Electronic Supporting Information

L-Proline-catalysed synthesis of chromeno[2,3-*b*]chromene from 4-hydroxy-2*H*-chromene-2-thione and some of their anti-proliferative study

Arnab Mandal,[a] Satyajit Singh,[a] Arisha Arora,[b] Sujisha S. Nambiar,[b] Siddhartha S. Ghosh,[b] Abu Taleb Khan*[a]

[a] Department of Chemistry, Indian Institute of Technology Guwahati, Guwahati-781039, Assam, India. E-mail: atk@iitg.ac.in

[b] Department of Bioscience and Bioengineering, Indian Institute of Technology Guwahati, Guwahati-781039, Assam, India.

Target prediction of compound 4a using Super PRED

SuperPred (https://prediction.charite.de/index.php) is a web-based tool for predicting potential targets of small-molecule compounds based on the similarity of the compounds to known bioactive compounds and the likelihood of the compounds interacting with various proteins. The SMILES of compound 4a were submitted to Super PRED. The highest-ranked protein targets were screened for further analysis.

Protein expression analysis using the UALCAN database

THE PREDICTED PROTEINS WERE UPLOADED using UALCAN (https://ualcan.path.uab.edu/), and expression analysis on breast cancer was performed.[1] Box plots were created to visually represent the expression levels of these proteins in comparison to normal tissues, and statistical significance was assessed using UALCAN's integrated tools.

Gene enrichment analysis

Gene enrichment analysis was performed on CTSD using the Enrichr web tool (https://maayanlab.cloud/Enrichr/) to identify biological pathways and functional annotations associated with CTSD. The significance levels and p-values were computed to assess the enrichment results.

Molecular docking

CTSD (PDB ID: 4OD9) crystal structure was obtained from the Protein Data Bank (https://www.rcsb.org/) and used as the receptor for molecular docking. The Swiss PDB viewer was used to perform protein-energy minimization. The protein and ligand structures were prepared with Autodock tools (MGLTools 1.5.7), and blind docking was performed for all compounds, as well as site-specific docking for the control inhibitor by specifying the region of the protein where the ligand is expected to bind using a grid box that encloses the search space. The docking simulation was run, and the AutoDock Vina search algorithm was used to generate potential binding poses for the ligand within the defined search space.[2] The docking simulation results are analysed to determine the ligand's best binding pose and binding affinity to the protein. Discovery Studio generated the 2D interaction profile of protein-ligand complexes.

Cell Lines and Culture Condition

Breast cancer cell line MCF7 was purchased from the National Centre of Cell Science (NCCS) cell repository in Pune, India. It was cultured in Dulbecco's Modified Eagle's medium-high glucose supplemented with l-glutamine and sodium pyruvate (Sigma Aldrich). Additionally, sodium bicarbonate, 10% fetal bovine serum (FBS) (Gibco, USA), 100 units/mL and 100 µg/mL penicillin and streptomycin, respectively (Thermo Fisher Scientific, USA) were added. The cells were maintained under humidified air conditions comprising 5% CO_2 at 37°C.

Cell viability assay

To assess the effect of the selected compounds on the viability of breast cancer cells, MTT (3-[4, 5-dimethylthiazol-2-yl]-2, 5-diphenyl tetrasodium bromide) cell viability assay was performed (ref 3). This assay's principle is based on converting coloured tetrazolium dye to purple-coloured insoluble formazan crystals by the metabolically active living cells, measured by the change in absorbance.[3] Standard protocol for MTT was used to measure cell viability. Briefly, $3x10^3$ cells/well were seeded in a 96-well plate (ThermoScientific, India) and grown for 24 hours at 37°C under 5% CO_2. Next, cells were treated with increasing concentrations of the compounds and incubated for 48 hours at 37°C under 5% CO_2. Following incubation, MTT solution was added to each well and incubated for 1.50 hours. The formazan crystals (MTT metabolic product) thus formed were dissolved in dimethyl sulfoxide (DMSO), and the absorbance was taken using a microplate reader (Infinite 200 PRO, Tecan, Germany) at 570 nm along with a reference wavelength at 630 nm. For the calculation, taking the cell viability of the untreated cells as 100%, the compounds' 50% inhibitory concentration (IC50) was determined from the *dose*-response curve using the GraphPad Prism. The formula used for the cell viability (%) determination is as follows:

$$\%\ Cell\ viability = \frac{[(Abs_{570} - Abs_{630})\ treated\ cells]}{[(Abs_{570} - Abs_{630})\ control\ cells]} X\ 100$$

