Supporting Information

Exploiting spirooxindoles for dual DNA targeting/CDK2 inhibition and simultaneous mitigation of oxidative stress towards selective NSCLC therapy; Synthesis, evaluation, and molecular modelling studies

Mohammad Shahidul Islam[1], Refaah M. Al-Jassas[1], Abdullah Mohammed Al-Majid[1,*], Matti Haukka[2], Mohamed S. Nafie[3,4], Marwa M. Abu-Serie[5], Mohamed Teleb[6,*], Amira El-Yazbi[7], Abdul Majeed Abdullah Alayyaf[1], Assem Barakat[1,*] , and Marwa M. Shaaban[6]

[1] Department of Chemistry, College of Science, King Saud University, P. O. Box 2455, Riyadh 11451, Saudi Arabia. shahid.10amui@gmail.com (M.S.I.); 442203285@student.ksu.edu.sa (R.M.A.-J.); amajid@ksu.edu.sa (A.M.A.-M.); mayyaf@ksu.edu.sa (A.M.A.A.)

[2] Department of Chemistry, University of Jyväskylä, P.O. Box 35, FI-40014 Jyväskylä, Finland; matti.o.haukka@jyu.fi

[3] Department of Chemistry, College of Sciences, University of Sharjah, Sharjah (P.O. Box 27272), United Arab Emirates. mohamed.ElSayed@sharjah.ac.ae

[4] Chemistry Department, Faculty of Science, Suez Canal University, Ismailia, 41522, Egypt. Email: mohamed_nafie@science.suez.edu.eg (M.S.N.)

[5] Medical Biotechnology Department, Genetic Engineering and Biotechnology Research Institute, City of Scientific Research and Technological Applications (SRTA-City), Egypt. marwaelhedaia@gmail.com

[6] Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Alexandria University, Alexandria 21521, Egypt.

[7] Department of Pharmaceutical Analytical Chemistry, Faculty of Pharmacy, Alexandria University, Alexandria, Egypt.

* Correspondence: E-mails: mohamed.t.ismail@alexu.edu.eg (M.T.); and ambarakat@ksu.edu.sa (A.B.).

Table of content

1. Chemistry

1.1. Method and materials

All chemicals were purchased from Aldrich, Sigma-Aldrich and Fluka, which were used without further purification unless otherwise stated. All melting points were measured using a Gallenkamp melting point apparatus in open glass capillaries and were uncorrected. Crude products were purified by column chromatography on silica gel of 100–200 mesh. IR spectra were measured as KBr pellets using a Nicolet 6700 FT-IR spectrophotometer. The NMR spectra were recorded using a Varian Mercury Jeol-400 NMR spectrometer. ^{1}H NMR (400, 500 or 700MHz) and ^{13}C NMR (101, 126 or 176 MHz) spectroscopy were performed in deuterated dimethylsulfoxide (DMSO-d_6). Chemical shifts (d) are reported in terms of ppm and coupling constants J are given in Hz. High-resolution mass spectra were recorded on LC/Q-TOF, 6530 (Agilent Technologies, Santa Clara, CA, USA) using the electrospray ionization method.

1.2. Single crystal X-ray diffraction analysis

The crystal of **6f** was immersed in cryo-oil, mounted in a loop, and measured at a temperature of 120 K. The X-ray diffraction data was collected on a Rigaku Oxford Diffraction XtaLAB Synergy R diffractometer using Cu Kα radiation. The *CrysAlisPro* [1] software package was used for cell refinement and data reduction. A multi-scan absorption correction (*CrysAlisPro* [1]) was applied to the intensities before the structure solution. The structure was solved by the intrinsic phasing (*SHELXT* [2]) method. Structural refinement was carried out using *SHELXL* [3] software with *SHELXLE* [4] graphical user interface. The NH hydrogen atom, involved in hydrogen bonding, was located from the difference Fourier map and refined isotropically. All other hydrogen atoms were positioned geometrically and constrained to ride on their parent atoms, with C-H = 0.95 -1.00 Å and U_{iso} = 1.2-1.5·U_{eq} (parent atom). The crystallographic details are summarized in Table X1.

Table S1. Crystal Data.

	6f
CCDC	2306410
empirical formula	$C_{41}H_{37}N_5O_5$
fw	679.75
temp (K)	120(2)
λ(Å)	1.54184
cryst syst	*Orthorhombic*
space group	$P2_12_12_1$
a (Å)	11.1512(4)
b (Å)	11.4201(5)
c (Å)	27.4907(13)
V (Å^3)	3500.9(3)
Z	4
ρ_{calc} (Mg/m^3)	1.290
μ(Mo Kα) (mm^{-1})	0.696
No. reflns.	39093
Unique reflns.	7536
Completeness to θ=67.684°	99.9%

GOOF (F^2)		1.052
R_{int}		0.1265
Absolute parameter	structure	0.12(17)
R_1[a] ($I \geq 2\sigma$)		0.0552
wR_2[b] ($I \geq 2\sigma$)		0.1225

[a] $R_1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$. [b] $wR_2 = \{\Sigma[w(F_o^2 - F_c^2)^2]/ \Sigma[w(F_o^2)^2]\}^{1/2}$

Table S2. Selected bond lengths [Å] and angles [°] for **6f**.

O(1)-C(26)	1.206(5)
O(2)-C(30)	1.375(6)
O(2)-C(31)	1.410(6)
O(3)-C(35)	1.222(5)
O(4)-N(5)	1.203(6)
O(5)-N(5)	1.227(6)
N(1)-N(2)	1.359(5)
N(1)-C(7)	1.363(5)
N(1)-C(6)	1.417(6)
N(2)-C(9)	1.345(5)
N(3)-C(24)	1.468(5)
N(3)-C(25)	1.475(5)
N(3)-C(17)	1.475(5)
N(4)-C(26)	1.362(6)
N(4)-C(27)	1.406(6)
N(5)-C(39)	1.467(6)

Symmetry transformations used to generate equivalent atoms:

Table S3. Hydrogen bonds for **6f** [Å and °].

D-H...A	***d(D-H)***	***d(H...A)***	***d(D...A)***	***<(DHA)***
C(5)-H(5)...O(1)#1	0.95	2.55	3.376(6)	146.0
C(38)-H(38)...O(1)#2	0.95	2.59	3.239(6)	126.0
C(40)-H(40)...O(3)#3	0.95	2.59	3.323(6)	133.7
N(4)-H(4)...N(2)#4	0.89(6)	2.27(6)	3.148(5)	170(5)

Symmetry transformations used to generate equivalent atoms:
#1 x+1,y,z #2 -x+1,y-1/2,-z+3/2 #3 -x+1,y+1/2,-z+3/2
#4 x-1,y,z

