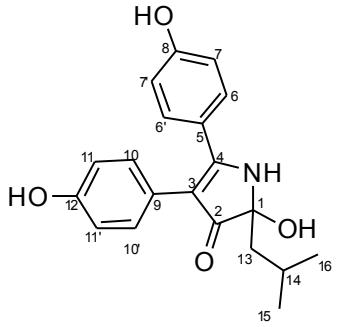


Discoipyrrole B (**2**)

Table S3. 1D and 2D NMR data of discoipyrrole C (**3**) in CD_3OD and DMSO-d_6

no.	δ_H, mult. (*J* in Hz) [a, d]	δ_C [b, d]	COSY [a, d]	HMBC [a,d]	δ_H, mult. (*J* in Hz) [a, e]	δ_C [c, e]	HMBC [e]
1		89.0				86.9	
2		202.3				199.9	
3		109.0				105.9	
4		174.6				170.1	
5		123.2				121.5	
6/6'	7.36, d (8.8)	131.9	7/7'	6, 8, 4	7.28, d (8.8)	130.2	6, 8, 4
7/7'	6.73, d (8.8)	116.4	6/6'	5, 8, 7/7’	6.76, d (8.8)	115.3	5, 8, 7/7'
8		162.3				160.6	
9		125.1				123.7	
10/10'	6.94, d (8.7)	132.1	11/11'	10, 12, 3, 11	6.86, d (8.7)	130.1	10, 12, 3
11/11'	6.68, d(8.7)	116.3	10/10'	9, 12, 11/11’	6.62, d (8.7)	114.7	9, 11/11', 12
12		157.0				155.2	
13	1.85 d (6.2)	46.9	14	15, 16, 1, 2	1.68, m	46.2	15, 16, 1, 2
14	1.67, m	25.2	13, 15, 16	1, 13, 15/16	1.64, m	23.1	1, 13, 15/16
15	0.94, d (6.6)	24.8	14	13, 16, 14	0.87, d (6.2)	24.1	13, 16,14
16	0.93, d (6.6)	24.7	14	13, 15, 14	0.93, d (6.2)	23.8	13, 15, 14
NH					7.96, s		1, 2, 3, 4, 5, 13
1-OH					9.29, s		
8-OH					10.53,s		
12-OH					6.01, s		

[a] recorded at 600 MHz; [b] recorded at 100 MHz; [c] derived from HSQC and HMBC experiments; [d] in CD_3OD; [e] in DMSO-d_6

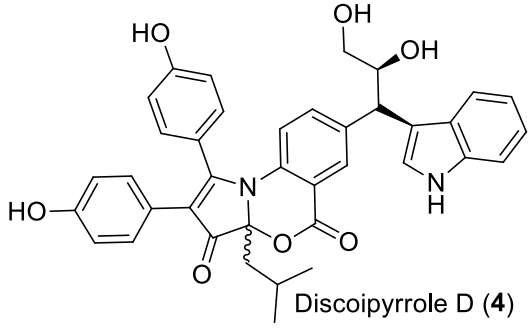


Discoipyrrole C(**3**)

Table S4. 1D and 2D NMR data of discoipyrrole D (**4,** mixture of diastereomers at C-1) in CD_3OD

no.	δ_H, mult. (*J* in Hz)	δ_H, mult. (*J* in Hz)[a]	δ_C [b]	COSY	HMBC	NOE
1			92.3			
2			196.4			
3			115.8			
4			172.4			
5			120.9			
6/6'	7.03, d (8.7)	7.01, d (8.7)	131.6	7/7'	6/6', 8, 4	7/7', 20
7/7'	6.82, d (8.7)	6.79, d (8.7)	116.8	6/6'	5, 7/7', 8	6/6'
8			160.9			
9			121.6			
10/10'	6.93, d (8.7)	6.92, d (8.7)	131.4	11/11'	10/10', 12, 3	11, 15
11/11'	6.62, d (8.7)	6.61, d (8.7)	115.7	10/10'	11/11', 9, 12	10
12			157.4			
13α	2.14, m		42.9	13b; 14	1, 16, 15, 14	13b, 15
13β	1.97, m			13a, 14	1, 2, 14, 15, 16	13a; 16
14	1.72, m		25.1	15, 16, 13α, 13β	1, 13, 15, 16	15; 16
15	0.91, d (6.7)		24.2	14	16, 14, 13	14; 10
16	0.81, d (6.7)		23.3	14	15, 14, 13	14; 20
17			163.7			
18			118.5			
19			136.4			
20	6.38, d (7.8)	6.39, d (7.8)	122.8	21	18, 22	21, 16, 6
21	7.53 dd (7.8, 2.0)		137.1	20, 23	23, 19, 32	20, 32
22			141.7			
23	8.12, d (2.0)	8.09, d (2.0)	131.8	21	19, 21, 17	32
24	7.31, s;	7.30, s	123.2		25, 26, 31	33, 34α. 34β
25			116.5			
26			127.8			
27	7.42, d (8.0)	7.40, d (8.0)	119.5	28	29, 31, 25	28; 32
28	6.95, m		119.7	27, 29	26, 30	27
29	7.06, dd (8.0, 8.0)		122.5	28, 30	27, 31	30
30	7.33, d (8.0)	7.32, d (8.0)	112.3	29	26, 28	29
31			137.7			
32	4.36, d (7.8)		46.0	33	34, 33, 24, 25, 26, 21, 22, 23	34β, 23, 27, 21, 34α(w)
33	4.40, m		75.1	32, 34	4	24, 34α, 34β(w)
34α	3.55, m		65.9	33, 34b		34β, 33, 24
34β	3.44, m			33, 34a		34α, 32, 24

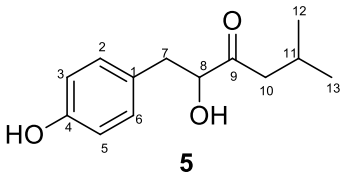
[a] different δ_H of another diastereoisomer; [b] derived from HSQC/ HMBC experiments; (w). signal is weak



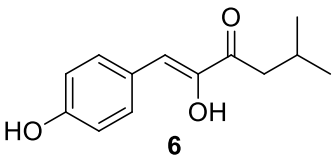
Discoipyrrole D (**4**)

Table S5. NMR data of **5**–**6** in CD_3OD

no	**5**		**6**			
	δ_H, mult. (*J* in Hz)	δ_C	δ_H, mult. (*J* in Hz)	δ_C	COSY	HMBC
1		129.6		127.8		
2, 6	7.03, d (8.4)	131.6	7.74, d (8.8)	133.3	3/5	6/2, 4, 7
3, 5	6.68, d (8.4)	116.2	6.78, d (8.8)	116.4	2/6	3/5, 1, 4
4		157.2		159.1		
7	2.91, dd (14.1, 4.8)	40.3	6.53, s	115.6		2/6, 8, 9
	2.69,dd (14.1, 7.8)					
8	4.19, dd (7.8, 4.8)	79.4		147.9		
9		214.5		199.0		
10	2.35, m	48.8	2.68, d (7.1)	45.2	11	9, 11, 12/13
11	2.06, m (6.7)	25.2	2.19, m (6.7)	27.7	10, 12, 13	12, 13, 10, 9
12	0.87, d (6.7)	23.0	0.99, d (6.7)	23.1	11	13, 11, 10
13	0.85, d (6.7)	23.1	0.98, d (6.7)	23.1	11	12, 11, 10



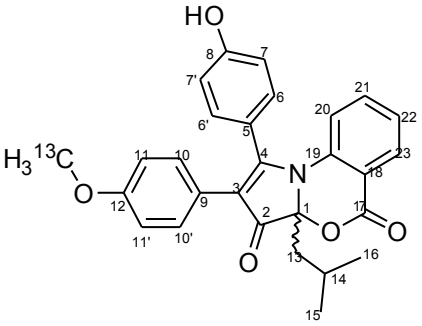
5



6

Table S6. NMR data of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)

no.	δ_H, mult. (*J* in Hz)	δ_C [a]	COSY	HMBC
1		92.6		
2		196.6		
3		116.5		
4		172.3		
5		121.1		
6, 6'	7.07, dd (7.5, 1.4)	131.6	7/7'	6, 8, 4
7, 7'	6.85, dd (7.5, 1.4)	116.9	6/6'	5, 7
8		160.3		
9		123.1		
10, 10'	7.05, d (8.9)	131.4	11/11'	10, 12, 3
11, 11'	6.77, d (8.9)	114.5	10/10'	9, 11, 12
12		160.2		
13a	2.20 dd (14.2 6.1)	43.1	13b, 14	14, 15, 16, 1, 2
13b	1.99, dd (14.2, 6.5)		13a, 14	14, 15, 16, 1
14	1.76, m	25.3	13a, 15, 16	
15	0.95, d (6.7)	24.4	14	13, 14, 16
16	0.84, d (6.7)	23.6	14	13, 14, 15
17		163.5		
18		119.3		
19		138.8		
20	6.47, d (8.0)	123.6	21	22, 18
21	7.43 ddd (7.5, 7.4, 1.6)	136.2	22, 24	19, 23
22	7.28, ddd (7.9, 7.4, 1.0)	126.4	21, 23	20, 18
23	8.04, dd (7.9, 1.1),	131.8	22	19, 21, 17
CH_3O	3.74, d (144)	55.8		11/11’



[a] data derived from HSQC and HMBC experiment

Table S7. NMR data of 4-[^{13}C]-12-*O*-[^{13}CH$_3$]- methyl-discoipyrrole A (**14**) in CD_3OD (600 HMz)

no.	δ_H, mult. (*J* in Hz)	δ_C [a]	COSY	HMBC
1		92.6		
2		196.6		
3		116.5		
4		172.3		
5		121.1		
6, 6'	7.07, dd (7.5, 1.4)	131.6	7/7'	8, 4
7, 7'	6.85, dd (7.5, 1.4)	116.9	6/6'	5, 7
8		160.3		
9		123.1		
10, 10'	7.05, d (8.9)	131.4	11/11'	10, 12, 3
11, 11'	6.77, d (8.9)	114.5	10/10'	9, 11, 12
12		160.2		
13a	2.20 dd (14.2 6.1)	43.1	13b, 14	14, 15, 16, 1
13b	1.99, dd (14.2, 6.5)		13a, 14	14, 15, 16, 1, 2
14	1.76, m	25.3	13a, 15, 16	
15	0.95, d (6.7)	24.4	14	13, 14, 16
16	0.84, d (6.7)	23.6	14	13, 14, 15
17		163.5		
18		119.3		
19		138.8		
20	6.47, d (8.0)	123.6	21	22, 18
21	7.43 ddd (7.5, 7.4, 1.6)	136.2	22, 24	19, 23
22	7.28, ddd (7.9, 7.4, 1.0)	126.4	21, 23	20, 18
23	8.04, dd (7.9, 1.1),	131.8	22	19, 21, 17
CH_3O	3.74, d (144)	55.8		11/11’

[a] data derived from HSQC and HMBC experiment

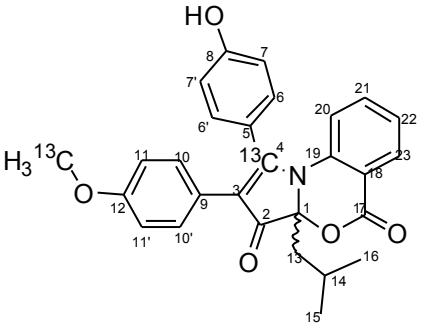


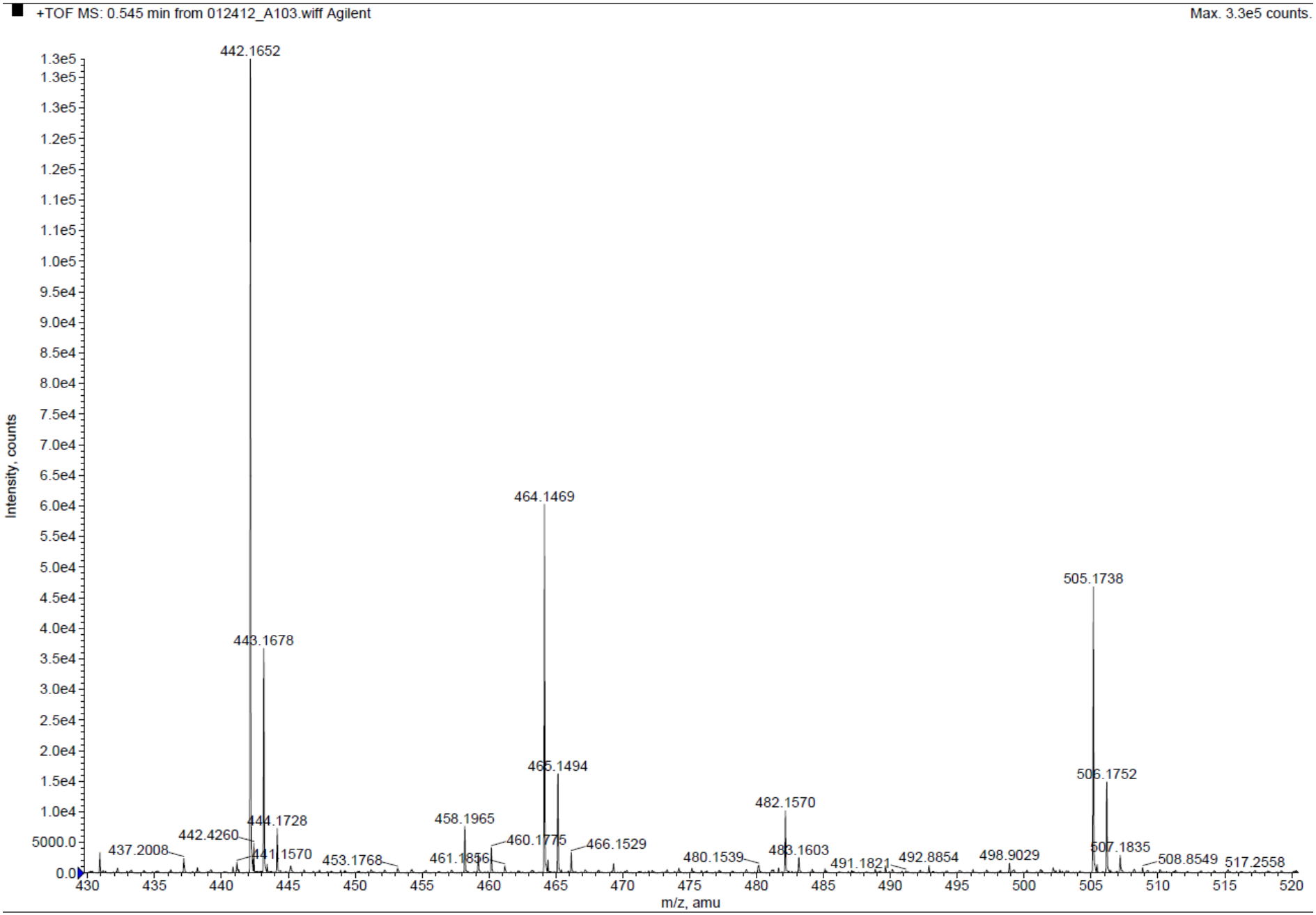
Table S8. Cytotoxicity data of **1-4** against DDR2 mutant and DDR2 wild-type NSCLA celll-lines

Compounds	HCC366 [a]	A549 [a]
1	0.120	10.2
2	0.190	8.8
3	0.712	19.6
4	0.275	13.4

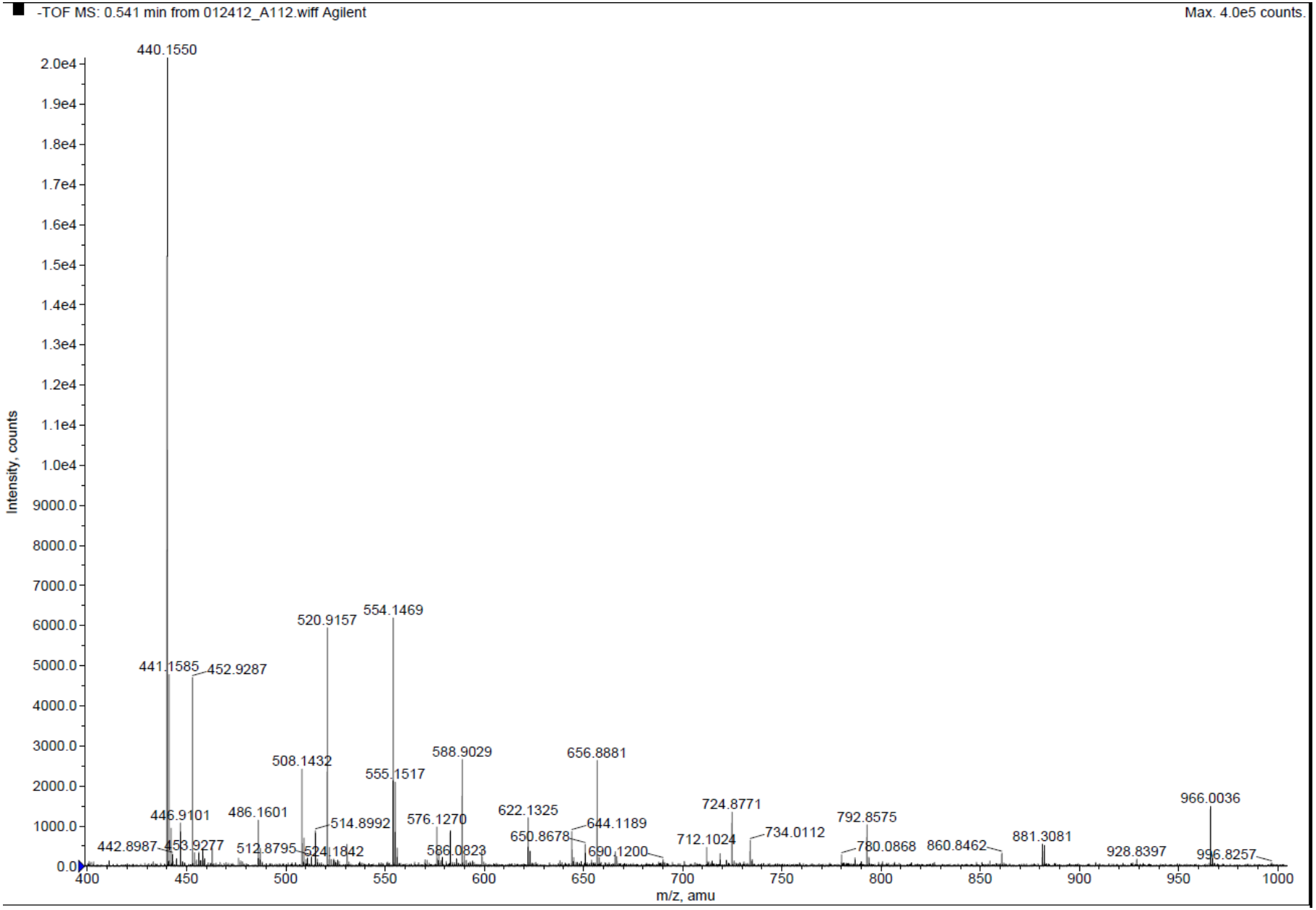
[a] Values in μM

18. HR-ESIMS and NMR spectra of discoipyrrole A (**1**)

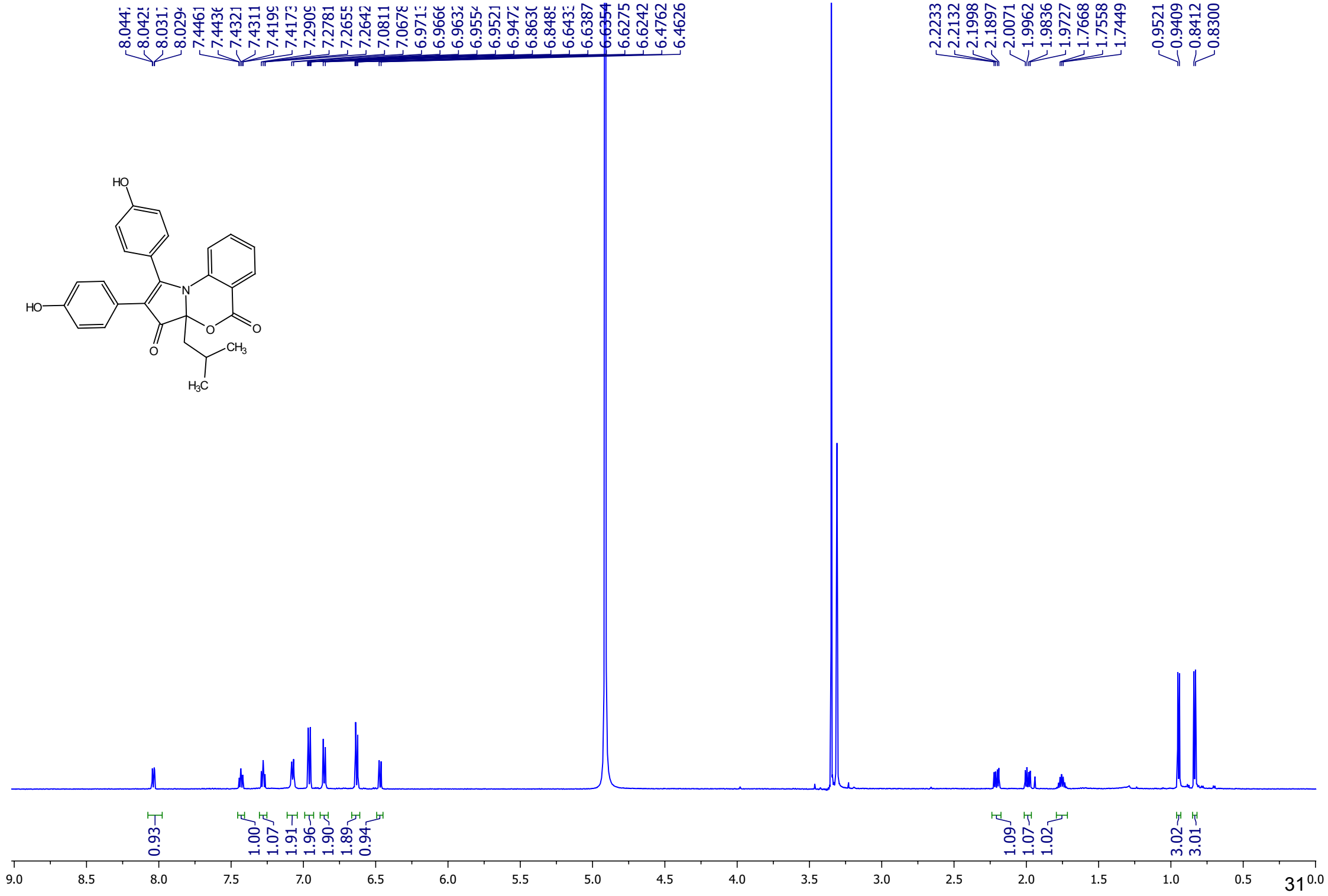
18.1 HR-ESIMS of discoipyrrole A (**1**)-positive