Reactive Oxygen Species Detection Assay

The level of reactive oxygen species (ROS) has been known to affect various aspects of tumour development, progression, and apoptosis. For the study, cellular ROS generation has been assessed using the DCFDA (2′,7′-dichlorofluorescein diacetate) dye. DCFDA is a non-fluorescent, ROS-sensitive dye and the precursor of DCF (2′,7′-dichlorofluorescein). On entering the cell, DCFDA is oxidized to DCF in the presence of several cellular reactive oxygen species within the cells, and DCF could be monitored by an increase in fluorescence.[4] For the study, cells were treated with the compounds for 6h and incubated with 2μM DCFDA for 30 min. Following this, the cells were trypsinized, washed once with 1X PBS, and resuspended in PBS. The generation of ROS was identified by flow cytometry analysis in the FITC channel using the CytoFLEX (Beckman Coulter) flow cytometer.

Live-dead cell imaging

For live-dead cell imaging, MCF-7 cells were treated with the parent compound 4a and the selected compounds 4d, 4l and 4p at their respective IC_{50} concentration for 48h. Following treatment, cells were incubated for 30 mins with 2 μM calcein-AM and 4 μM PI. After the incubation, the live-dead populations were imaged using a ZOE Fluorescent Cell Imager fluorescence microscope (Bio-Rad).

Table S1. Crystallographic parameters for **4d** and **4m**.

	4d	**4m**
CCDC No.	2284128	2284121
Empirical formula	$C_{24}H_{19}ClO_4$	$C_{25}H_{21}ClO_4$
Formula weight	406.868	420.895
Temperature/K	296.00	298
Crystal system	monoclinic	triclinic
Space group	$P2_1/c$	P-1
a/Å	9.1059(15)	9.3610(7)
b/Å	21.916(4)	11.2049(8)
c/Å	10.1282(17)	11.6661(9)
α/°	90	103.780(3)
β/°	105.051(5)	93.986(3)
γ/°	90	114.062(2)
Volume/Å³	1951.9(6)	1065.83(14)
Z	4	2
ρcalc(g/cm³)	1.385	1.311
μ/mm⁻¹	0.225	0.208
F(000)	849.1	440.6
Radiation	Mo Kα (λ = 0.71073)	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	3.72 to 52.46	4.16 to 51.34
Index ranges	$-11 \le h \le 11$, $-26 \le k \le 27$, $-12 \le l \le 12$	$-11 \le h \le 11$, $-13 \le k \le 13$, $-14 \le l \le 14$
Reflections collected	41936	25693
Independent reflections	3852	4032
Data/restraints/parameters	3852/0/264	4032/0/274
Goodness-of-fit on F²	1.049	1.027
Final R indexes [I>=2σ (I)]	$R_1 = 0.0896$, $wR_2 = 0.2752$	$R_1 = 0.0538$, $wR_2 = 0.1478$
Final R indexes [all data]	$R_1 = 0.1056$, $wR_2 = 0.2848$	$R_1 = 0.0698$, $wR_2 = 0.1572$

