Synthesis

Reviews and Full Papers in Chemical Synthesis

Supporting Information

for DOI: 10.1055/a-2181-9709

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Thieme

Supplementary Information

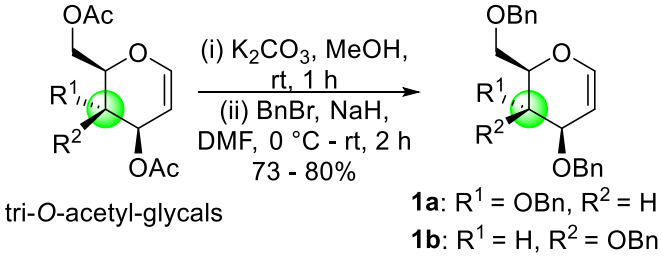
An efficient synthesis of natural product-inspired naphthoquinone fused glycohybrids and their *in-silico* docking studies

Ashish Khanna,[a] Ghanshyam Tiwari,[a] Vinay Kumar Mishra,[a] Kavita Singh,[b] Ram Sagar[*,a,b]

[a]*Department of Chemistry, Institute of Science, Banaras Hindu University, Varanasi – 221005.*
[b]*Glycochemistry Laboratory, School of Physical Sciences, Jawaharlal Nehru University, New Delhi – 110067*
** Corresponding Author*
Email id: ram.sagar@jnu.ac.in

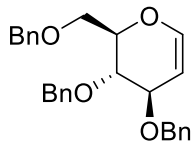
1. General experimental procedure for the synthesis of 3,4,6-tri-*O*-benzyl glycals (**1a** & **1b**):

To synthesize 2-*C*-formyl glucal, we initially carried out peracetylation on commercially available D-Glucose using acetic anhydride and perchloric acid as a catalyst which produces penta-*O*-acetylated glucopyranoside. Then, using penta-*O*-acetylated glucopyranoside, we performed a bromination reaction to produce a bromine group at the anomeric site to form 1-bromo-tetra-*O*-acetylated glucopyranoside. Further, we performed reductive elimination of 1-bromo-tetra-O-acetylated glucopyranoside under nitrogen atmosphere in presence of zinc and ammonium chloride with dry acetonitrile to produce tri-*O*-acetyl glucal with a good isolated yield using reported modified Fischer Zachs method. Then, benzylation of tri-*O*-acetyl glucal is achieved by deprotecting of acetyl group followed by benzylation with benzyl bromide in presence of strong base to give tri-*O*-benzyl glucal **1a** in a good isolated yield (73 %). The similar synthetic approach has been adopted for the synthesis of 2-*C*-formyl tri-*O*-benzyl-galactal **1b** with an excellent isolated yield (80 %) (Scheme S1). The tri-*O*-acetyl-glycals (500 mg, 1 equiv.), was stirred with K_2CO_3 (0.1 equiv.) in MeOH for an hour at room temperature, until complete deacetylation (TLC). Then reaction mixture was filtered using celite-sintered funnel and filtrate was concentrated to get decetyated product as crude residue which was used in next step without further purification. The crude residue was dissolved in dry DMF followed by addition of NaH (3 equiv.) at 0 °C. The BnBr (3.3 equiv.) was added drop wise at same temperature and the reaction mixture was further stirred for 3 hours at rt. The completion of reaction was monitored by the TLC. After the completion, the reaction mixture was quenched by adding MeOH and reaction mixture was extracted with EtOAc (4 × 10 ml) and combined organic layer washed with brine solution and concentrated to get crude benzylated product. After flash column chromatography the pure benzyated glucal product **1a** is good isolated yield (600 mg, 73%). The similar reaction protocol was adopted for the synthesis of bezylated galactal **1b** and pure product was obtained in very good isolated yield (610 mg, 75%).



Scheme S1: Synthesis of tri-*O*-benzyl glycals **1a** & **1b**.

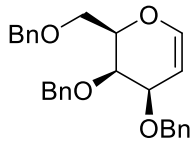
(2R,3S,4R)-3,4-bis(benzyloxy)-2-((benzyloxy)methyl)-3,4-dihydro-2H-pyran (1a):



1a

The product **1a** obtained as a colourless syrup (600 mg, 73%). R_f = 0.53 (1:9 EtOAc: Hex) visualized by charring with 10 % H_2SO_4 methanolic solution. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.39-7.25 (m, 15H, Ar-*H*), 6.36 (d, *J* = 7.6 Hz, 1H, H-1), 4.88-4.84 (m, 1H, H-2), 4.70-4.54 (m, 4H, Ar-CH_2), 4.50-4.41 (m, 2H, Ar-C*H*$_2$), 4.18-4.09 (m, 1H, H-5), 3.94 (s, 1H, H-6), 3.77 (dd, *J* = 10.5, 7.6 Hz, 1H, H-6), 3.64 (q, *J* = 5.1 Hz, 1H, H-3), 3.55 (s, 1H, H-4). ^{13}C-NMR (126 MHz, $CDCl_3$) δ 144.3(C-1), 138.6, 138.4, 138.1(Ar-*C-1*), 129.1, 128.9, 128.7, 128.5, 128.5, 128.4, 128.2, 128.0, 127.9, 127.8, 127.8, 127.6, 127.5 (Ar – *C-2- 5*), 100.0 (C-2), 76.8 (C-3,4), 75.8(C-5), 73.5, 72.2, 71.4 (Ar-*CH*$_2$), 68.5 (C-6).

(2R,3R,4R)-3,4-bis(benzyloxy)-2-((benzyloxy)methyl)-3,4-dihydro-2H-pyran (1b):



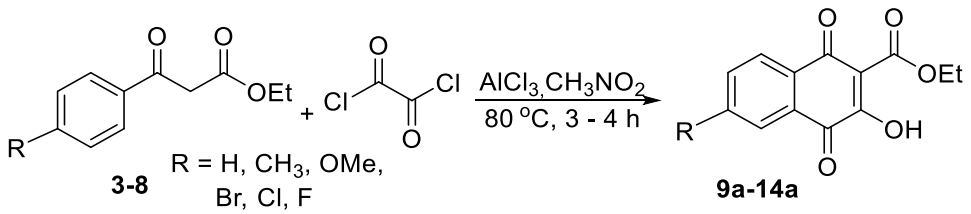
1b

The product **1b** obtained as colourless syrup (610 mg, 75%). R_f = 0.5 (1:9 EtOAc: Hex) visualized by charring with 10% H_2SO_4 methanolic solution. ^{1}H-NMR (500 MHz, $CDCl_3$) δ 7.39-7.25 (m, 15H, Ar-*H*), 6.31 (d, *J* = 7.6 Hz, 1H, H-1), 4.88-4.84 (m, 1H, H-2), 4.70-4.54 (m, 4H, Ar-CH_2), 4.50-4.41 (m, 2H, Ar-C*H*$_2$), 4.11 (m, 1H, H-5), 3.94 (s, 1H, H-6), 3.77 (dd, *J* = 10.1 Hz, 1H, H-6), 3.74 (d, *J* = 4.4 Hz, 1H, H-4), 3.64 (q, *J* = 5.1 Hz, 1H, H-3). ^{13}C-NMR (126 MHz, $CDCl_3$) δ 144.3(C-1), 138.6, 138.4, 138.1(Ar-*C-1*), 129.1, 128.9, 128.7, 128.5, 128.5, 128.4, 128.2, 128.0, 127.9, 127.8, 127.8, 127.6, 127.5 (Ar – *C-2- 5*), 100.0 (C-2), 77.2 (C-3,4), 75.8(C-5), 73.5, 72.2, 71.4 (Ar-*CH*$_2$), 68.5 (C-6).