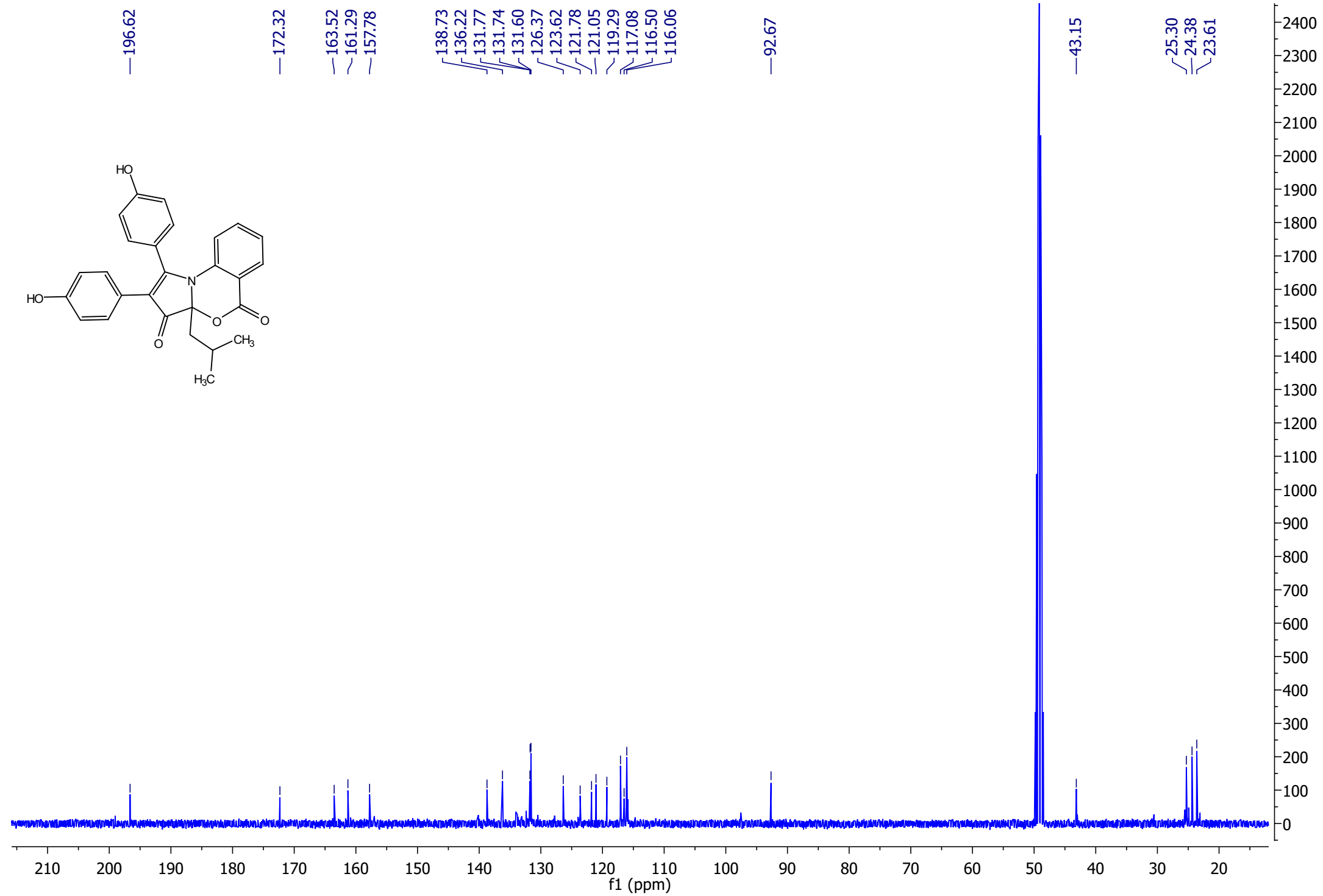
18.1 HR-ESIMS of discoipyrrole A (**1**)-negative



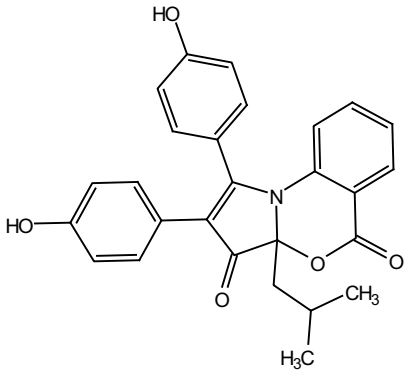
18.2. ^{1}H NMR of discoipyrrole A (**1**) in CD_3OD (600 HMz)



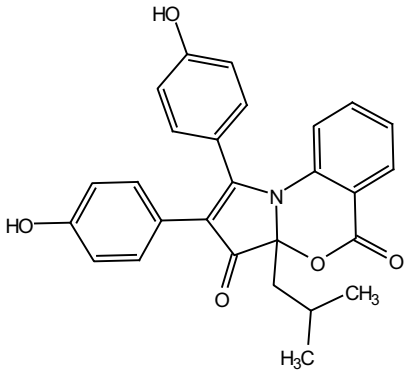
18.3. ^{13}C NMR of discoipyrrole A (**1**) in CD_3OD (100 HMz)



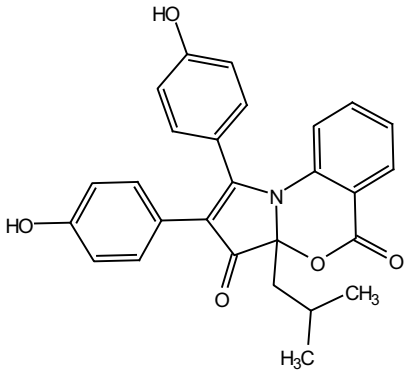
18.4. ^{1}H-^{1}H COSY spectrum of discoipyrrole A (**1**) in CD_3OD (600 HMz)



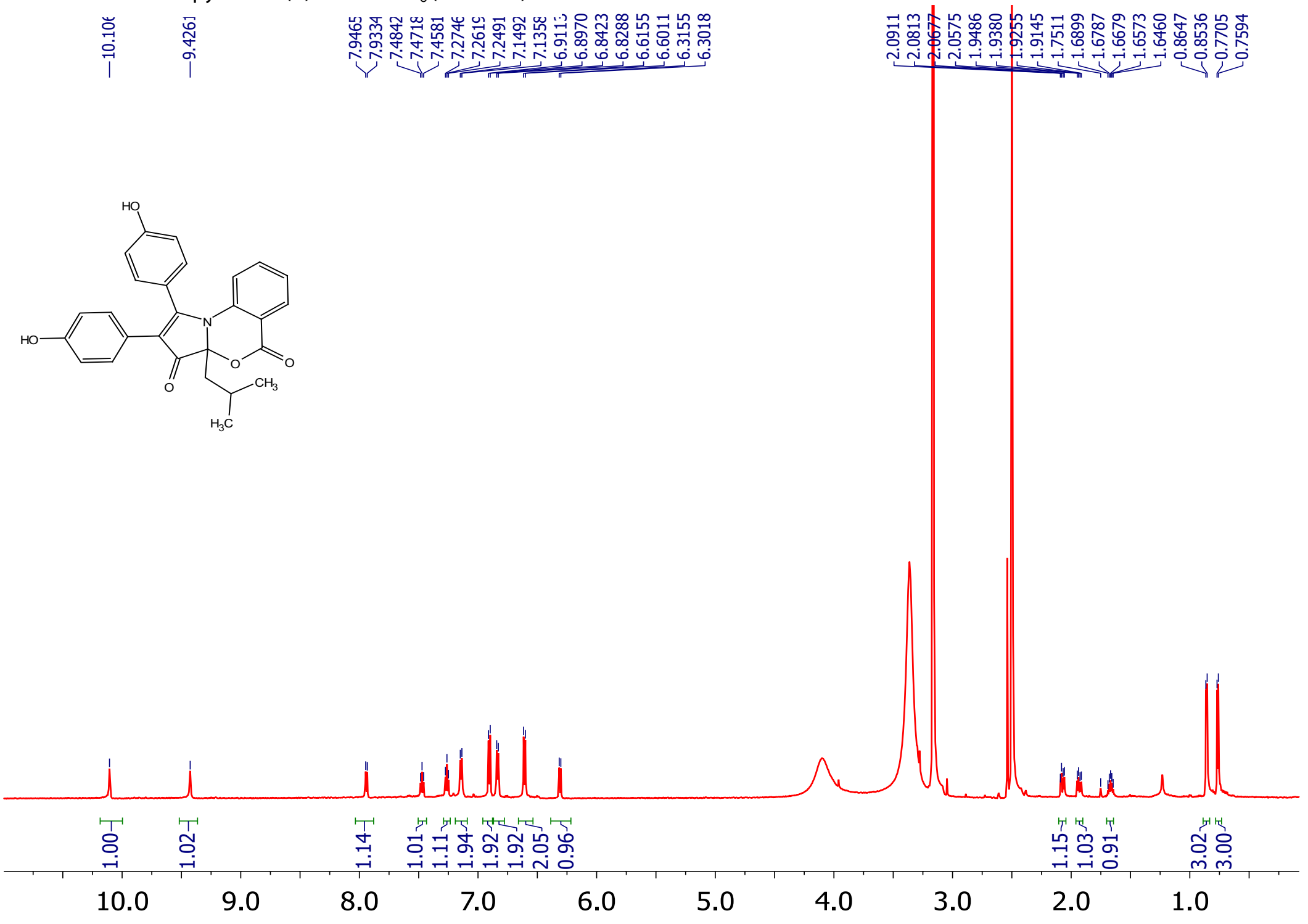
18.5. HSQC spectrum of discoipyrrole A (**1**) in CD_3OD (600 HMz)



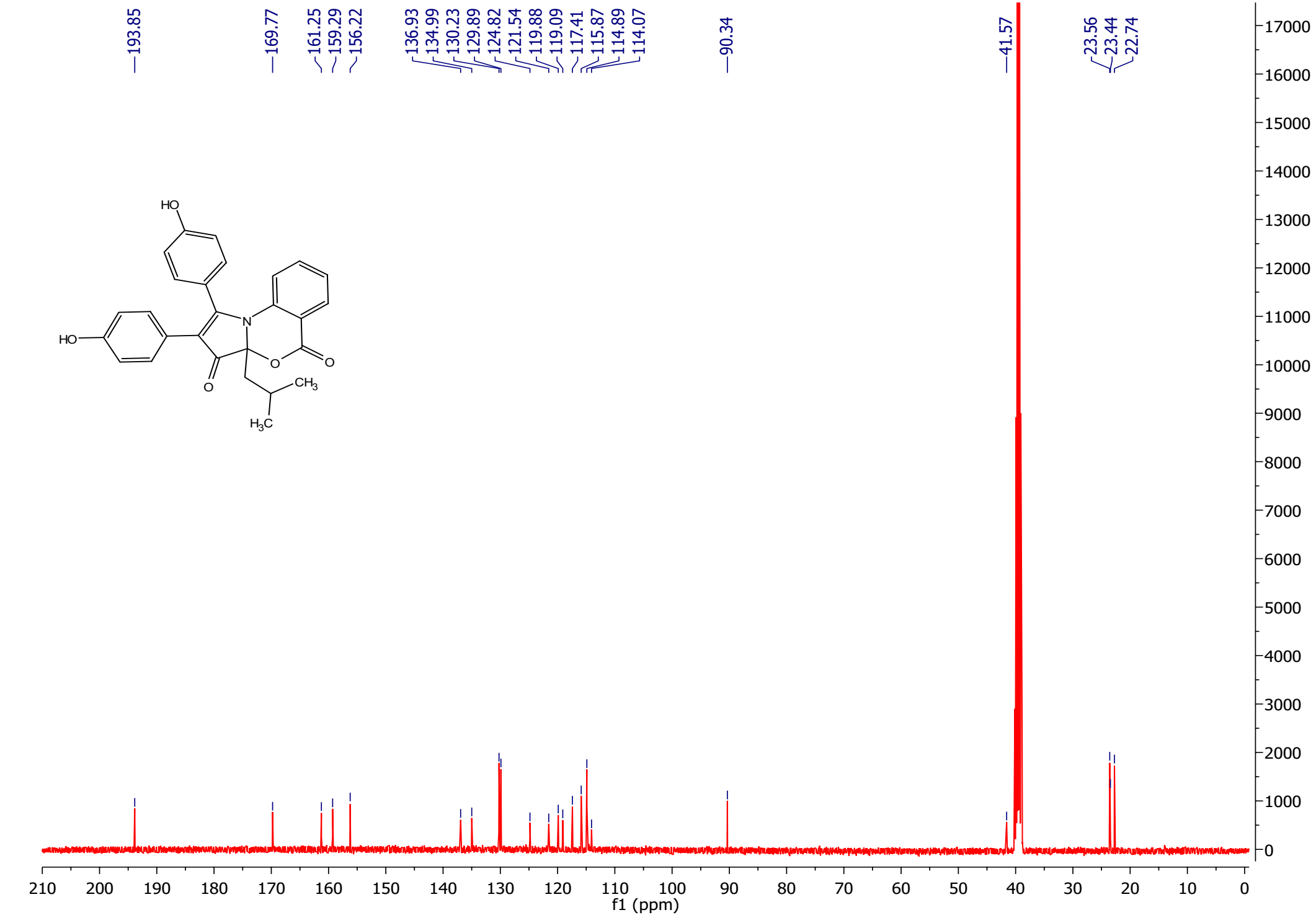
18.6. HMBC spectrum of discoipyrrole A (**1**) in CD_3OD (600 HMz)



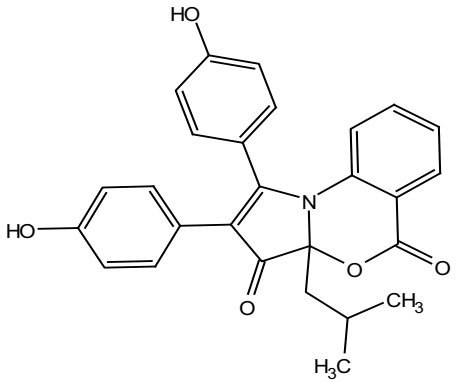
18.7. ^{1}H NMR of discoipyrrole A (**1**) in DMSO-d_6 (600 HMz)



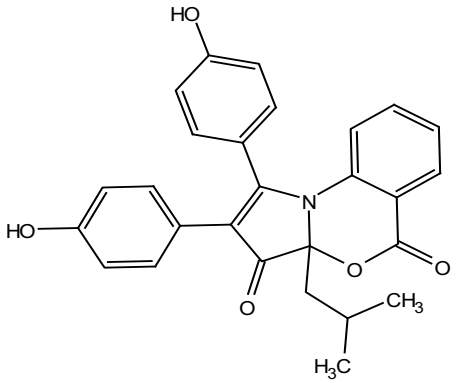
18.8. ^{13}C NMR of discoipyrrole A (**1**) in DSMO-d_6 (100 HMz)



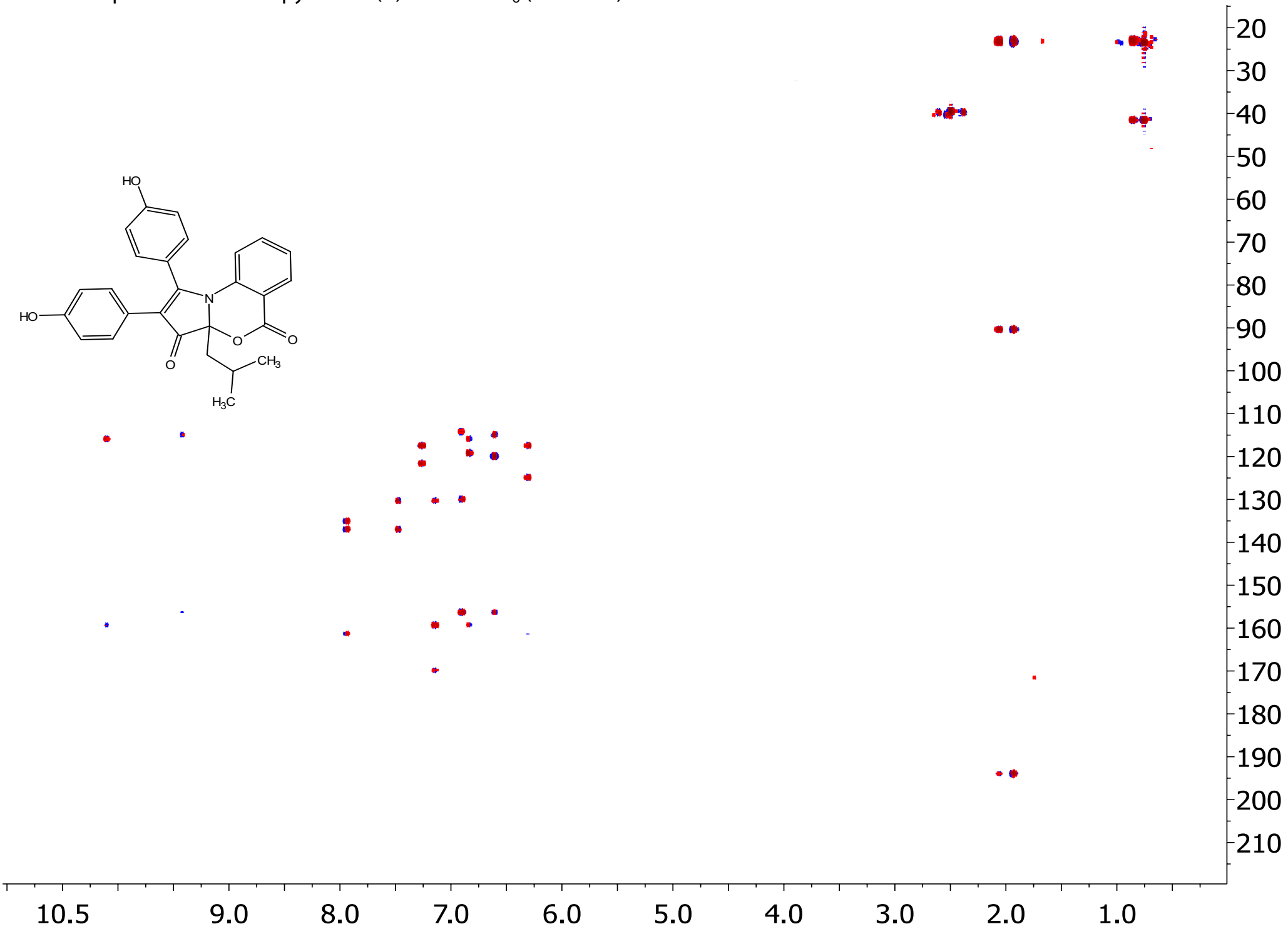
18.9. ^{1}H-^{1}H COSY spectrum of discoipyrrole A (**1**) in DSMO-d_6 (600 HMz)