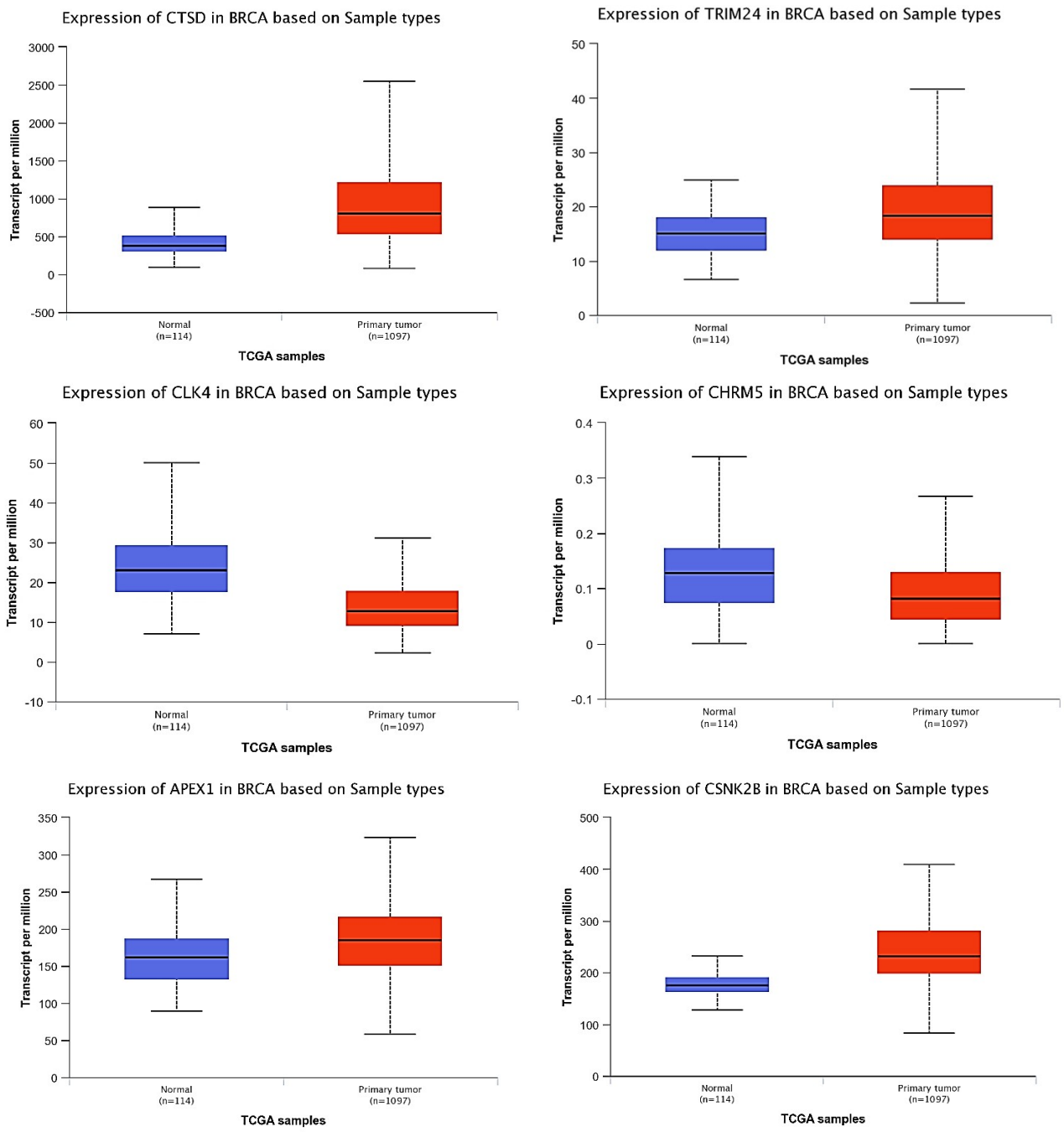


Fig. S1 Gene expression analysis of predicted targets.