1.3. NMR spectra

1.3.1. NMR spectra of chalcones 3a-d

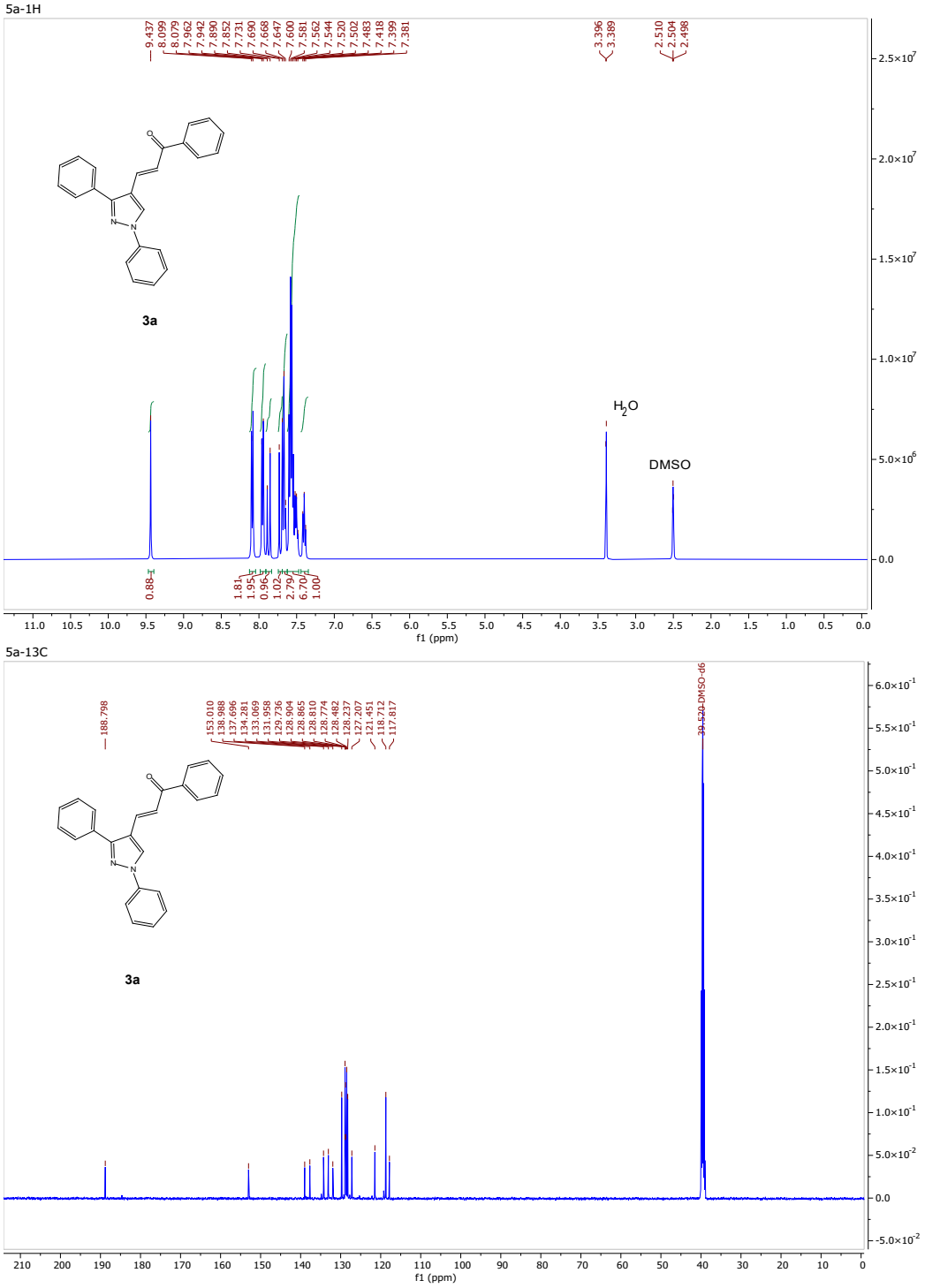


Figure S1: ^{1}H-NMR and ^{13}C-NMR of chalcone-**3a**

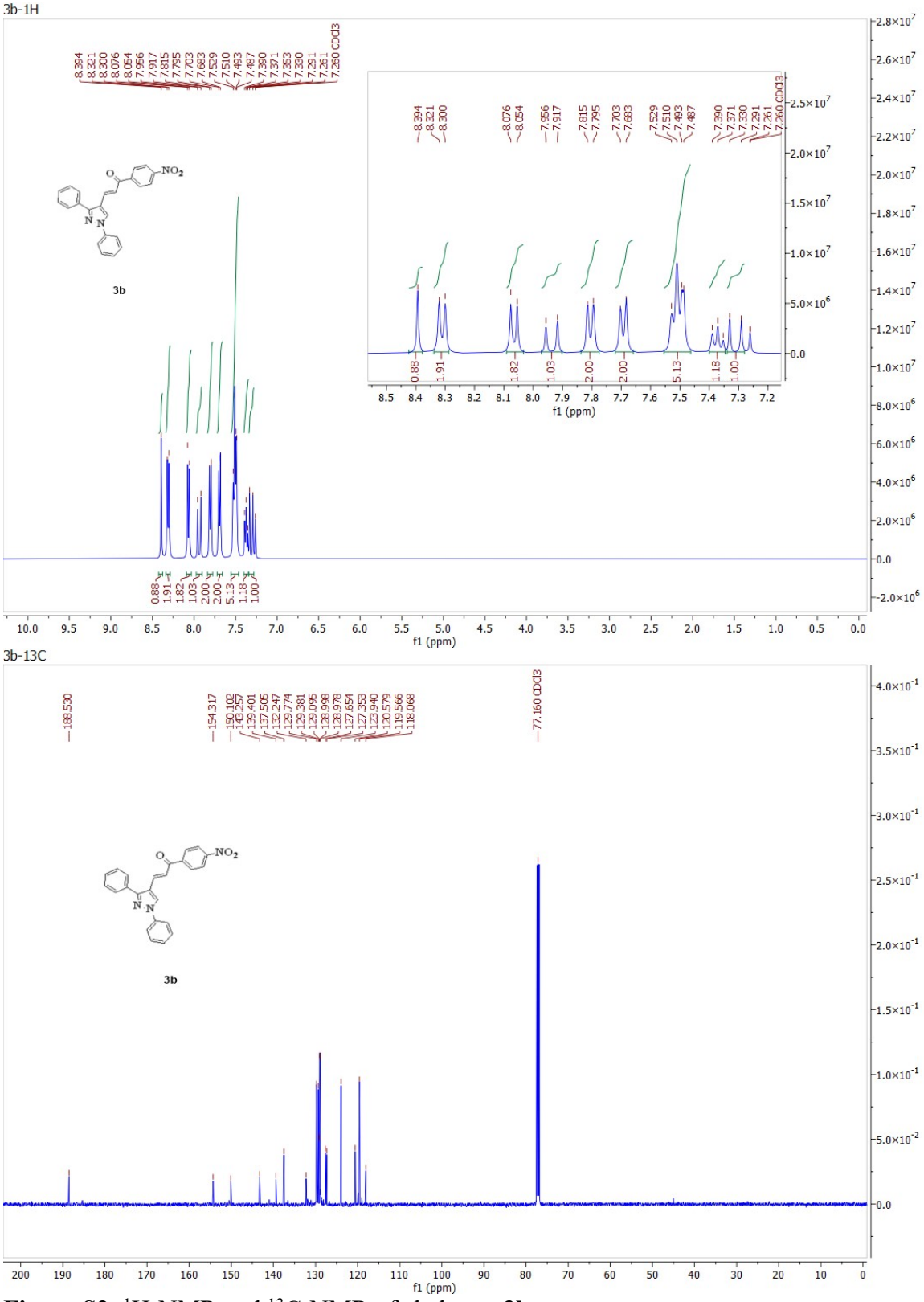


Figure S2: ^{1}H-NMR and ^{13}C-NMR of chalcone-**3b**

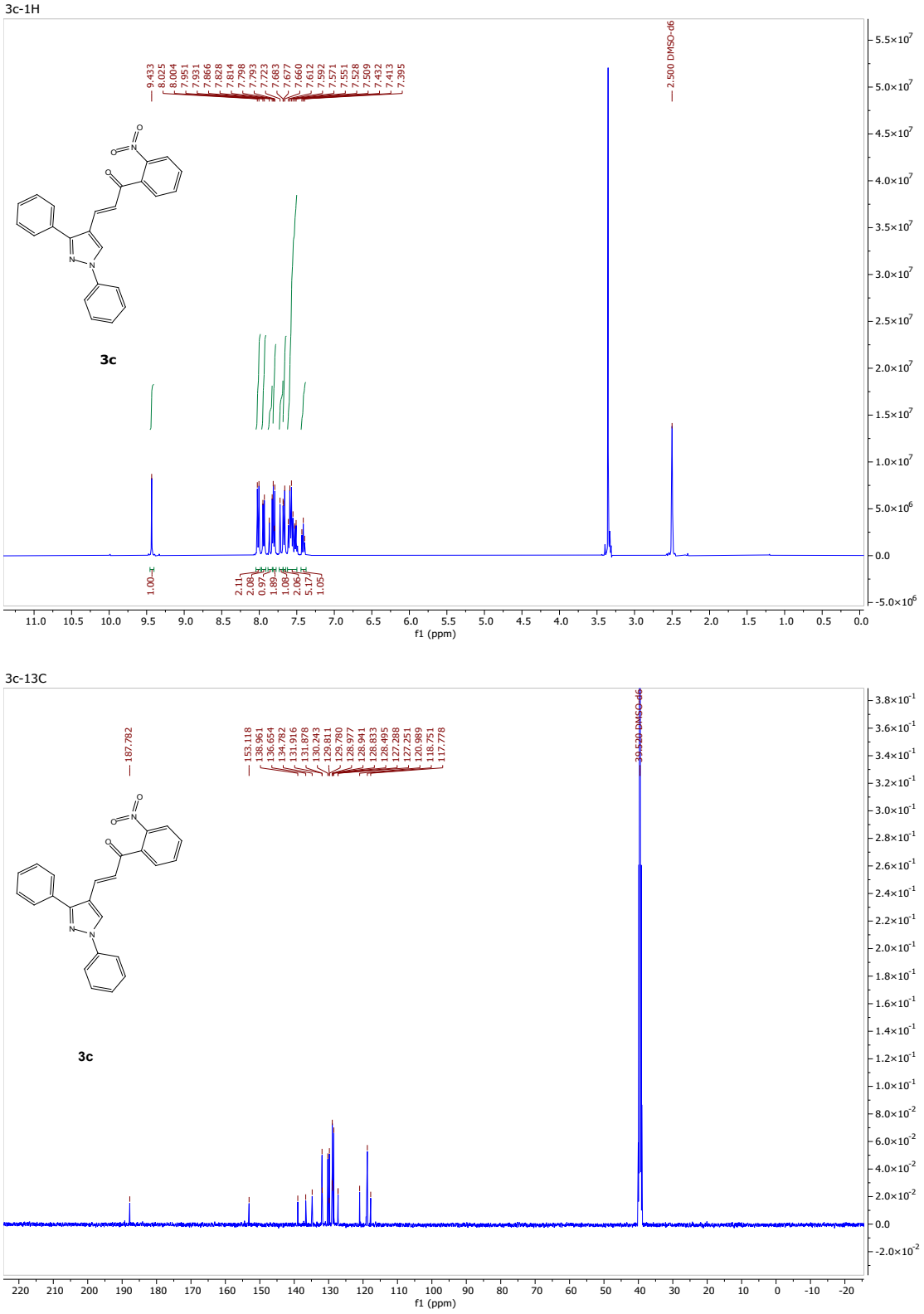


Figure S3: ^{1}H-NMR and ^{13}C-NMR of chalcone-**3c**

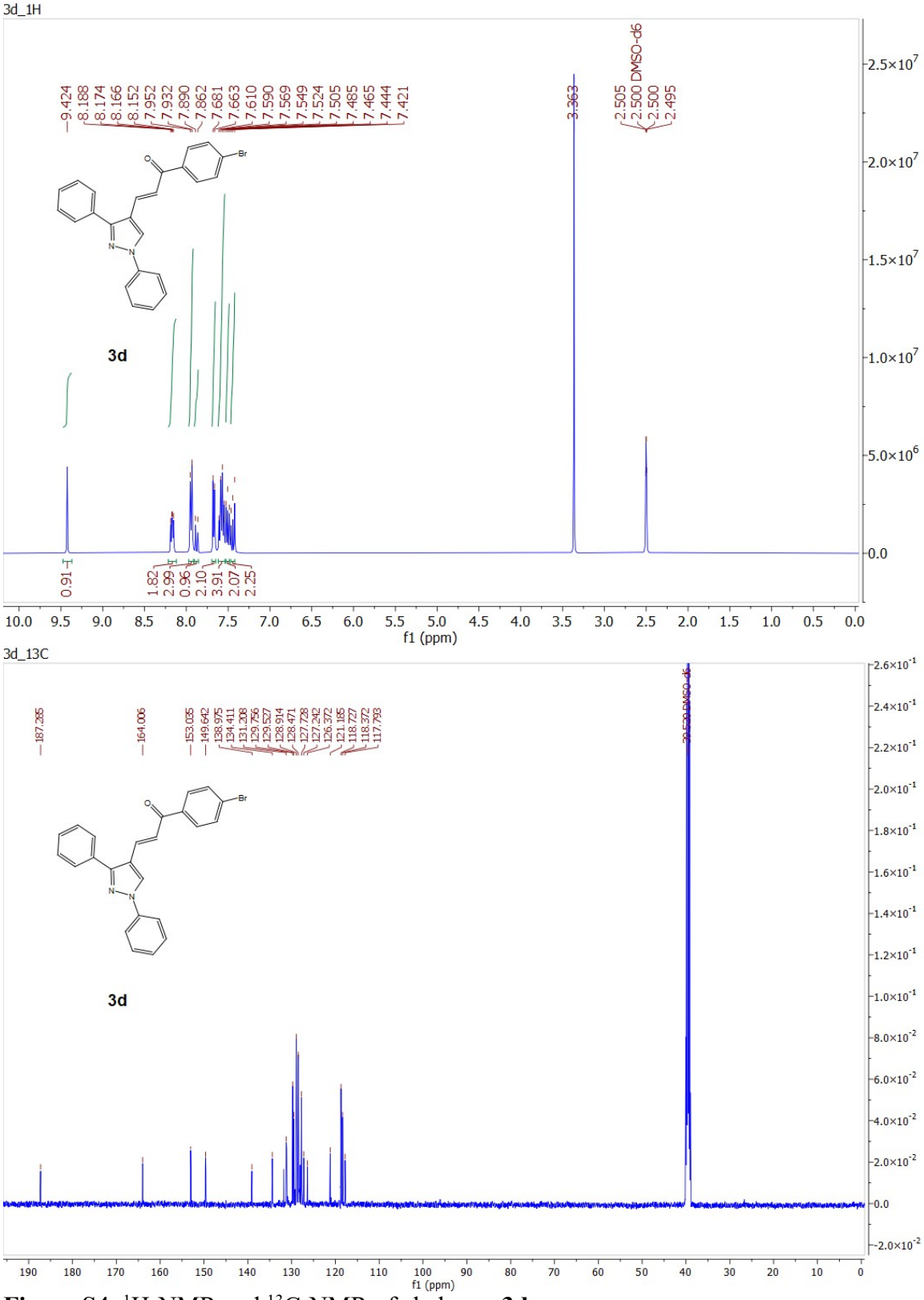


Figure S4: ^{1}H-NMR and ^{13}C-NMR of chalcone-**3d**

1.3.2. NMR spectra of spiro compounds 6a-p

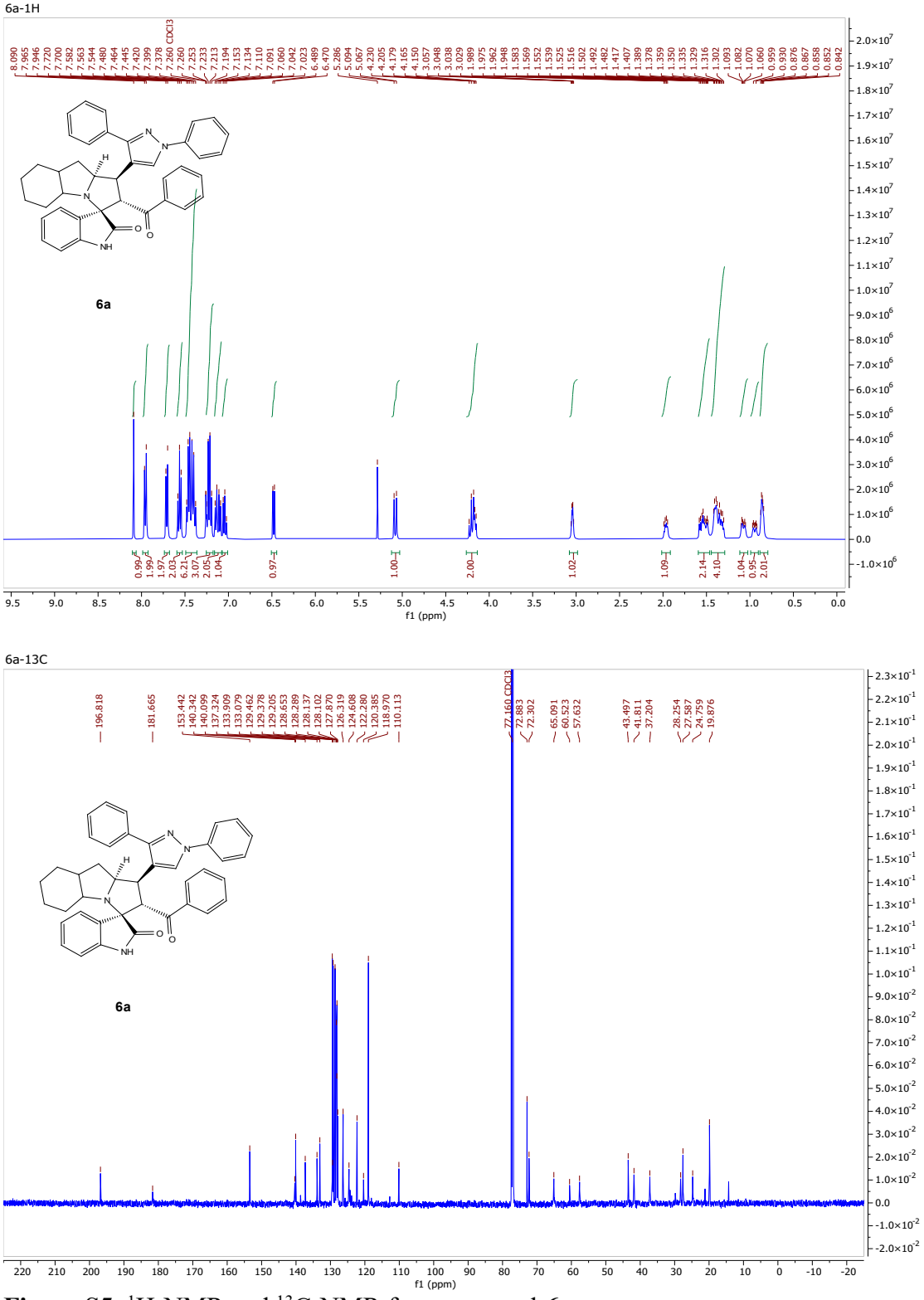


Figure S5: ^{1}H-NMR and ^{13}C-NMR for compound-**6a**

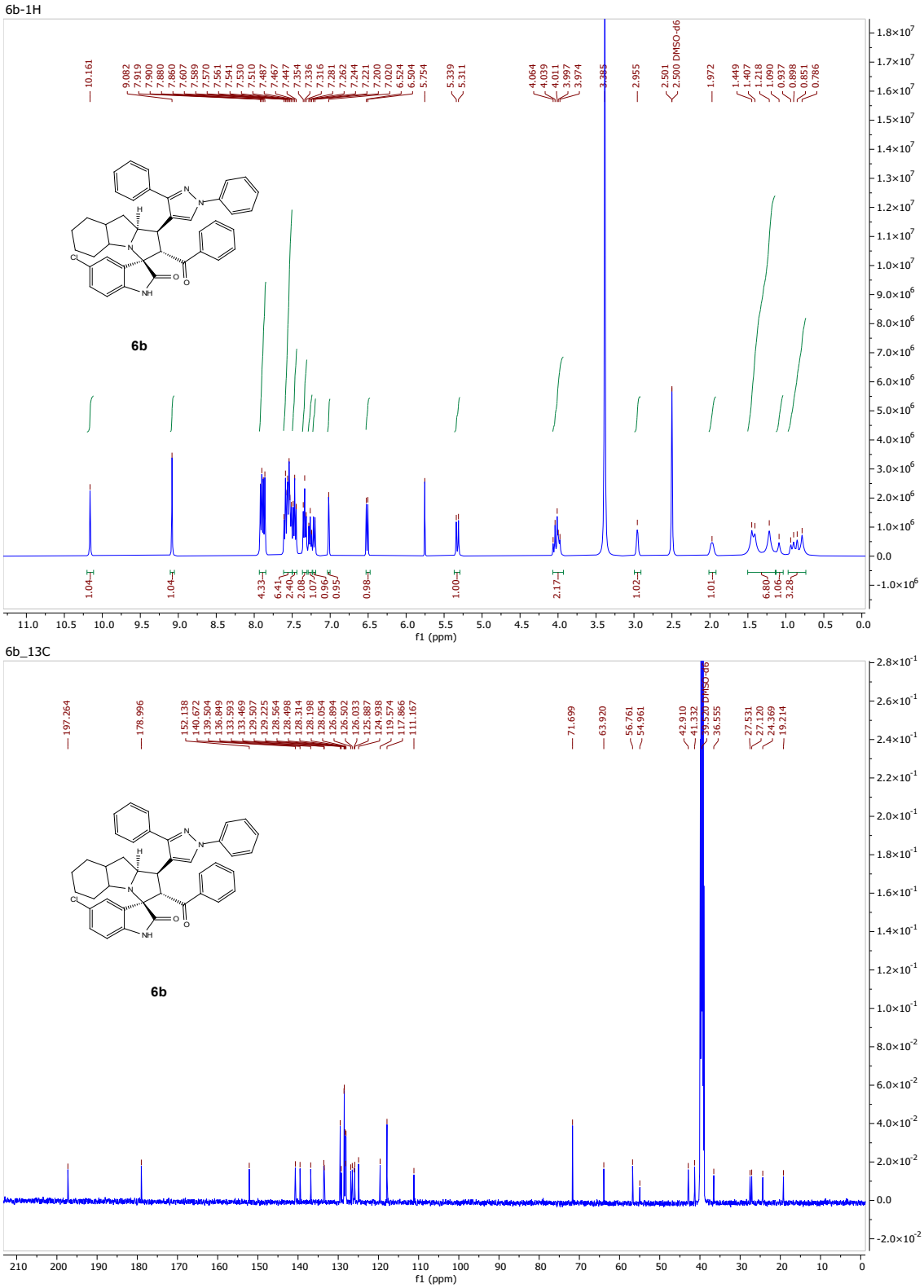


Figure S6: ^{1}H-NMR and ^{13}C-NMR for compound-**6b**

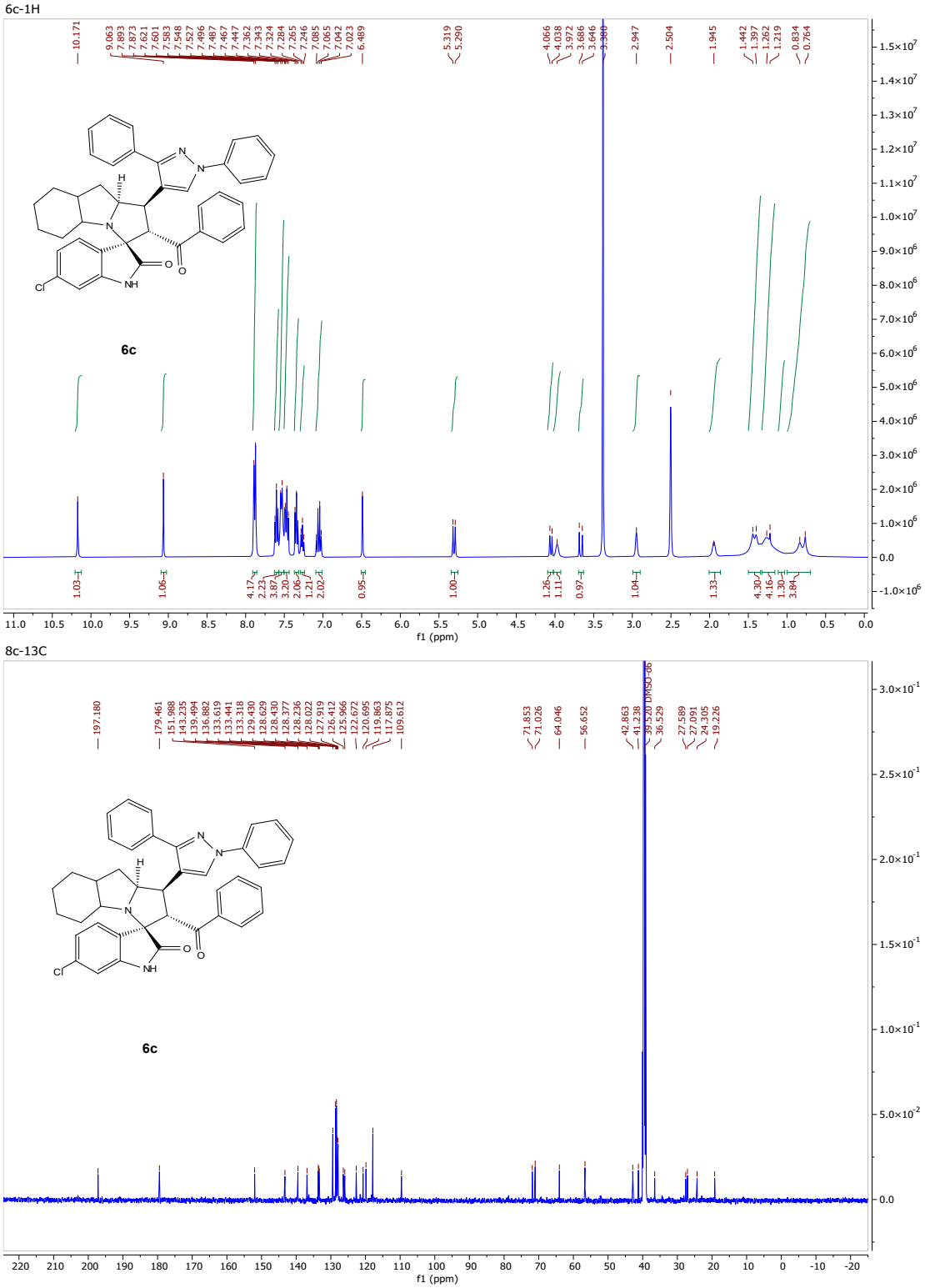


Figure S7: ^{1}H-NMR and ^{13}C-NMR for compound-**6c**

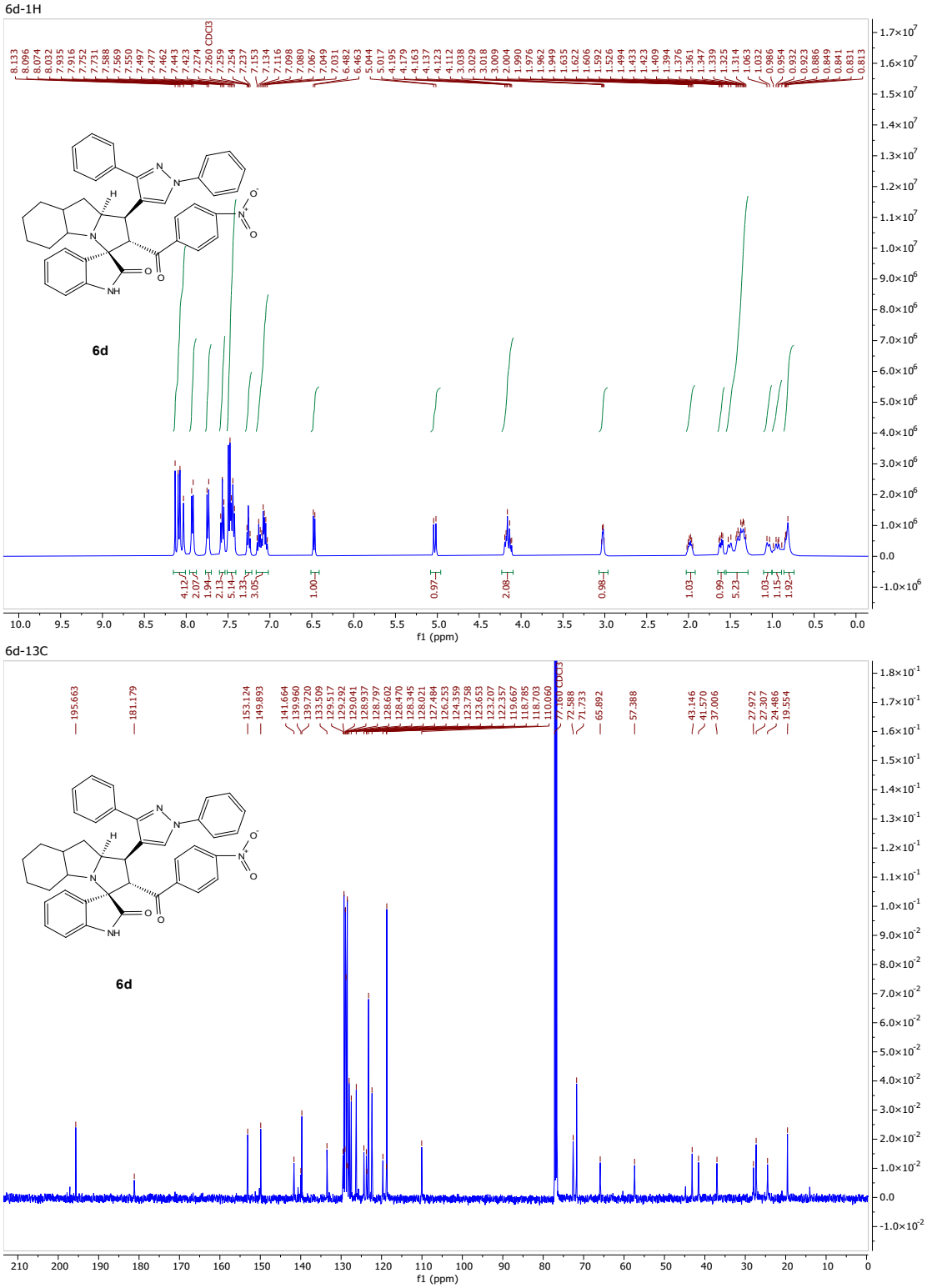


Figure S8: ^{1}H-NMR and ^{13}C-NMR for compound-**6d**

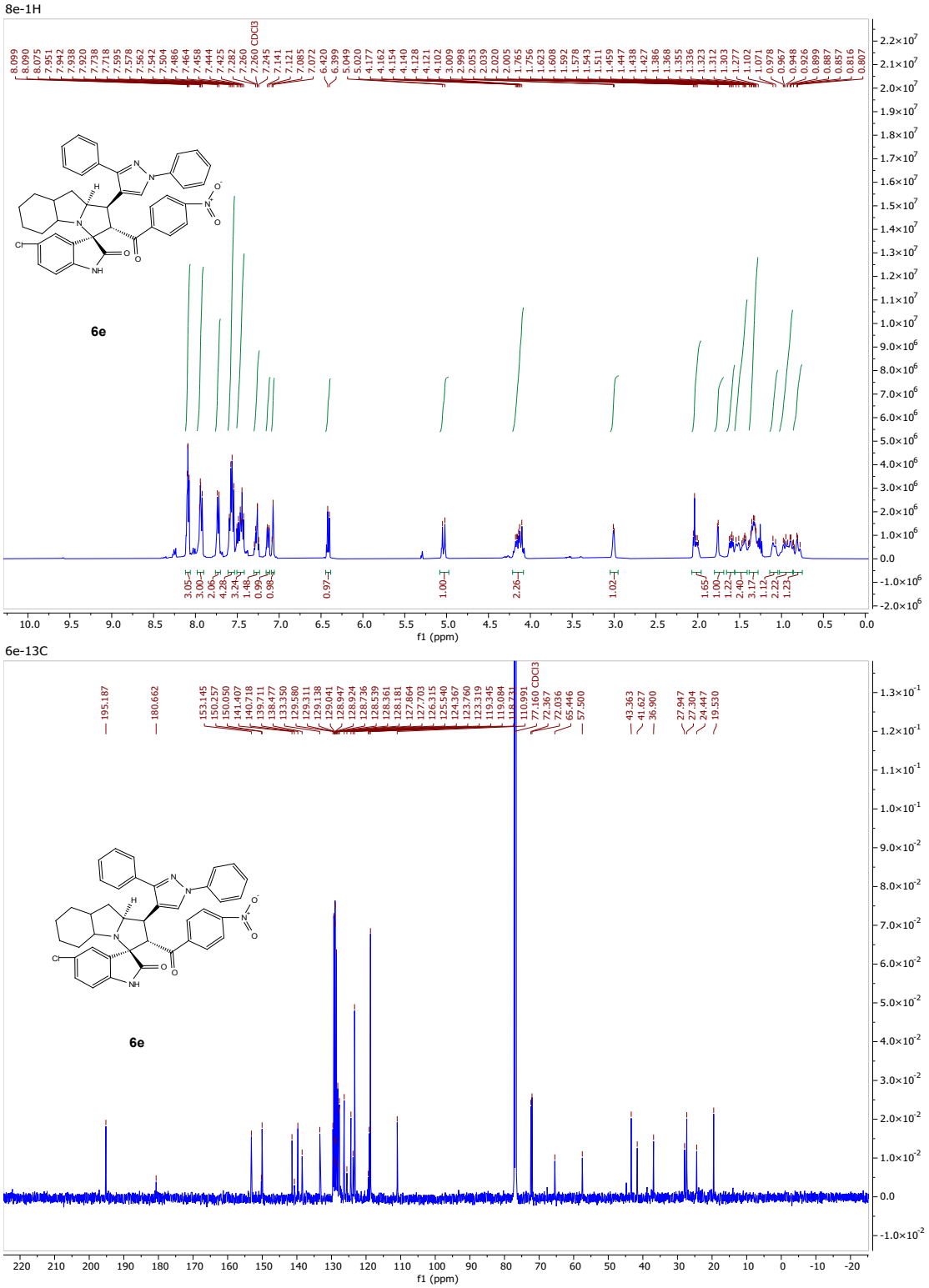


Figure S9: ^{1}H-NMR and ^{13}C-NMR for compound-**6e**

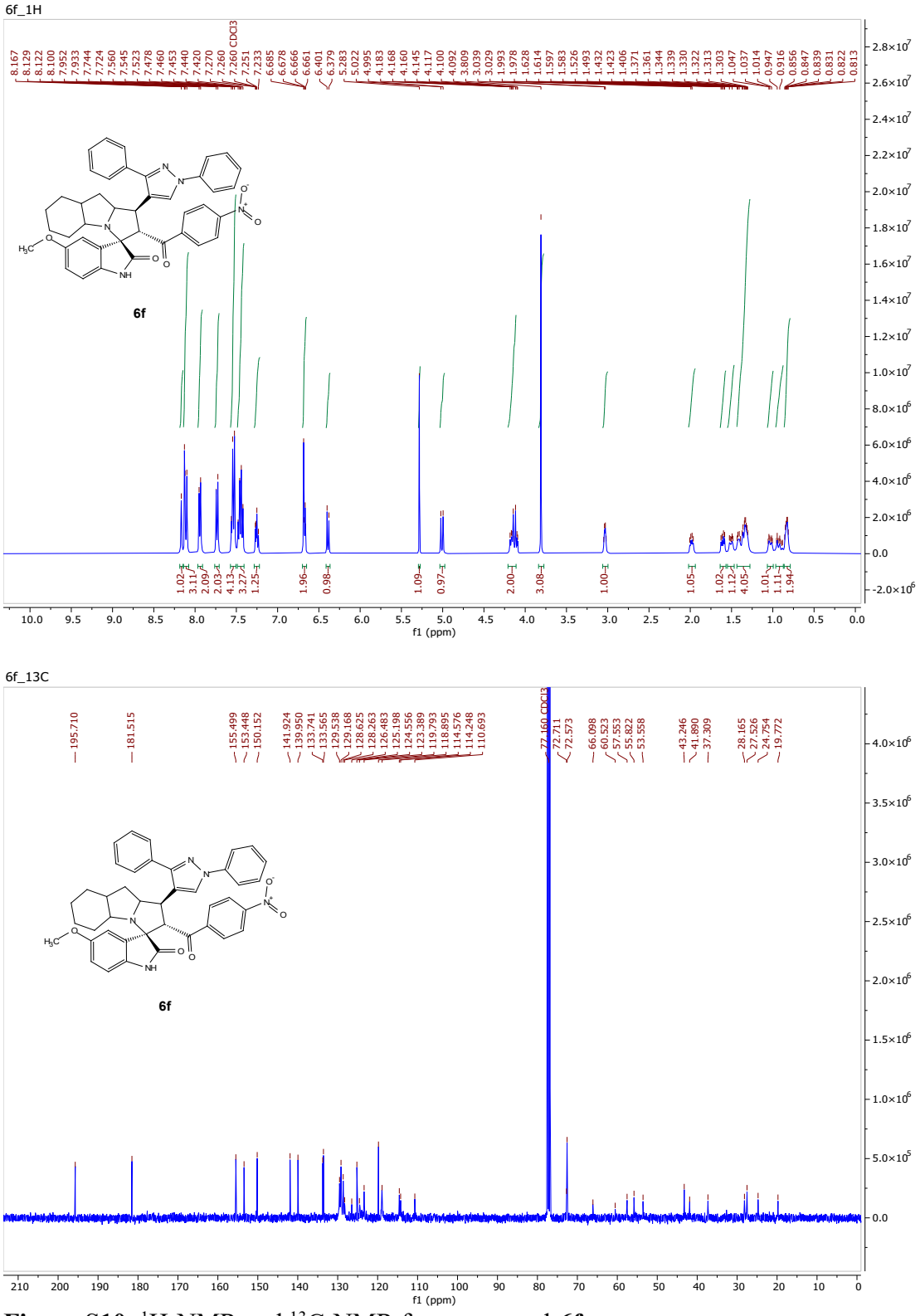


Figure S10: ^{1}H-NMR and ^{13}C-NMR for compound-**6f**

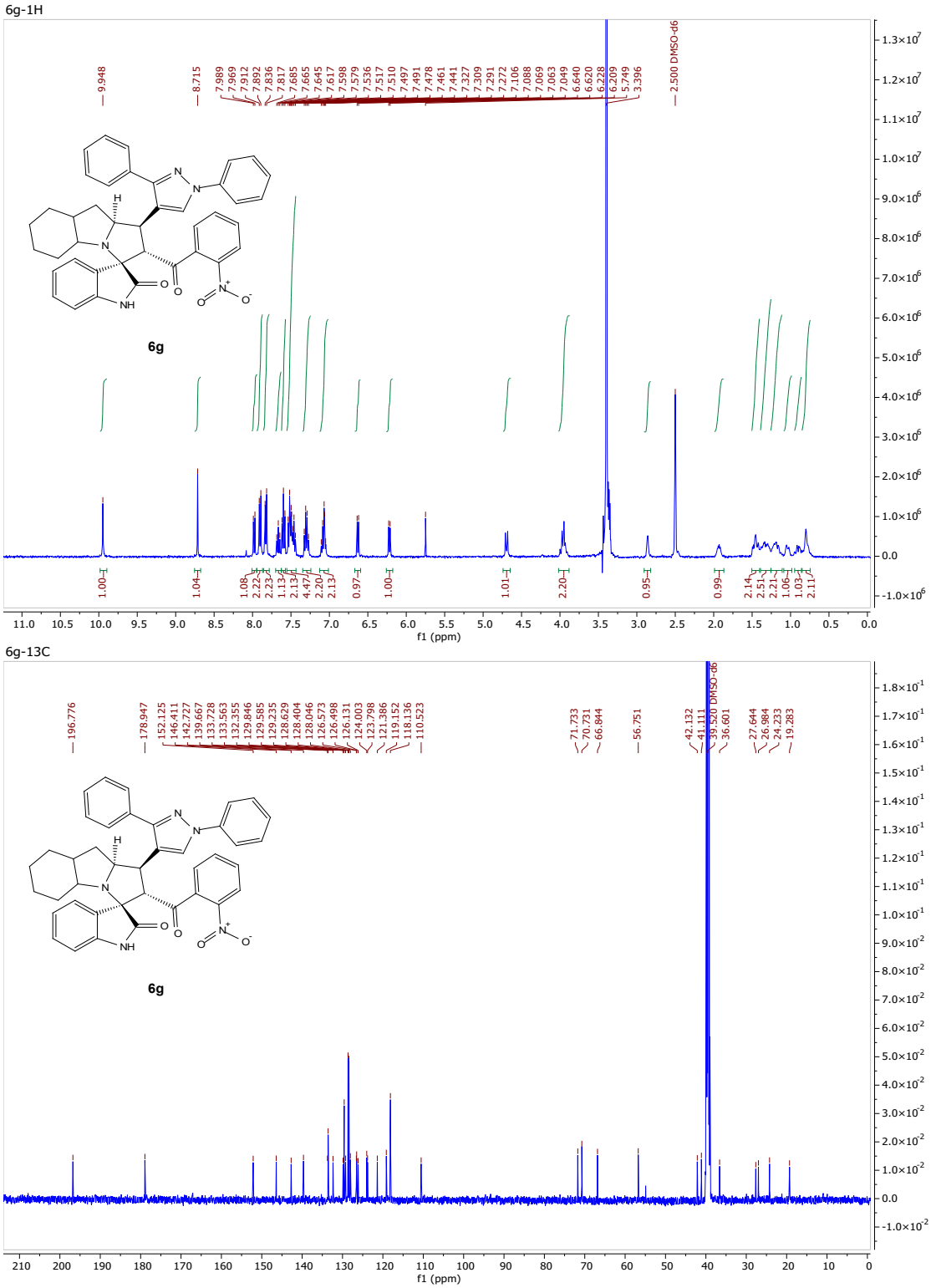


Figure S11: ^{1}H-NMR and ^{13}C-NMR for compound-**6g**

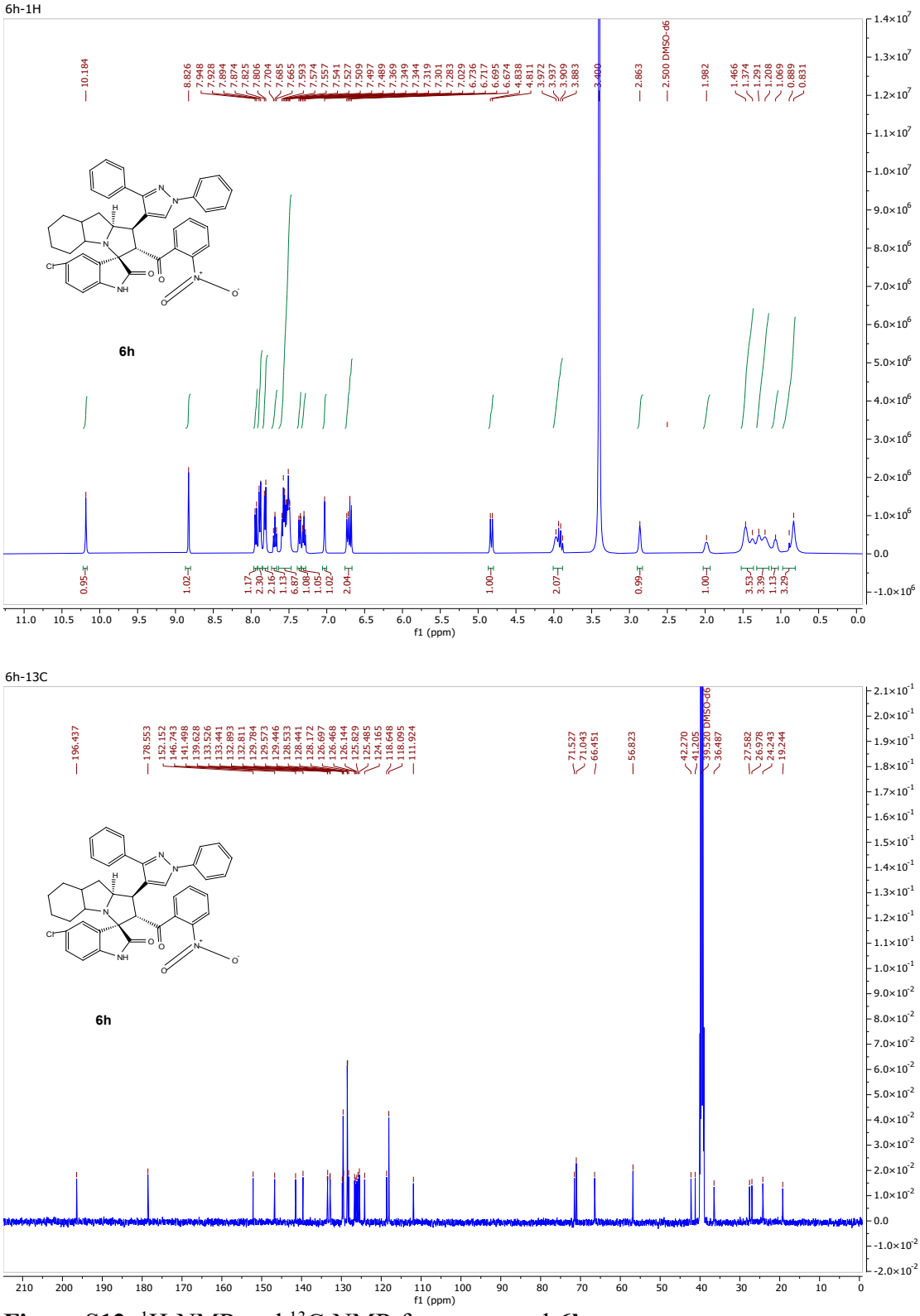


Figure S12: ^{1}H-NMR and ^{13}C-NMR for compound-**6h**

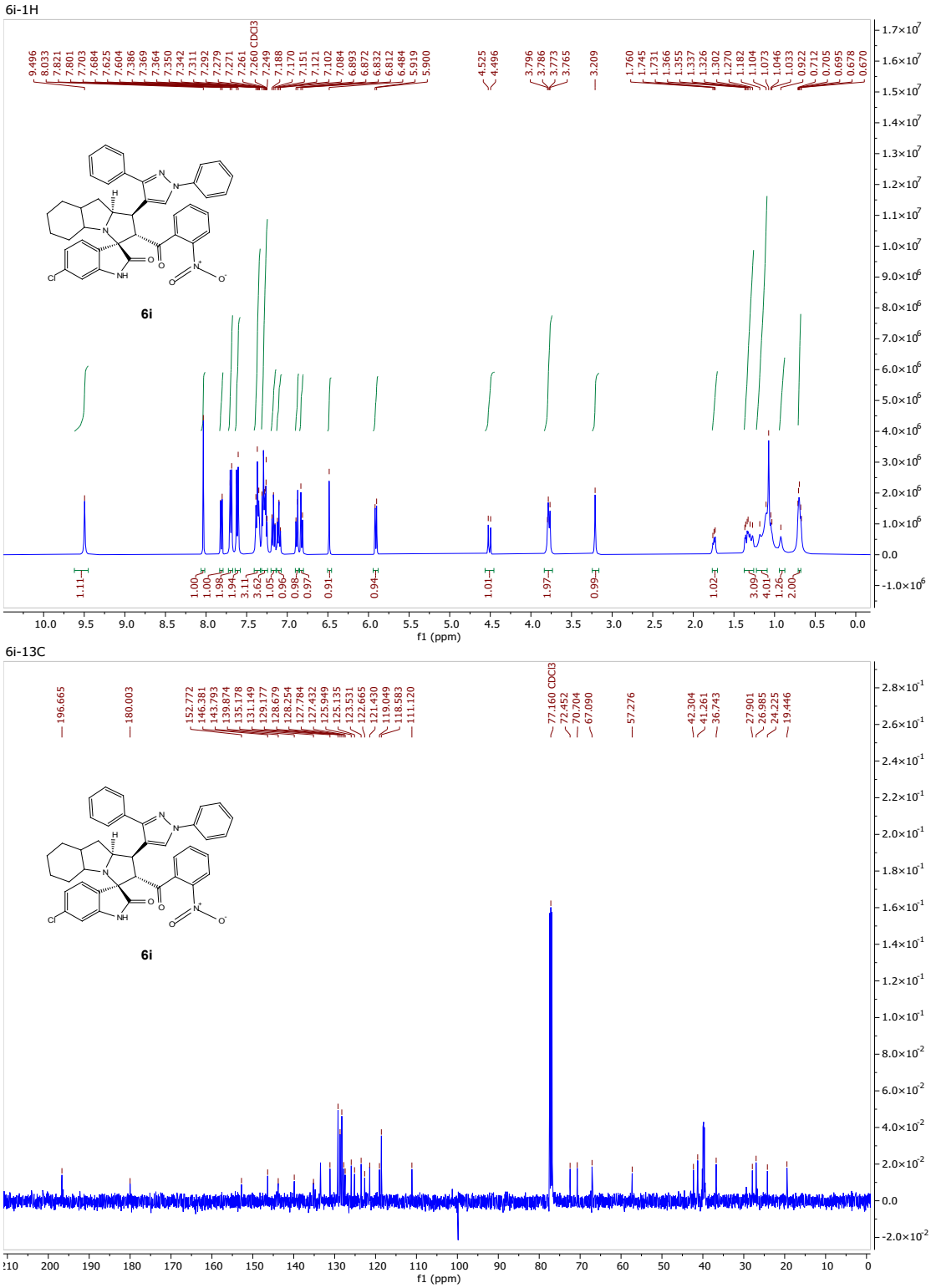


Figure S13: ^{1}H-NMR and ^{13}C-NMR for compound-**6i**

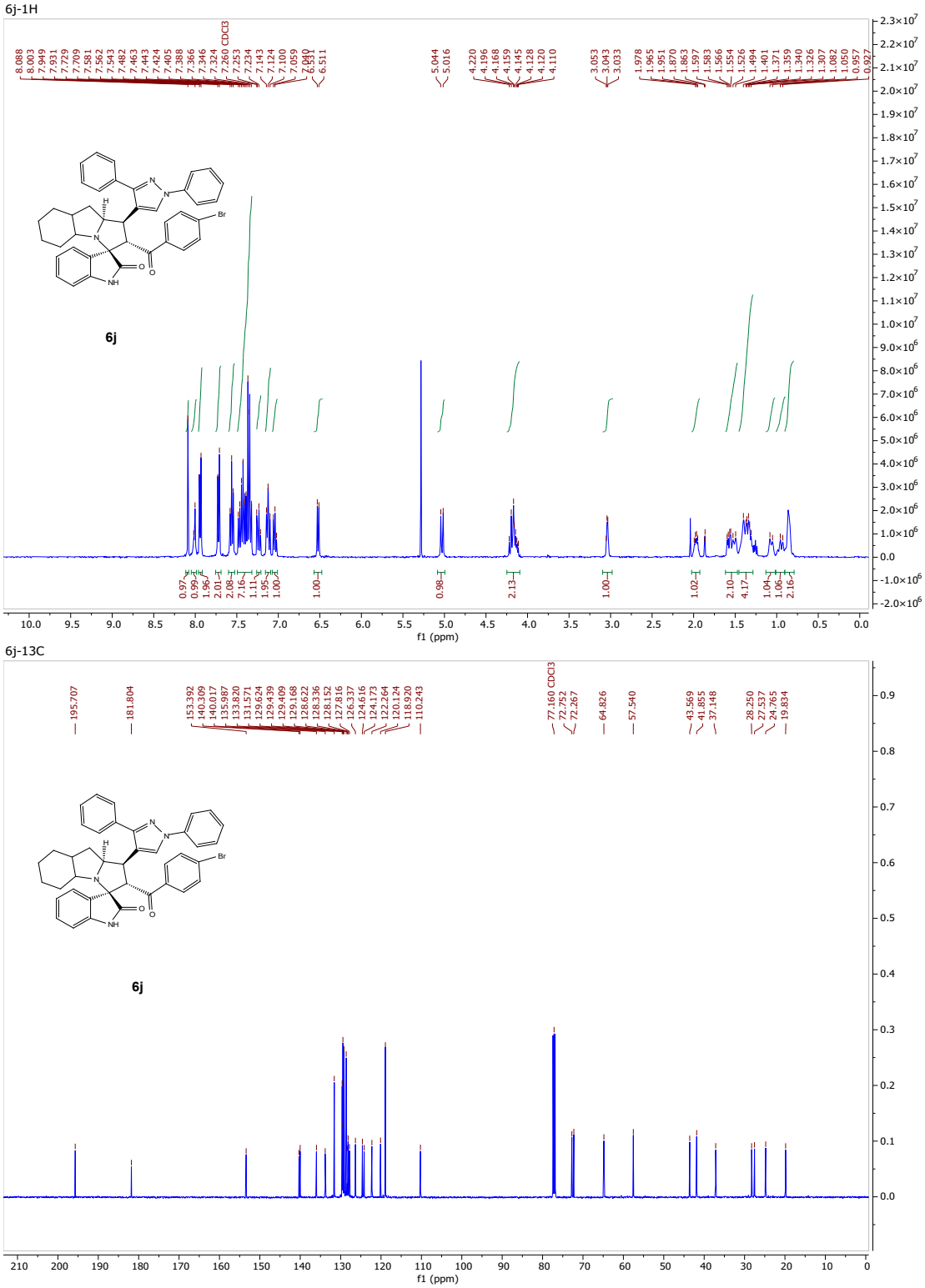


Figure S14: ^{1}H-NMR and ^{13}C-NMR for compound-**6j**

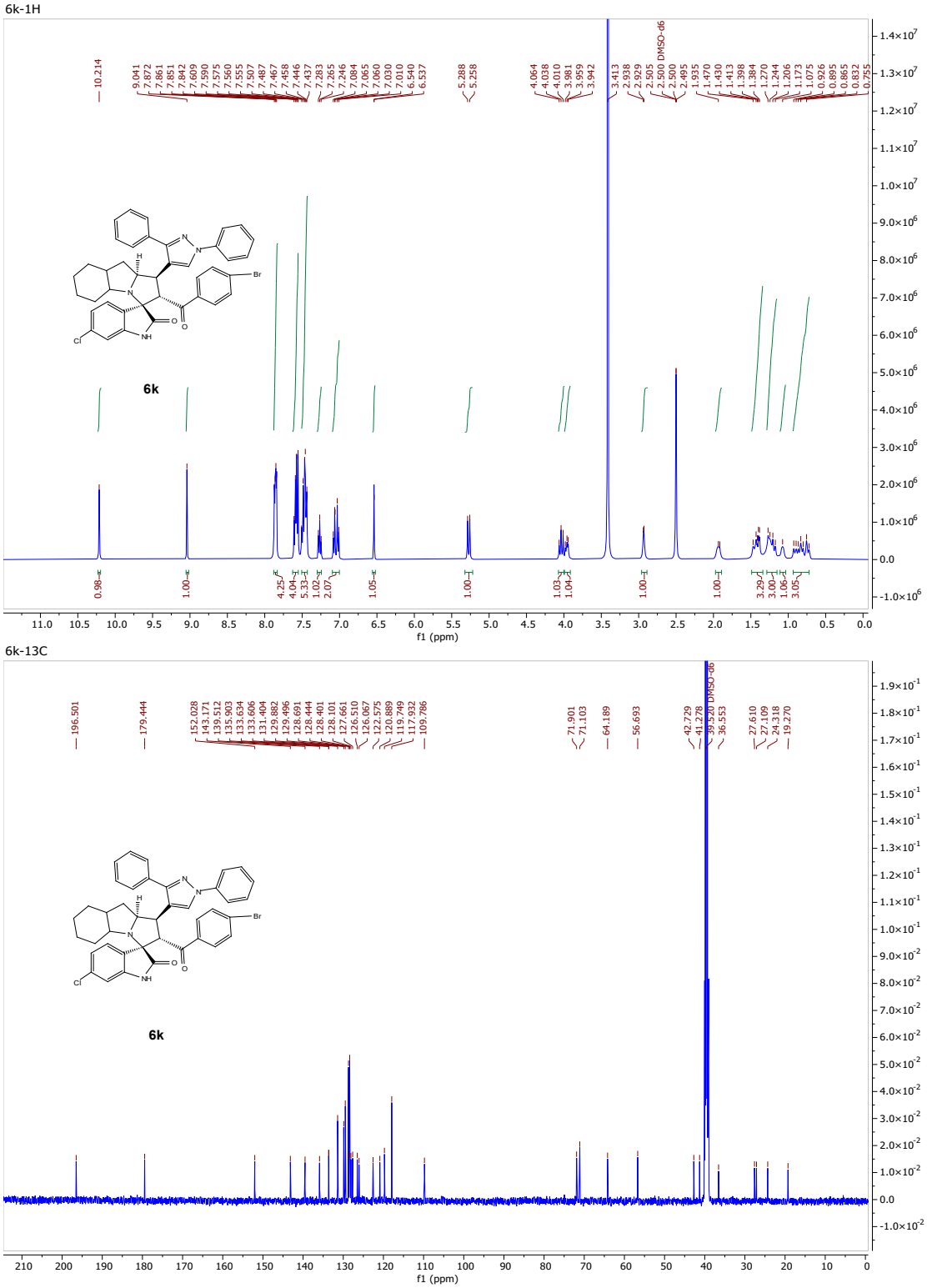


Figure S15: ^{1}H-NMR and ^{13}C-NMR for compound-**6k**

1.3.3. NMR spectra of spiro compounds 8a-e

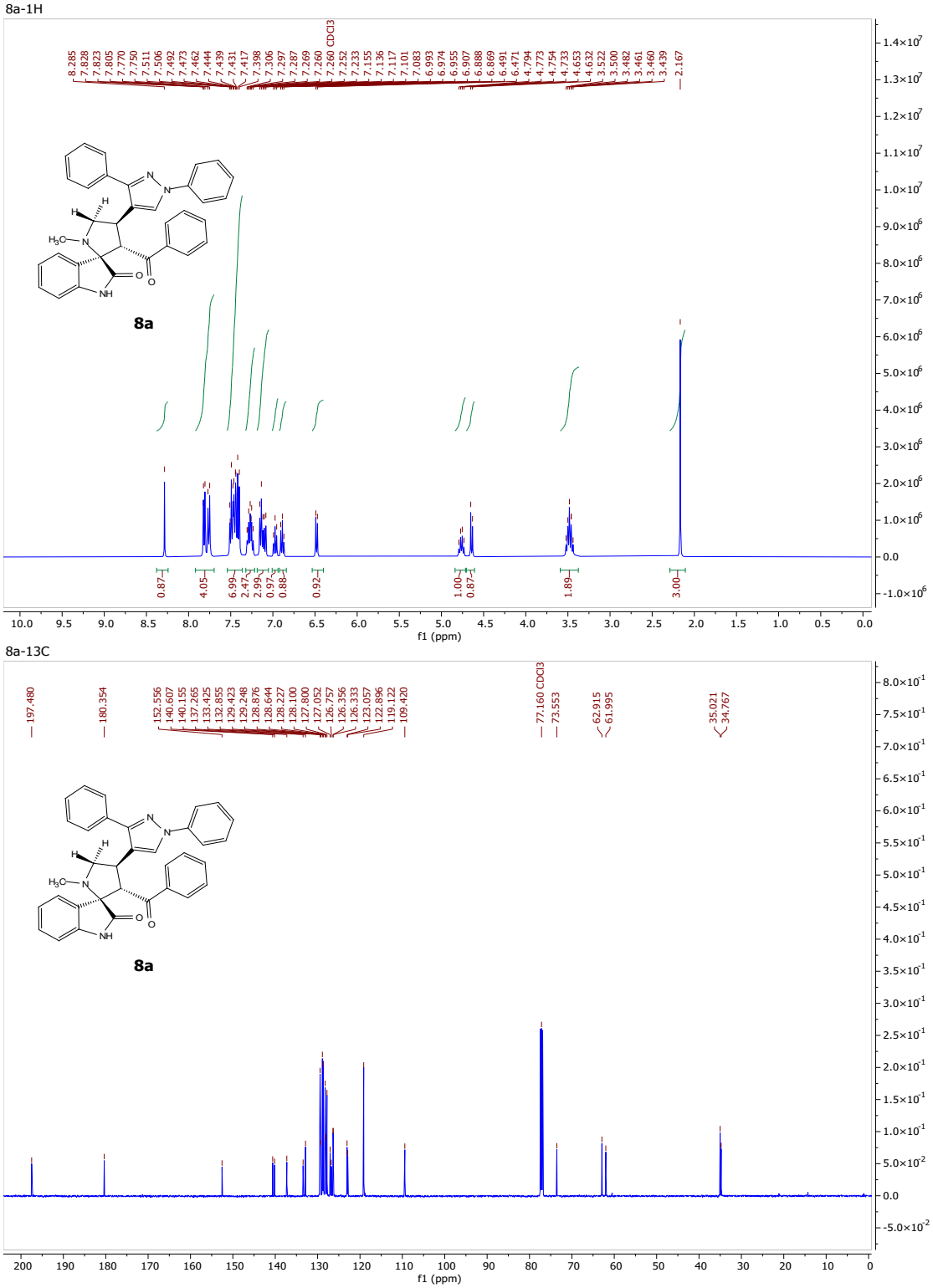


Figure S16: ^{1}H-NMR and ^{13}C-NMR for compound-**8a**

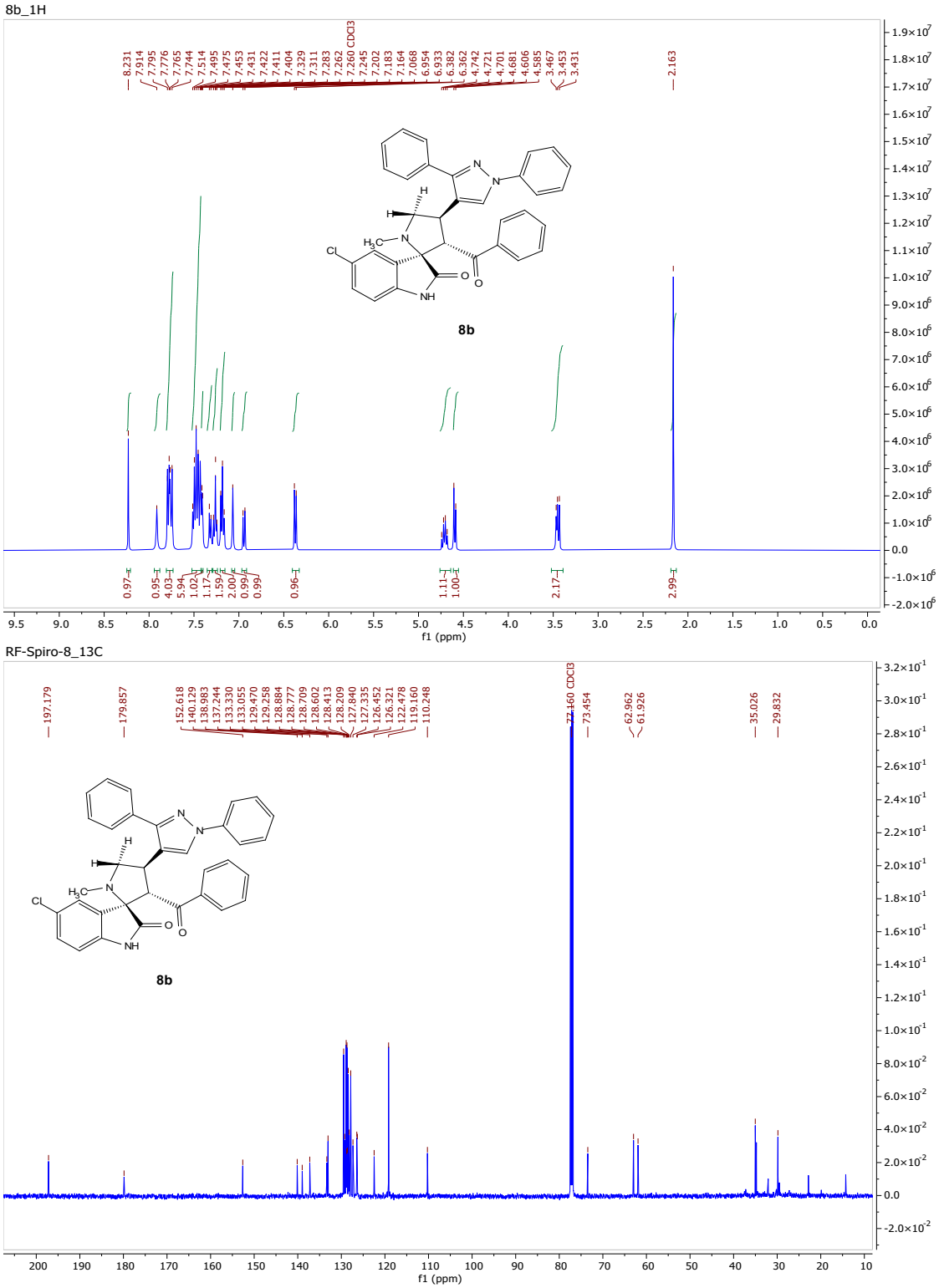


Figure S17: ^{1}H-NMR and ^{13}C-NMR for compound-**8b**

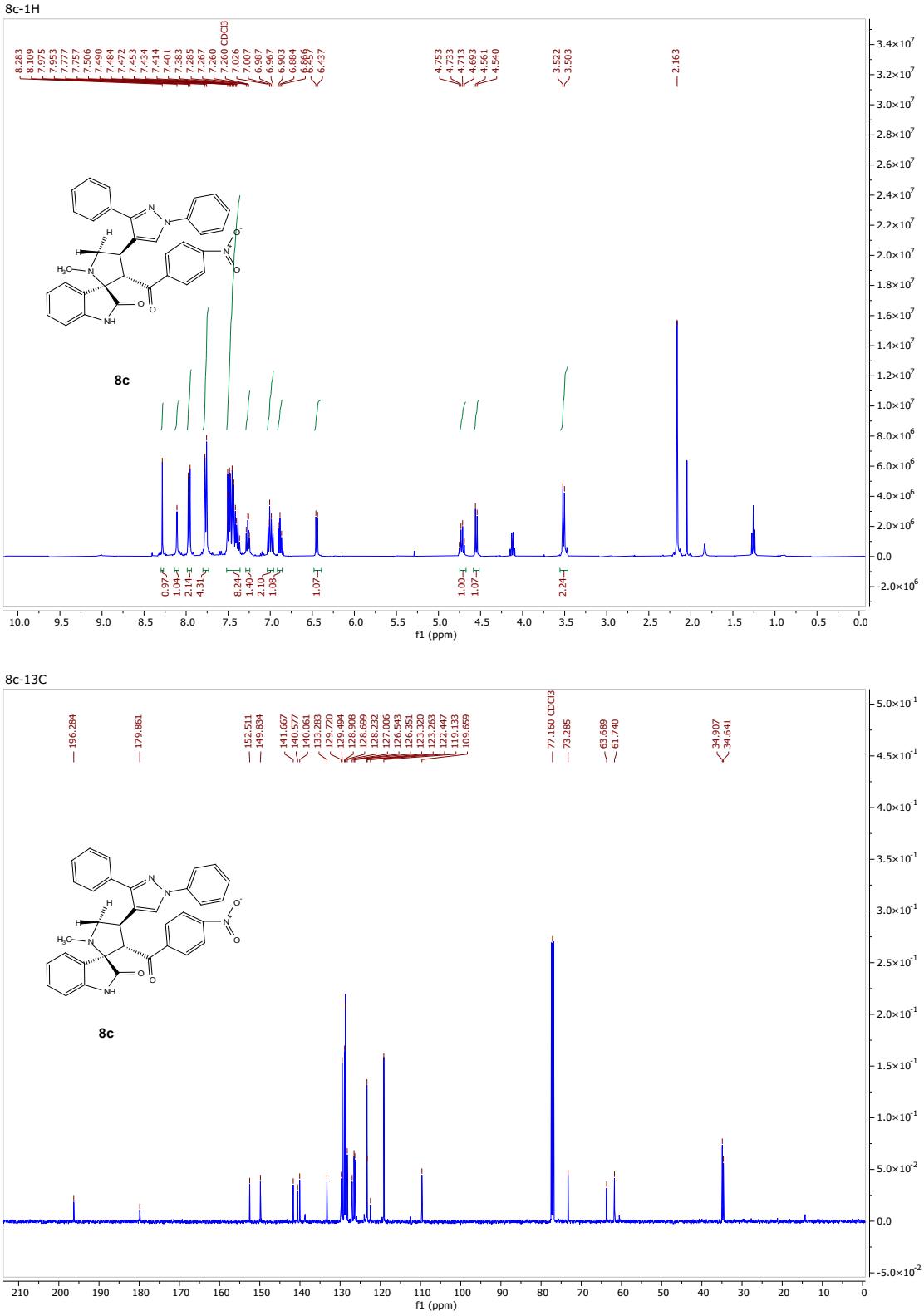


Figure S18: ^{1}H-NMR and ^{13}C-NMR for compound-**8c**

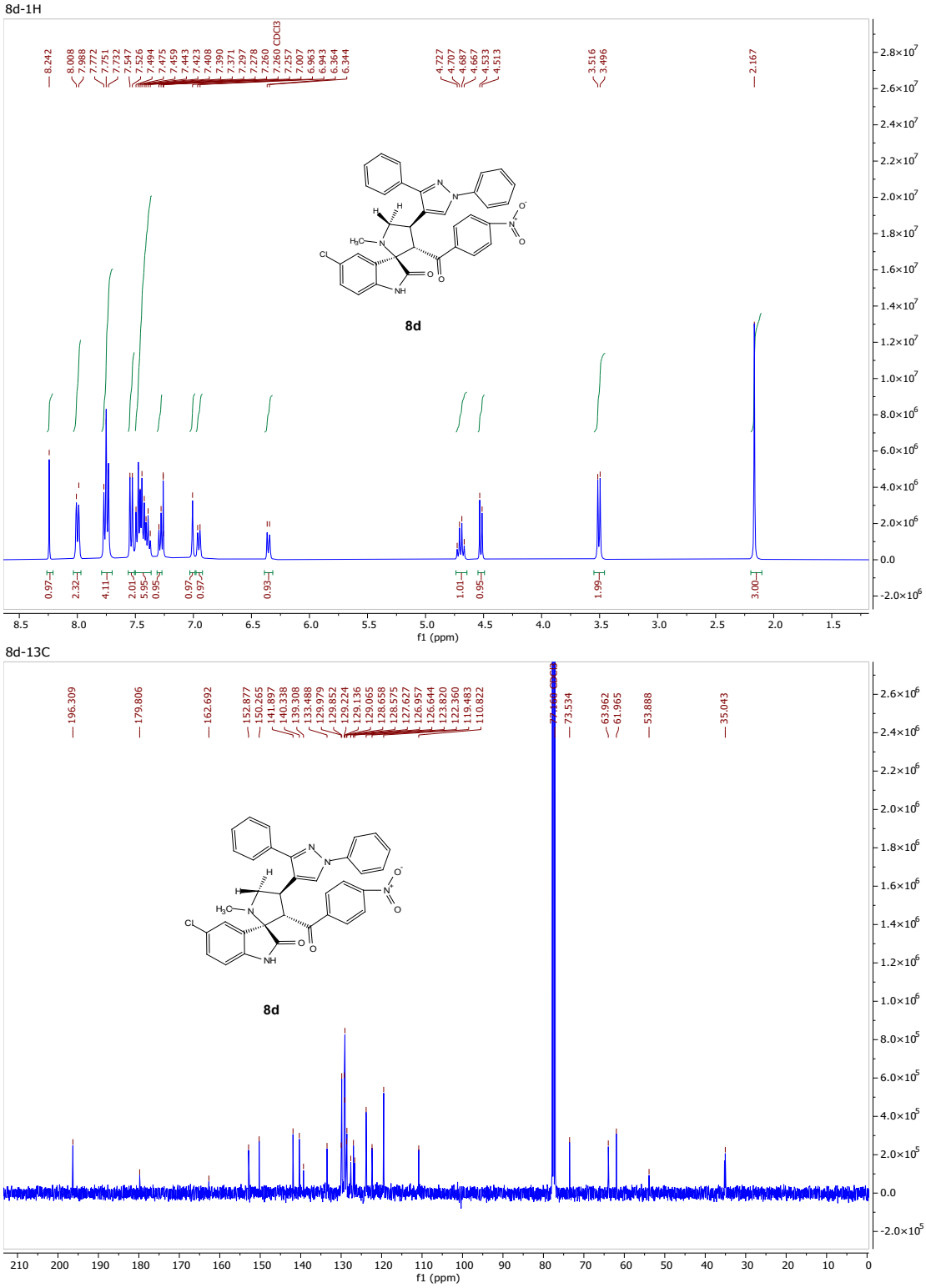