18.10. HSQC spectrum of discoipyrrole A (**1**) in DSMO-d_6 (600 HMz)

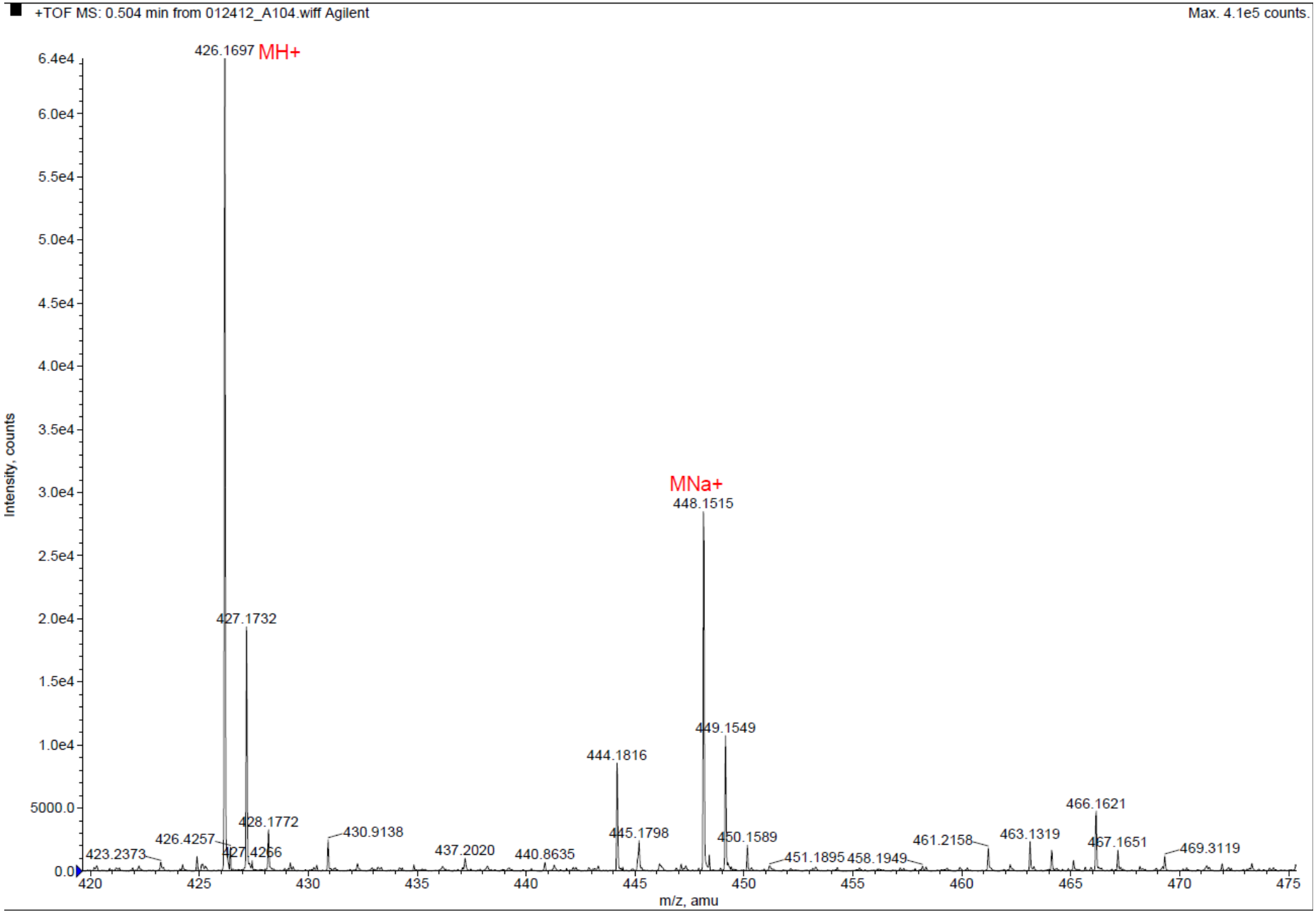


18.11. HMBC spectrum of discoipyrrole A (**1**) in DSMO-d_6 (600 HMz)

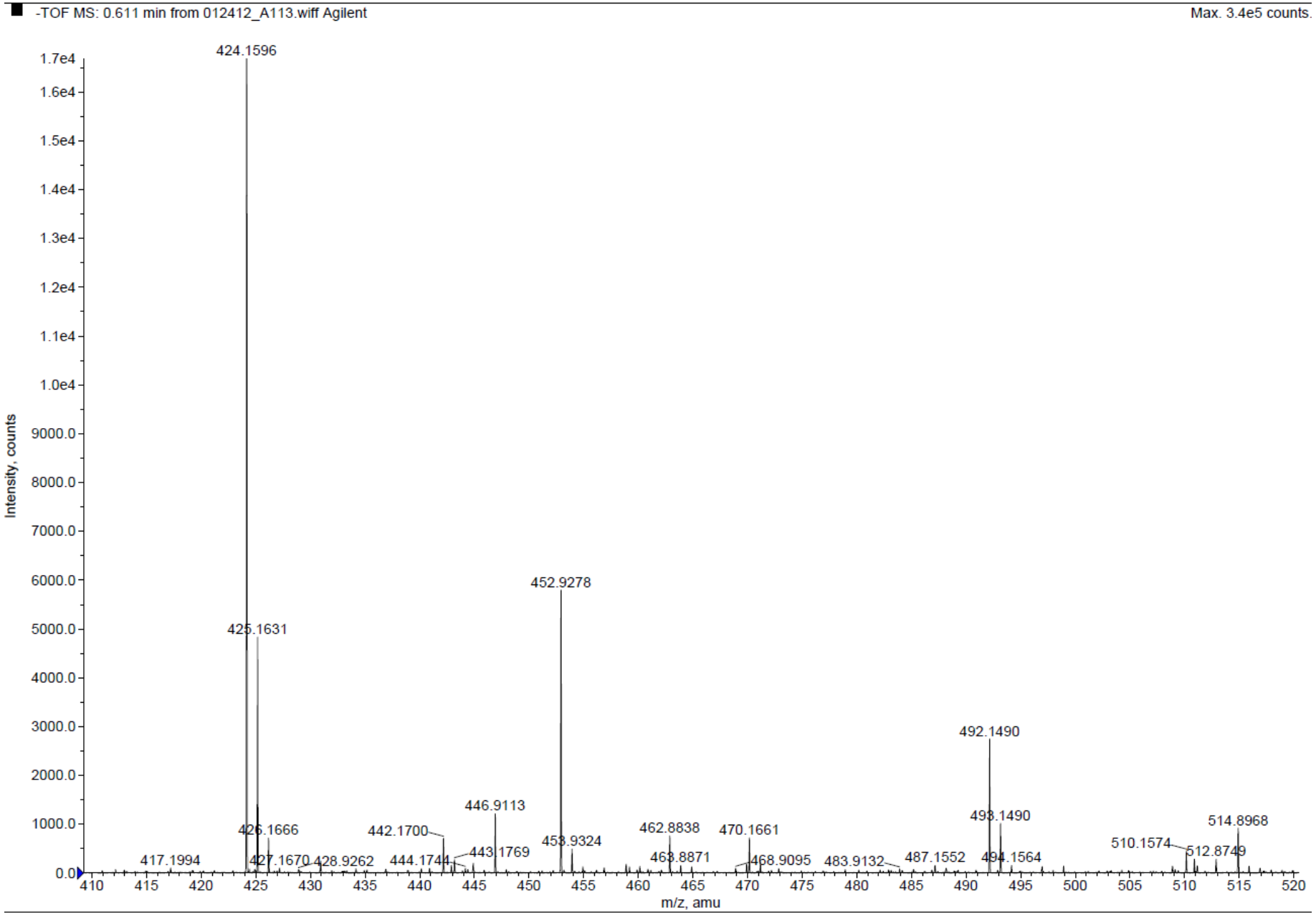


19. HR-ESIMS and NMR data of discoipyrrole B (**2**)

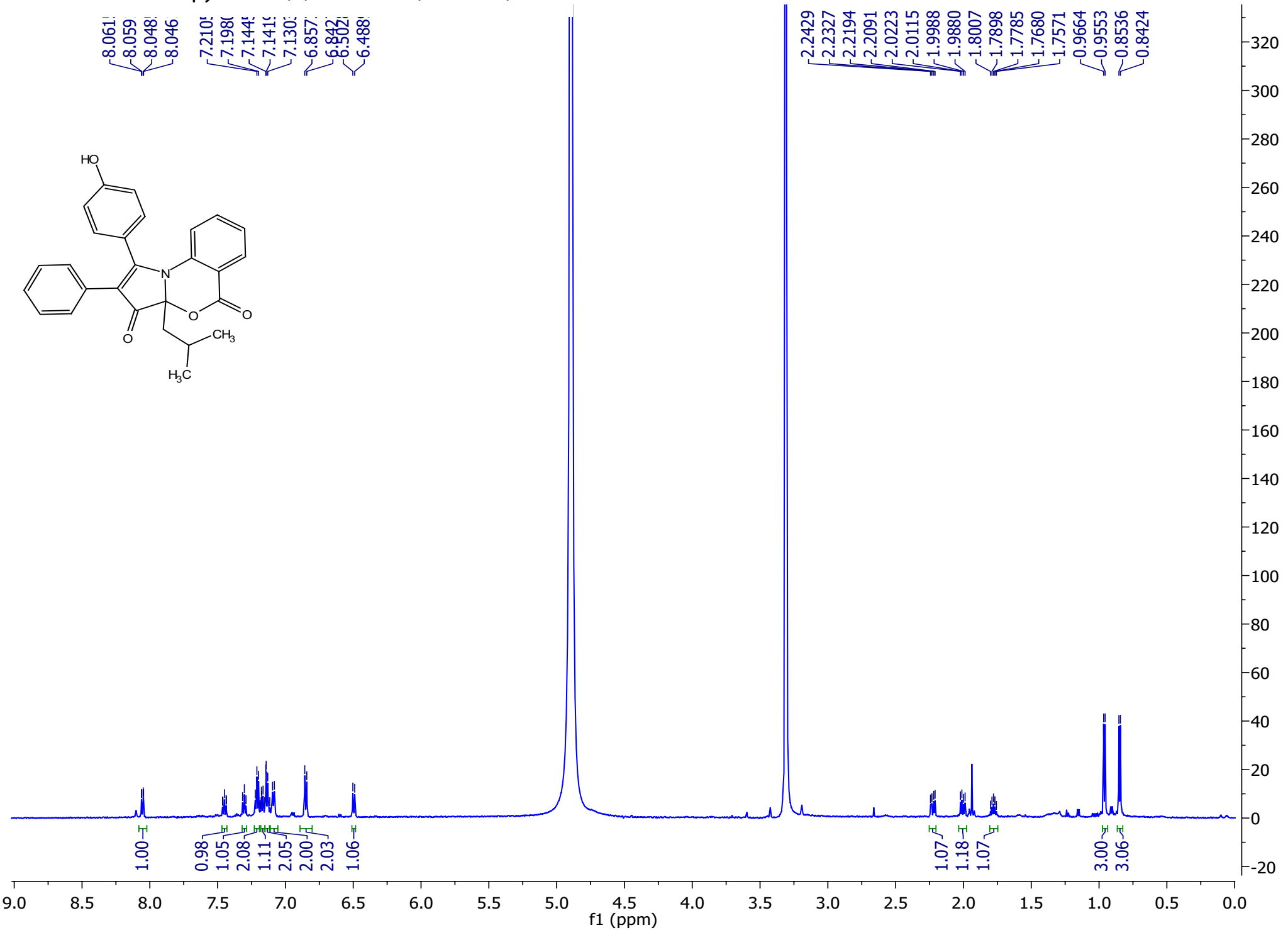
19.1 HR-ESIMS of discoipyrrole B (**2**)-positive



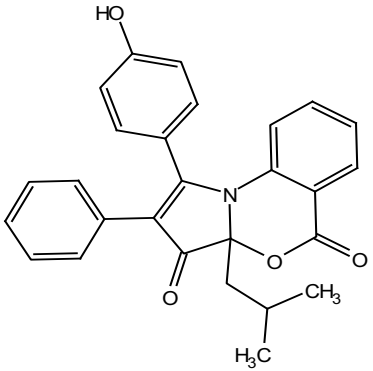
19.1. HR-ESIMS of discoipyrrole B (**2**)-negative



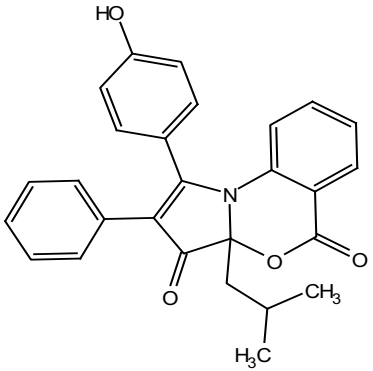
19.2 ^{1}H NMR of discoipyrrole B (**2**) in CD_3OD (600 HMz)



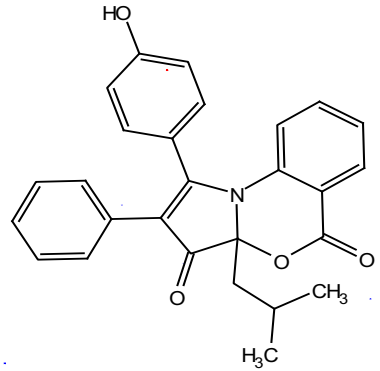
19.3 ^{1}H-^{1}H COSY spectrum of discoipyrrole B (**2**) in CD_3OD (600 HMz)



19.4 HSQC spectrum of discoipyrrole B (**2**) in CD_3OD (600 HMz)

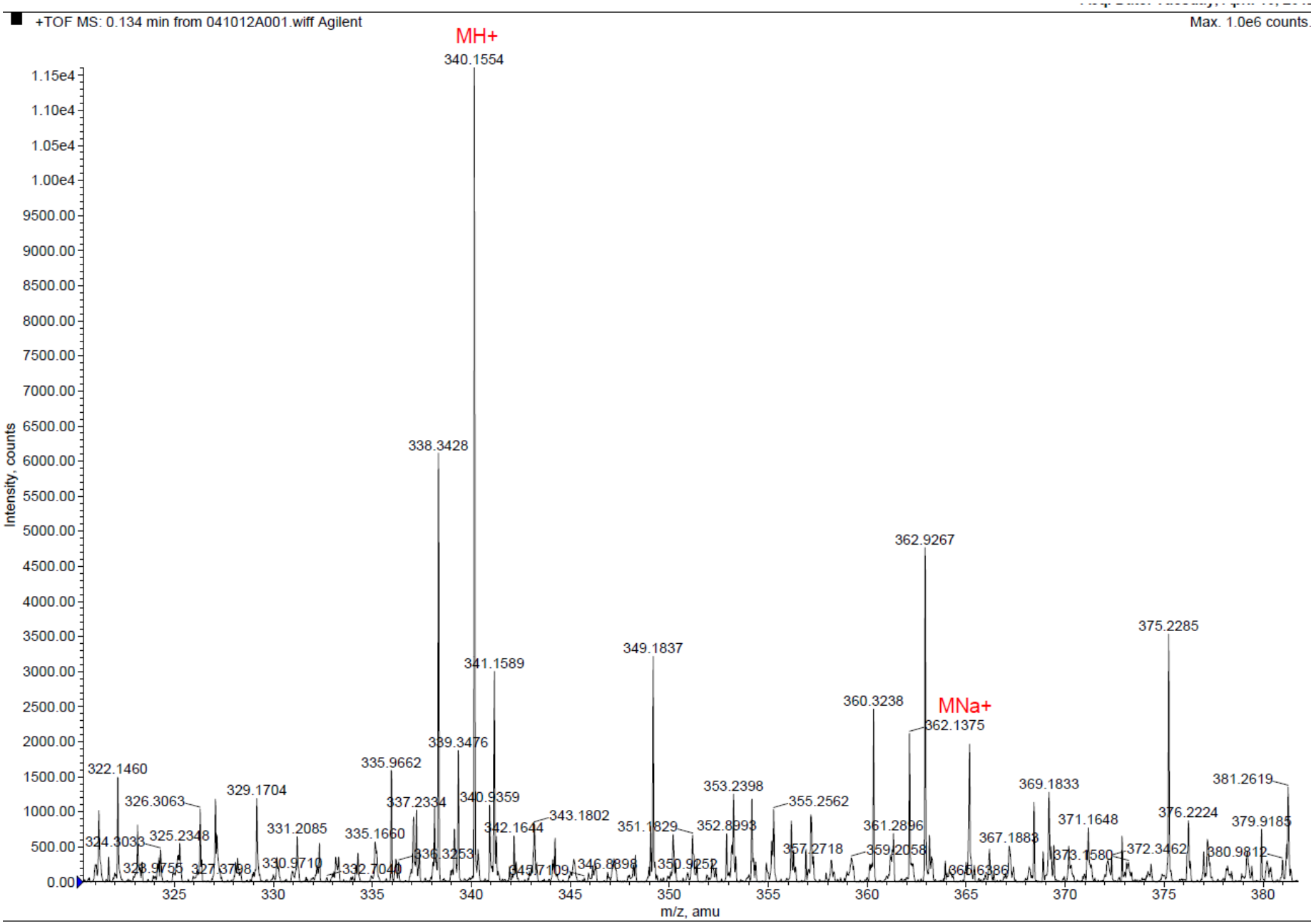


19.5 HMBC spectrum of discoipyrrole B (**2**) in CD_3OD (600 HMz)

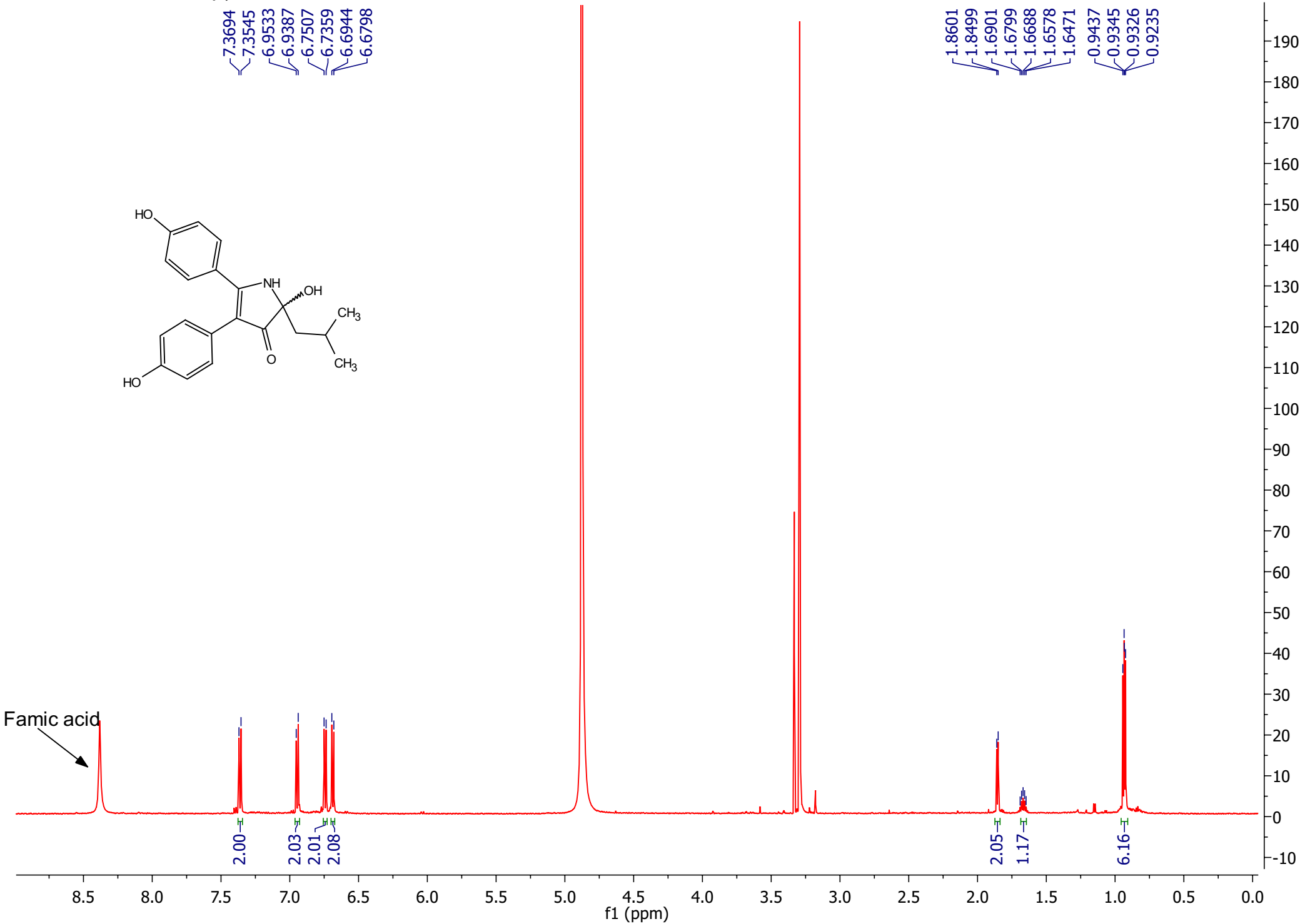


20. HR-ESIMS and NMR data of discoipyrrole C (**3**)

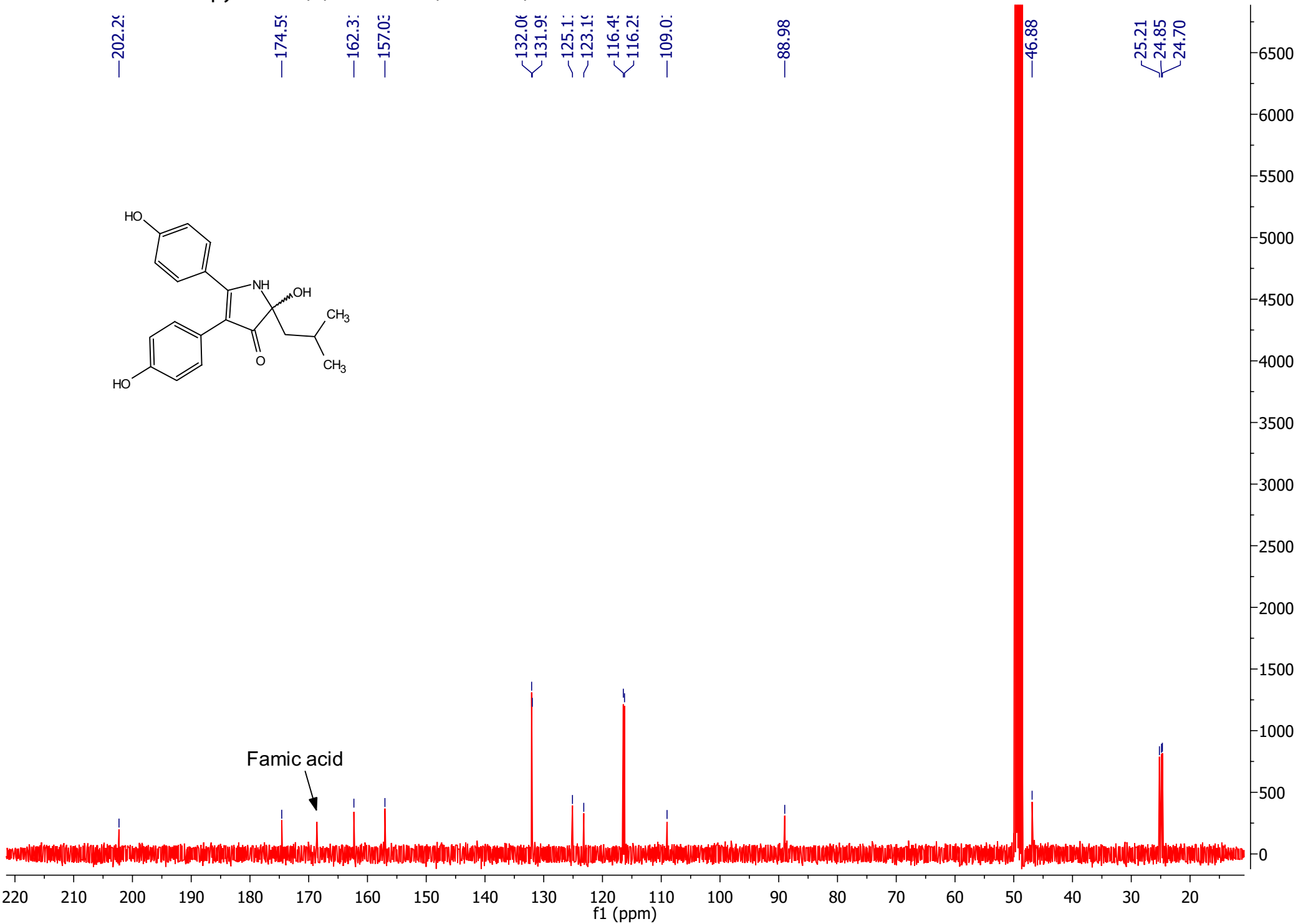
20.1. HR-ESIMS of discoipyrrole C (**3**)-positive