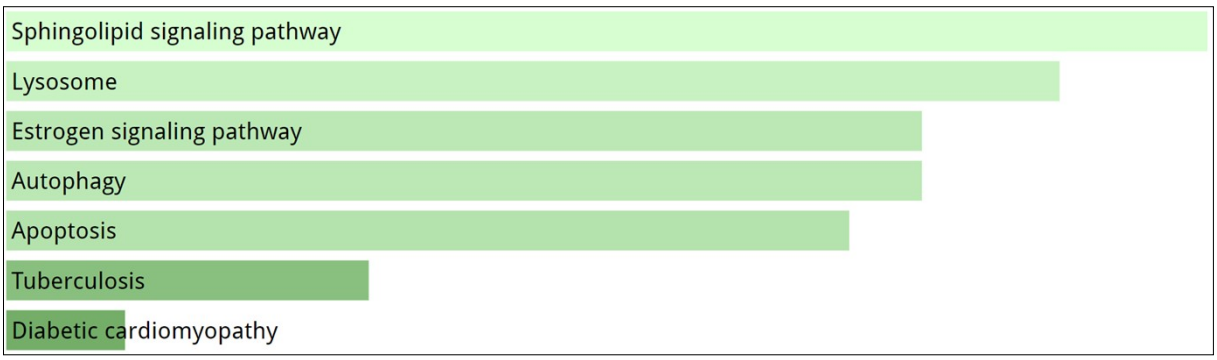
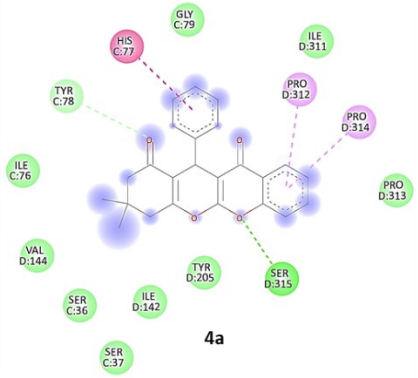
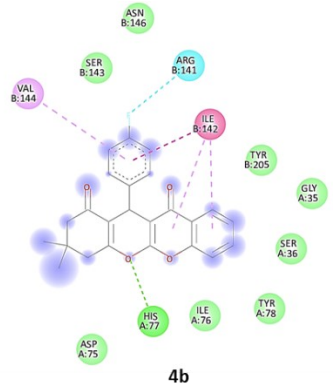


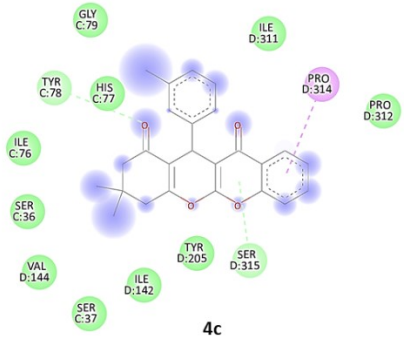
Fig. S2 gene enrichment analysis of CTSD.