2. Modified experimental procedure and spectral data for the synthesis of 2-hydroxy naphthoquinone esters (9a – 14a):

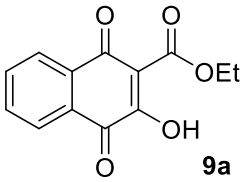
To a dried two neck round bottom flask, a solution of β- ketoester **3** (500 mg, 1.0 equiv.) in dry nitromethane (5 ml) followed with the batchwise addition of dry $AlCl_3$ (475 mg, 1.5 equiv.) at low temperature to the reaction mixture. After 15 min of stirring, oxalyl chloride (350 mg, 1.1 equiv.) was added dropwise, followed with the addition of remaining batch of $AlCl_3$ (475 mg, 1.5 equiv.) .Then after 15 min of stirring, entire reaction mixture batch was then transferred to 80 °C for 3 - 4 h. A solution of 10 % oxalic acid (10 ml) was added under stirring at r.t to quench. The resulting reaction mixture was extracted by EtOAc and dried over high vacuum, purified over flash column chromatography to obtain pure 2-hydroxy naphthoquinone esters **9a** in good yield (Scheme S2). The similar reaction protocol was adopted for the preparation

of **10a-14a** compounds and they were directly used as a substrate in next step without any further purification.



Scheme S2: Modified method for the synthesis of 2-hydroxy naphthoquinone esters **9a–14a**.

Ethyl 3-hydroxy-1,4-dioxo-1,4-dihydronaphthalene-2-carboxylate (9a):



Compound 9a: The product obtained as a yellow semi solid (384 mg, 60%). R_f = 0.1 (9:1 EtOAc: CH_3OH) visualized under UV chamber. ^{1}H-NMR (500 MHz, DMSO-d_6) δ 7.85 (d, *J* = 7.9 Hz, 1H), 7.77 (d, *J* = 7.5 Hz, 1H), 7.67 (t, *J* = 7.7 Hz, 1H), 7.55 (t, *J* = 7.7 Hz, 1H), 4.09 - 4.02 (q, *J* = 7.7, 7.2 Hz, 2H), 1.15 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (126 MHz, DMSO-d_6) δ 184.5, 181.2, 165.1, 159.8, 134.7, 133.4, 132.0, 130.5, 126.1, 125.7, 111.2, 61.3, 14.8.

3. SAR studies:

Our investigation revealed that the molecule exhibits multiple constitutional molecular descriptors, including aromatic rings (AR_1-AR_4), hydrogen bond acceptor sites (A_1-A_7), and hydrophobic sites (H_1-H_2). When we introduced different substituents (R) on the aromatic ring (AR_1), encompassing both electron-donating and electron-withdrawing groups, we observed a diverse range of ligand potency and intrinsic activity. This indicates that these substitutions play a crucial role in modulating the molecule's biological activity.

Moreover, the other aromatic rings (AR_2-AR_4) contribute to effective binding within the hydrophobic pocket of our target protein (Figure S1). This information will guide us in further optimizing our molecule's design for enhanced binding and potential therapeutic applications.

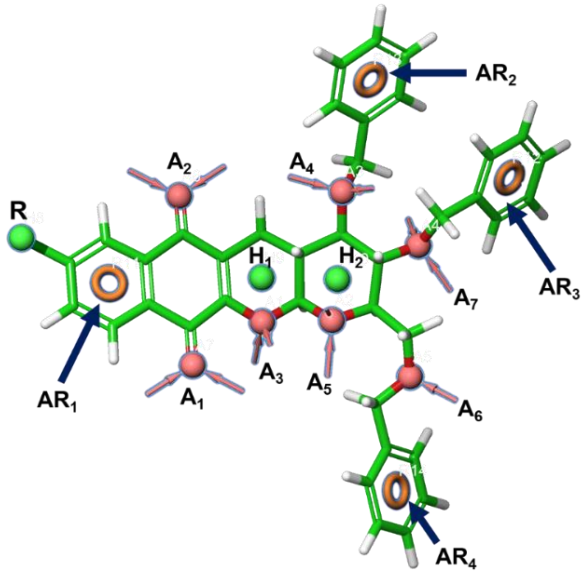


Figure S1. Pharmacophore model showing different constitutional molecular descriptors on naphthoquinone derived glycohybrids.

4. *In-silico* Molecular Interactions and Docking Analysis:

The docking studies were carried out using various derived naphthoquinone glycohybrids with proposed binding pocket of X-ray crystallographic structure (PDB ID: **6BFN**, resolution: 2.4Å) obtained from rcsb protein data bank. Docking was carried out with Autodock Vina 4.0, further for visualization and analysis of the interaction of ligands with the protein after docking carried out with PyMol software. 2D visualization and detailed ligand interaction visualization was carried out with Biovia Discovery Studio Visualizer v20.1.0.19295. The QSAR and SAR studies was carried out using Schrödinger Maestro tool.

(a) Ligand Preparation:

The 3D structure of designed naphthoquinone based glycohybrids was drawn using Chem3D software and ran MM2 for the minimization of energy which is then converted in SDF (structure data file) format. These SDF files were converted to PDB (protein data bank) format using OpenBabel tool (OPENBABEL - Chemical file format converter, cheminfo.org). Further PDB file format of ligand was then imported to Maestro tool and optimized through prescribed command lines for the ligand preparation. The processed data file of prepared ligand was then saved in the workspace as MAEGZ file format. Having possible rotatable bonds and minimized conformational structure.

(b) Protein Preparation:

It was selected and downloaded from RCSB-PDB database (https://www.rcsb.org/) using the PDB ID: **6BFN** and was optimized along with the ligands using Autodock Tools 4.0. For docking, both the protein and ligands were converted to .pdbqt format in Autodock Tools 4.0. Docking performed using Autodock Vina program 4.0 (http://vina.scripps.edu/) using grid dimensions of 40×40×40 cm^3 and their binding energies (dG) were generated. The generated protein-ligand interactions were analyzed and illustrated using LIGPLOT+ v1.4.5 and Discovery studio to generate 2D plots of the interactions. The generated plots showed hydrophobic interactions and hydrogen bonds between the ligand and binding site of protein. For 3D visualisation, the output file was opened in PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.sx) and interactions between the ligand and protein were visualised in their respective binding pockets. 1,4-naphthoquinone was also docked as standard.

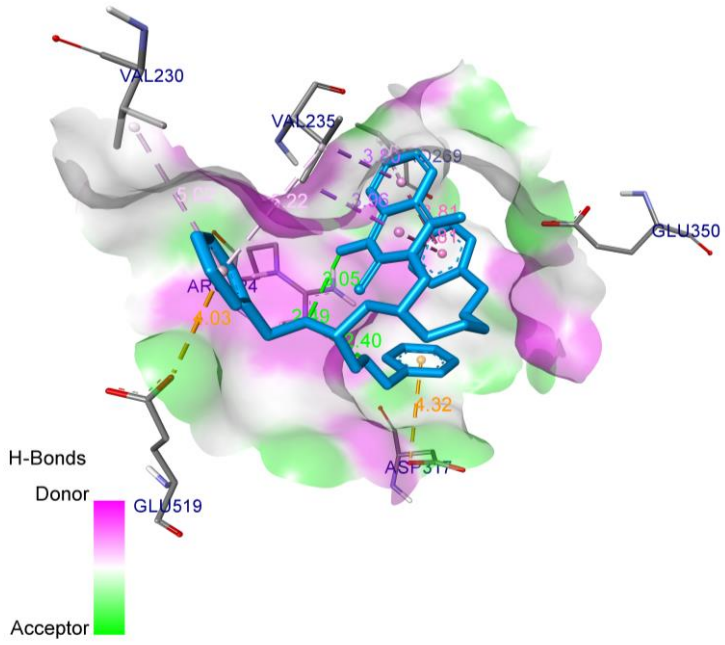
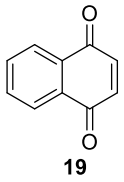
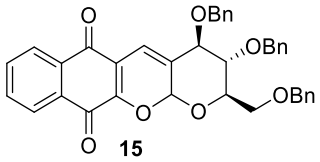
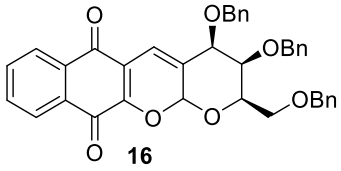
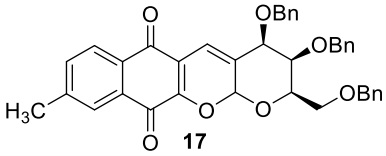
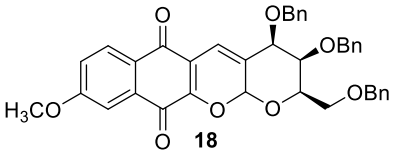