Figure S19: ^{1}H-NMR and ^{13}C-NMR for compound-**8d**

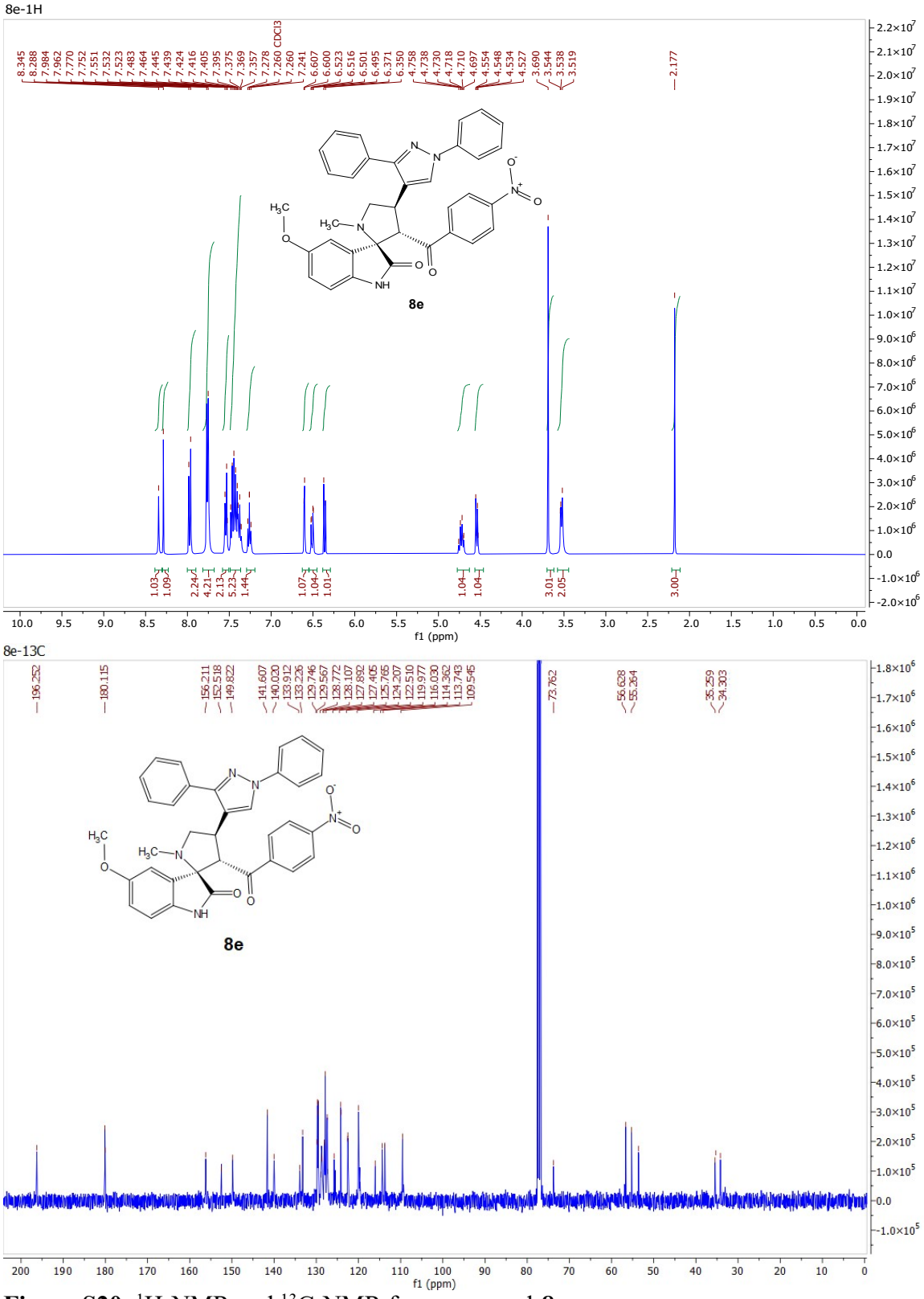


Figure S20: ^{1}H-NMR and ^{13}C-NMR for compound-**8e**

1.4. IR spectra

1.4.1. IR-spectroscopy of chalcones -3a-d

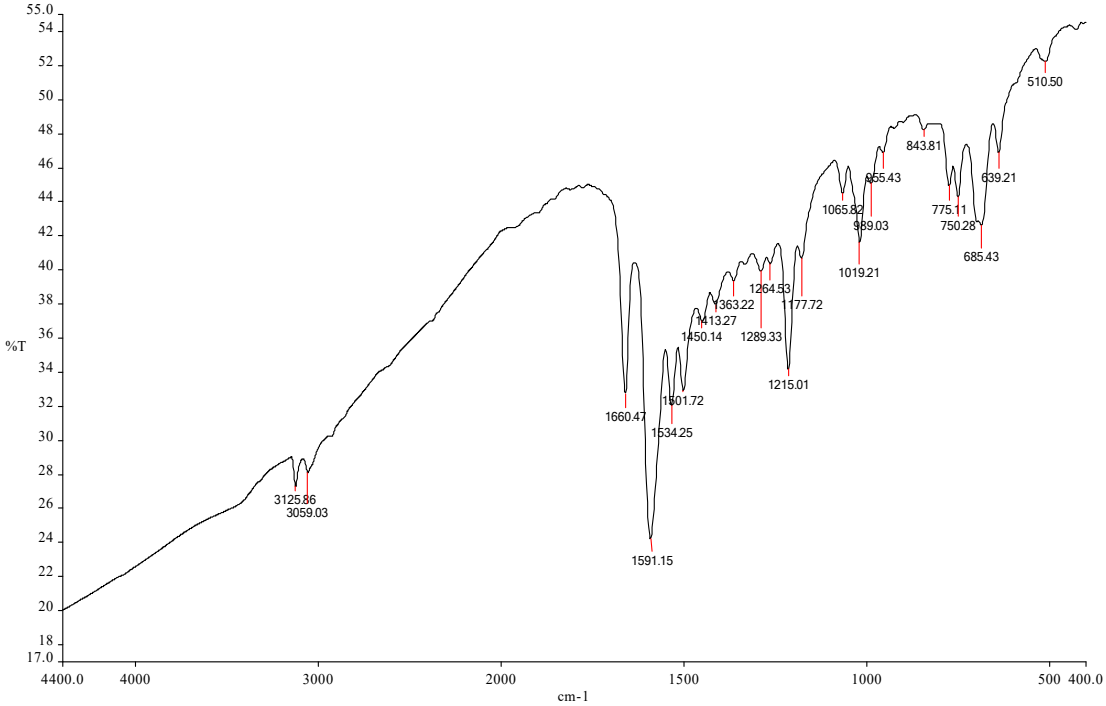


Figure S21: IR-spectroscopy for compound-**3a**

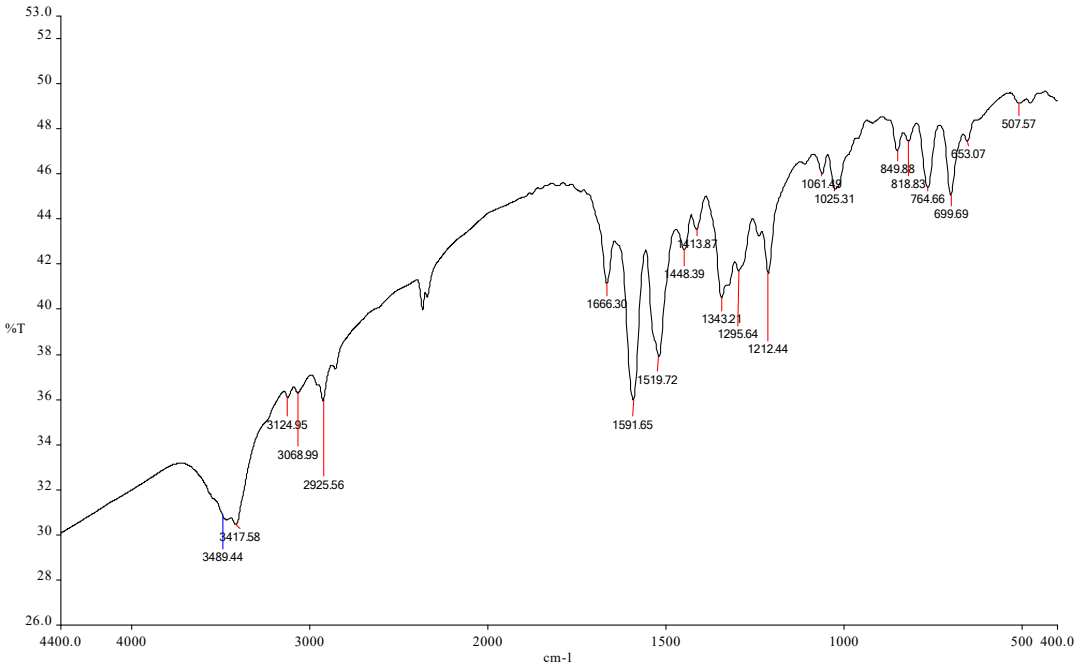


Figure S22: IR-spectroscopy for compound-**3b**

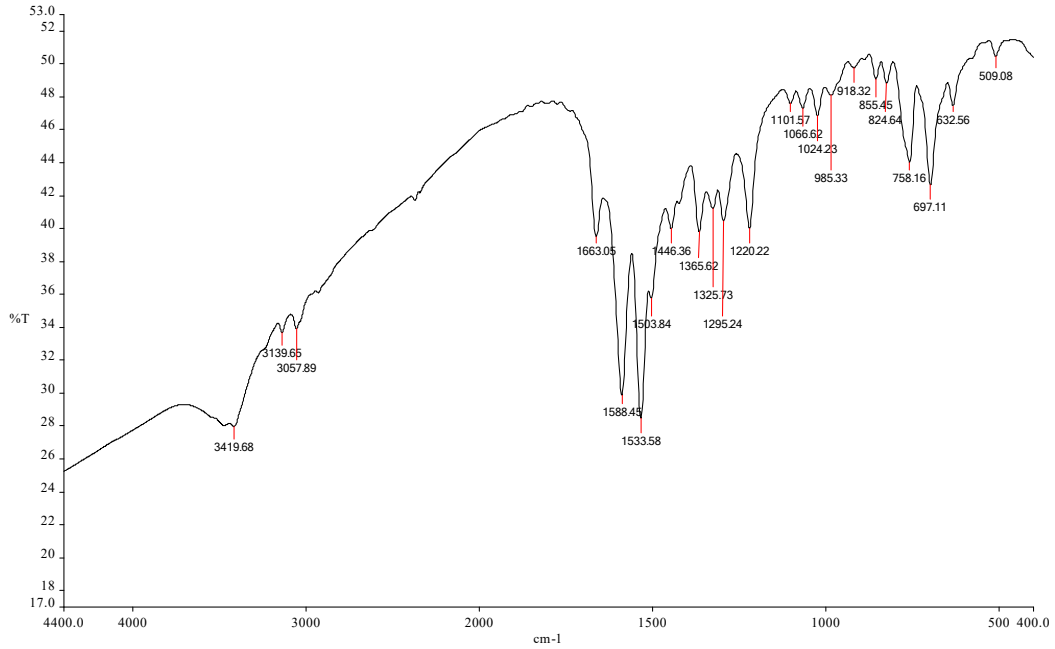


Figure S23: IR-spectroscopy for compound-**3c**

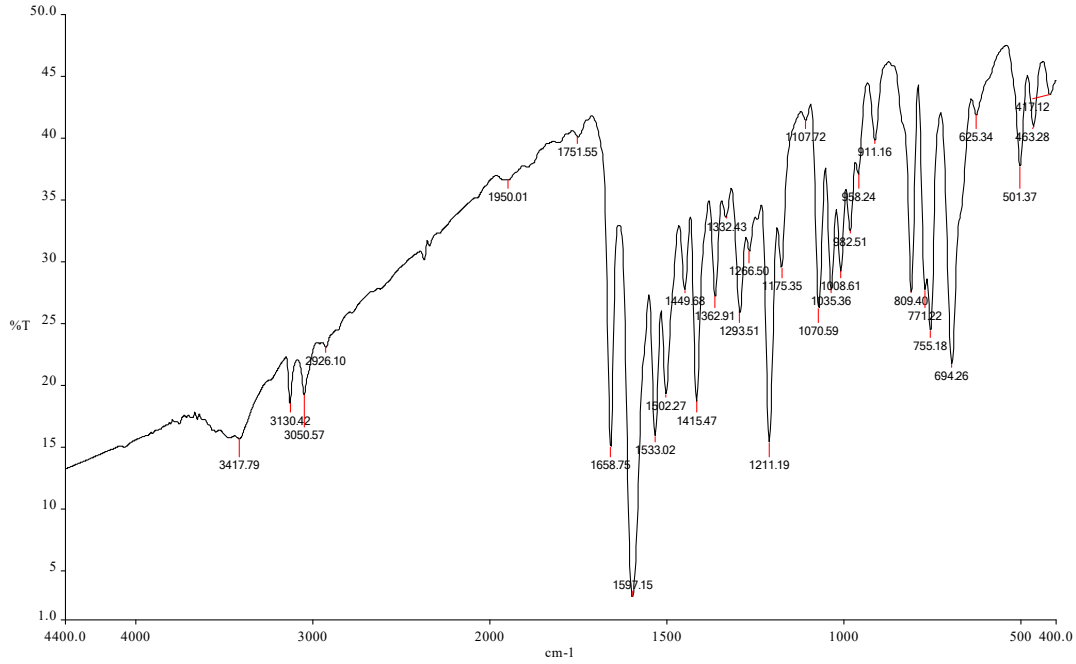


Figure S24: IR-spectroscopy for compound-**3d**

1.4.2. IR-spectroscopy of spiro compounds-6a-k

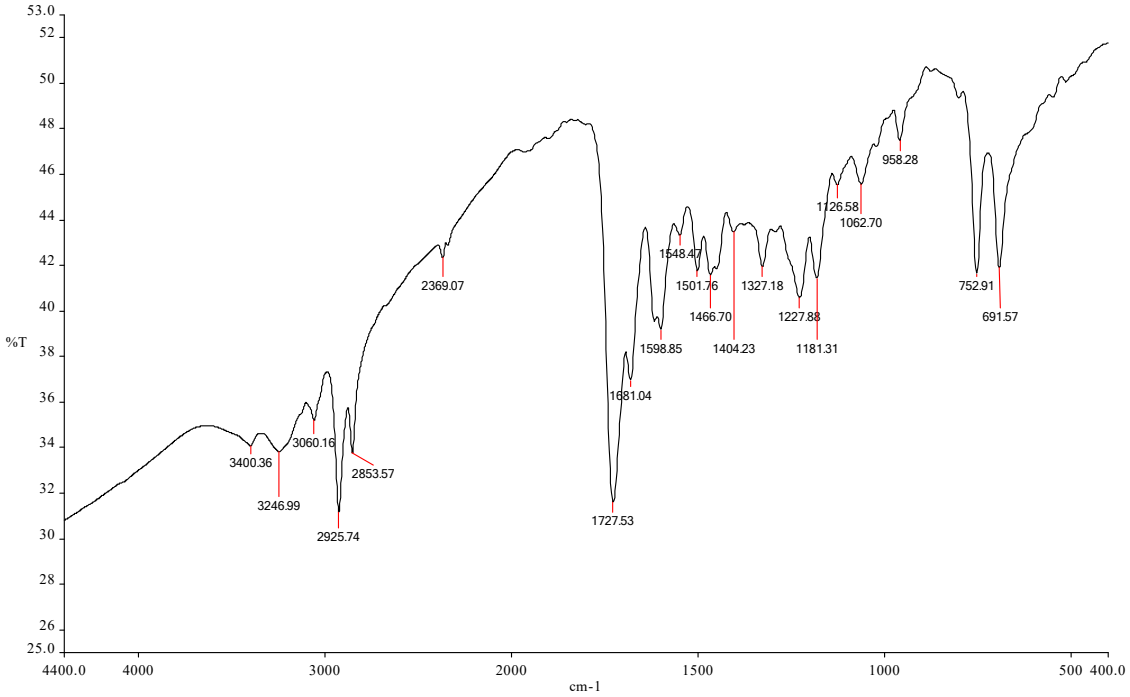


Figure S25: IR-spectroscopy for compound-**6a**

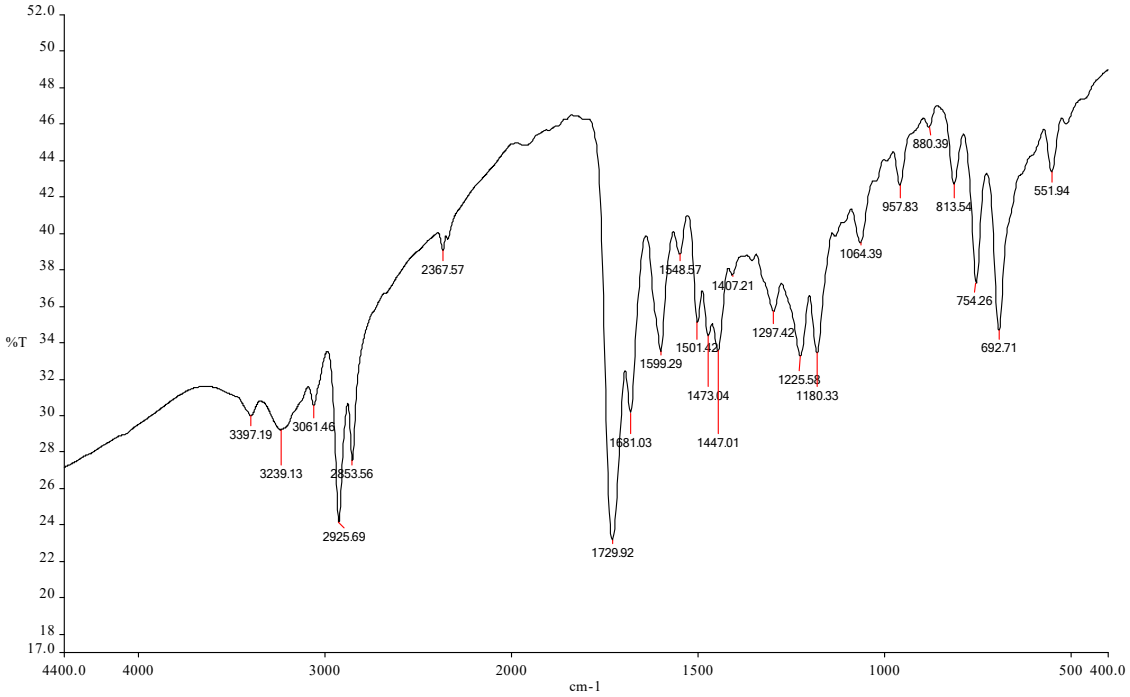


Figure S26: IR-spectroscopy for compound-**6b**

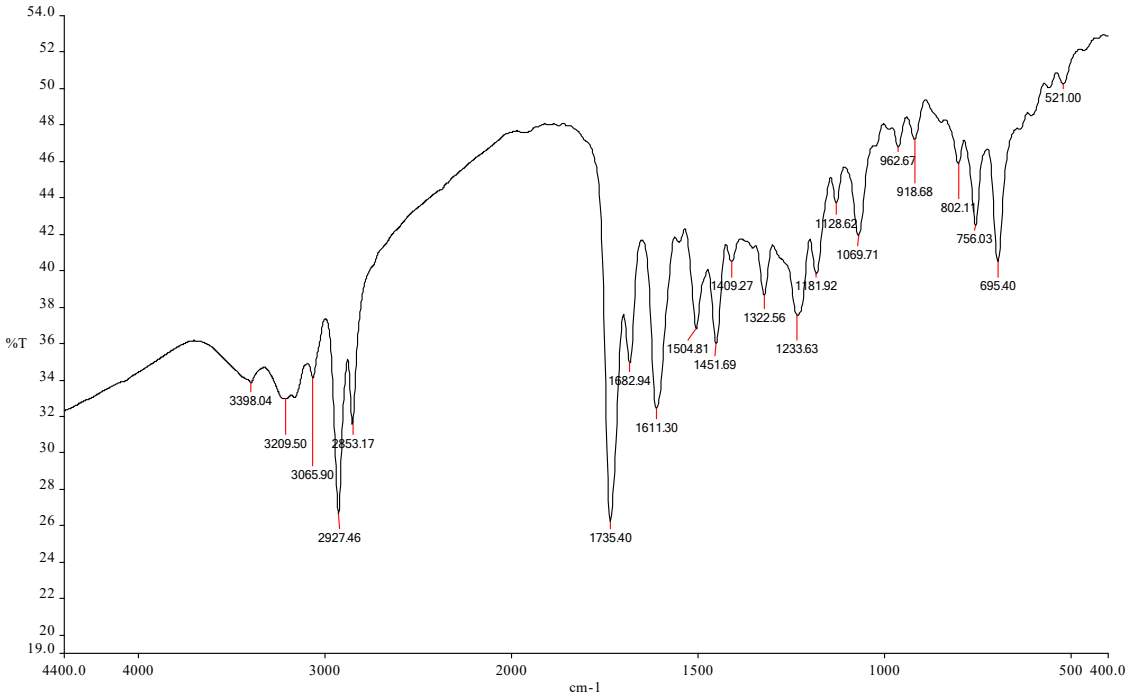


Figure S27: IR-spectroscopy for compound-**6c**

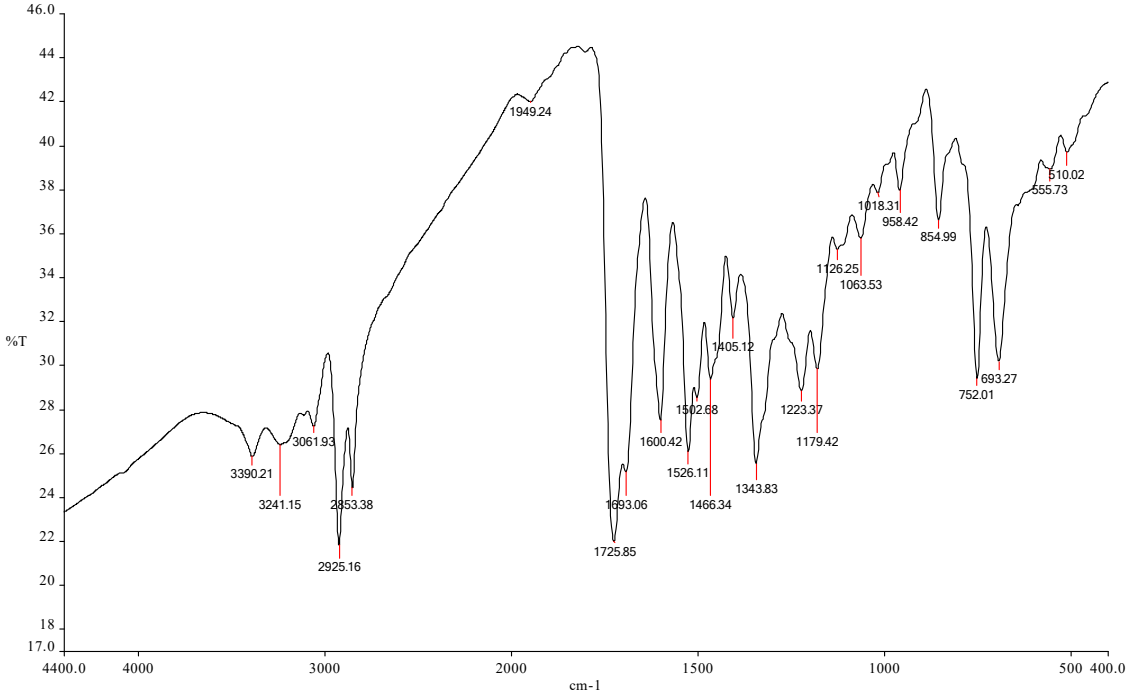


Figure S28: IR-spectroscopy for compound-**6d**

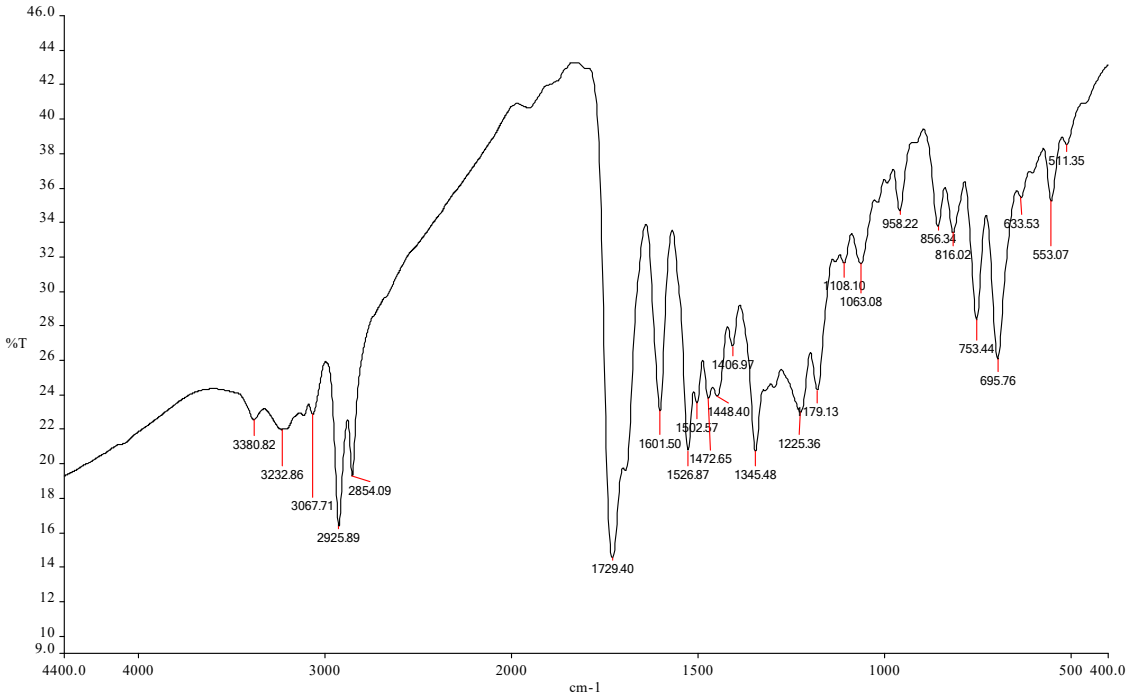


Figure S29: IR-spectroscopy for compound-**6e**

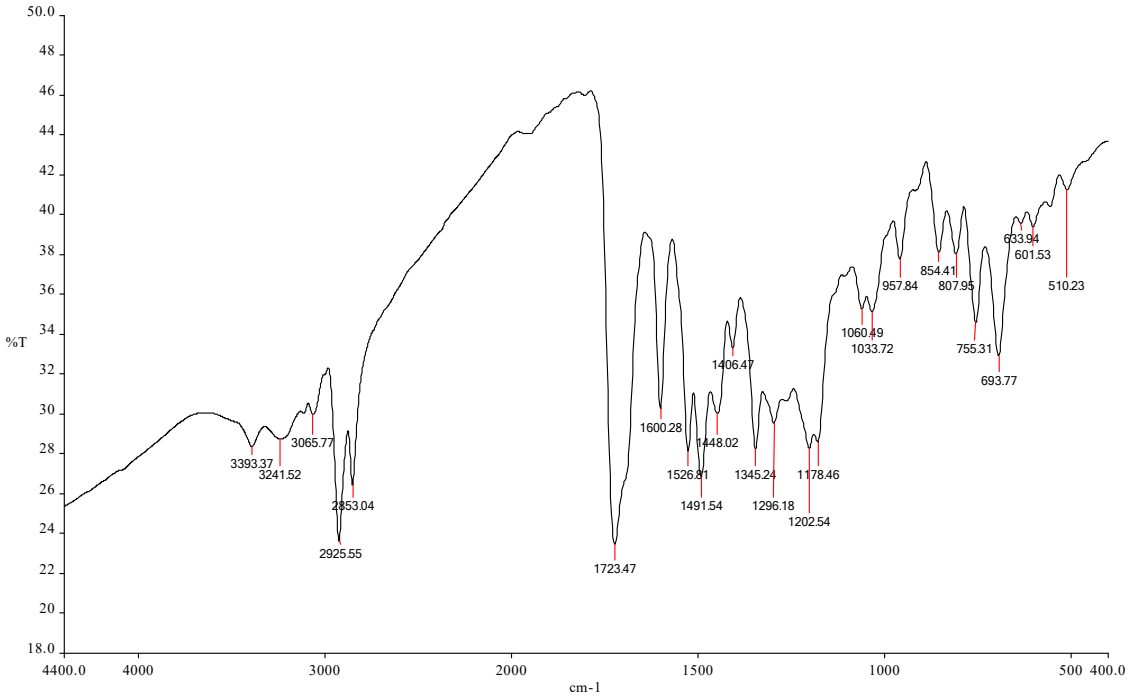


Figure S30: IR-spectroscopy for compound-**6f**

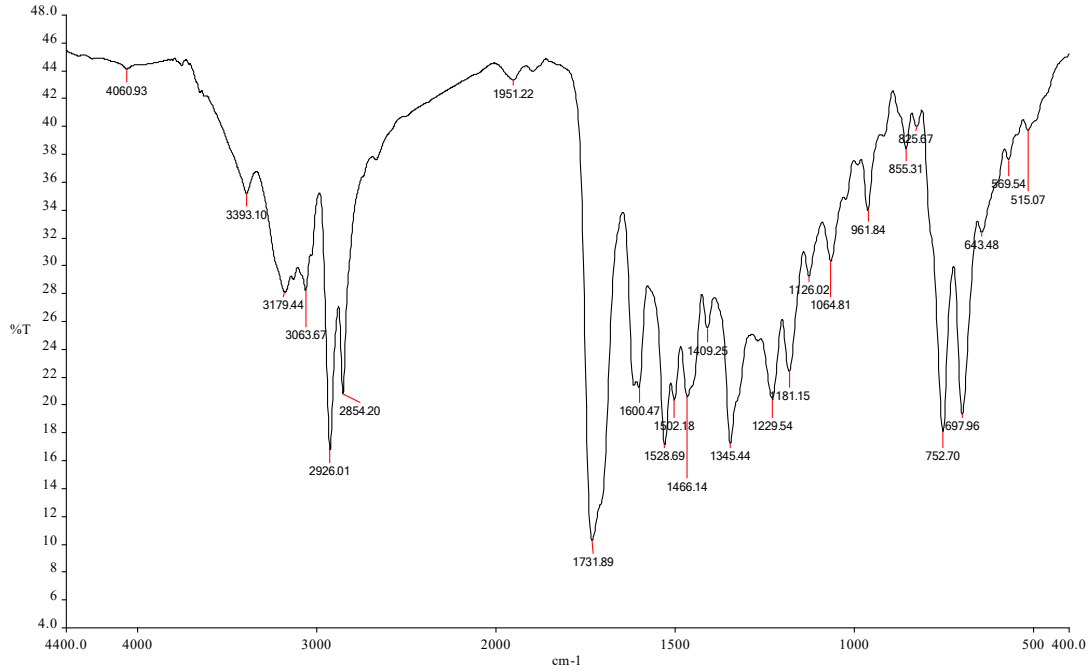


Figure S31: IR-spectroscopy for compound-**6g**

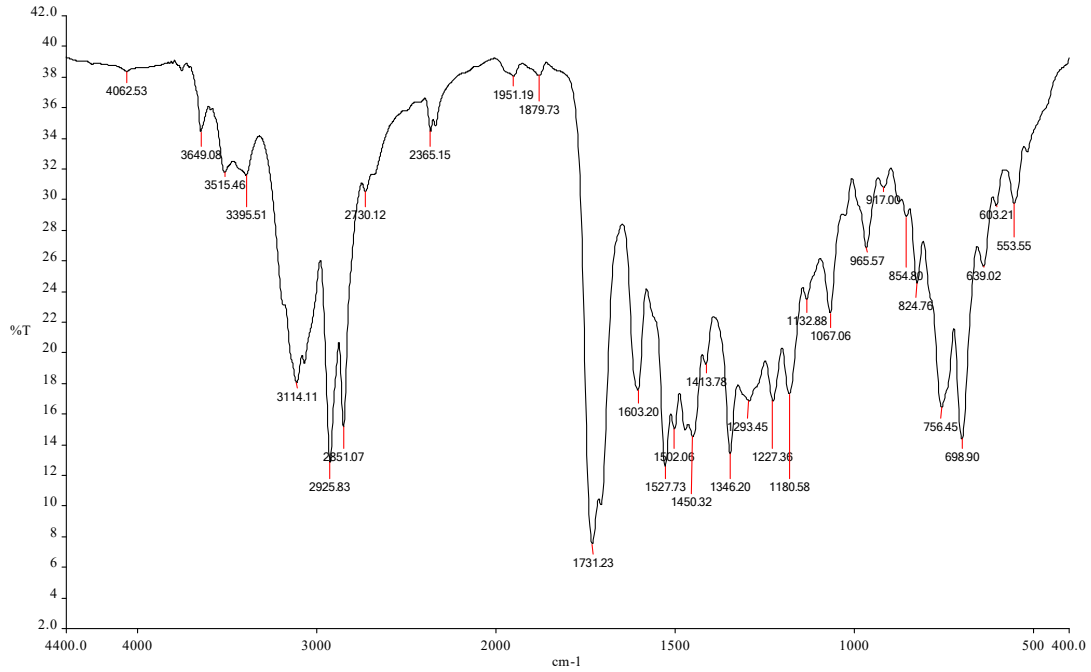


Figure S32: IR-spectroscopy for compound-**6h**

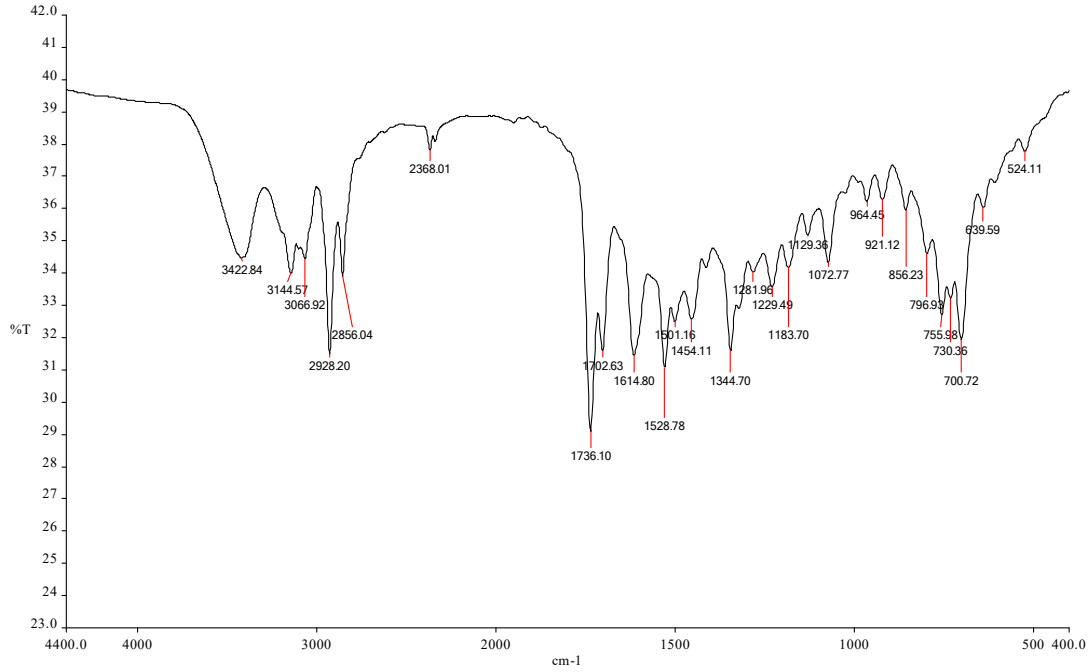


Figure S33: IR-spectroscopy for compound-**6i**

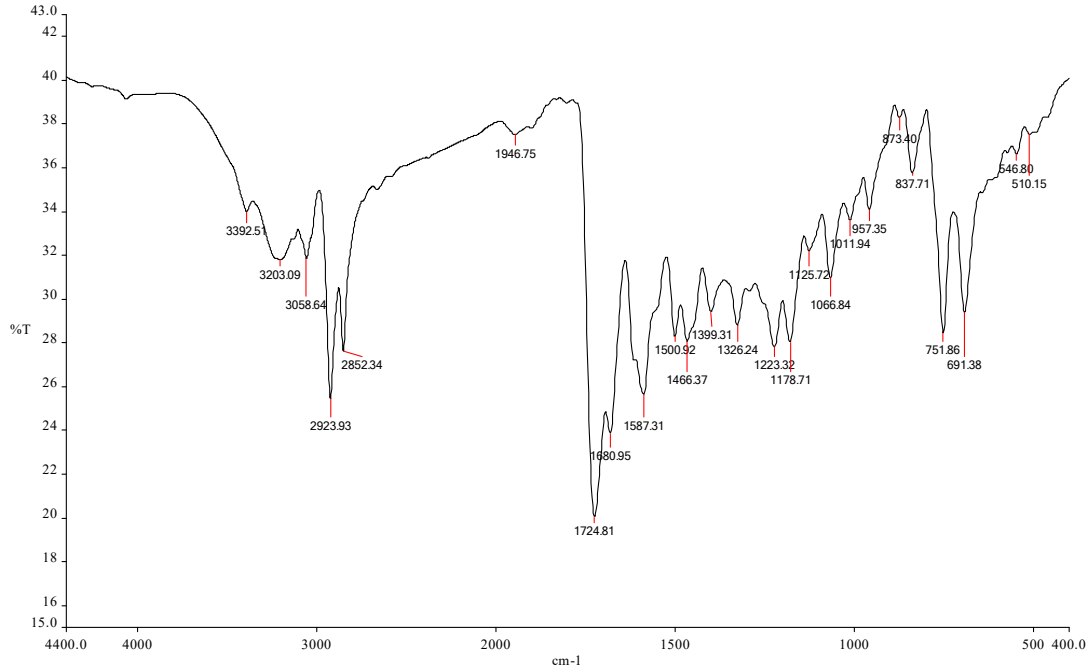


Figure S34: IR-spectroscopy for compound-**6j**

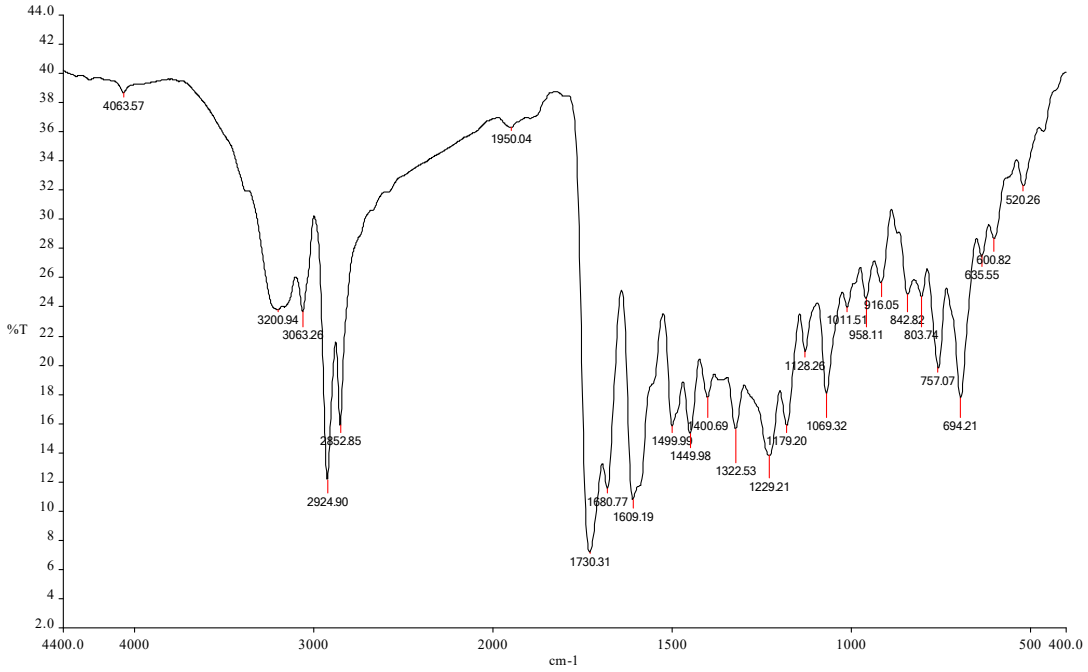


Figure S35: IR-spectroscopy for compound-**6k**

1.4.3. IR-spectroscopy of spiro compounds-8a-e

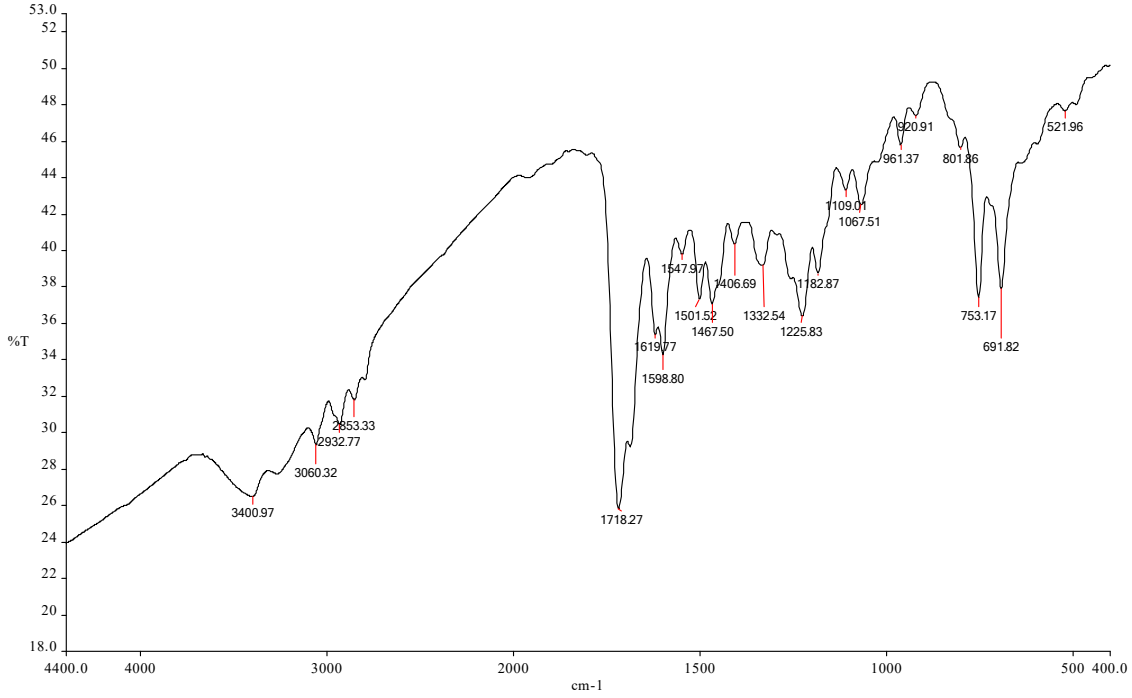


Figure S36: IR-spectroscopy for compound-**8a**

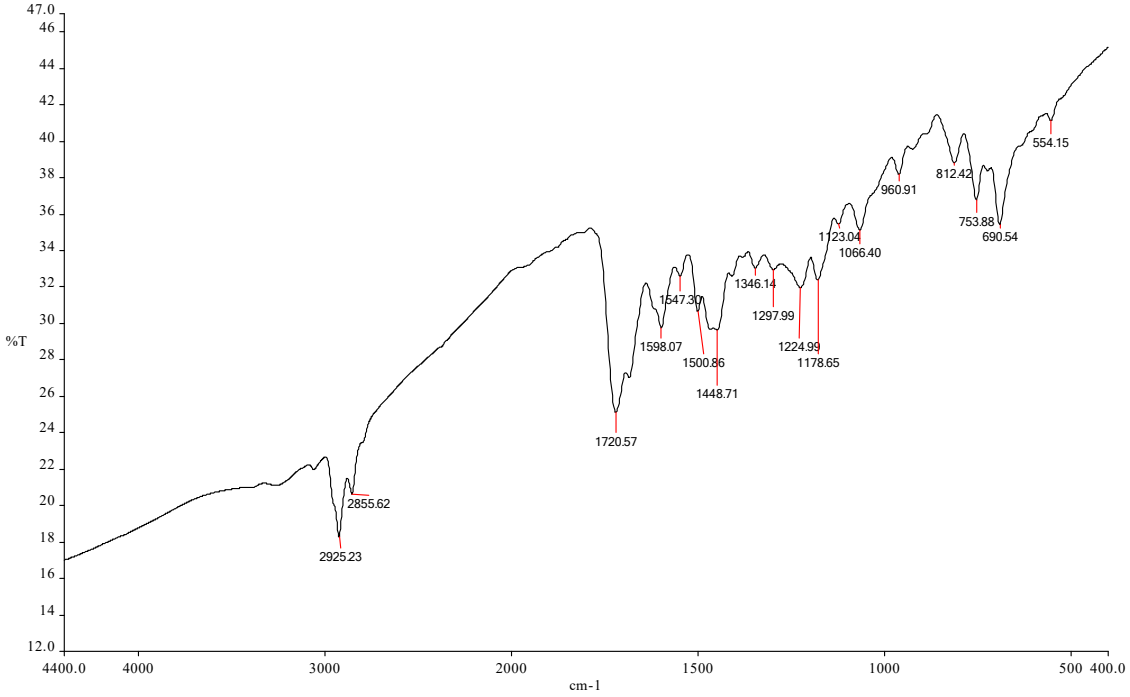


Figure S37: IR-spectroscopy for compound-**8b**

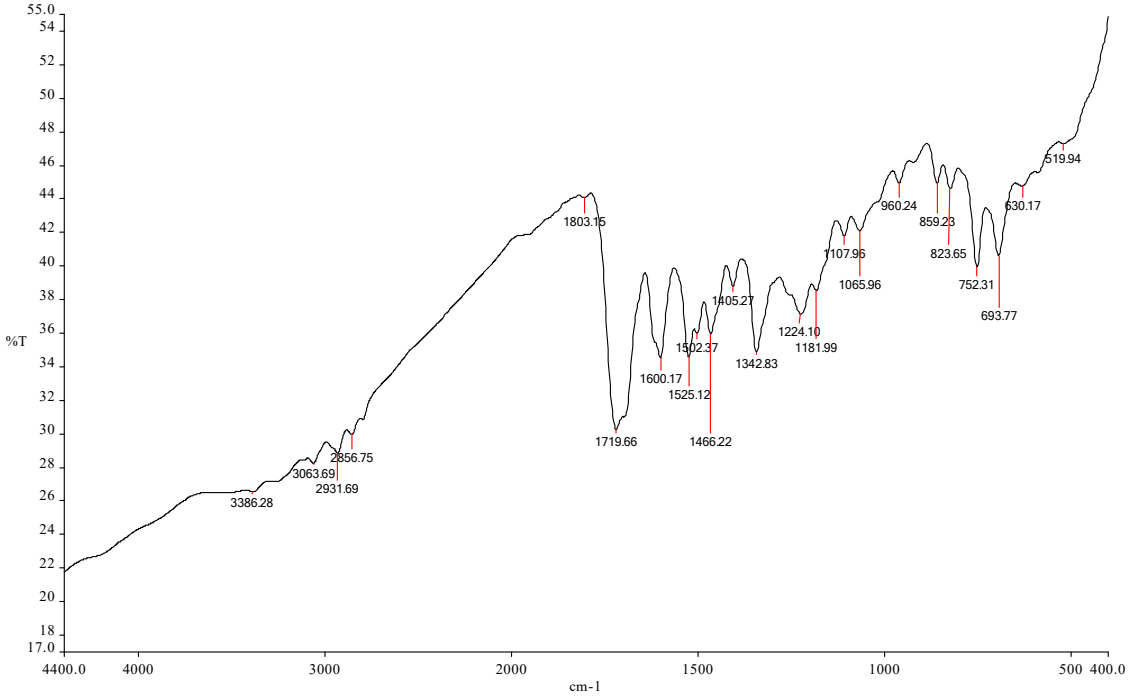


Figure S38: IR-spectroscopy for compound-**8c**

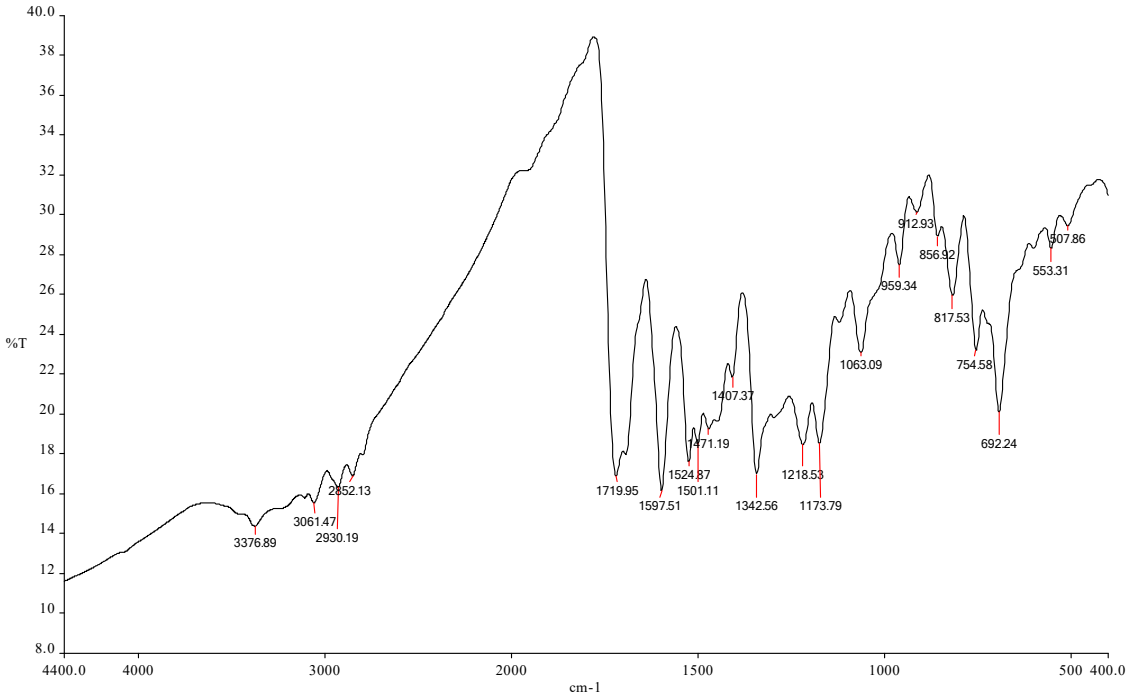


Figure S39: IR-spectroscopy for compound-**8d**