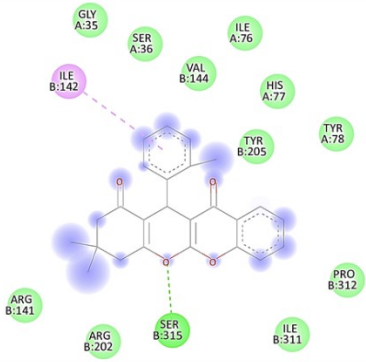
4a



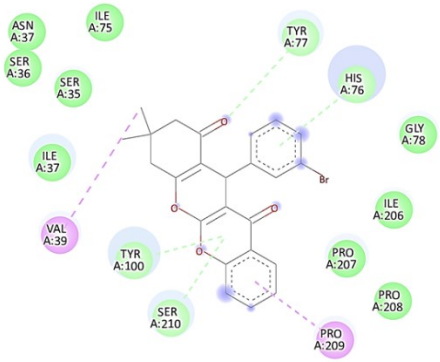
4b



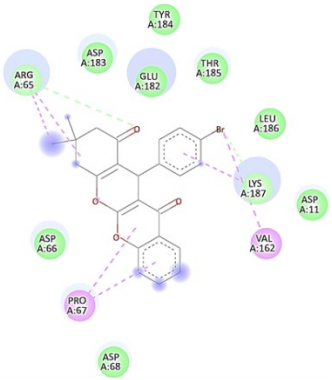
4c



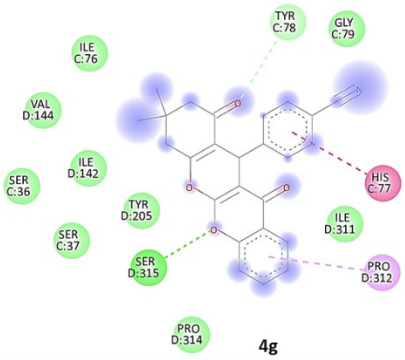
4d



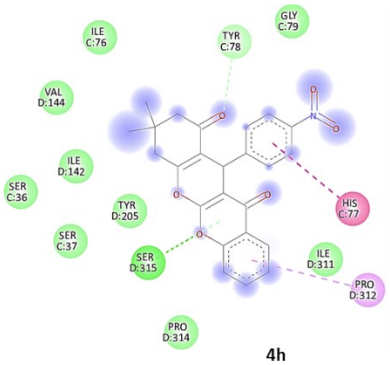
4e



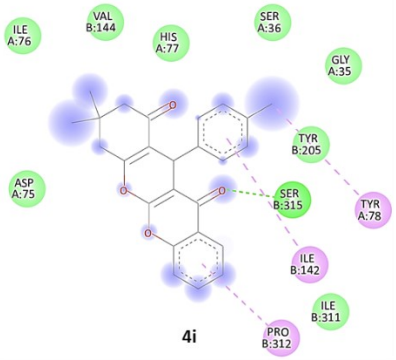
4f



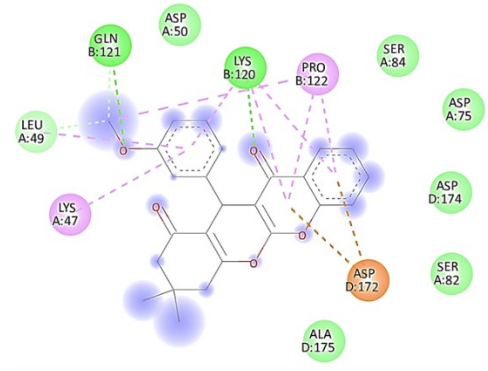
4g



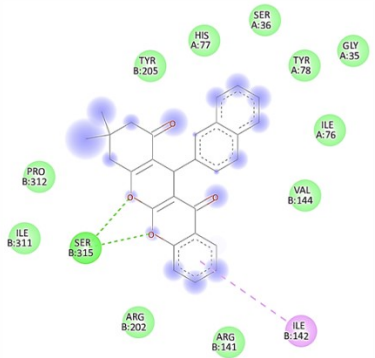