Figure S2: Docking analysis of naphthoquinone based glycohybrid **15** and catalytic site of IRAK1 protein with 3D representation of hydrogen bond donor/acceptor surface (shown in pink and green color), hydrogen bond(green color). The docking analysis was conducted using the collected set of compounds (cyan) into the proposed binding pocket of the X-ray crystallographic structure of IRAK1 protein (Interleukin Receptor associated kinase-1) (PDB ID: **6BFN**, resolution: 2.4 Å).

Table S1: Binding energies of naphthoquinone-derived glycohybrids **15-18** (ligands) against 6BFN (Protein):

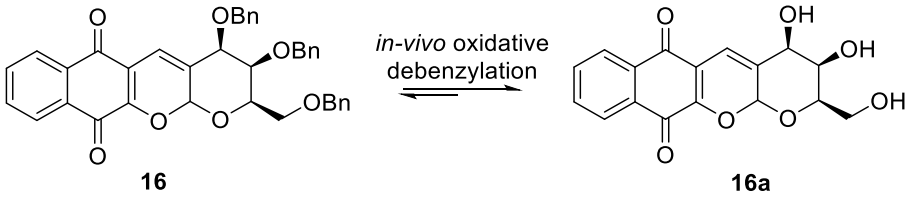
S. No.	Ligands	Binding affinity (ΔG) (Kcal/mol)
1.	 **19**	-8.1
2.	 **15**	-9.4
3.	 **16**	-9.9
4.	 **17**	-9.0
5.	 **18**	**-10.1**

5. Pharmacokinetic (ADMET) analysis:

The synthesized naphthoquinones based glycohybrids were analysed for their drug likeness studies, bioavailability and ADME (absorption, distribution, metabolism and excretion) using qikprop tool (Schrodinger Maestro) and SwissADME. Toxicity of the compounds were tested using ProTox-II (https://tox-new.charite.de/protox_II/).

In order to anticipate the crucial pharmacokinetic characteristics of our designed glycohybrids, various properties were assessed, including absorption, permeability, distribution, metabolism, excretion, and toxicity. These evaluations were aimed at understanding their effects on human body targets, as well as ensuring safety assessments. Thus, we performed ADMET analysis using qikprop tool and swissADME tool for the ADME analysis and Protox II for the prediction of oral toxicity of the designed glycohybrids. ADME analysis for benzyl protected

naphthoquinone based glycohybrids was carried out, the observation showed that it follows the Lipinski rule of five and thus may possess drug likeness. The pharmacokinetic report suggests that compound have high GI(Gastro-intestinal) absorption. The high value of $logP_{o/w}$ (6.45) suggests the high hydrophobicity of the compound which allows effective penetration through certain protein barriers in cellular absorption. Thus, it allows, effective transport of the compound around the body. It has been during the QSAR and QSPR studies that, upon absorption of designed naphthoquinone based glycohybrid, it gets metabolized through initial oxidative debenzylation which furnished the hydroxyl derivative of the naphthoquinone based glycohybrid. Upon the ADME analysis it has been found that the hydroxyl derivative of naphthoquinone based glycohybrid **16a** follows Lipinski, Ghose, Veber, Egan and Muegge model (used as an indicator of druglikeness of the molecule). Along with this due to reduction in size and increased solubility ($logP_{o/w}$ = 1.62) its efficacy and bioavailability has been increased (Scheme S3).

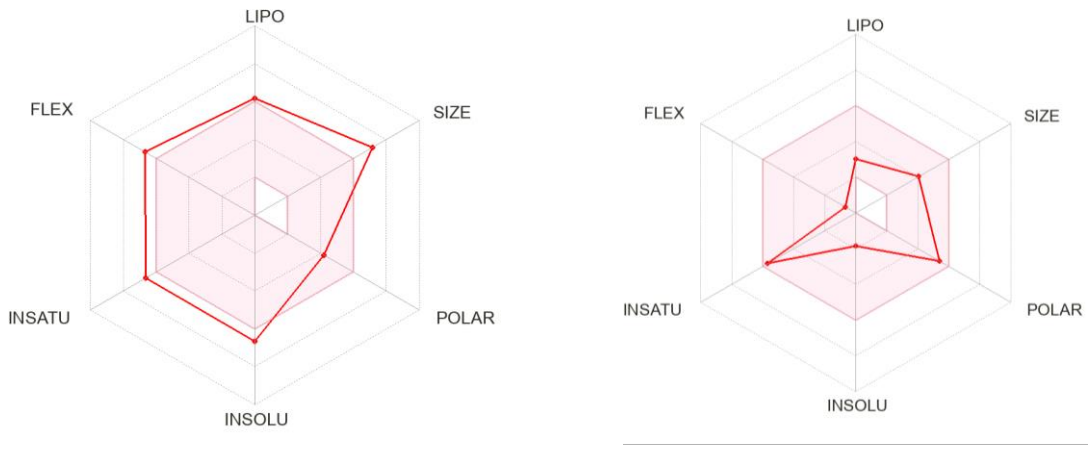


Scheme S3: Expected enzyme catalyzed *in-vivo* oxidative debenzylation of compound **16**.

The ADME analysis with other substituted naphthoquinone based glycohybrids shows increase in the $logP_{o/w}$ value for the compound having electron withdrawing substituents and decrease in the $logP_{o/w}$ value for the compound having electron donating substituents. This suggests the role of diverse substitution on the druggability of the synthesized compound.

The radar diagram (Figure S3) shows the suitable physiochemical space for the bioavailability of the benzyl protected naphthoquinone based glycohybrid I and metabolized debenzylated hydroxyl derivative of naphthoquinone based glycohybrid II. The light pink coloured hexagon indicates the optimum range of bioavailability and red coloured line indicates values of the derived compounds. The significant increase in bioavailability of metabolized compound was clearly visible in plot II. This infers that our designed glycohybrid also acts as a prodrug candidate.

The toxicity of compound **16** and **16a** were predicted to be of class 4 with predicted LD_{50} of 560 mg/kg and 1000 mg/kg. The increased LD_{50} value also suggest that the compound may not show cytotoxicity when gets metabolized inside the body.