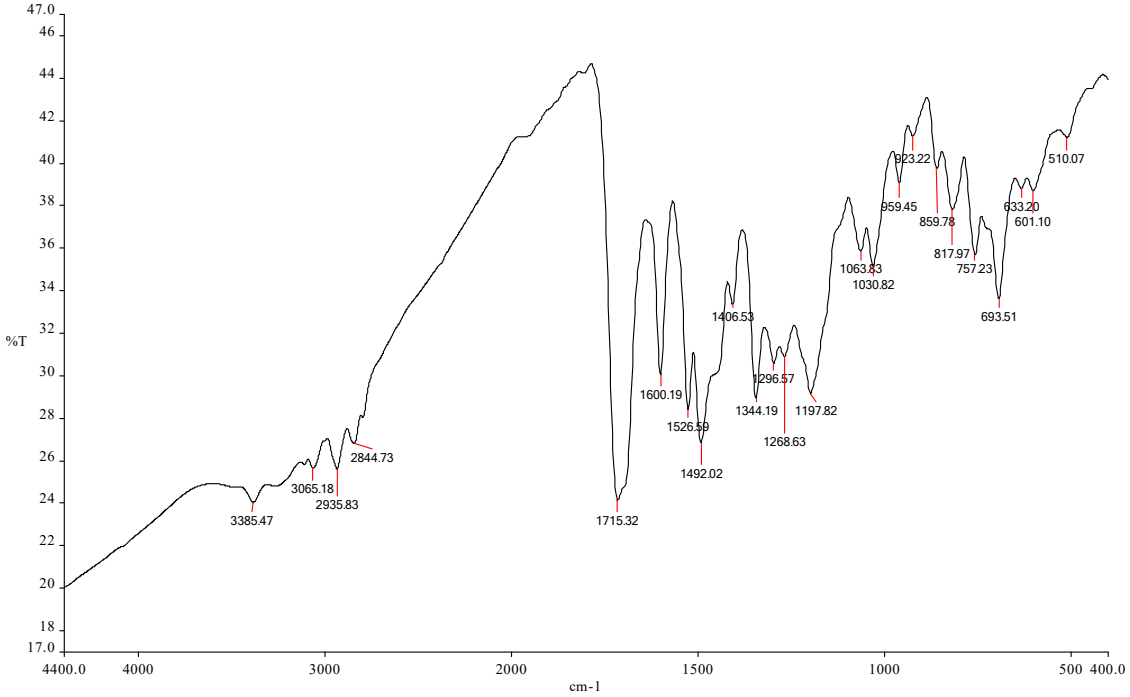


Figure S40: IR-spectroscopy for compound-**8e**

1.5. Mass spectra

1.5.1. LCMS spectra of compounds 6a-k

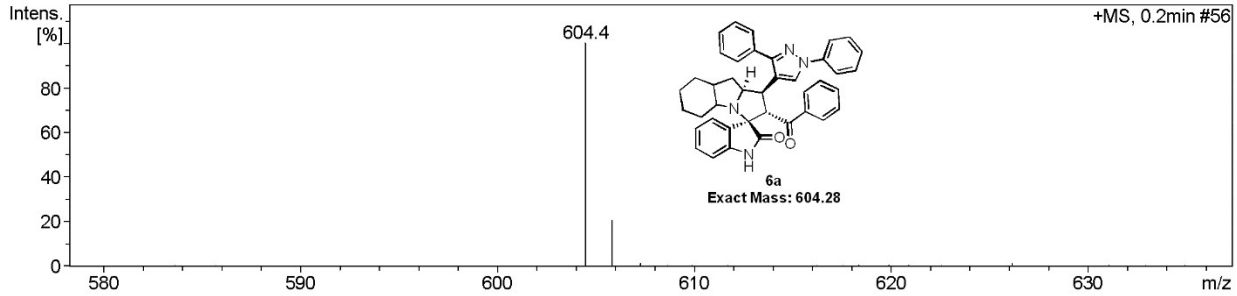


Figure S41: LCMS spectra for compound-**6a**

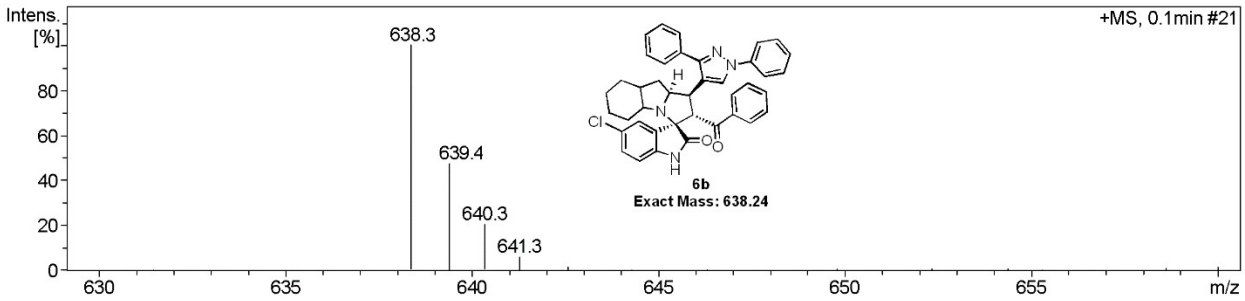


Figure S42: LCMS spectra for compound-**6b**

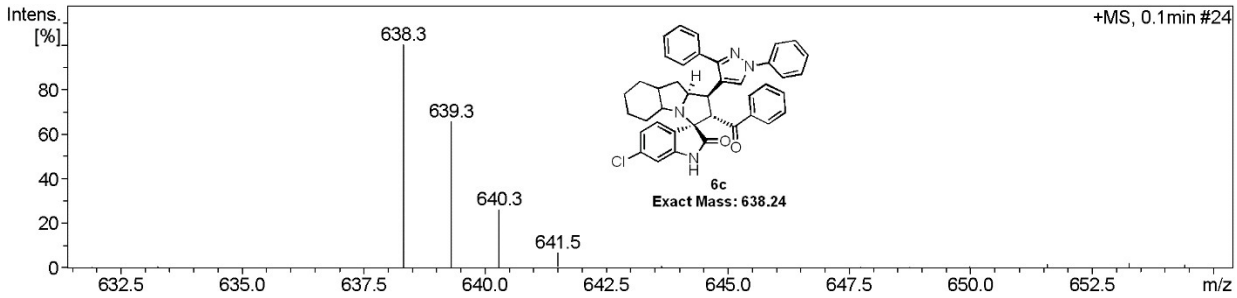


Figure S43: LCMS spectra for compound-**6c**

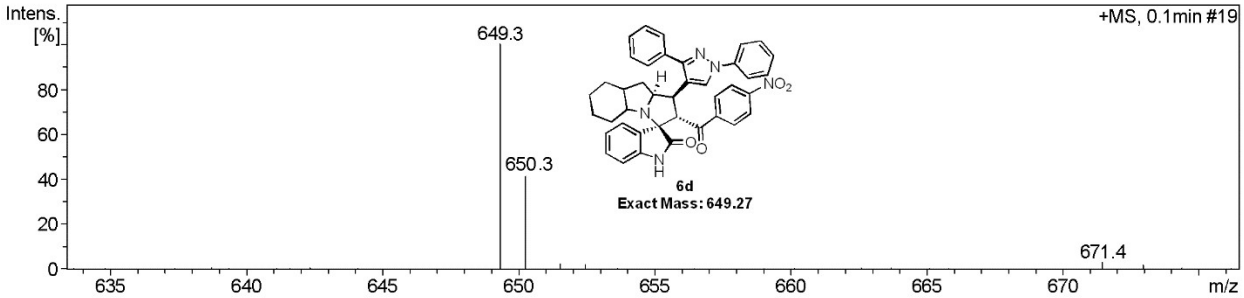


Figure S44: LCMS spectra for compound-**6d**

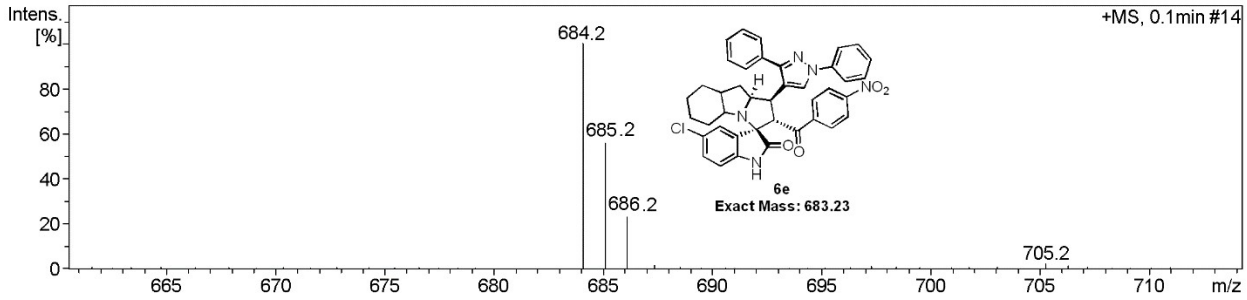


Figure S45: LCMS spectra for compound-**6e**

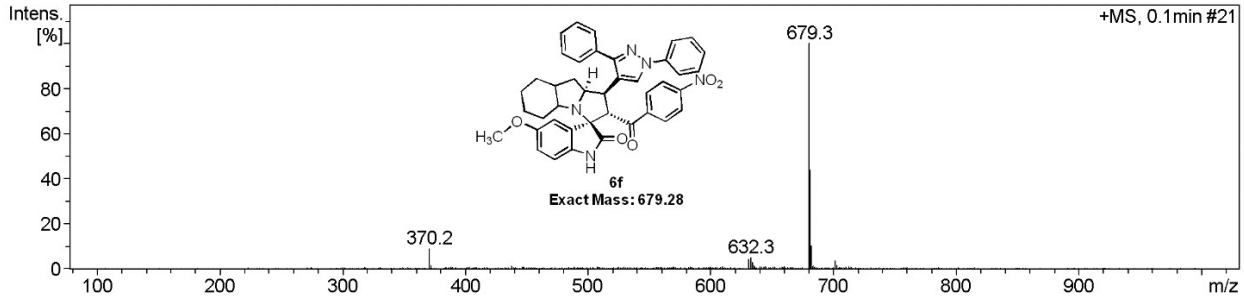


Figure S46: LCMS spectra for compound-**6f**

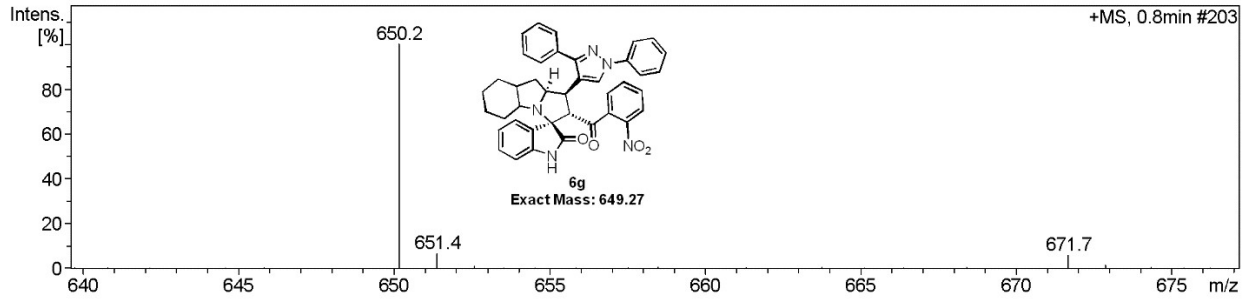


Figure S47: LCMS spectra for compound-**6g**

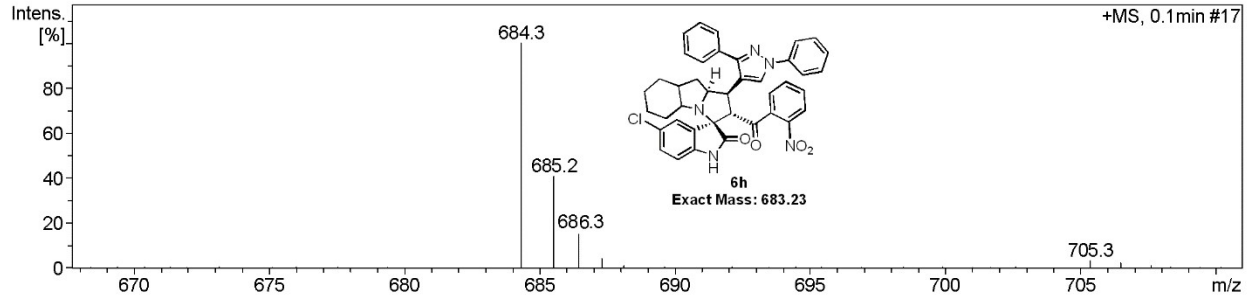


Figure S48: LCMS spectra for compound-**6h**

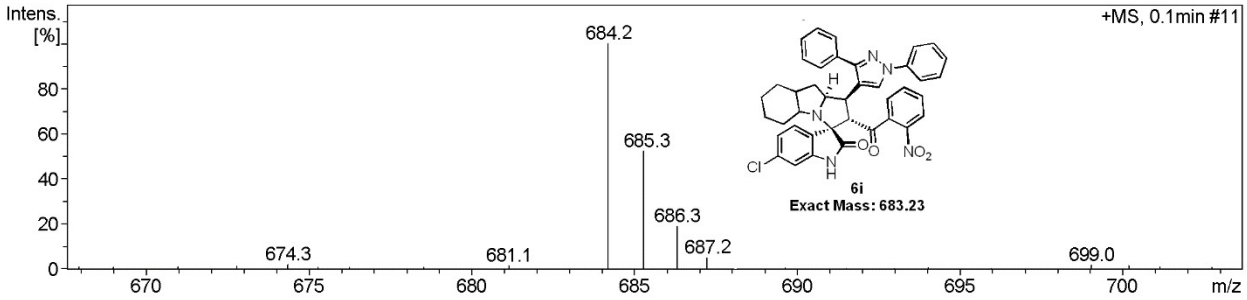


Figure S49: LCMS spectra for compound-**6i**

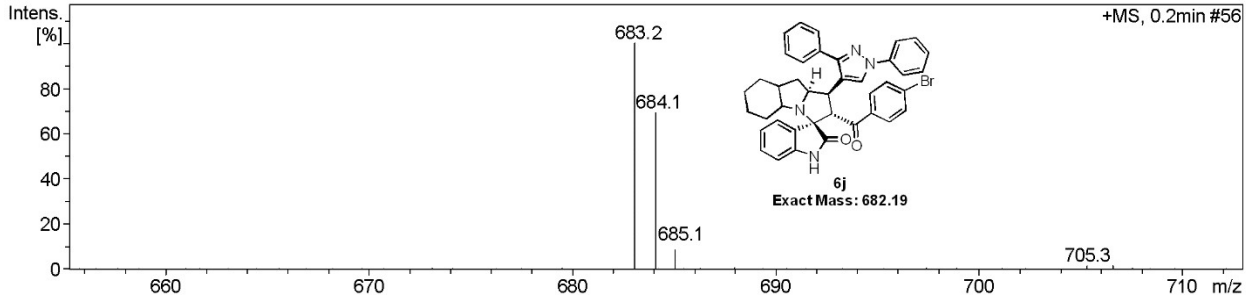


Figure S50: LCMS spectra for compound-**6j**

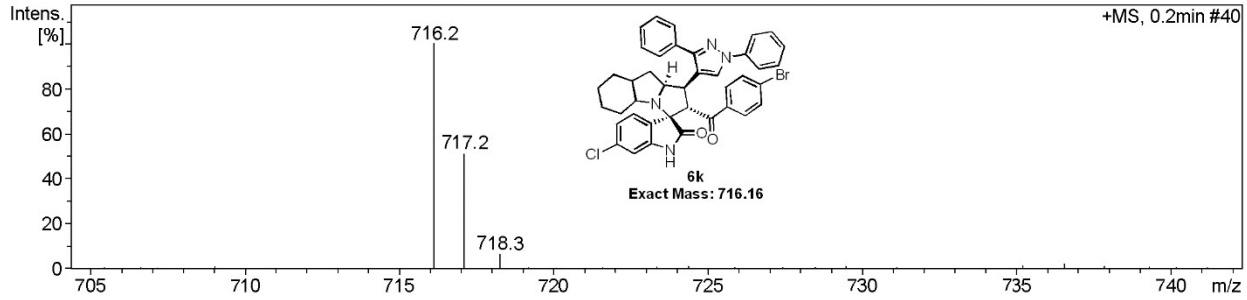


Figure S51: LCMS spectra for compound-**6k**

1.5.2. LCMS spectra of compounds 8a-e

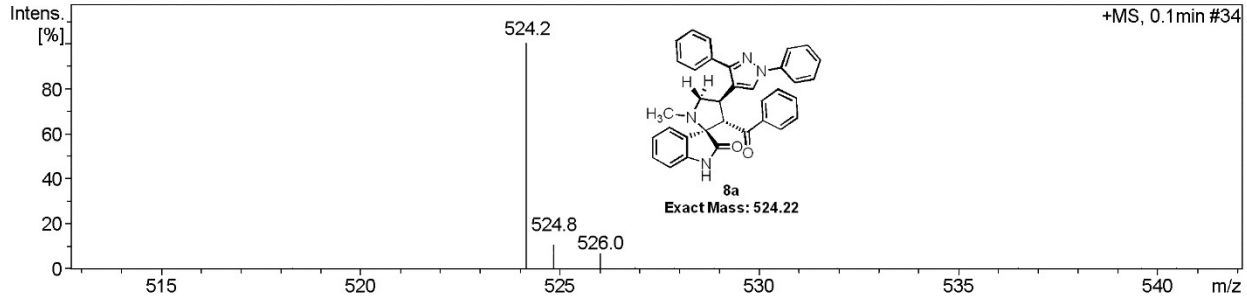


Figure S52: LCMS spectra for compound-**8a**

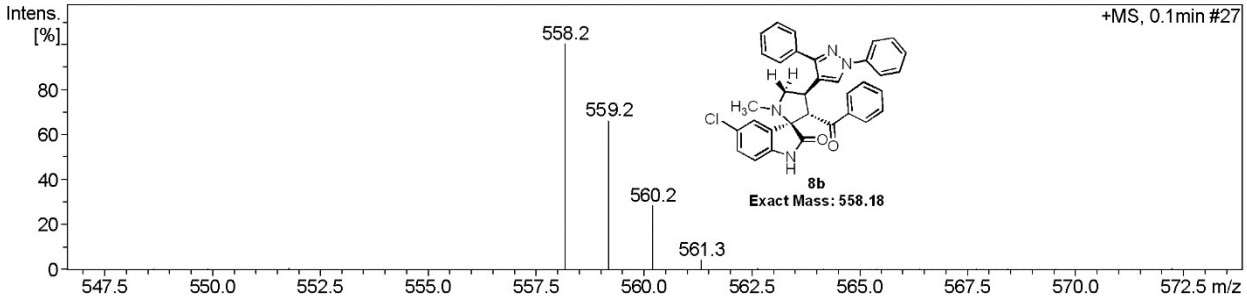


Figure S53: LCMS spectra for compound-**8b**

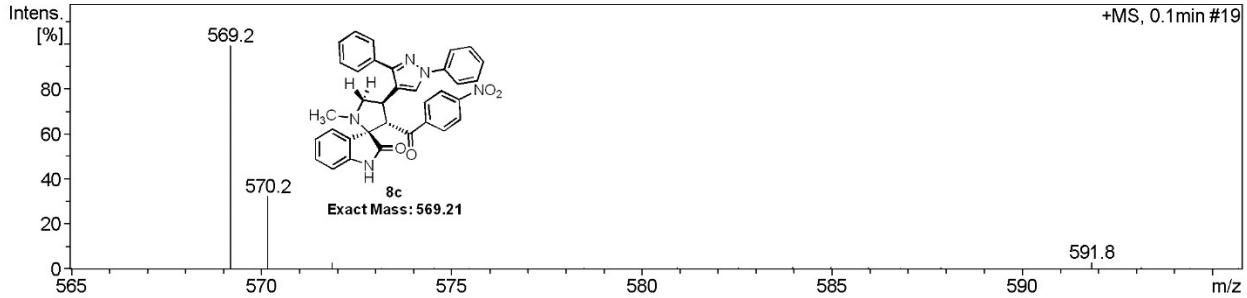


Figure S54: LCMS spectra for compound-**8c**

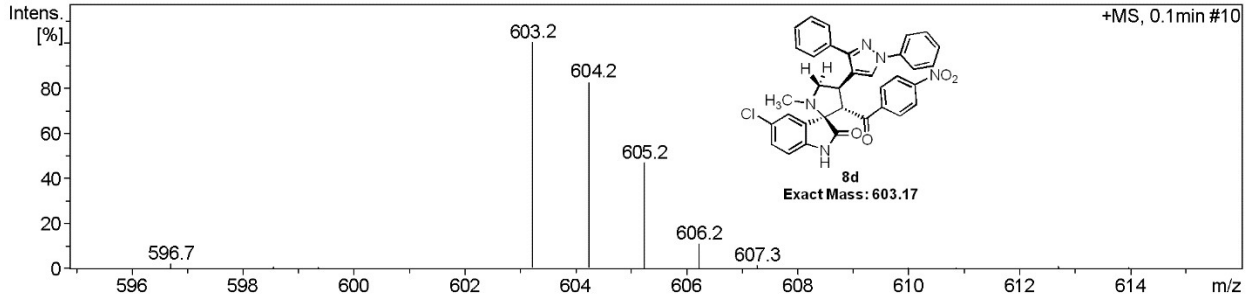


Figure S55: LCMS spectra for compound-**8d**

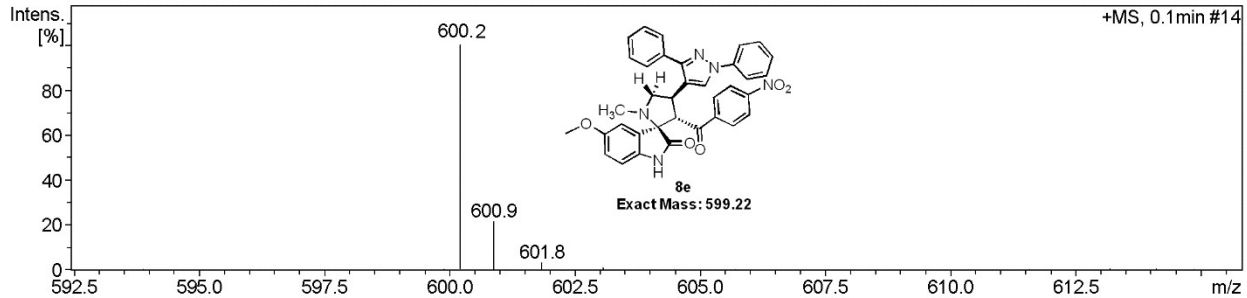


Figure S56: LCMS spectra for compound-**8e.**

1.5.3. HRMS spectra of compounds 6e and eh

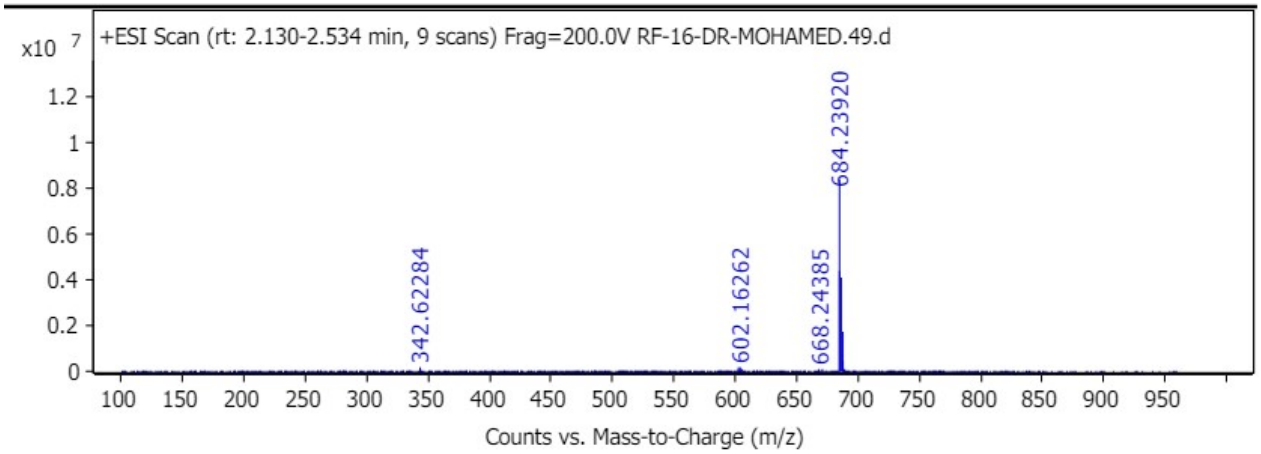


Figure S57. HRMS spectrum of compound 6e.

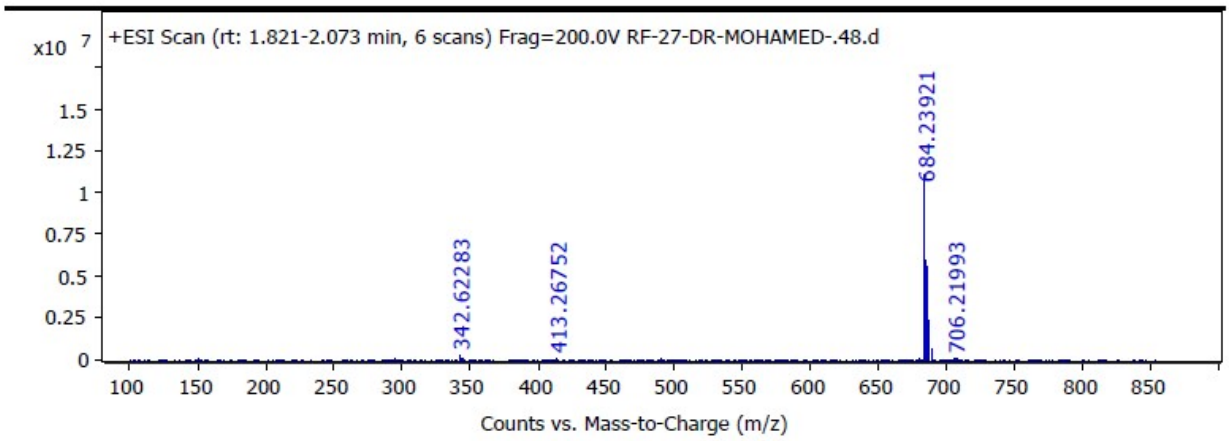
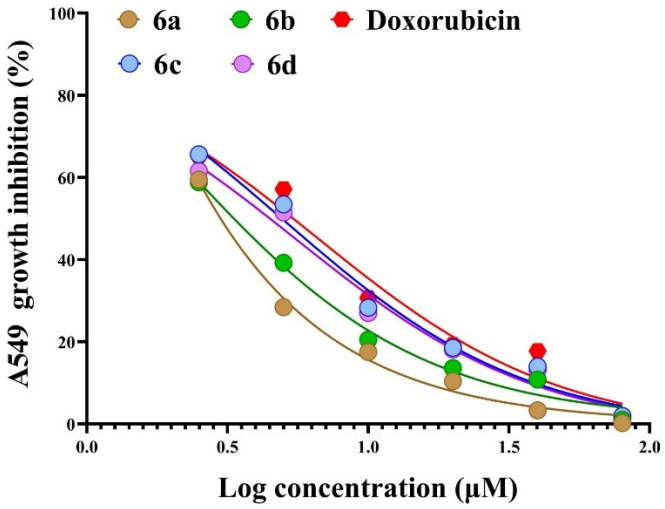


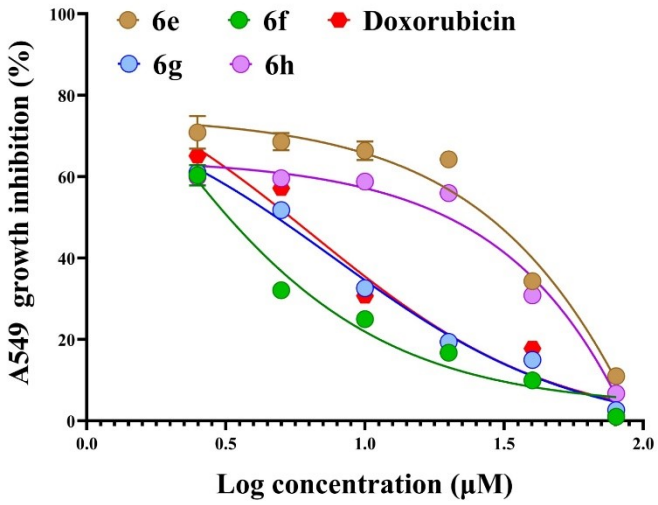
Figure S58. HRMS spectrum of compound 6h.

2. Biological screening

2.1. Cytotoxicity screening