4h



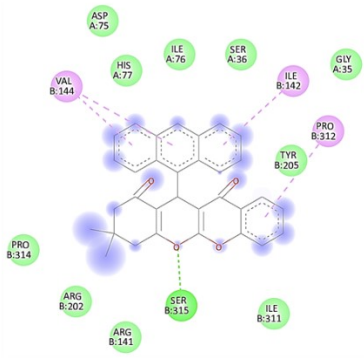
4i



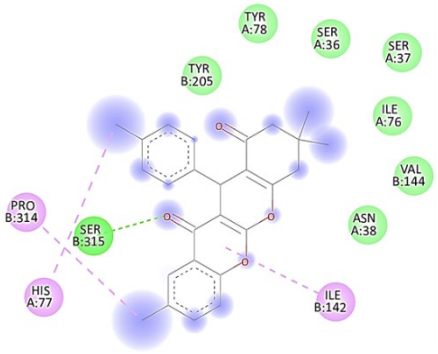
4j



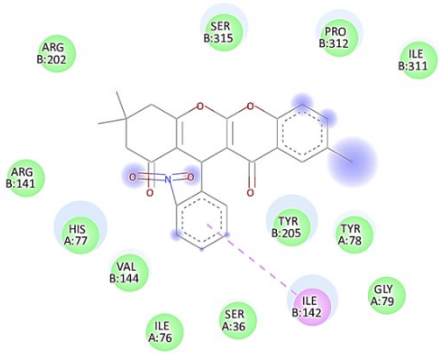
4k



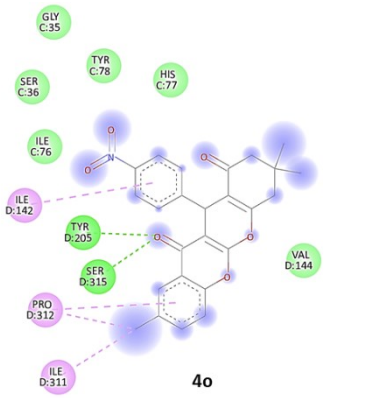
4l



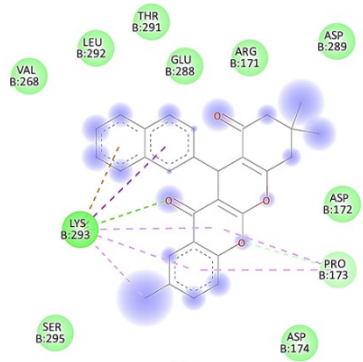
4m



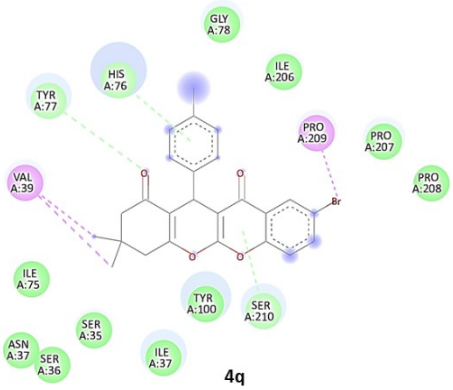
4n



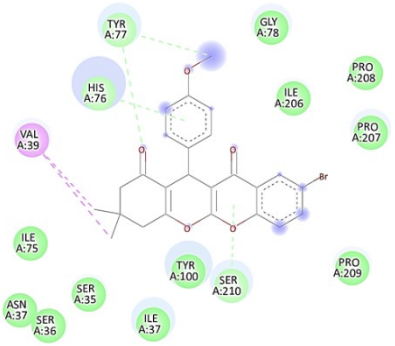
4o



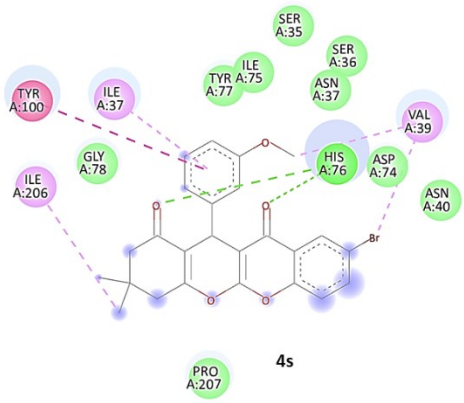