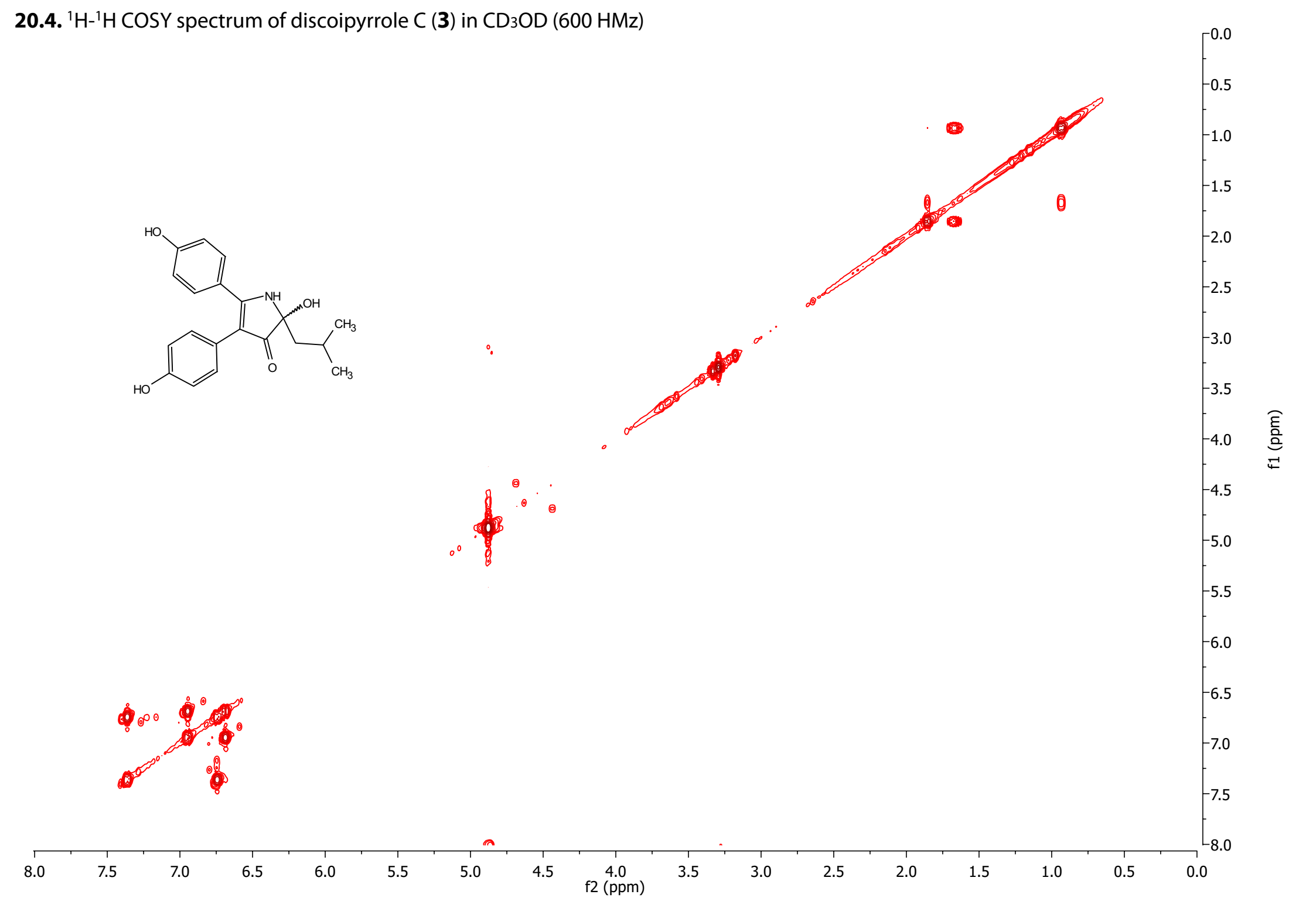
20.2. ^{1}H NMR of discoipyrrole C (**3**) in CD_3OD (600 HMz)



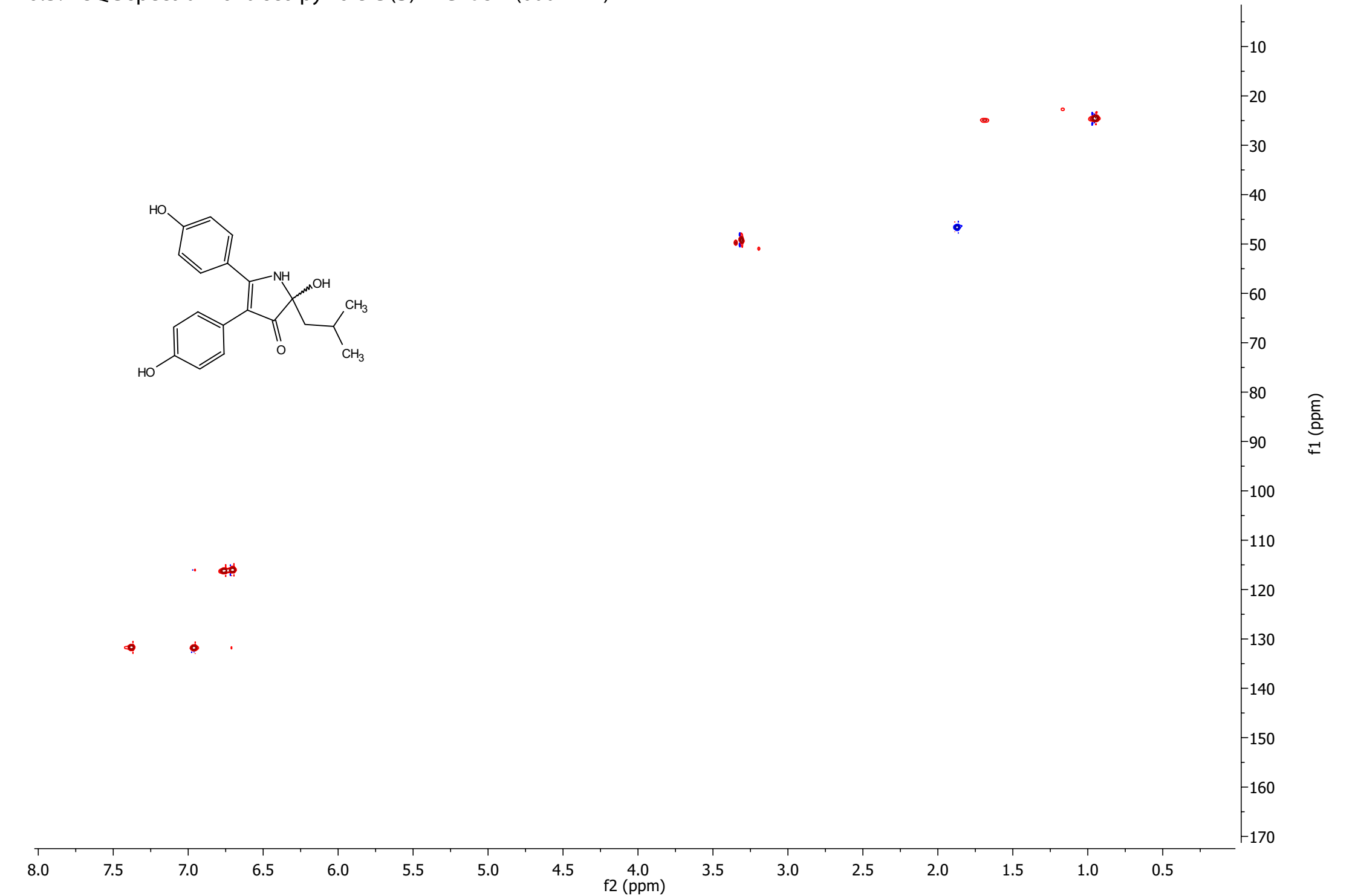
20.3. ^{13}C NMR of discoipyrrole C (**3**) in CD_3OD (100 HMz)



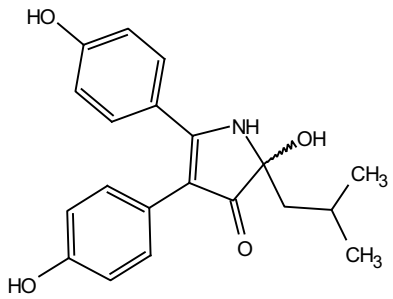
20.4. ^{1}H-^{1}H COSY spectrum of discoipyrrole C (**3**) in CD_3OD (600 HMz)



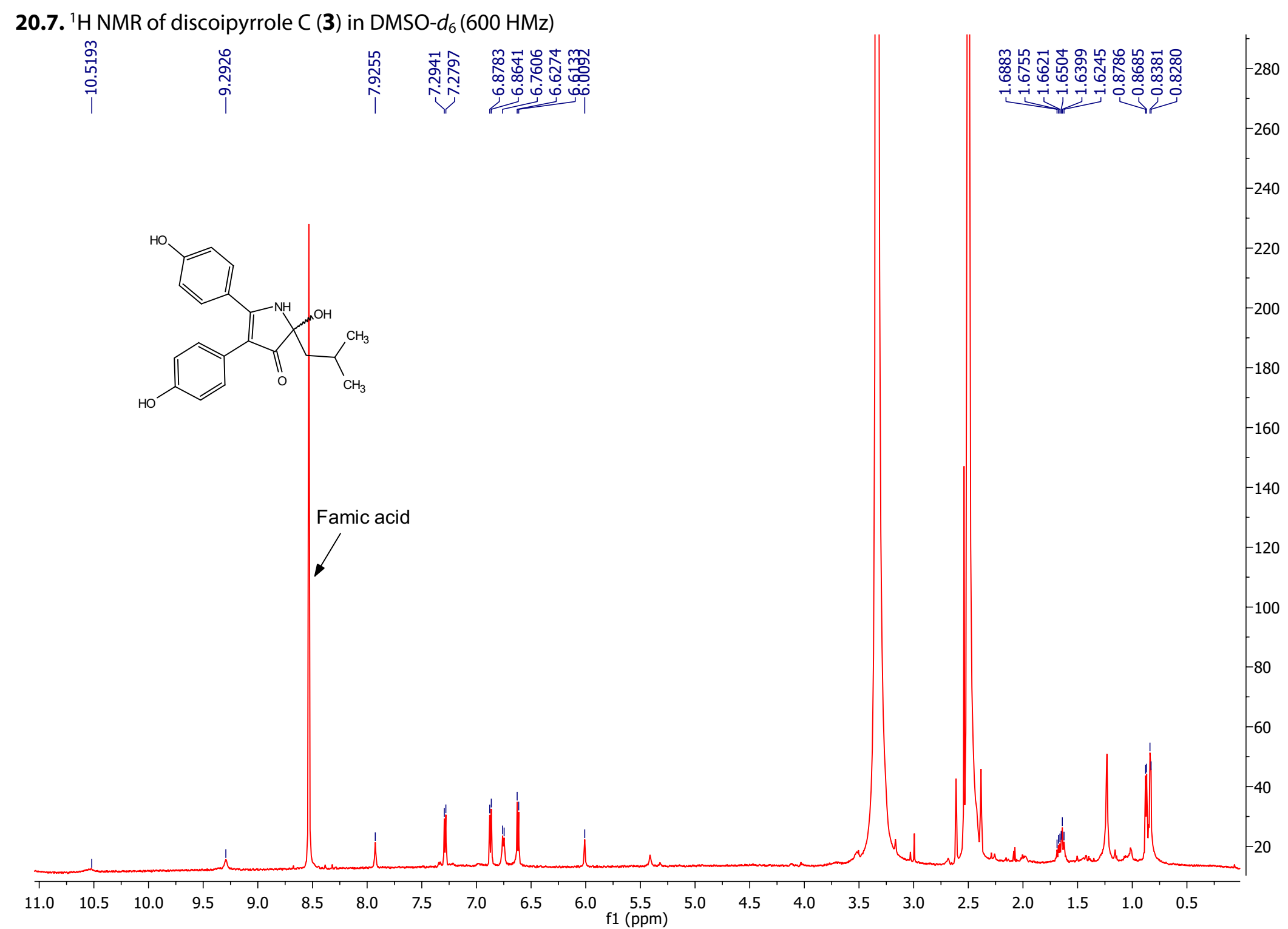
20.5. HSQC spectrum of discoipyrrole C (**3**) in CD_3OD (600 HMz)



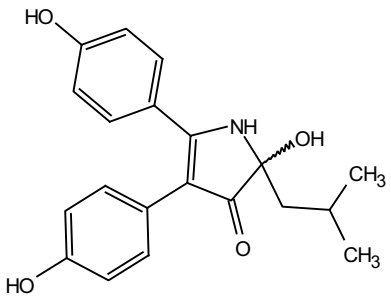
20.6. HMBC spectrum of discoipyrrole C (**3**) in CD_3OD (600 HMz)



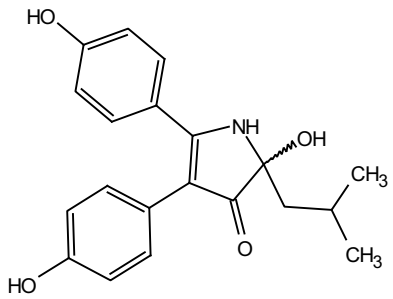
20.7. ^{1}H NMR of discoipyrrole C (**3**) in DMSO-d_6 (600 HMz)



20.8. HSQC spectrum of discoipyrrole C (**3**) in DSMO-d_6 (600 HMz)

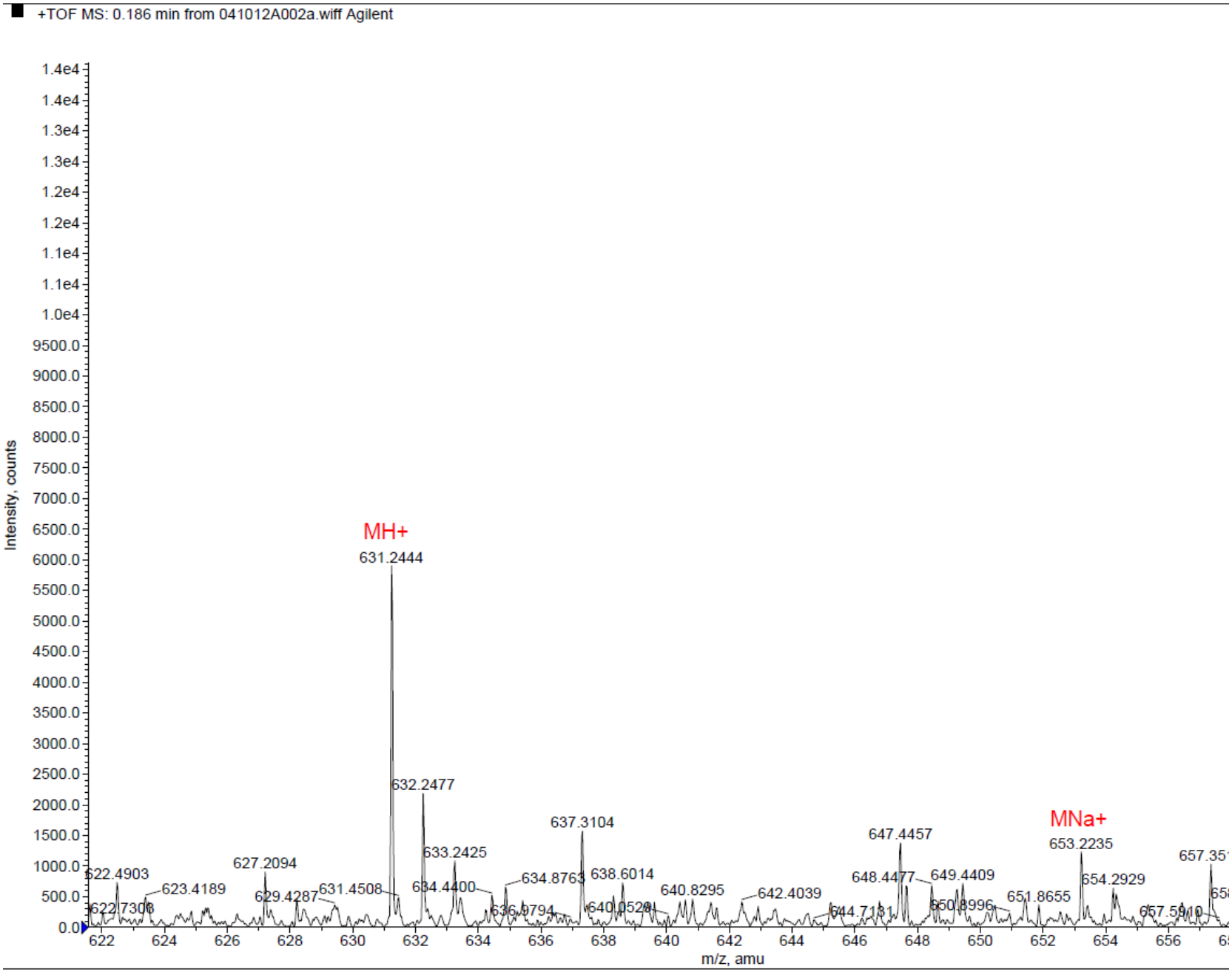


20.9. HMBC spectrum of discoipyrrole C (**3**) in DSMO-d_6 (800 HMz)

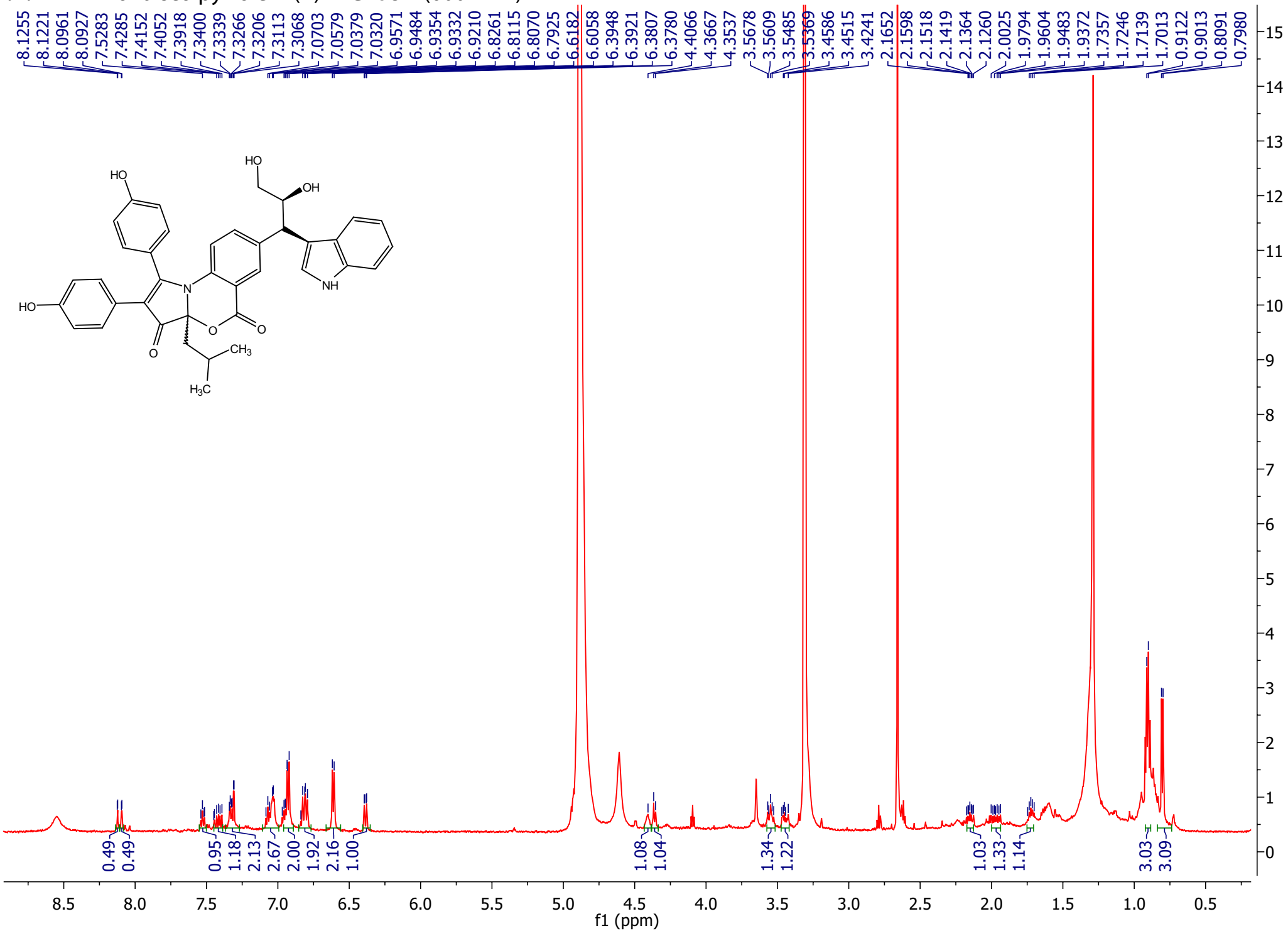


21. HR-ESIMS and NMR spectra of discoipyrrole D (**4**)

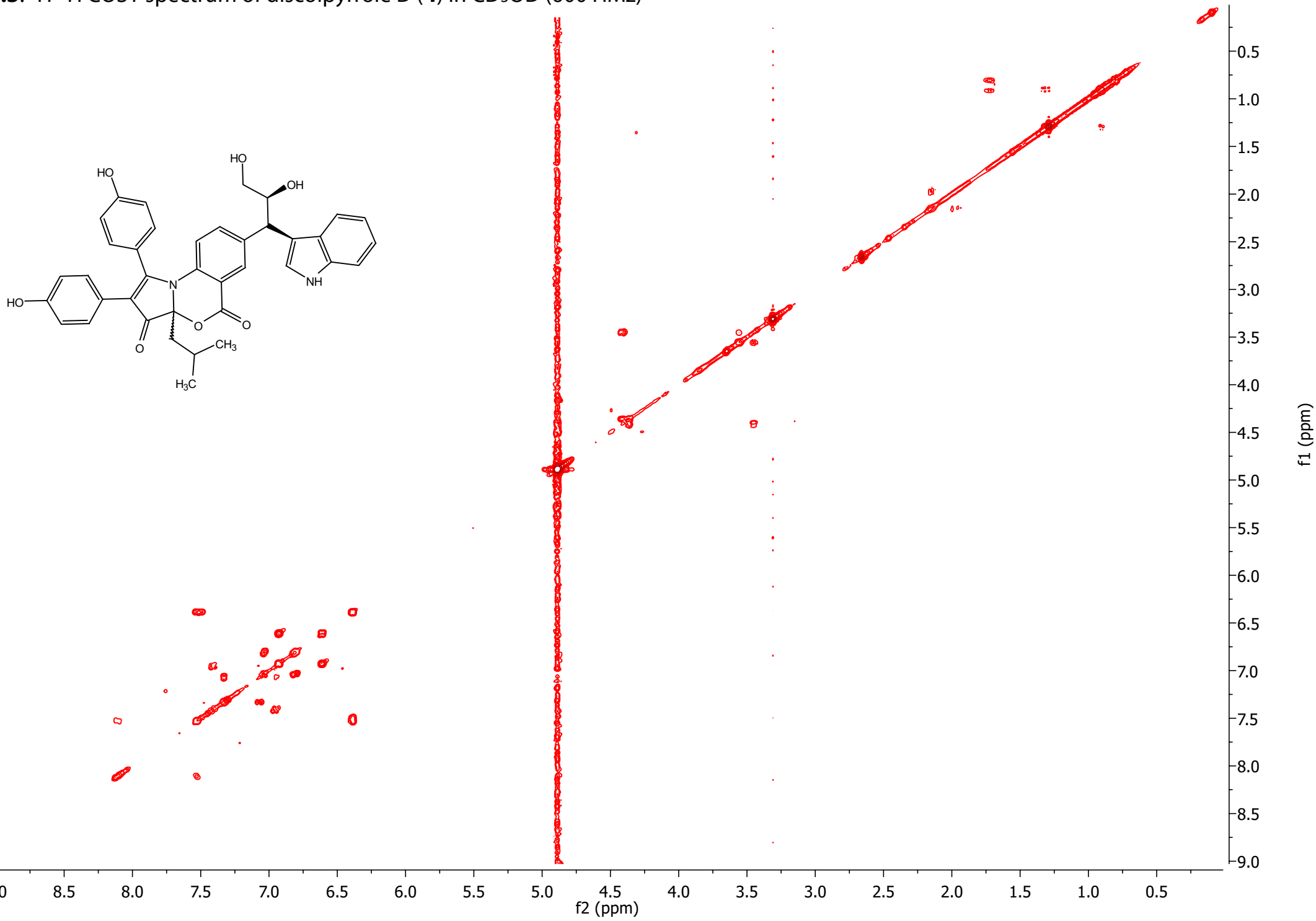
21.1. HR-ESIMS of discoipyrrole D (**4**)-positive