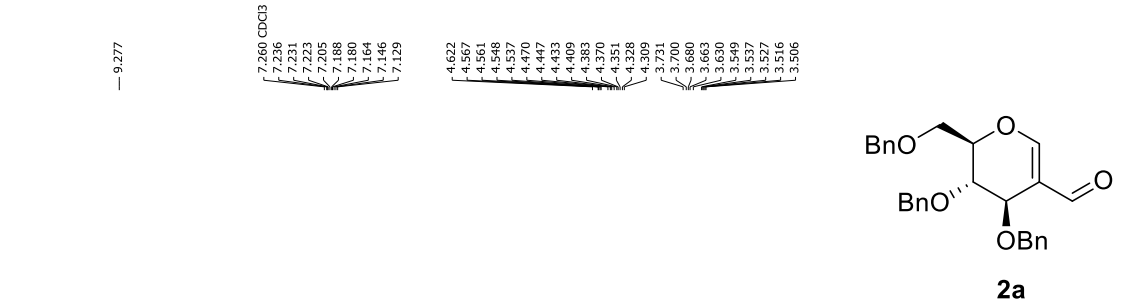


I **II**

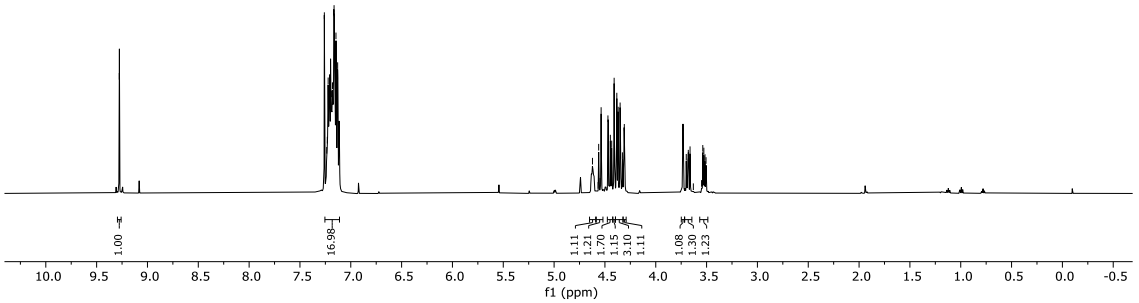
Figure S3: Radar plot for the representation of physiochemical space for the bioavailability of compound. Plot I indicate the bioavailability of compound **16** and plot **II** indicates the bioavailability of compound **16a**.

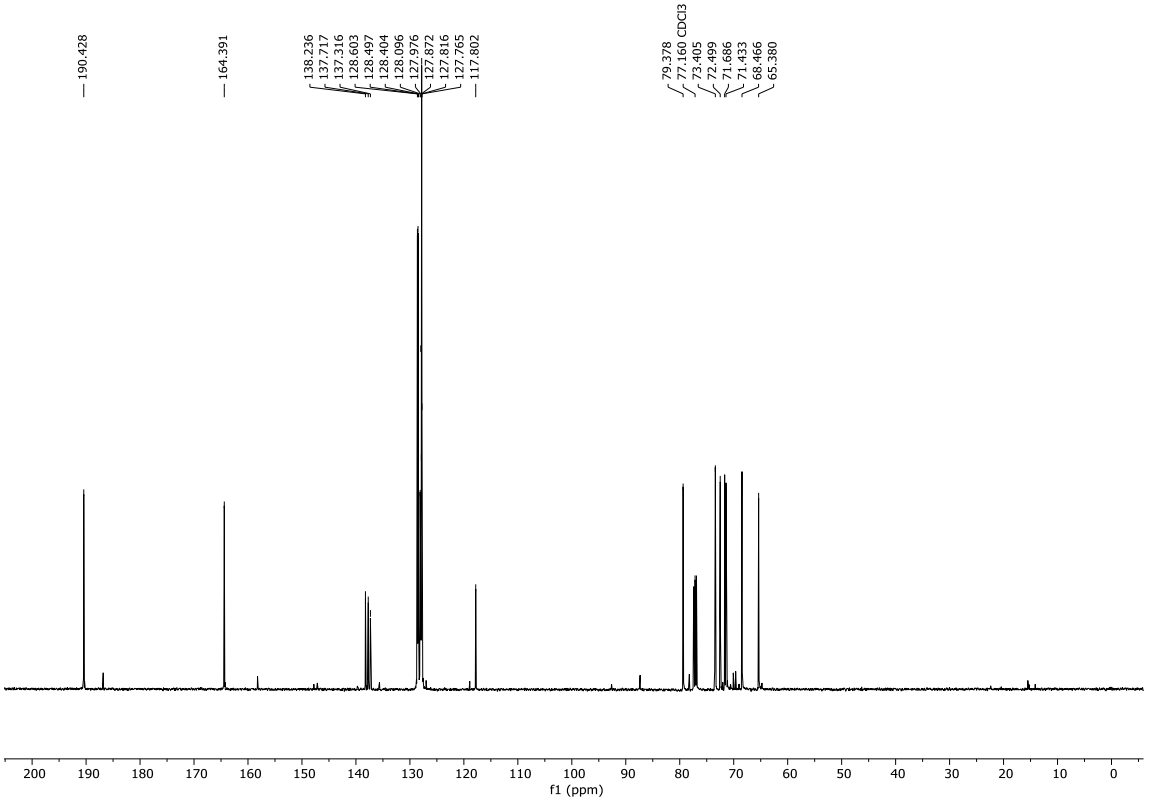
6. Copies of ^{1}H and ^{13}C NMR spectra of representative compounds