4p



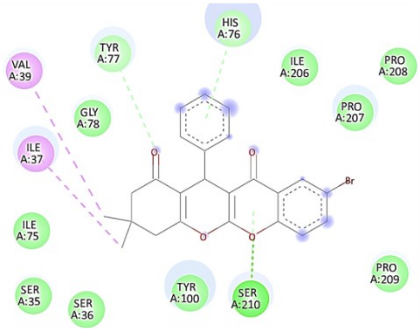
4q



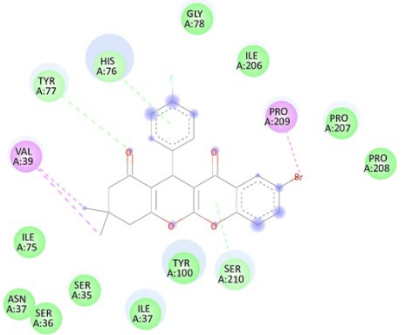
4r



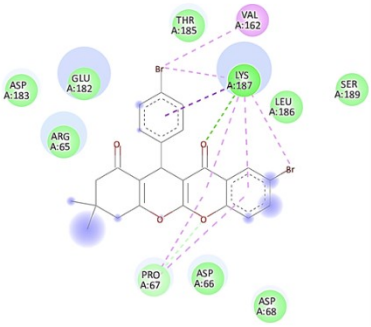
4s



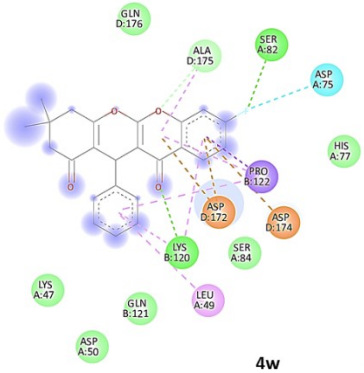
4t



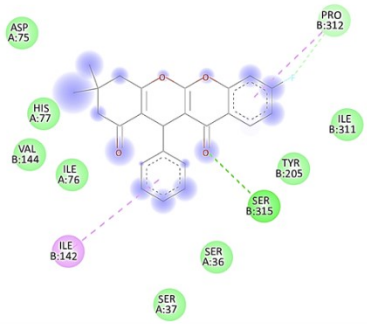
4u



4v



4w



4x

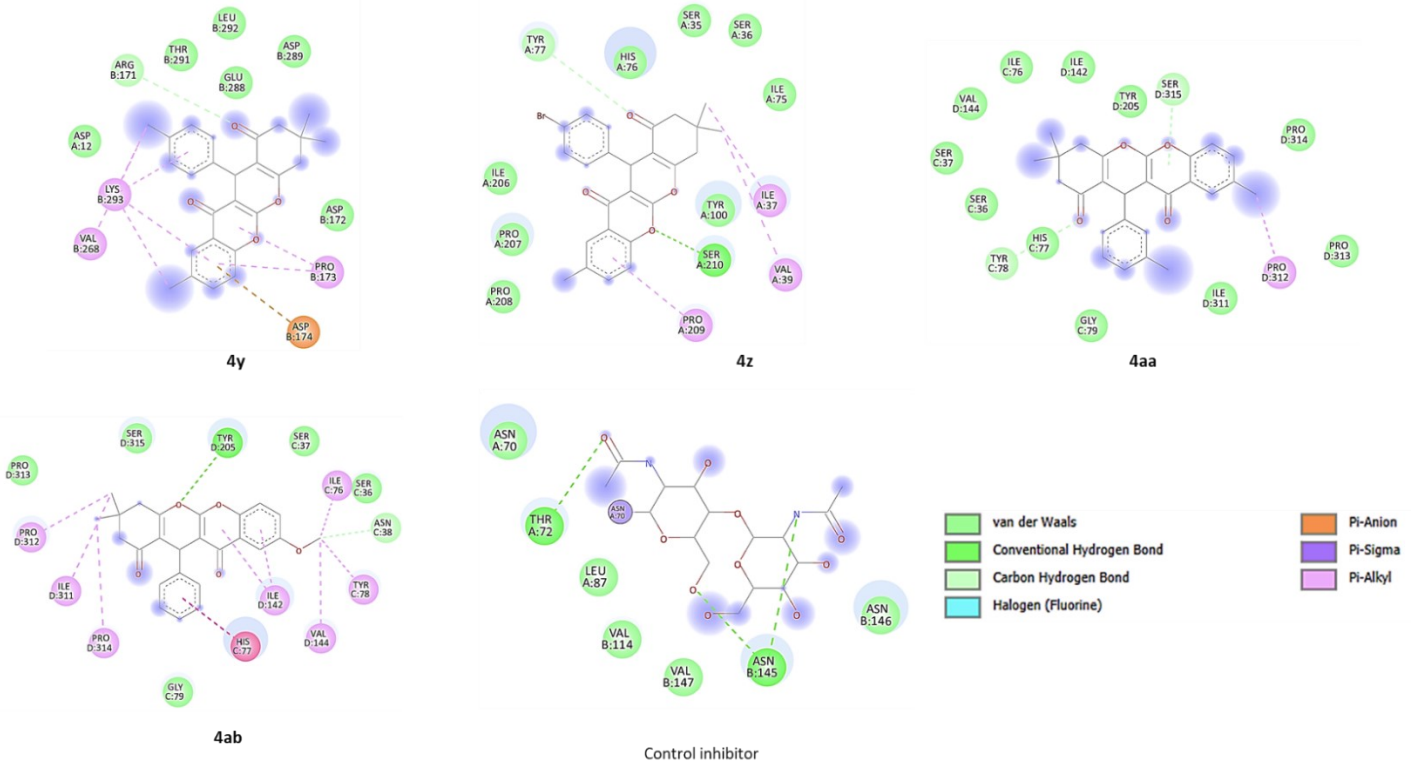


Fig. S3. 2D interaction profiles of compounds with CTSD.