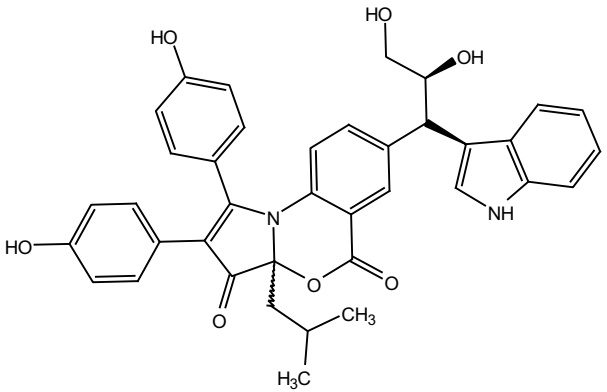
21.2. ^{1}H NMR of discoipyrrole D (**4**) in CD_3OD (600 HMz)



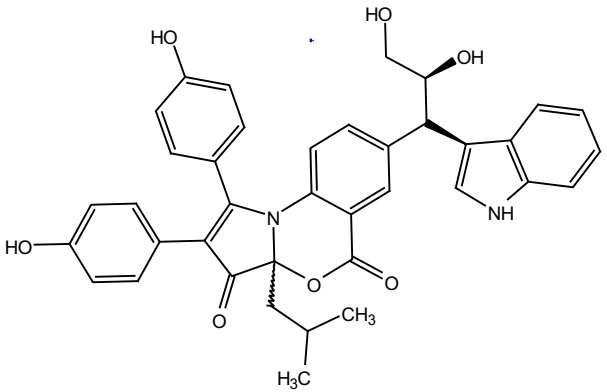
21.3. ^{1}H-^{1}H COSY spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)



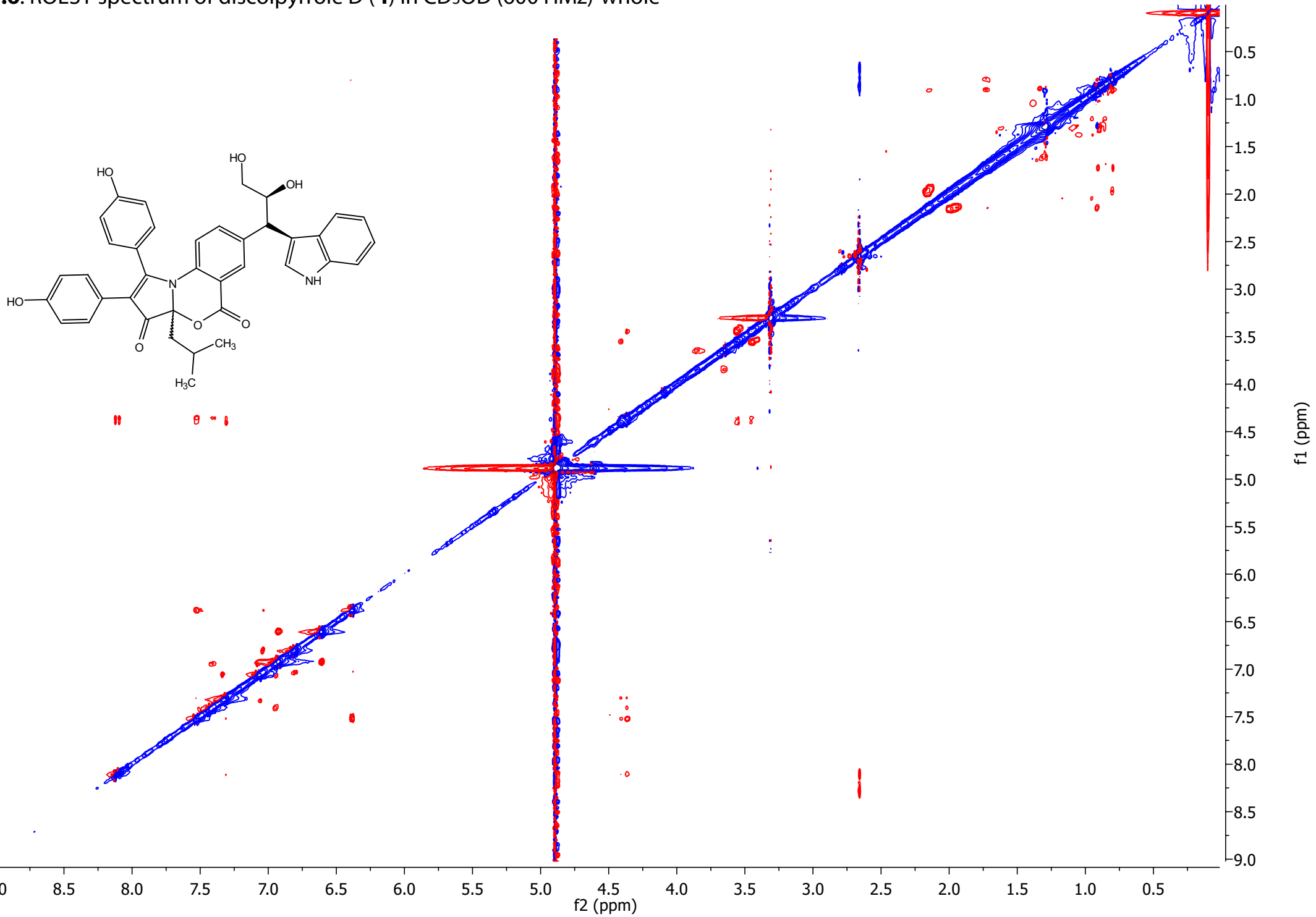
21.4. HSQC spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)



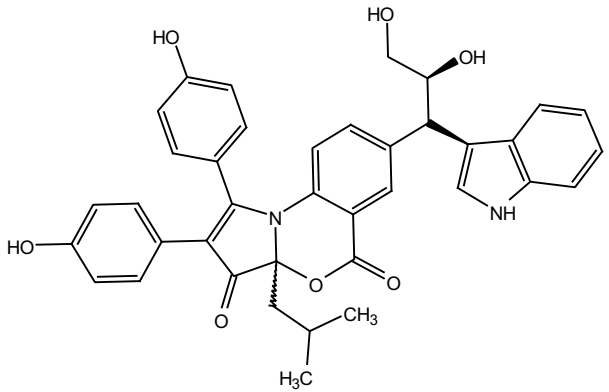
21.5. HMBC spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)



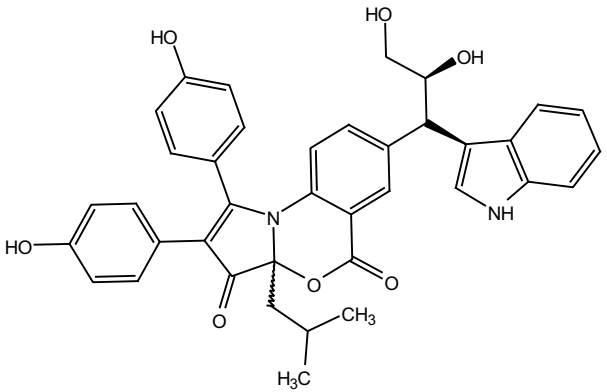
21.6. ROESY spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)-whole



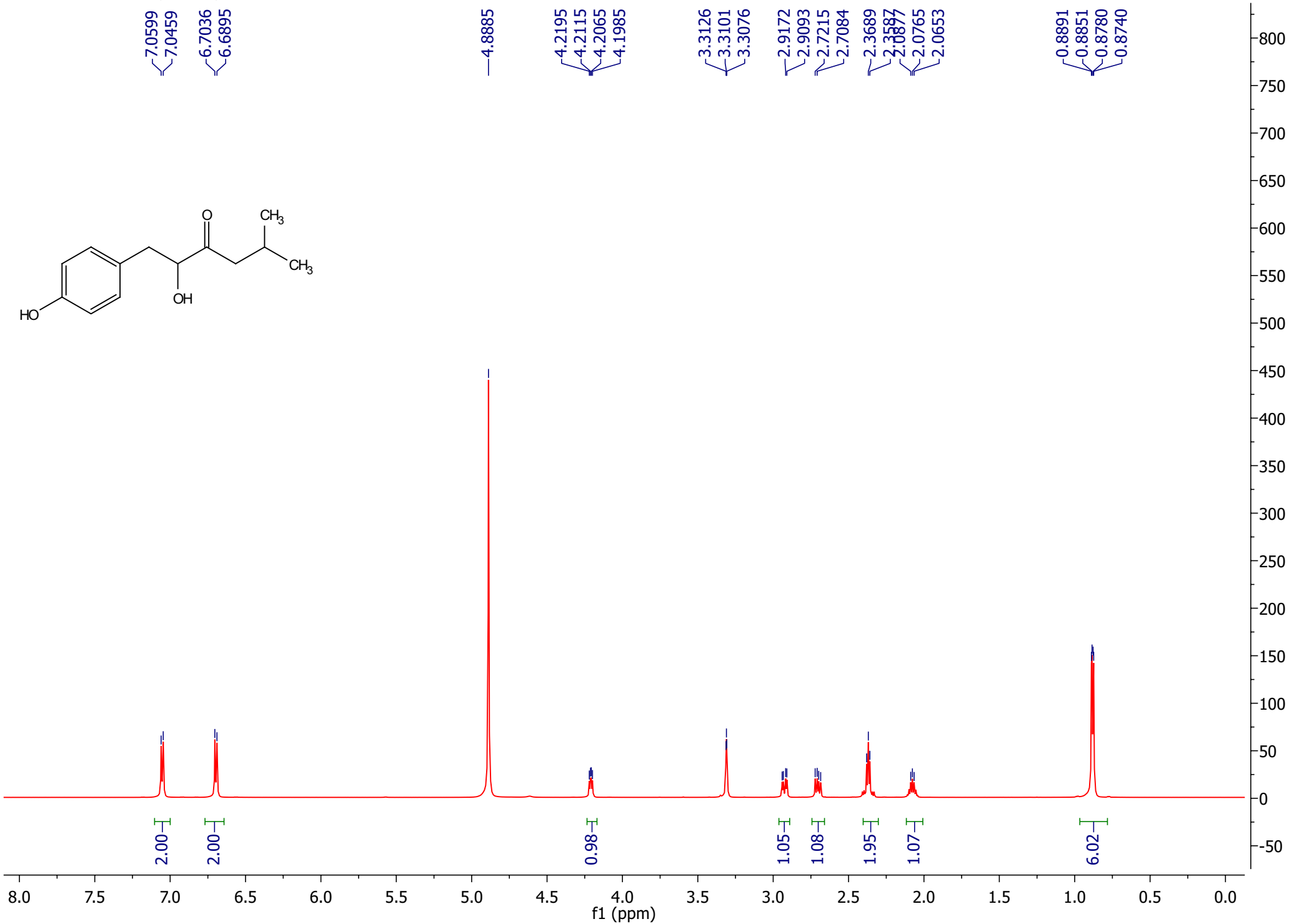
21.7. ROESY spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)-zoom in -1



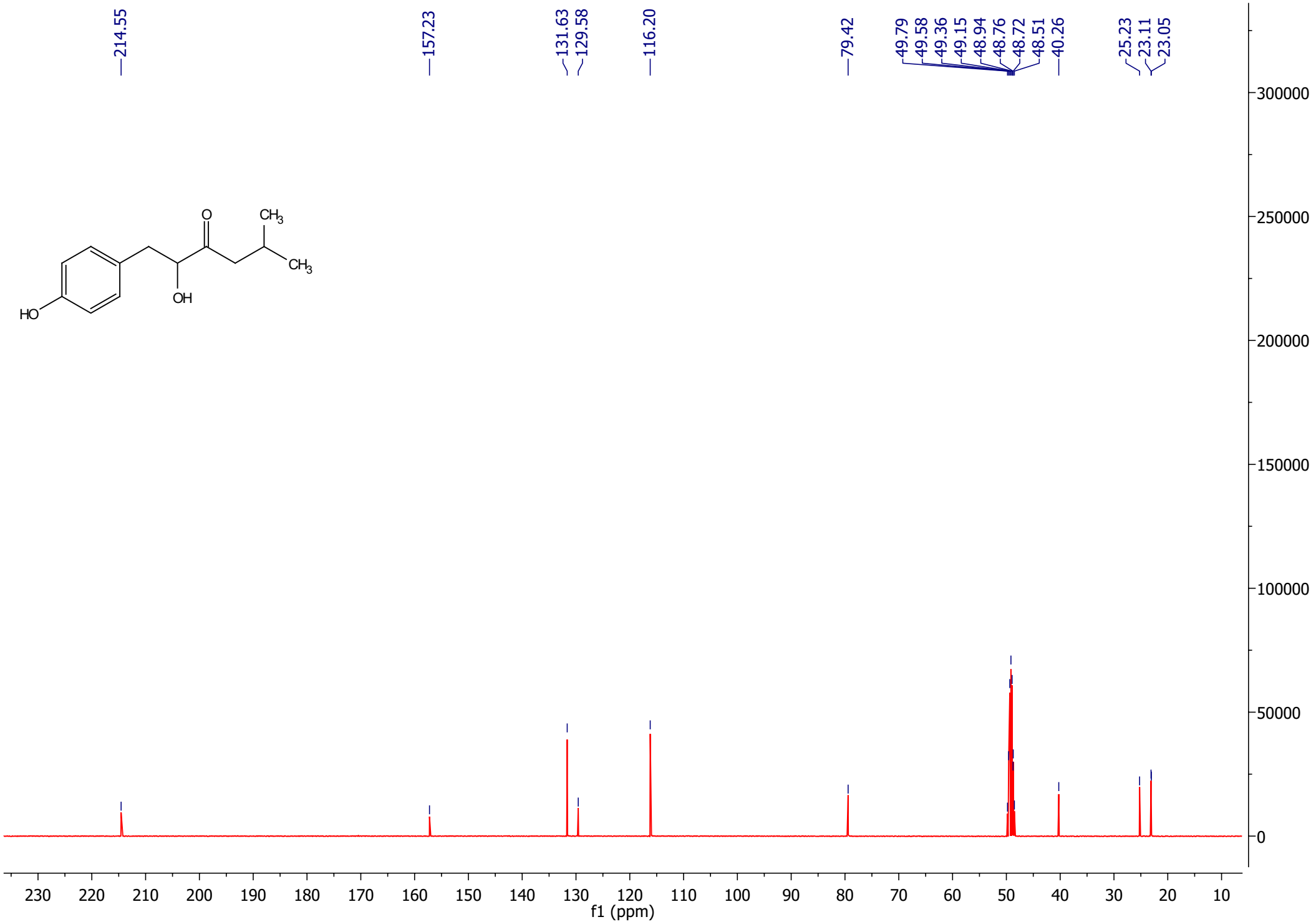
21.8. ROESY spectrum of discoipyrrole D (**4**) in CD_3OD (600 HMz)-zoom in -2



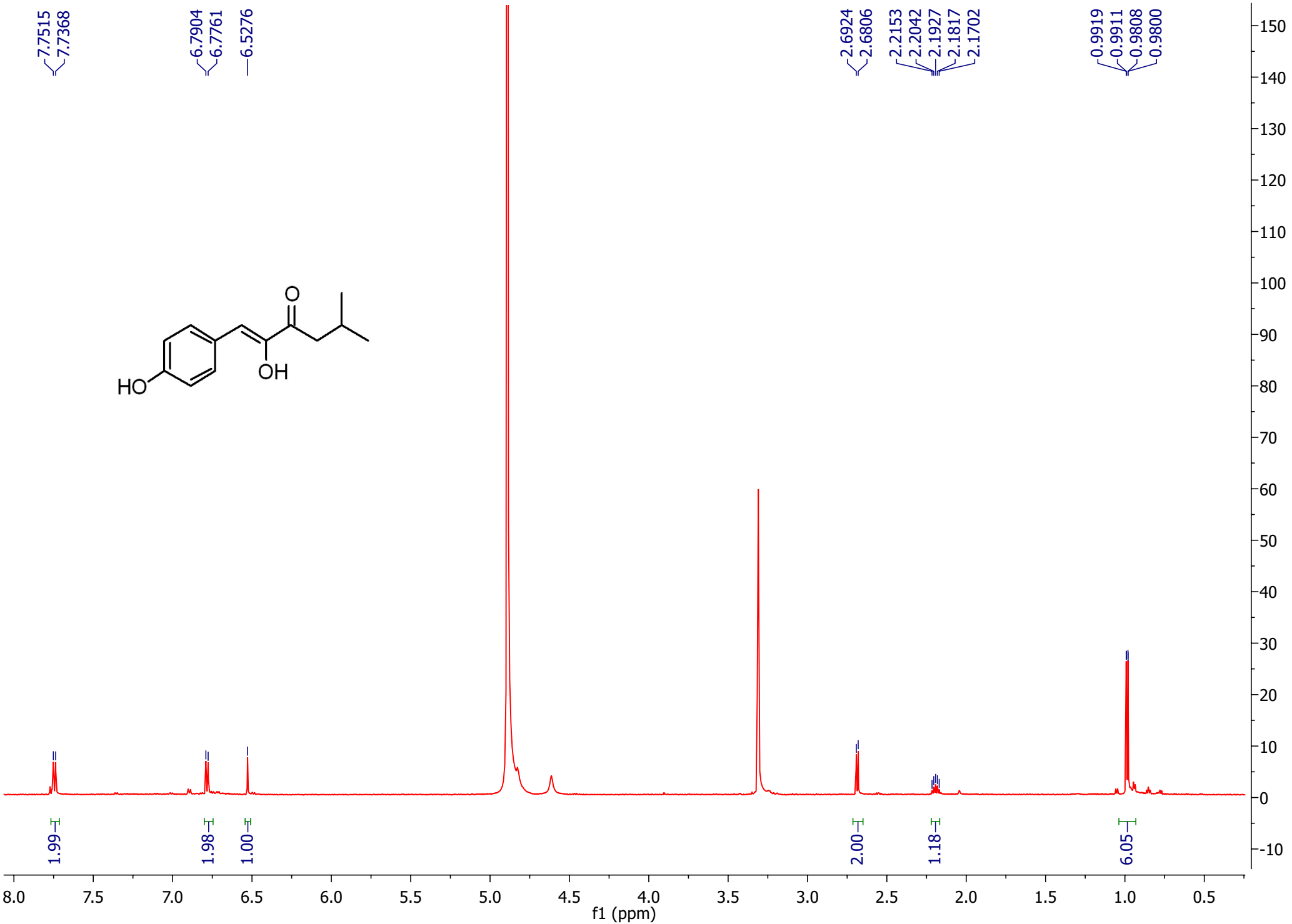
22.1 ^{1}H NMR of 4-hydroxy-sattabacin (**5**) in CD_3OD (600 HMz)