Normal human lung fibroblast Wi-38 cell line was used to detect cytotoxicity of the studied compounds. Wi-38 cell line was cultured in DMEM medium-contained 10% fetal bovine serum (FBS), seeded as $5x10^3$ cells per well in 96-well cell culture plate and incubated at 37ºC in 5% CO_2 incubator. After 24 h for cell attachment, serial concentrations of the synthetic compounds were incubated with Wi-38 cells for 72 h. Cell viability was assayed by MTT method. Twenty microliters of 5 mg/ml MTT (Sigma, USA) was added to each well and the plate was incubated at 37ºC for 3 h. Then MTT solution was removed, 100 µl DMSO was added and the absorbance of each well was measured with a microplate reader (BMG LabTech, Germany) at 570 nm. The dose (EC50 and EC_{100}) values (at 50% and100% cell viability, respectively) of the tested compounds was estimated by the Graphpad Instat software. The anticancer effect of the investigated compounds was assayed using human lung cancer cell lines A549 [Source: American Type Culture Collection (ATCC)] cultured in DMEM (Lonza, USA) supplemented with 10% FBS, and H460 cultured in RPMI-containing 10% FBS. All cancer cells ($5x10^3$ cells/well) were seeded in sterile 96-well plates. After 24h, serial concentrations of the tested compounds were incubated with two cancer cell lines for 72 h at 37ºC in 5% CO_2 incubator. MTT method was done as described above. The half maximal inhibitory concentration (IC_{50}) values were calculated using the Graphpad Instat software.





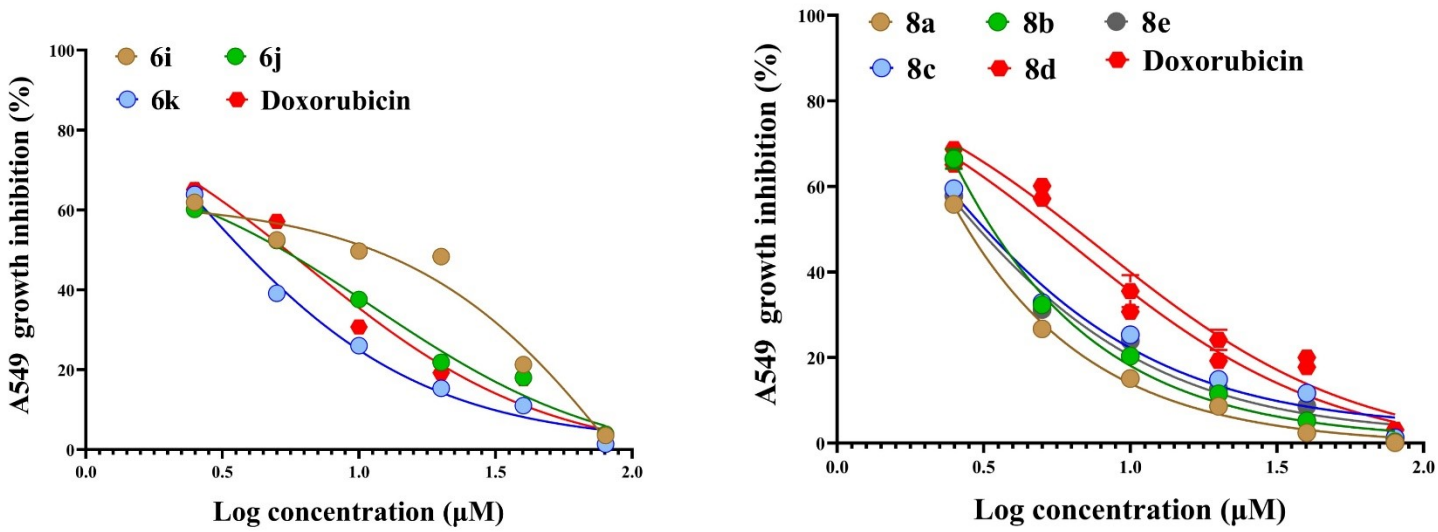


Figure S59. Growth inhibition curves of **6a-k**, **8a-e** and Doxorubicin against A549 cells.

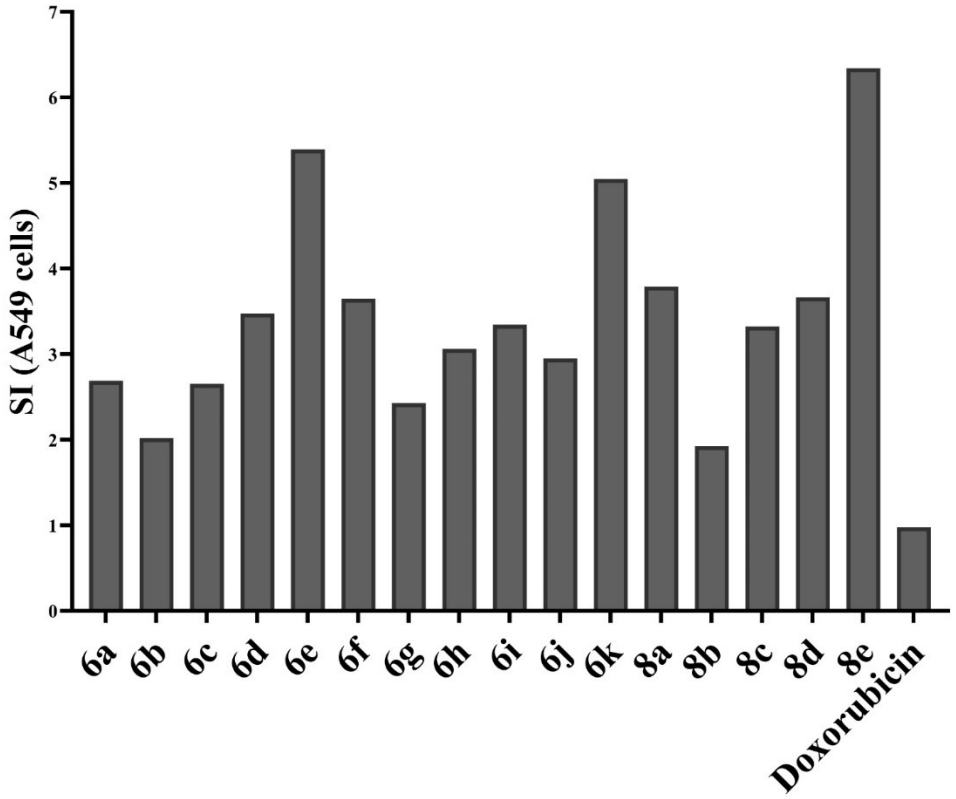


Figure S60. Selectivity indices of **6a-k**, **8a-e** and Doxorubicin towards A549 cells against Wi-38. Selectivity index (SI) is defined as the ratio between the compound's IC_{50} against normal cells and its IC_{50} against cancer cells. Generally, an anticancer compound with $SI \geq 3$ is considered as highly selective [5].