^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **2a**

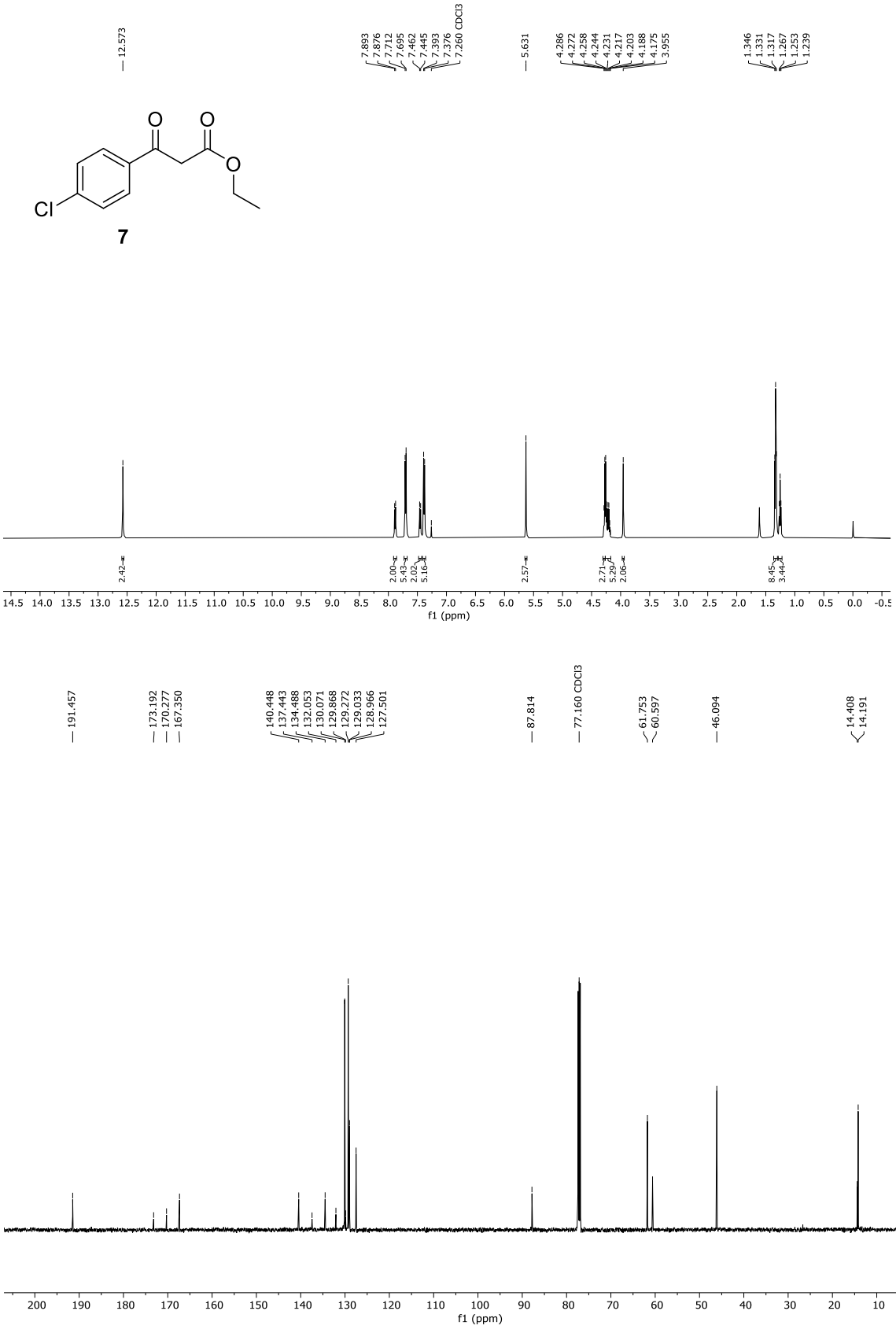


2a

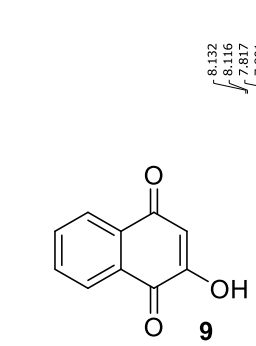




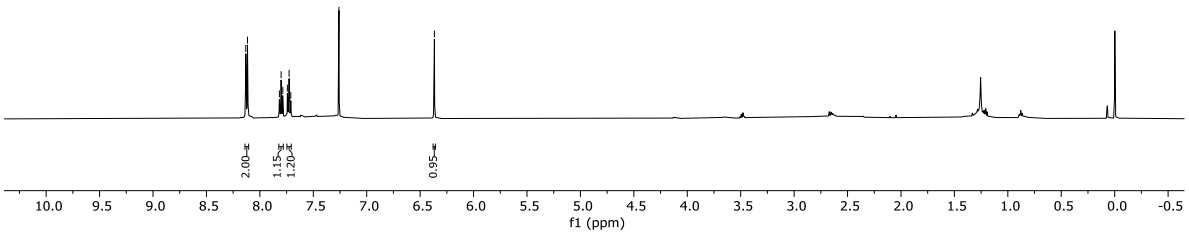
^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **7**