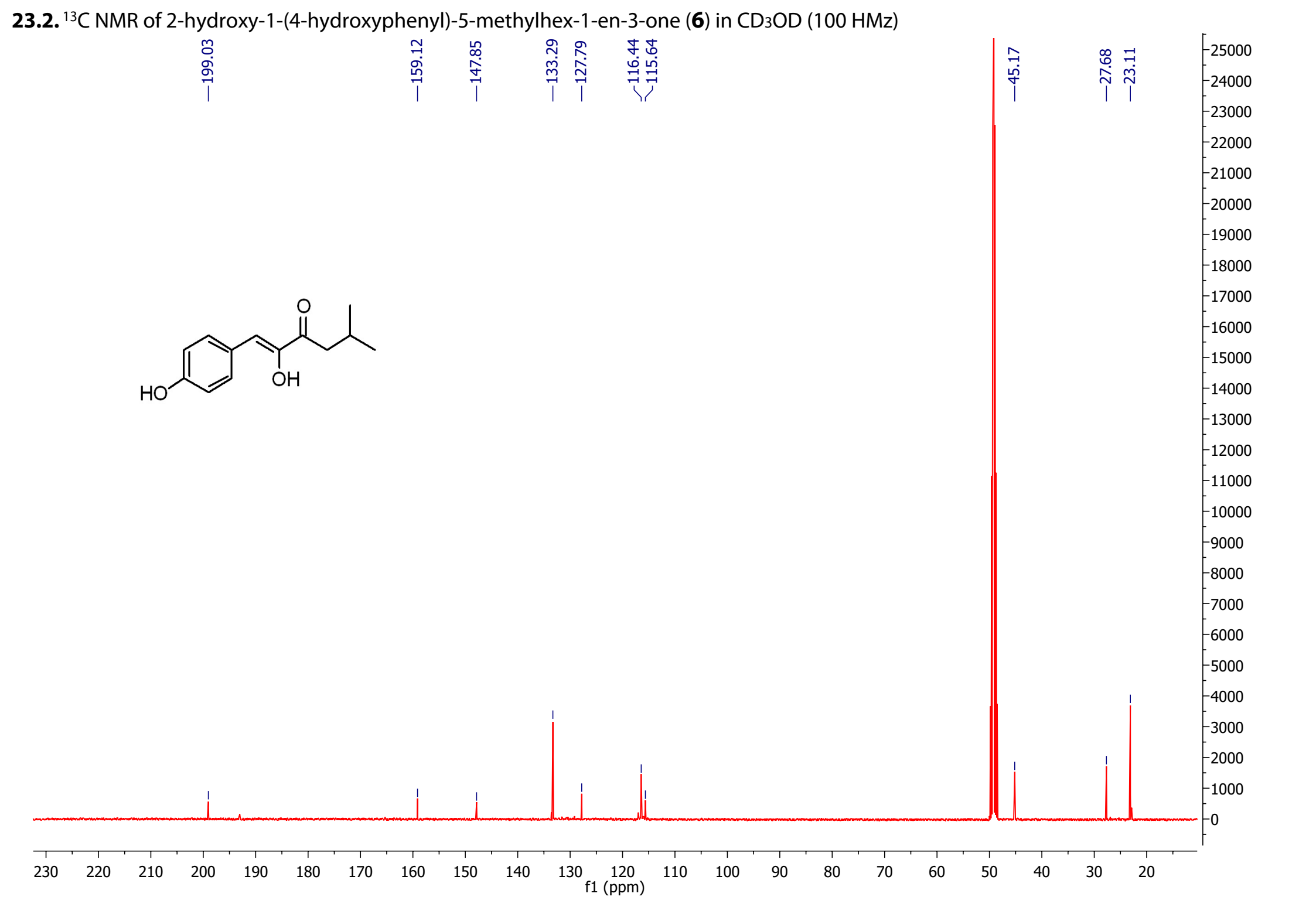
22.2 ^{13}C NMR of 4-hydroxy-sattabacin (**5**) in CD_3OD (100 HMz)



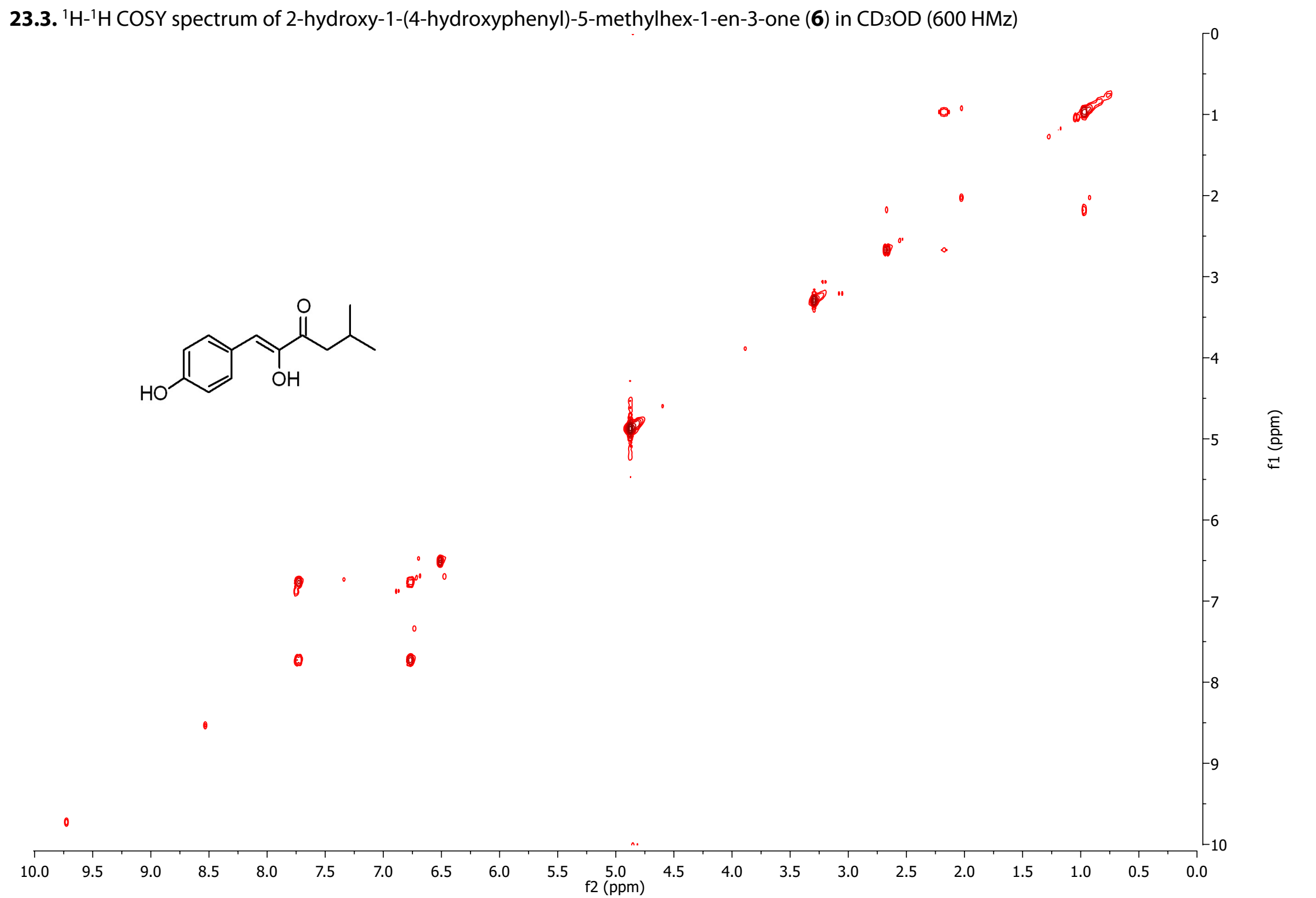
23.1 ^{1}H NMR of 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) in CD_3OD (600 HMz)



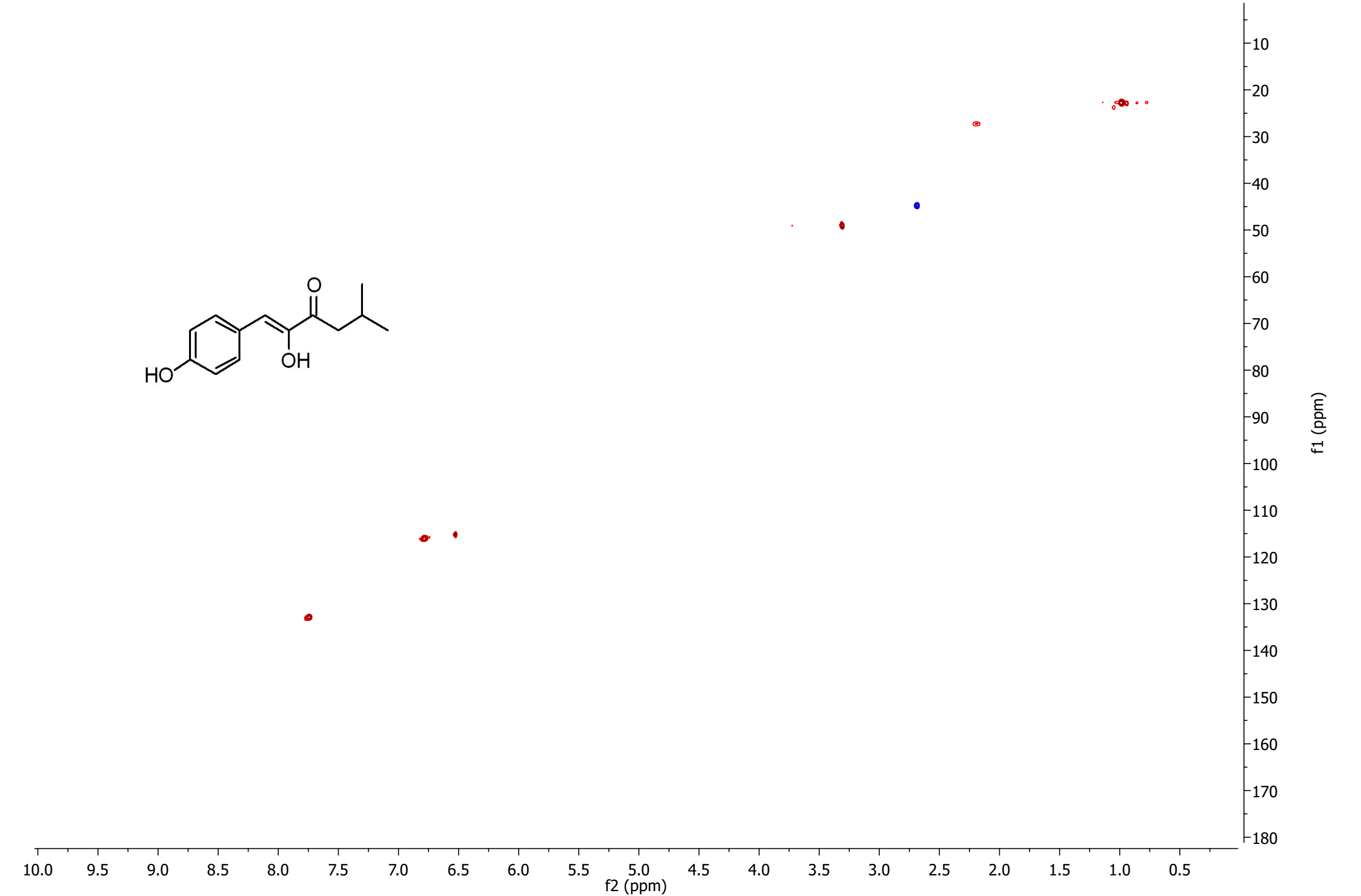
23.2. ^{13}C NMR of 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) in CD_3OD (100 HMz)



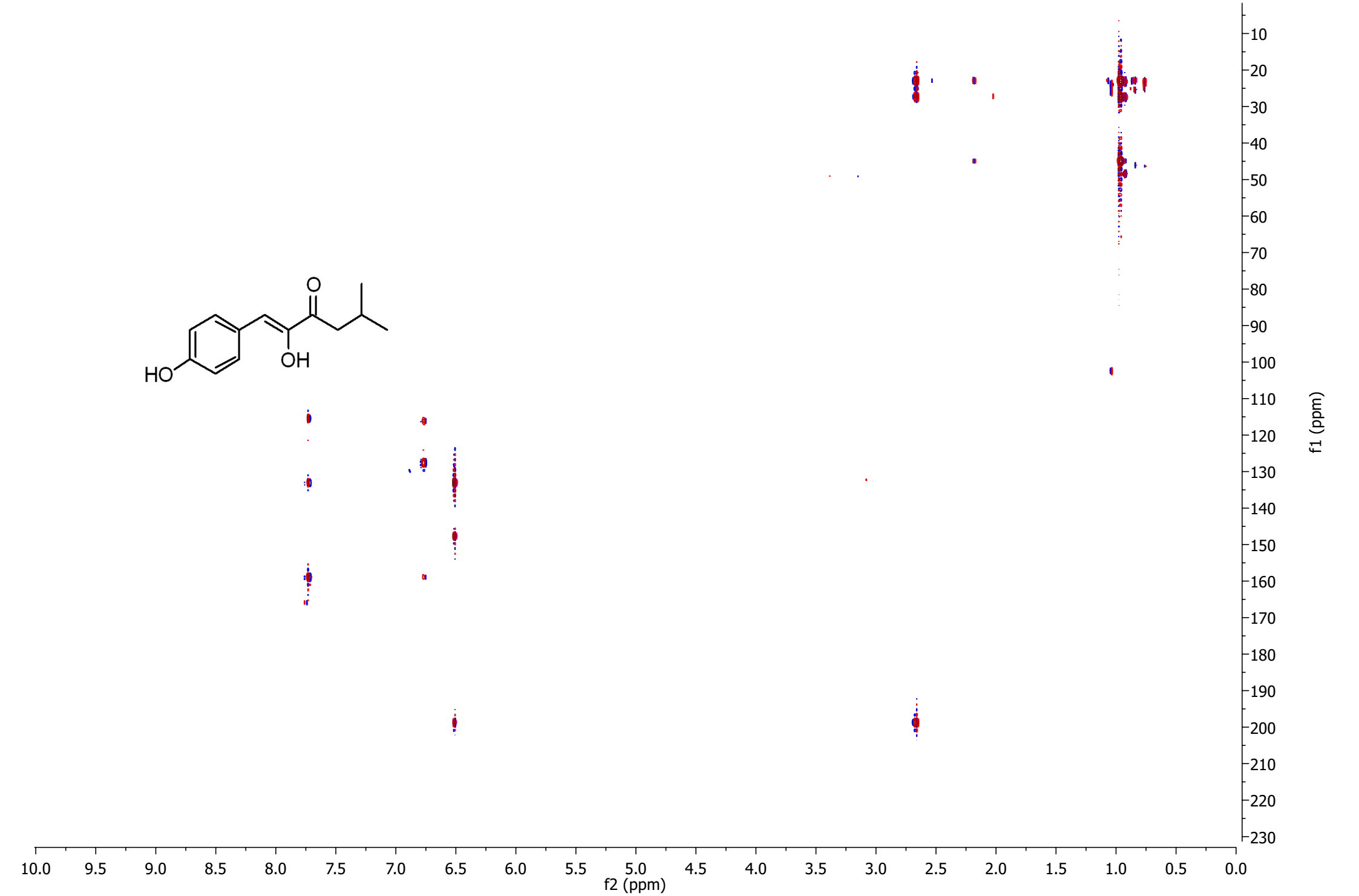
23.3. ^{1}H-^{1}H COSY spectrum of 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) in CD_3OD (600 HMz)



23.4. HSQC spectrum of 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) in CD_3OD (600 HMz)

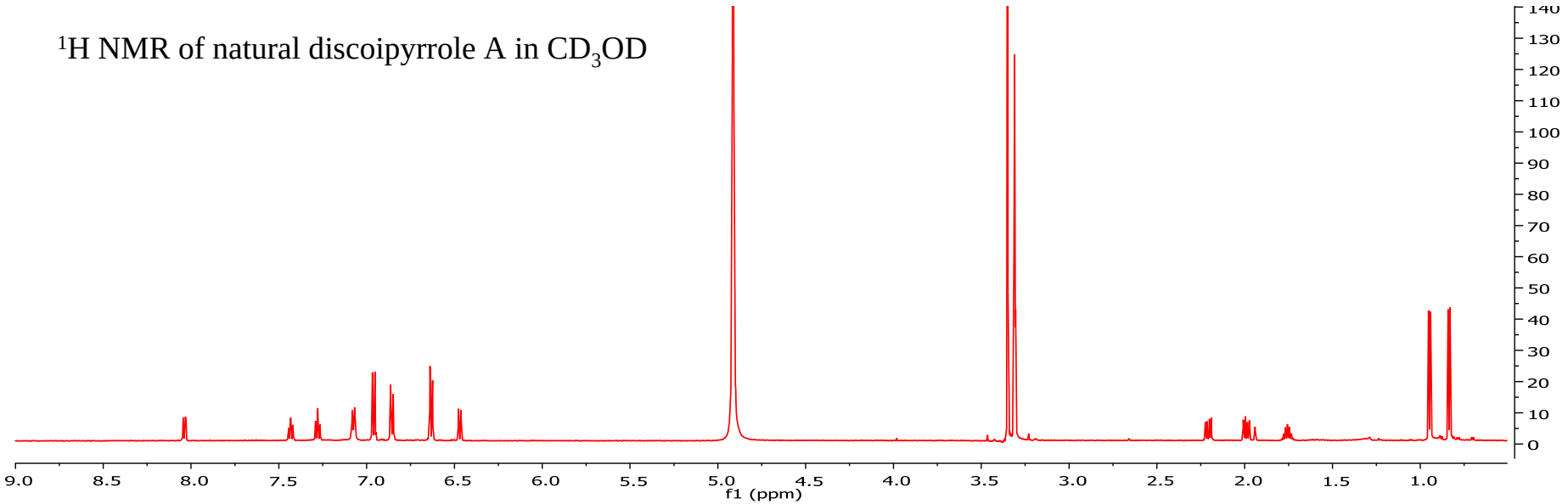


23.5. HMBC spectrum of 2-hydroxy-1-(4-hydroxyphenyl)-5-methylhex-1-en-3-one (**6**) in CD_3OD (600 HMz)

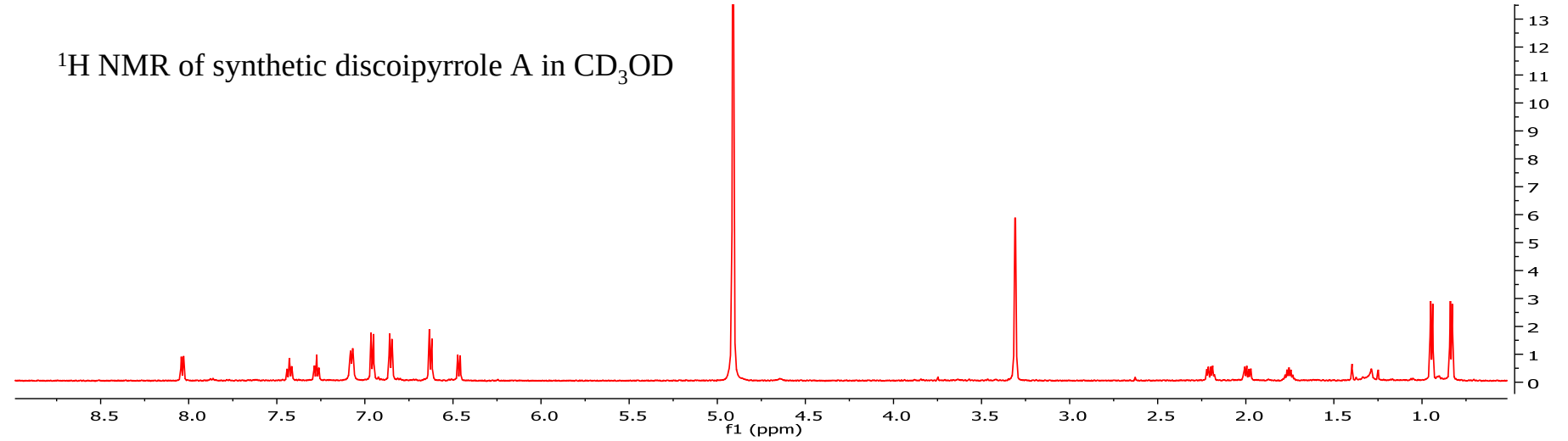


24. Comparison of ^{1}H NMR spectra of natural discoipyrrole A and synthetic discoipyrrole A in CD_3OD

^{1}H NMR of natural discoipyrrole A in CD_3OD

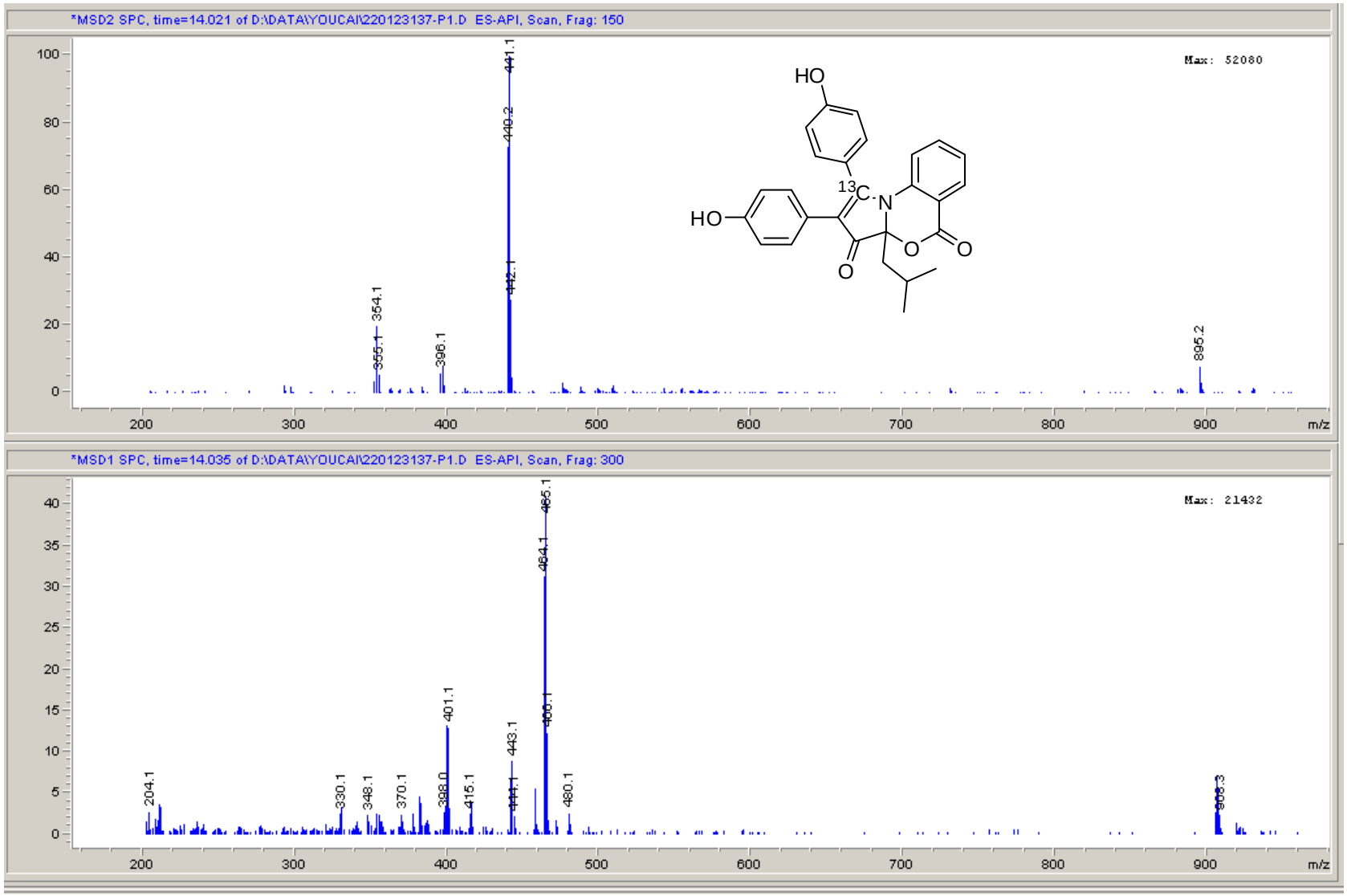


^{1}H NMR of synthetic discoipyrrole A in CD_3OD

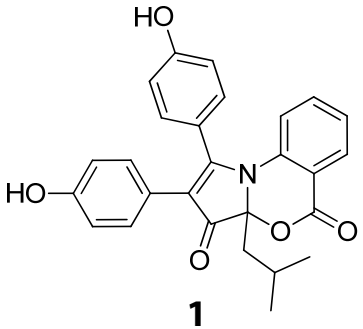


25. MS and NMR spectra of4-[^{13}C]-discoipyrrole A (**10)** from precursor-directed biosynthesis

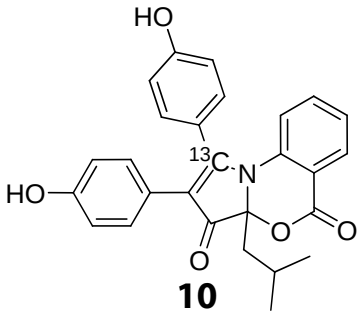
25.1 ESI-MS of 4-[^{13}C]-discoipyrrole A (**10)**