2.2. Morphological alteration of the treated cancer cells

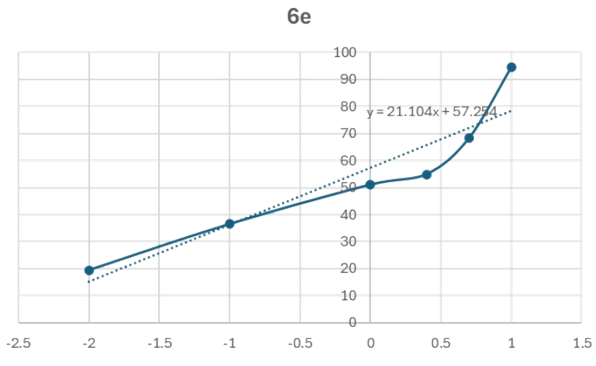
Furthermore, cellular morphological changes before and after treatment with the most effective and safest anticancer compounds were investigated using phase contrast inverted microscope with a digital camera (Olympus, Japan) (**Figure S61)**.

Figure S61. Morphological changes of lung cancer cells after 72 h treatment with **6e** and **6h.**

2.3. CDK2 inhibition

CDK2 luminescence kinase Assay kit (Catalog #79599, Kinase-Glo Plus, Promega, USA) were performed to evaluate the inhibitory potency of compounds **6e** and **6h** against CDK2. The autophosphorylation percentage inhibition by compounds was calculated using the following equation: $100 - [Control/Treated - Control)]$ using the curves of percentage inhibition of eight concentrations of each compound, IC_{50} was calculated using the GraphPad prism7 software.

[Conc]	10	5	2.5	1	0.1	0.01
log [Conc]	1	0.69897	0.39794	0	-1	-2
% inhbition	94.37793	68.19365	54.91931	51.06715	36.49141	19.41697



[Conc]	10	5	2.5	1	0.1	0.01
log [Conc]	1	0.69897	0.39794	0	-1	-2
% inhbition	92.0354	68.34982	64.75794	52.36856	31.28579	8.95367

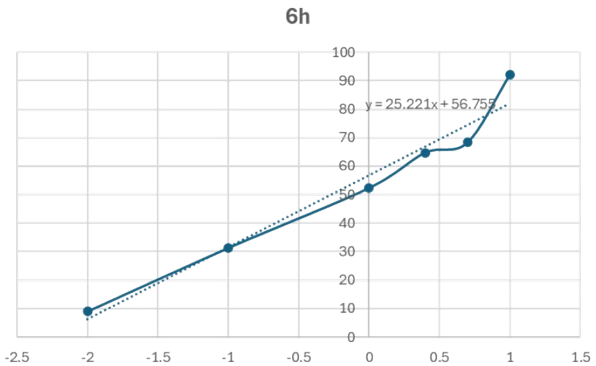


Figure S62. Dose-response curves for the CDK2 inhibition versus serial dilutions of compound **6e and 6h.**

3. Computational studies

3.1. Docking

The available three-dimensional crystal structures of DNA dodecamer (PDB ID: 2DND [5]) and CDK2 active domain (PDB code: 2A4L [6]) co-crystallized with their respective ligands were retrieved from the Protein Data Bank (PDB, www.rcsb.org) for docking employing Molecular Operating Environment (MOE) software package version MOE 2019.102, Chemical Computing Group, Montreal, Canada [7]. Unwanted residues and ligands were eliminated. The DNA dodecamer was prepared and refined with adjusting missing loops employing the default "Structure preparation" MOE setting, then energy minimized employing Amber10: EHT force field with reaction-field electrostatics (an