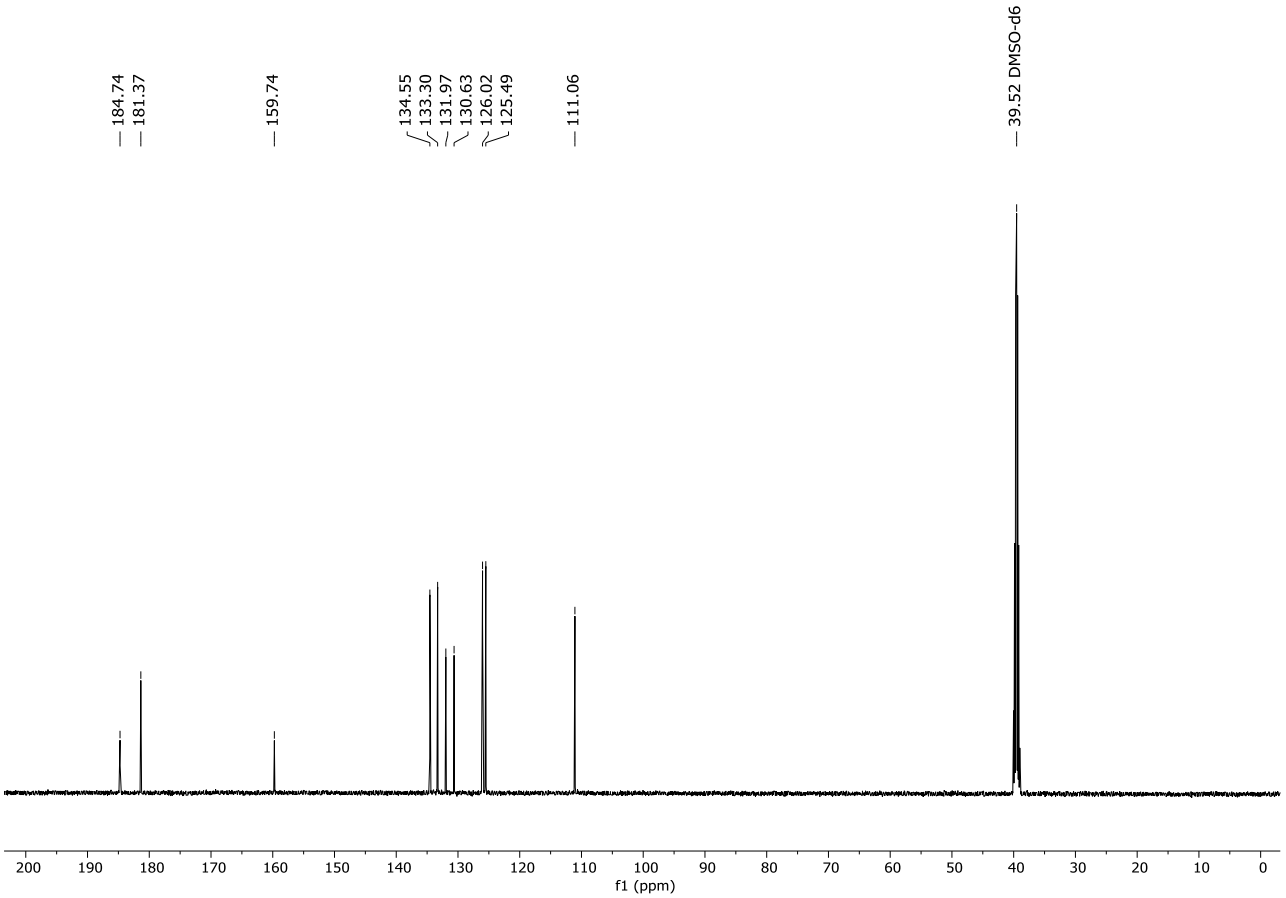


^{1}H NMR ($CDCl_3$, 500 MHz) ^{13}C NMR (DMSO-d_6, 126 MHz) of **9**

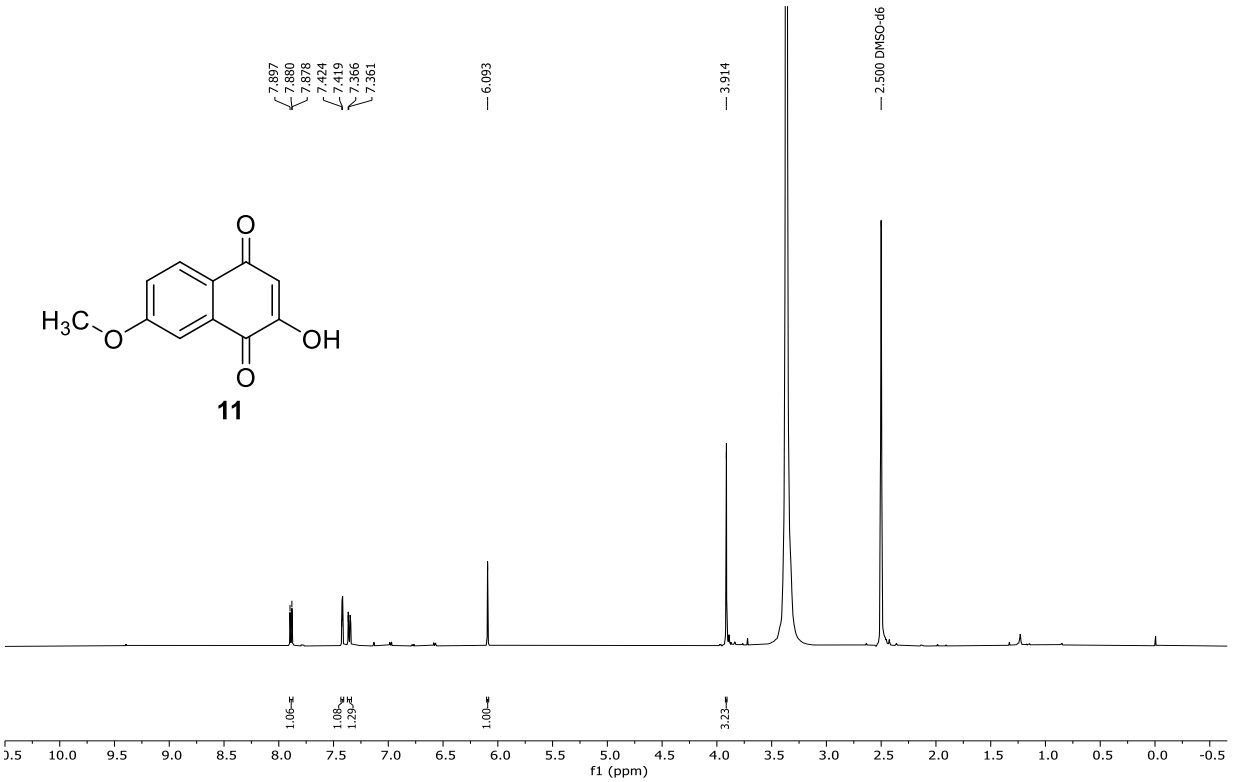


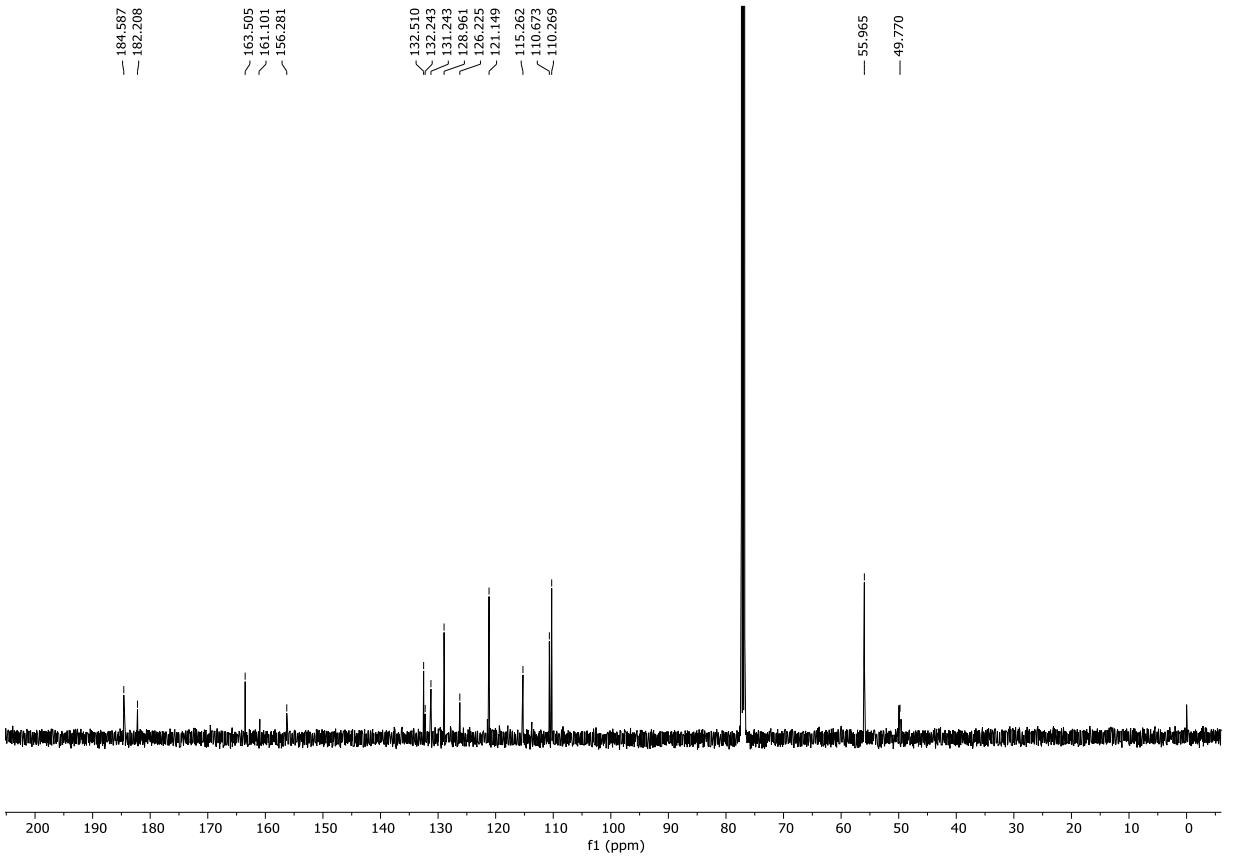
9



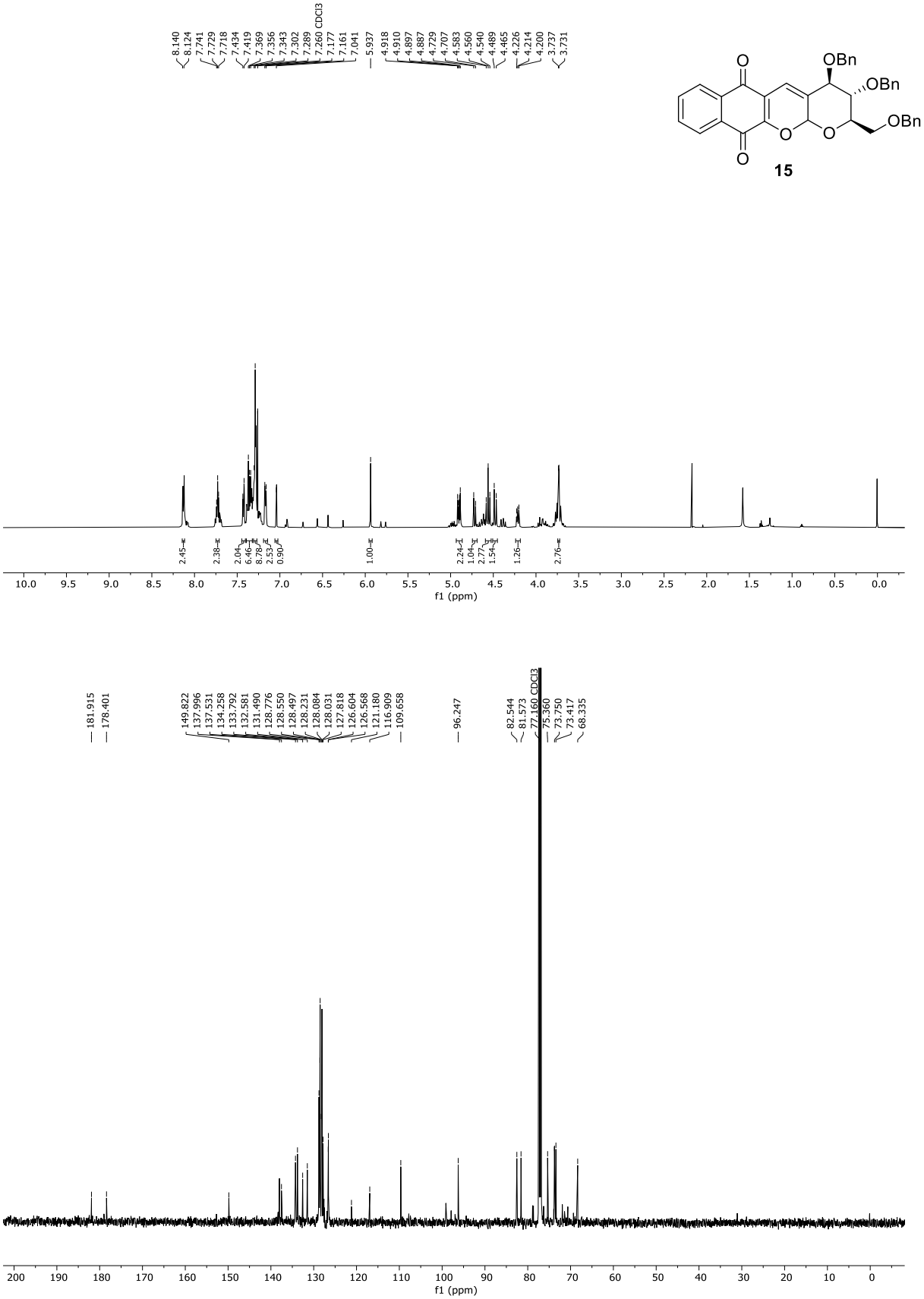


^{1}H NMR (DMSO-d_6, 500MHz) and ^{13}C NMR ($CDCl_3$, 126MHz) of **11**