25.2 Comparison of ^{1}H NMR spectra of discoipyrrole A (**1**) and 4-[^{13}C]-discoipyrrole A (**10**) in CD_3OD

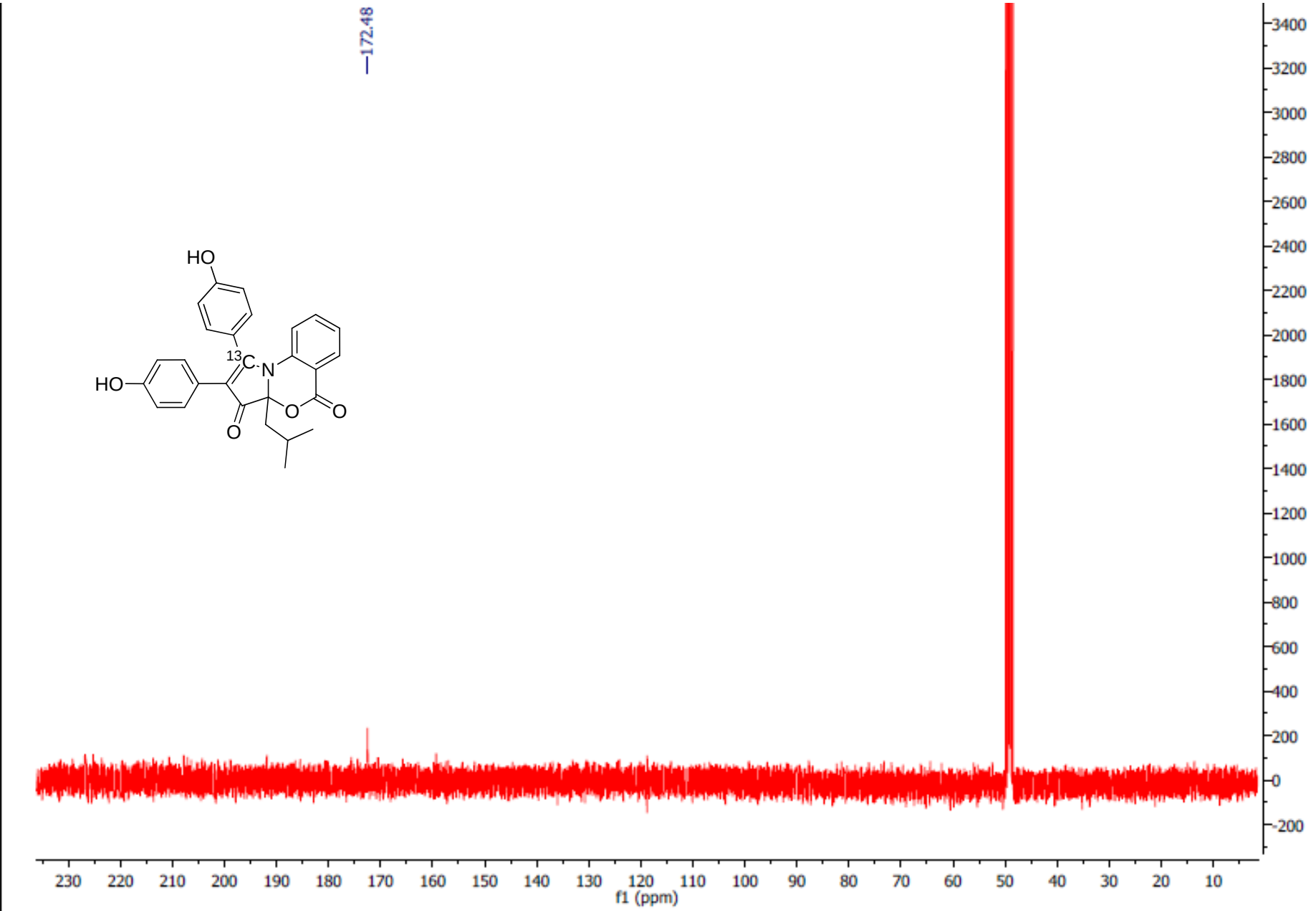


1



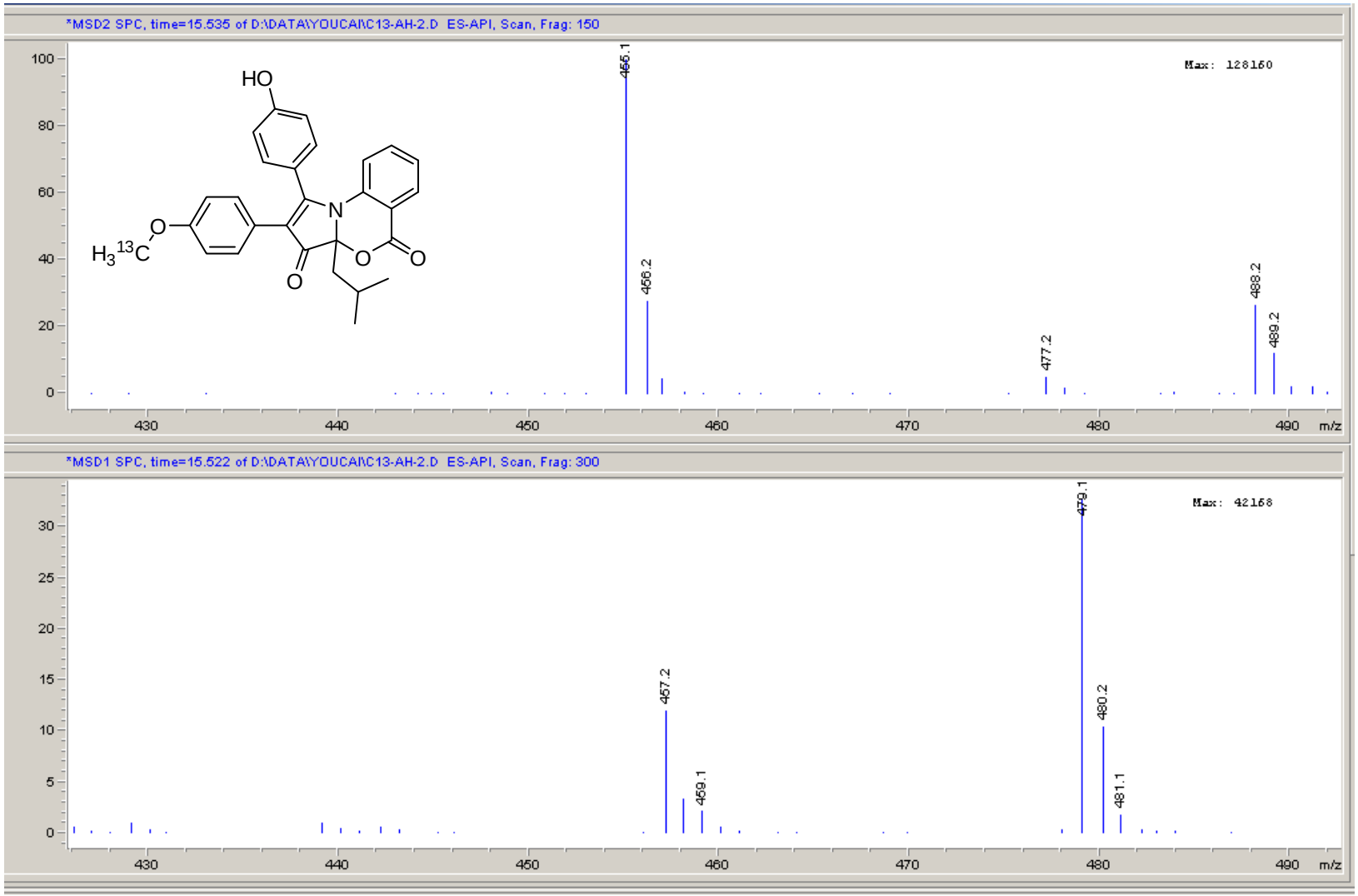
10

25.3 ^{13}C NMR spectrum of 4-[^{13}C]-discoipyrrole A (**10**) in CD_3OD

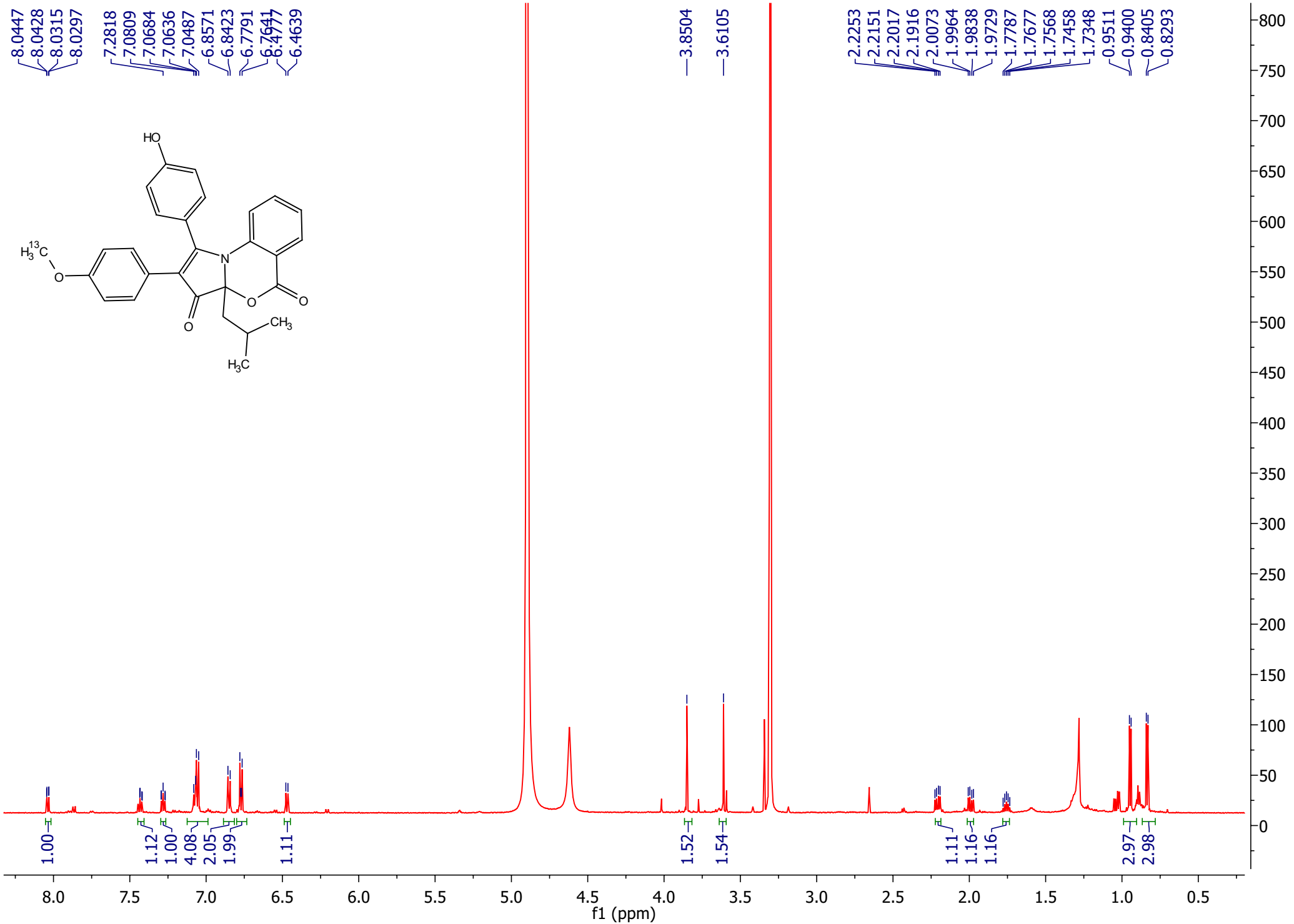


26. MS and NMR spectra of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**)

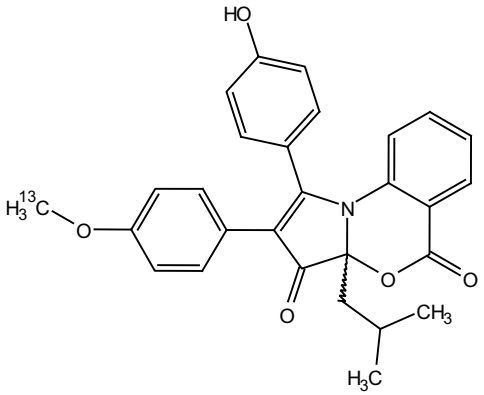
26.1 ESI-MS of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**)



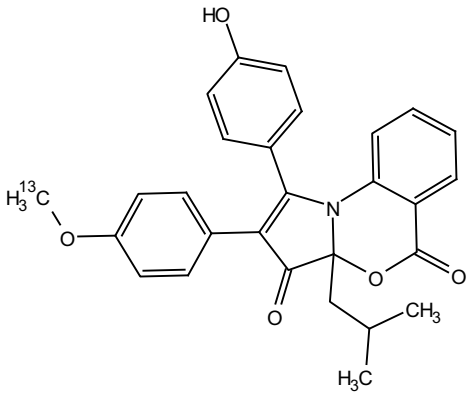
26.2. ^{1}H NMR of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)



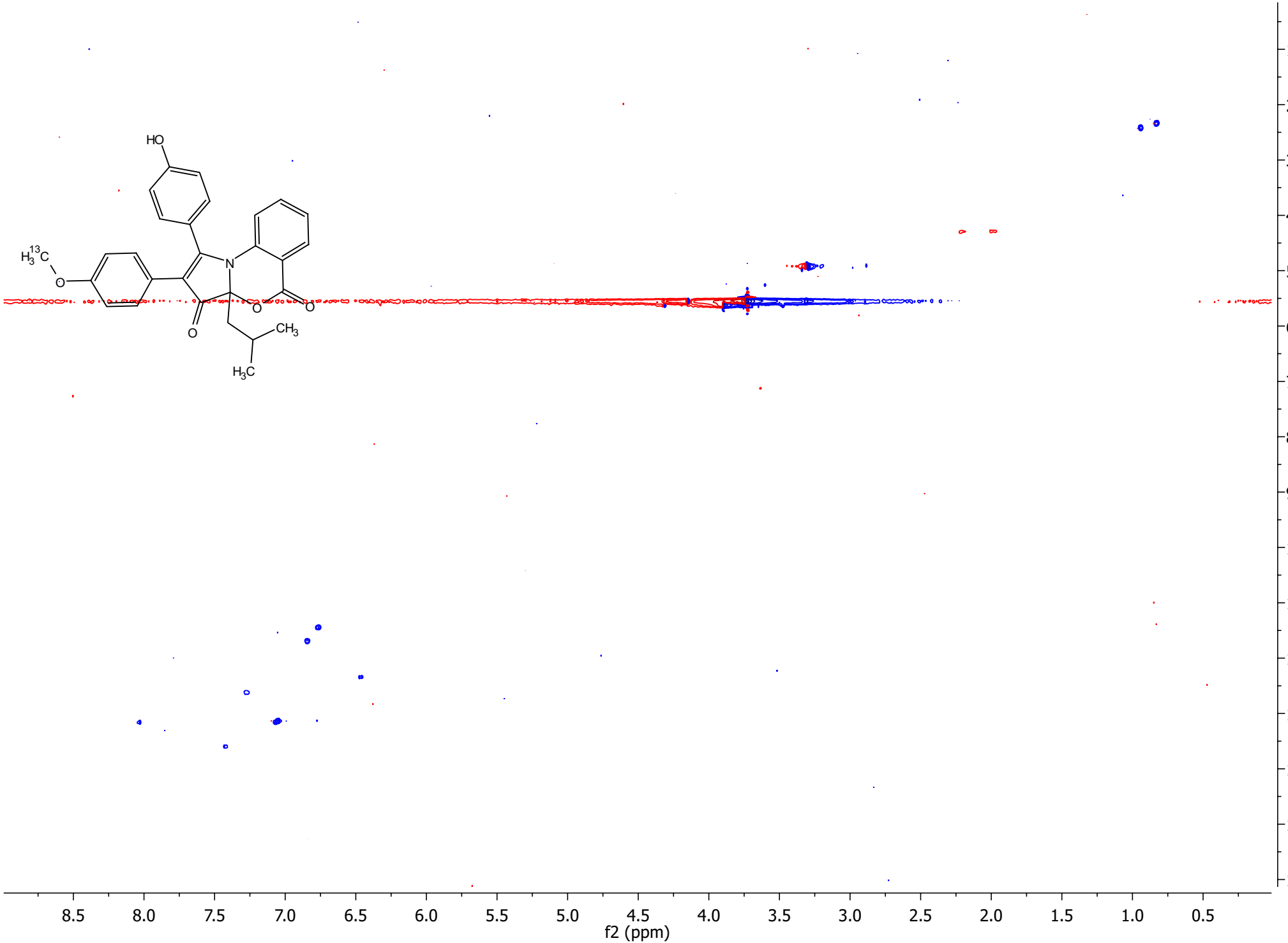
26.3. ^{13}C NMR of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)



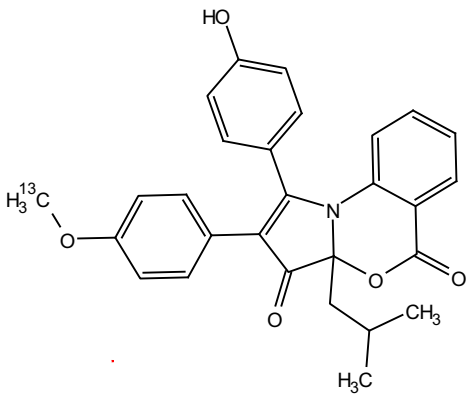
26.3. ^{1}H-^{1}H COSY spectrum of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)