interior dielectric constant of 1 and an exterior dielectric of 80) using an 8–10 Å cutoff distance. The active pocket encompassing distamycin was located. Validation was done by restoring most of the experimental binding interaction following redocking distamycin in the active pocket with RMSD < 2 Å. The complex as well as the precursor were energy minimized and docked employing the MOE dock tool into the active binding site using rigid sampling. Triangular matcher algorithm was set as the ligand placement method while London dG was the default scoring function to generate the top ten non-redundant poses for each ligand. Based on the energy binding score, the best-docked structure for each ligand was determined. Representations of the selected 2D and 3D molecular interactions were generated and inspected.

3.2. MD simulations

3.2.1. Trajectory analysis

GROMACS utilities were used for the analysis of the MD simulations. The root mean square deviation (RMSD) of atom position for ligand and protein was calculated by fitting protein backbone atom with the *gmx_rms* subprogram. Similarly, root mean square fluctuations (RMSF) based on the protein C-alpha atoms were calculated using *gmx_rmsf*. Radius of gyration of all protein atoms was calculated with the *gmx_gyrate* and number of hydrogen bonds were calculated (in-side the protein-ligand interface) with the *gmx_hbond*. The utility *gmx_distance* was used to calculate the center of mass distance between the protein and the ligand during the simulation. The VMD molecular graphics program was used for trajectory visualization, detailed hydrogen bonding information, secondary structure analysis (DSSP) and protein-ligand contact frequency (CF) analysis.

3.2.2. Binding Free Energy (MM/PBSA Calculations)

For systems which were chosen for further analysis, MM/PBSA (Molecular Mechanics/Poisson−Boltzmann Surface Area) calculations were done using g_mmpbsa, a GROMACS tool used to calculate an estimated binding affinity. In general terms, the binding free energy of the protein with ligand in solvent can be expressed as:

$$\Delta G_{binding} = \Delta G_{complex} - (\Delta G_{protein} + \Delta G_{ligand})$$

where, $\Delta G_{complex}$ is the total free energy of the protein–ligand complex, and $\Delta G_{protein}$ and ΔG_{ligand} are total free energies of the isolated protein and ligand in solvent, respectively. g_mmpbsa can also be used to estimate the energy contribution per residue to the binding energy. To decompose the binding energy, at first ΔE_{MM}, ΔG_{polar} and $\Delta G_{non-polar}$ were separately calculated for each residue and were then summed up to obtain the contribution of each residue to the binding energy. Considering that g_mmpbsa only read the files of some specific GROMACS versions, the binary run input file (.tpr) required for MM-PBSA calculation through the g_mmpbsa was regenerated by GROMACS 5.1.4. The molecular structure file (.gro), topology file (.top) and MD-parameter file (.mdp) were necessary to generate the binary run input file, and they all came from the MD process.

References

1. Rikagu Oxford Diffraction, CrysAlisPro, 2020, Rikagu Oxford Diffraction inc., Yarnton, Oxfordshire, England.
2. G. Sheldrick, Acta Crystallographica Section A, 2015, 71, 3-8.
3. G. Sheldrick, Acta Crystallographica Section C, 2015, 71, 3-8.
4. C.B. Hübschle, G.M. Sheldrick, B. Dittrich, ShelXle a Qt graphical user interface for SHELXL, Journal of Applied Crystallography 44(6) (2011) 1281-1284.
5. P. Prayong, S. Barusrux and N. Weerapreeyakul, Fitoterapia, 2008, 79, 598-601.
6. M. Coll, C. A. Frederick, A. Wang and A. Rich, Proceedings of the National Academy of Sciences, 1987, 84, 8385-8389.
7. W. F. De Azevedo, S. Leclerc, L. Meijer, L. Havlicek, M. Strnad and S. H. Kim, European journal of biochemistry, 1997, 243, 518-526.
8. Chemical Computing Group, Inc. Molecular Operating Environment [(MOE 2019.10) Chemical Computing Group, Inc.; Montreal, QC, Canada. http://www.chemcomp.com (accessed on 21 April 2023).