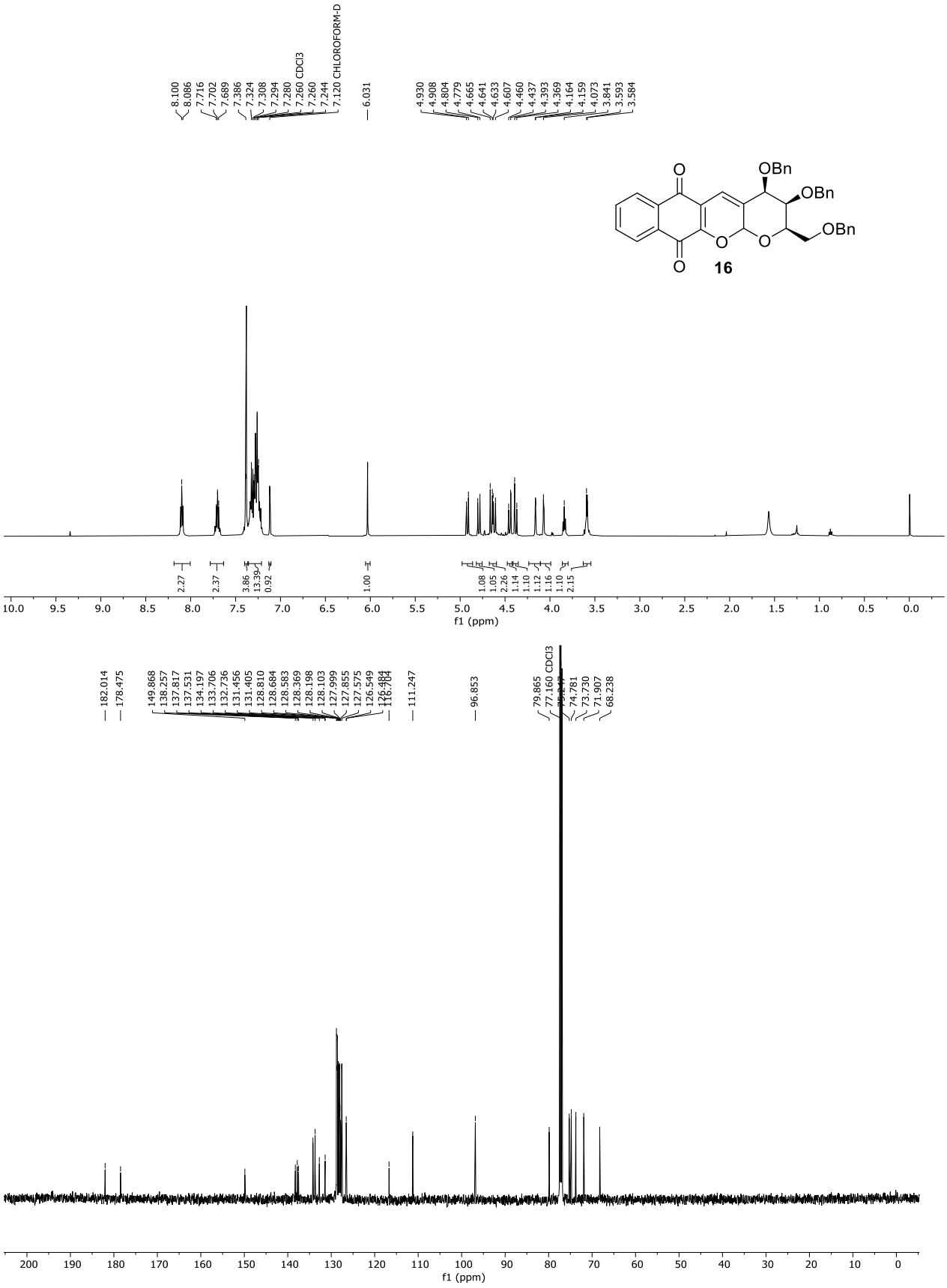




^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **15**

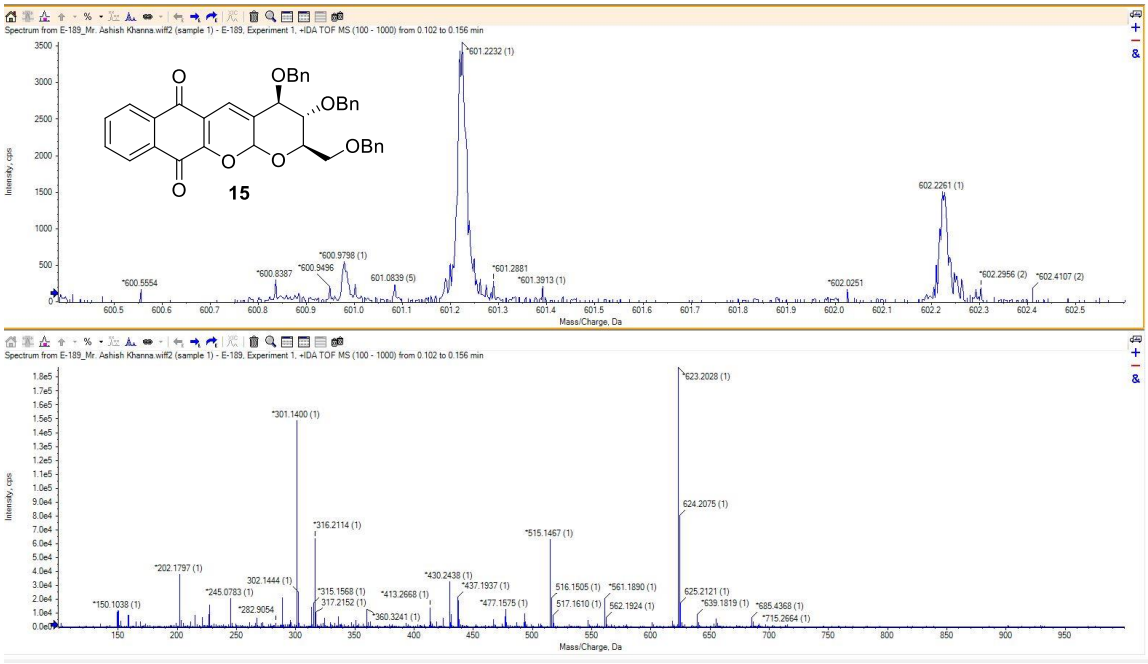


^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **16**

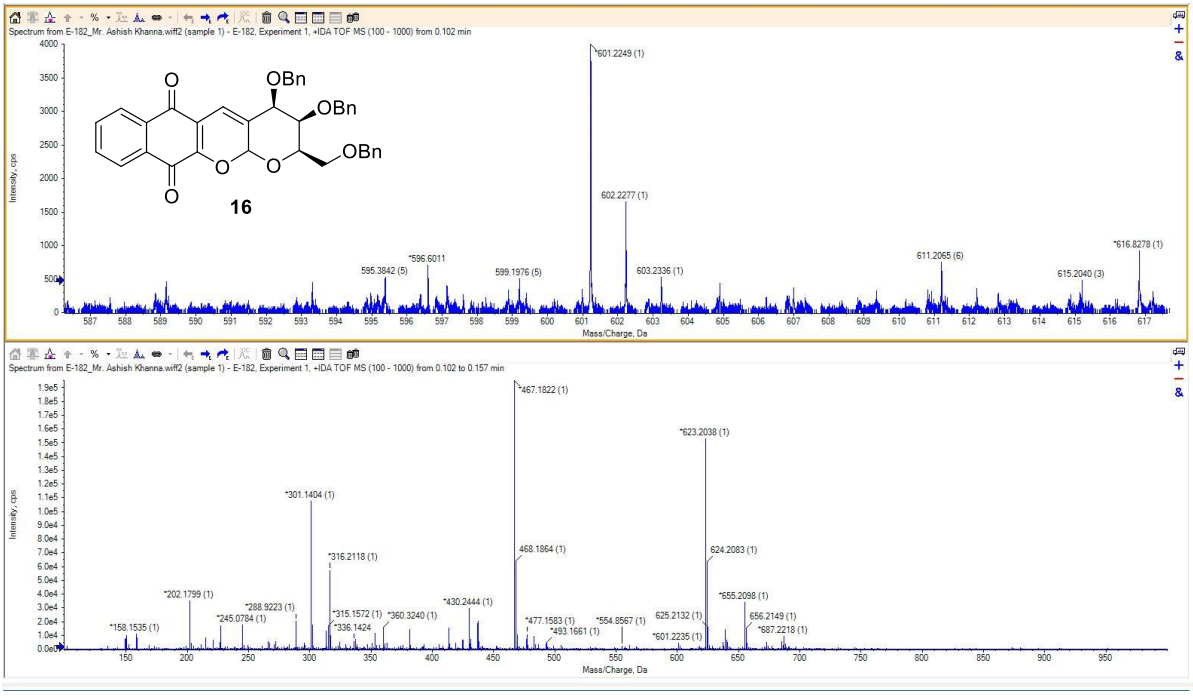


Copies of HRMS data for compound **15** and **16**

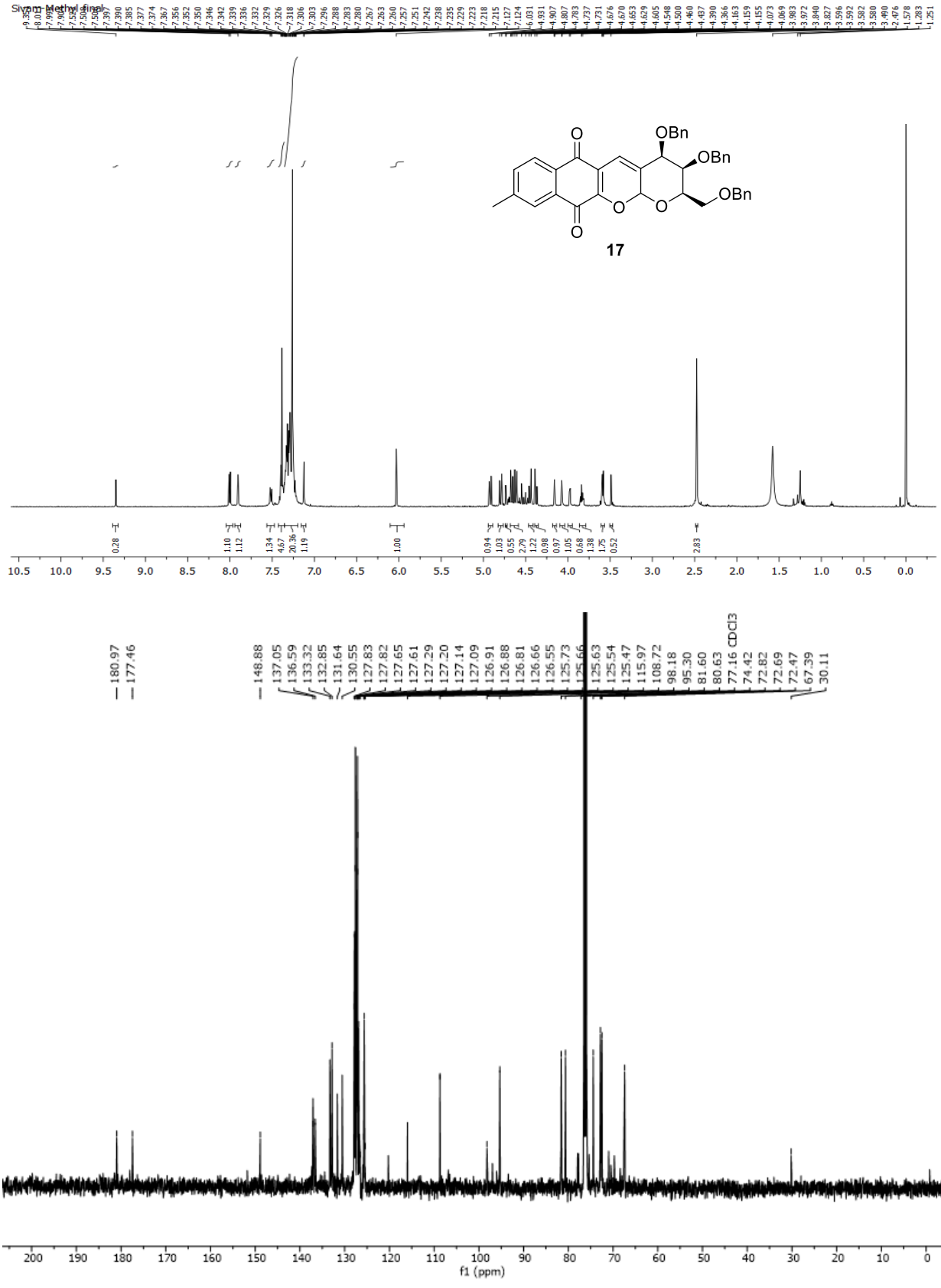
Compound 15: HRMS (ESI) *m/z*: Calcd. for $C_{38}H_{32}O_7$ $[M+Na]^+$ 623.2040, found 623.2028.



Compound 16: HRMS (ESI) *m/z*: Calcd. for $C_{38}H_{32}O_7$ $[M+Na]^+$ 623.2040, found 623.2038.



^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **17**



^{1}H NMR ($CDCl_3$, 500 MHz) and ^{13}C NMR ($CDCl_3$, 126 MHz) of **18**