26.4. HSQC spectrum of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)

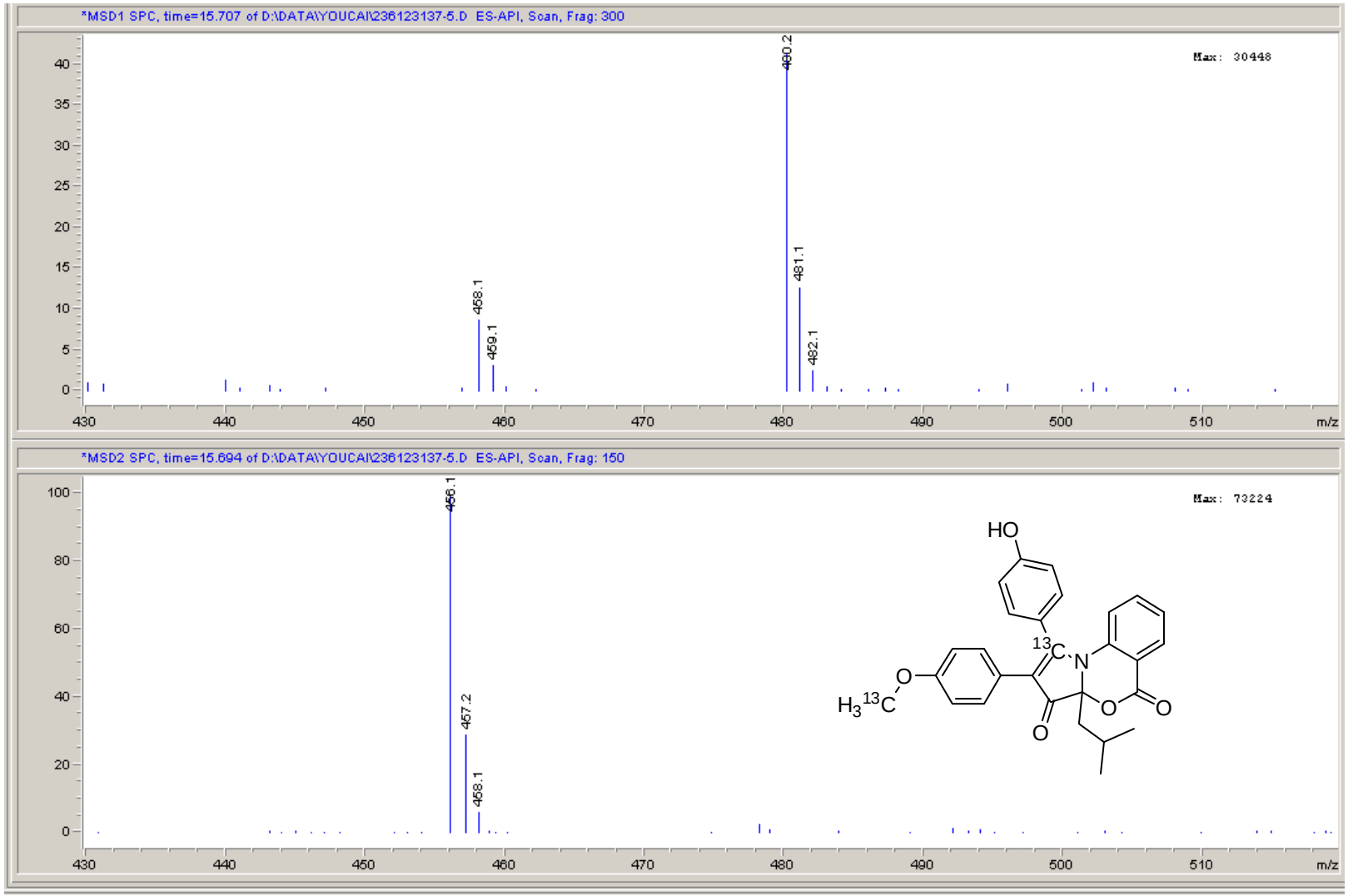


26.5 HMBC spectrum of 12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**12**) in CD_3OD (600 HMz)

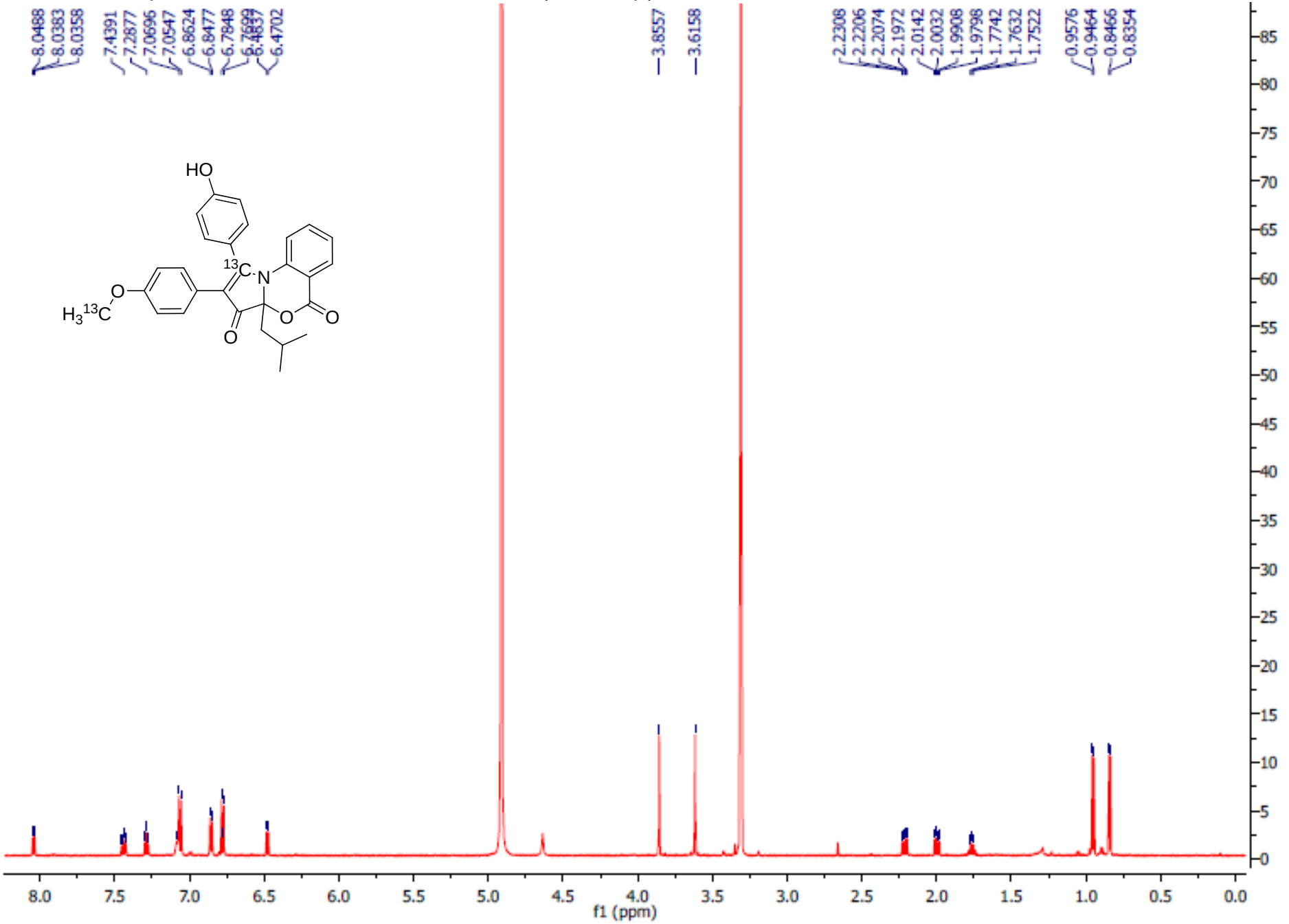


27. MS and NMR spectra of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)**

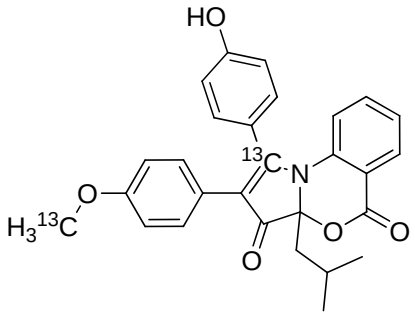
27.1 ESI-MS of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)**



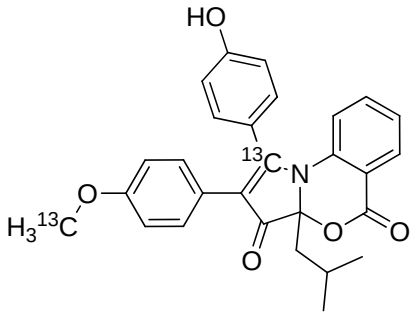
27.2 ^{1}H NMR spectrum of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)** in CD_3OD (600 HMz)



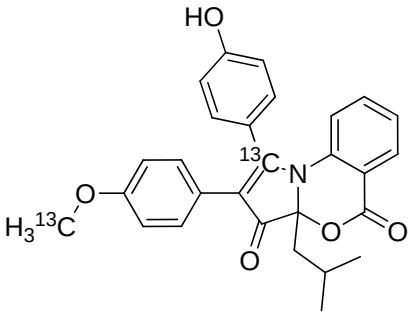
27.3 ^{13}C NMR spectrum of of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)** in CD_3OD (100 HMz)



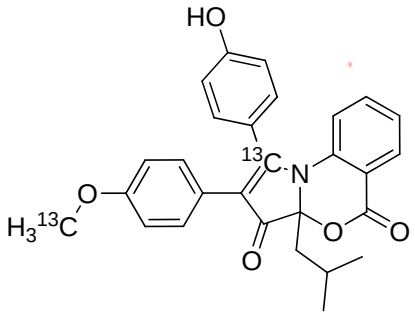
27.4 ^{1}H-^{1}H COSY of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A (**14**) in CD_3OD (600 HMz)



27.5 HSQC of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)** in CD_3OD (600 HMz)

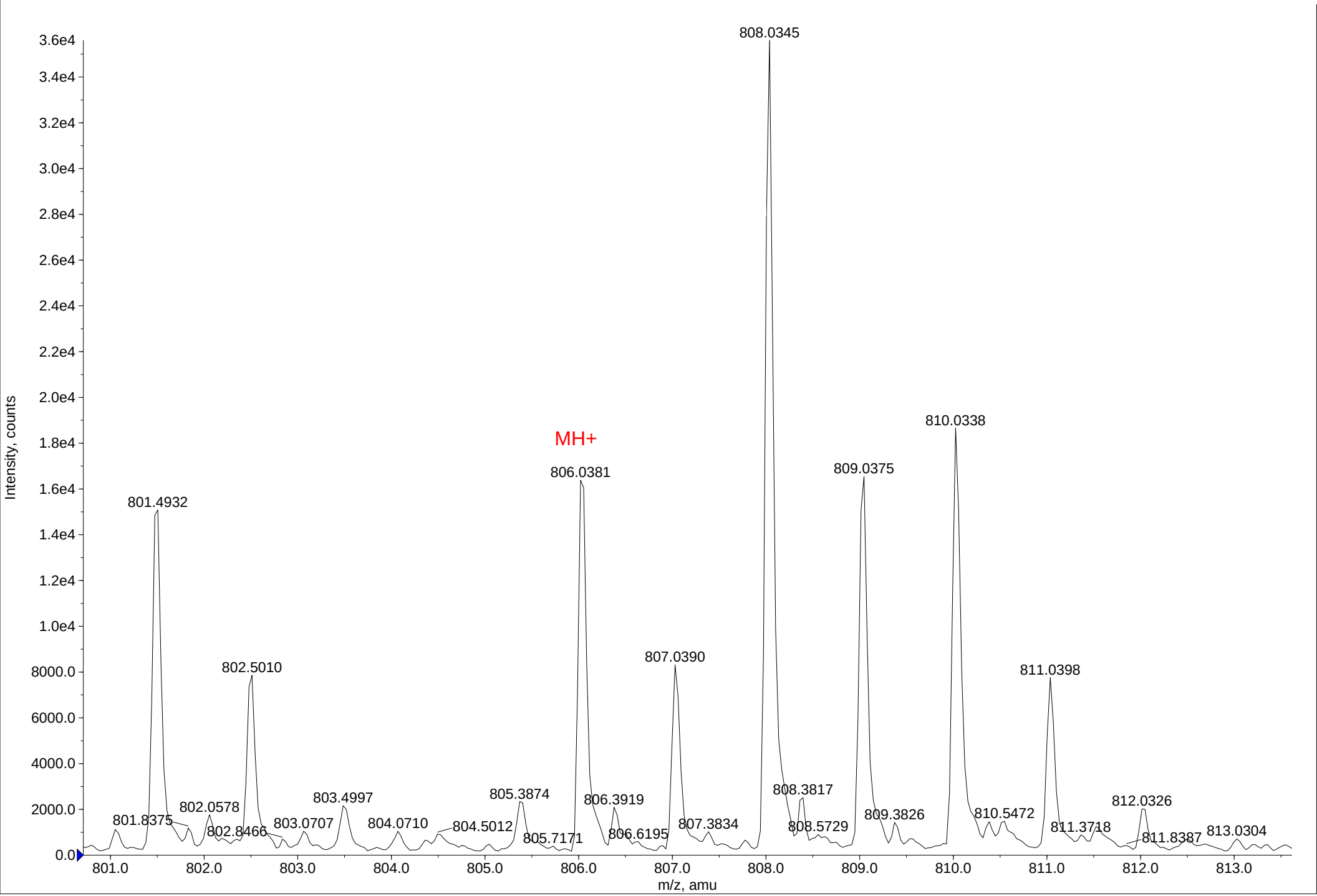


27.6 HMBC of 4-[^{13}C]-12-*O*-[$^{13}CH_3$]-methyl-discoipyrrole A **(14)** in CD_3OD (600 HMz)

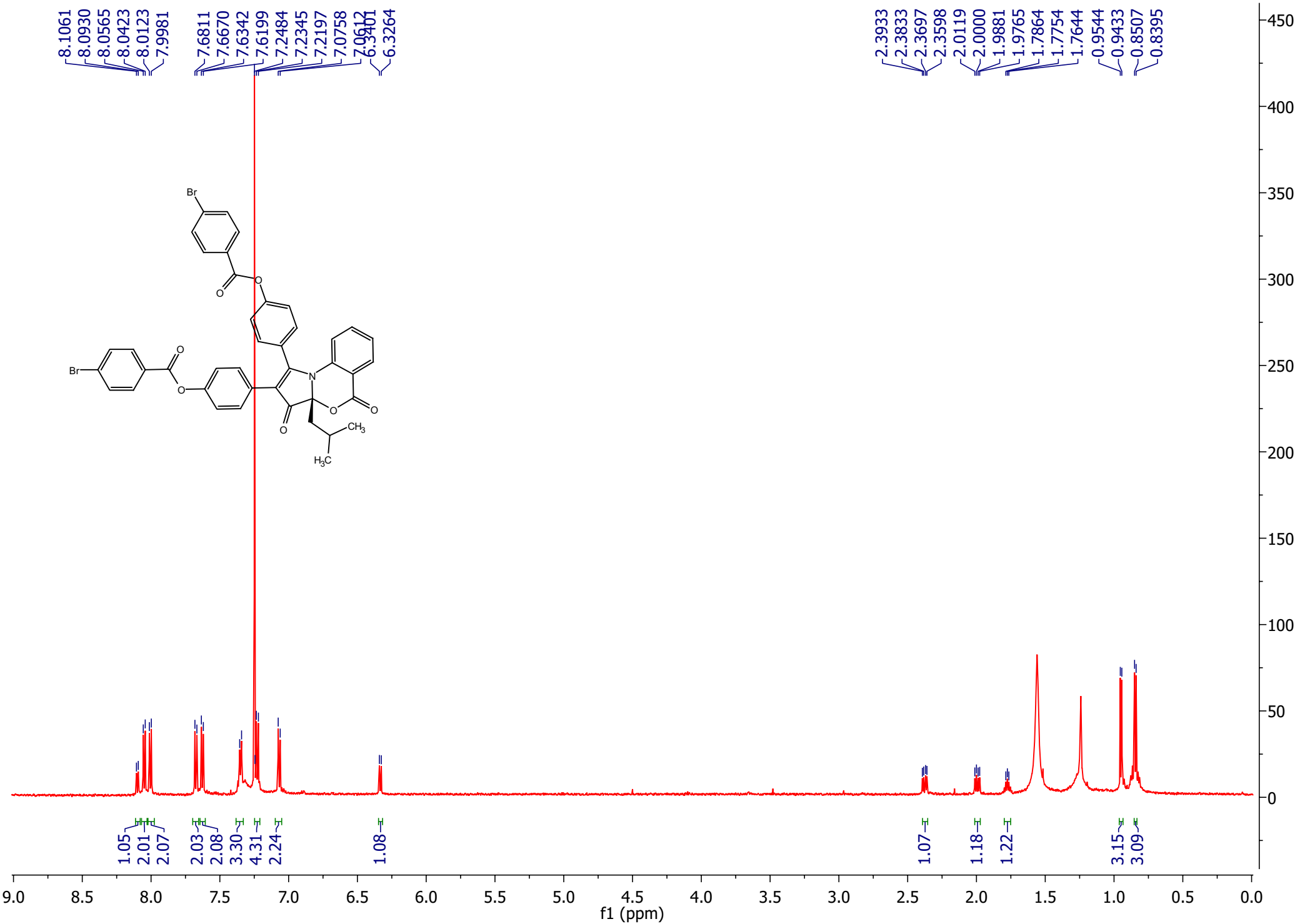


28. HR-ESIMS and NMR spectra of **1a**

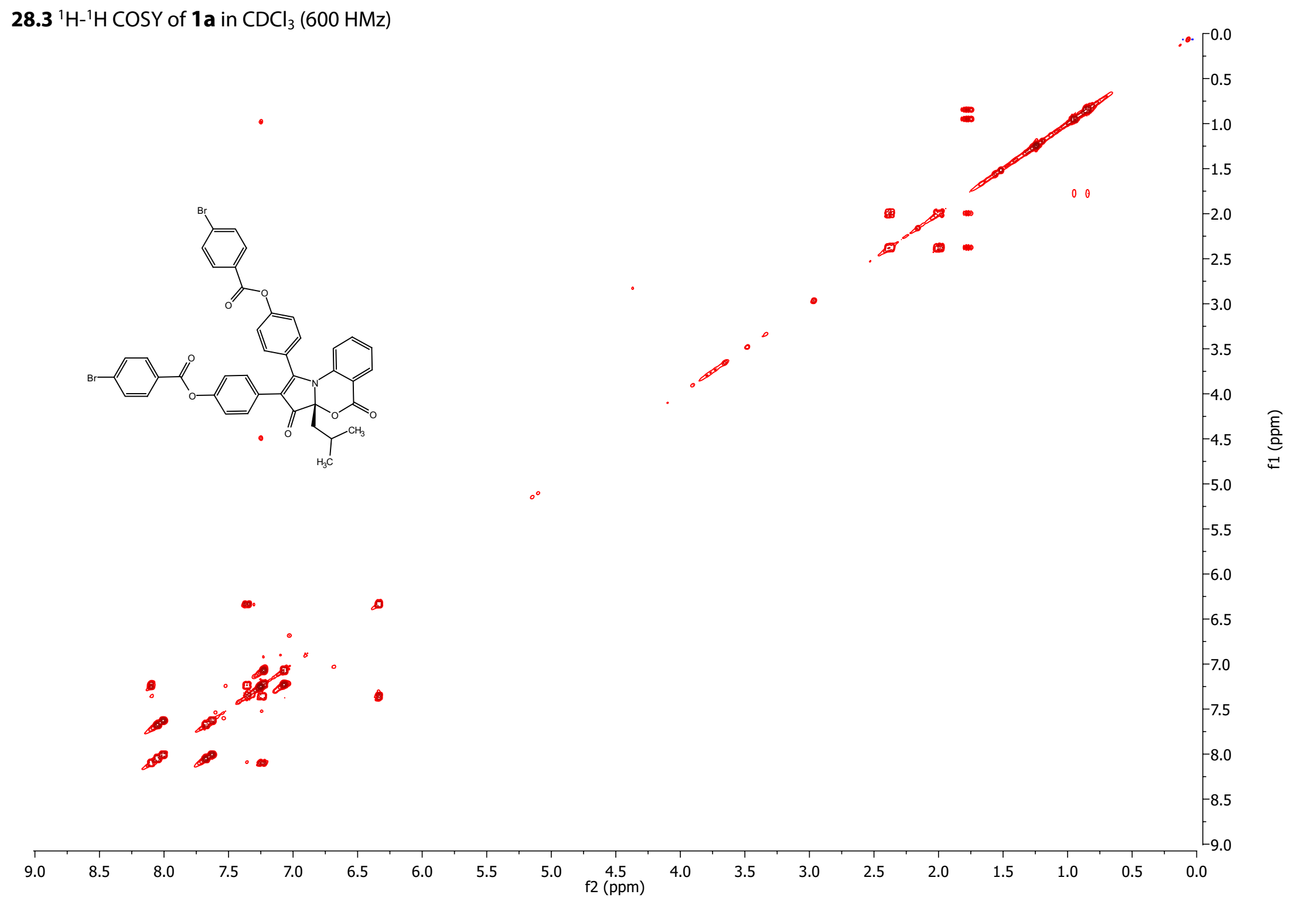
28.1. HR-ESIMS of **1a**-positive



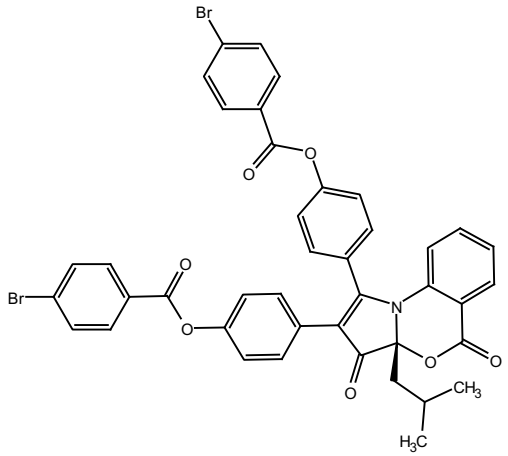
28.2 ^{1}H NMR of **1a** in $CDCl_3$ (600 HMz)



28.3 ^{1}H-^{1}H COSY of **1a** in $CDCl_3$ (600 HMz)



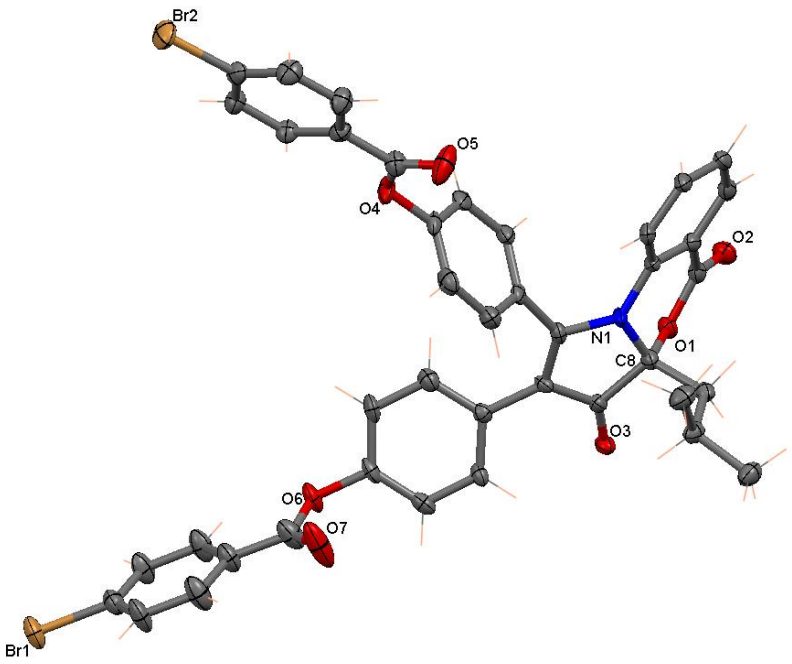
28.4 HSQC of **1a** in $CDCl_3$ (600 HMz)



Crystallographic data of 1a

Crystal data and structure refinement for C41H29Br2NO7.

Identification code	macmillan3	
Empirical formula	C41 H29 Br2 N O7	
Formula weight	807.47	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	P2(1)	
Unit cell dimensions	a = 15.289(3) Å	α= 90°.
	b = 9.4370(18) Å	β= 102.377(3)°.
	c = 25.086(5) Å	γ = 90°.
Volume	3535.5(12) $Å^3$	
Z	4	
Density (calculated)	1.517 Mg/m^3	
Absorption coefficient	2.346 mm^{-1}	
F(000)	1632	
Crystal size	0.25 x 0.20 x 0.02 mm^3	
Theta range for data collection	1.36 to 25.03°.	
Index ranges	-18<=h<=18, -11<=k<=11, -29<=l<=29	
Reflections collected	27424	
Independent reflections	12396 [R(int) = 0.0521]	
Completeness to theta = 25.03°	100.0 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.745 and 0.537	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	12396 / 1 / 923	
Goodness-of-fit on F^2	1.004	
Final R indices [I>2sigma(I)]	R1 = 0.0449, wR2 = 0.0902	
R indices (all data)	R1 = 0.0623, wR2 = 0.0967	
Absolute structure parameter	0.001(6)	
Largest diff. peak and hole	0.813 and -0.609 e.$Å^{-3}$	


Br2
O5
O4
O2
N1
O1
C8
O3
O6
O7
Br1