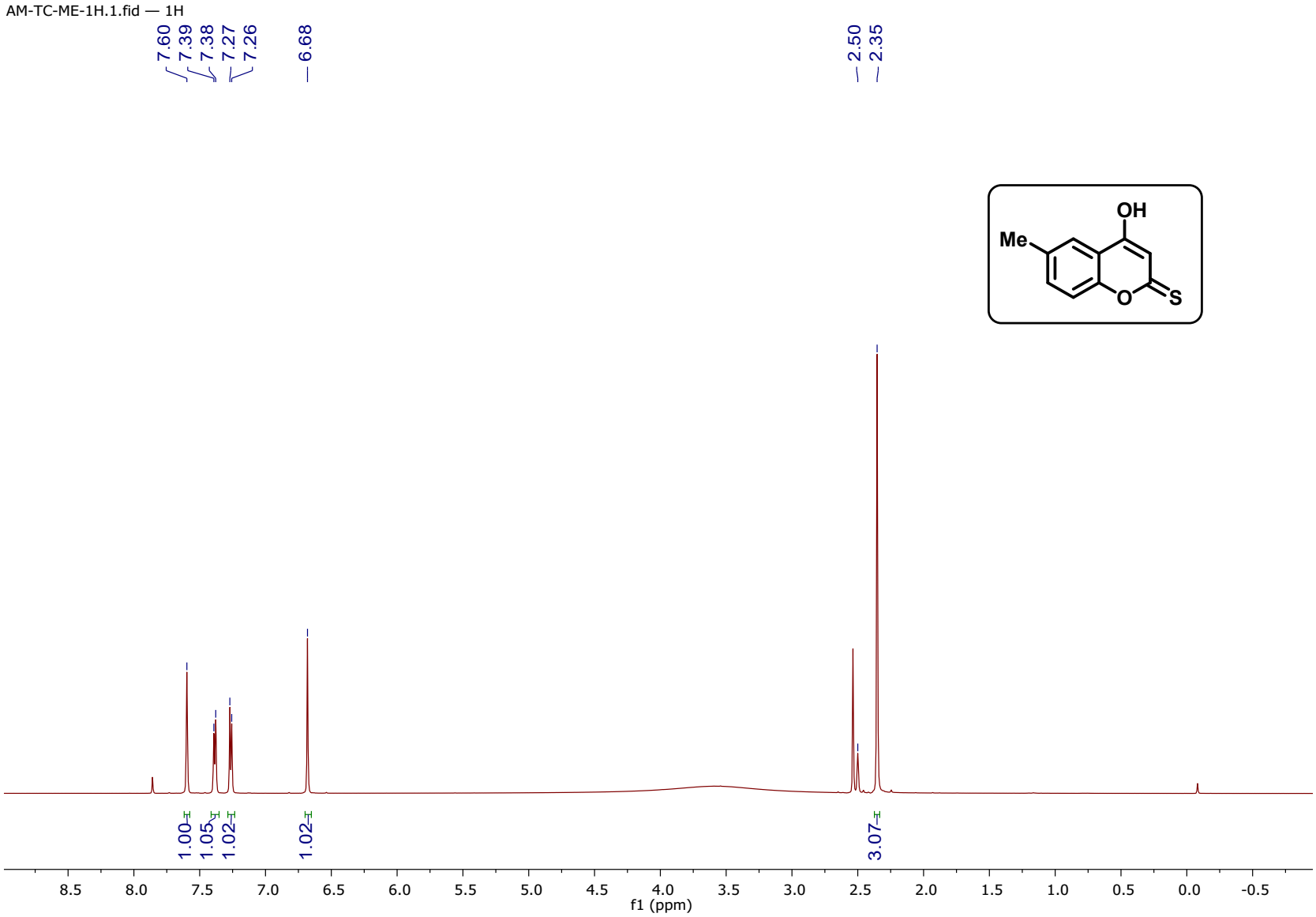
Table S2. Molecular docking analysis of compounds

Compound	Binding energy	No of H bonds	Interacting amino- acids in H bond	No of other bonds	Interacting amino-acids in the non-covalent interactions
4a	**-7.7**	1	SER315	13	PRO313, TYR205, 1LE142, SER37, SER36, VAL144, ILE76, TYR78, HIS77, GLY79, ILE311, PRO312, PRO314
4b	-8.6	1	HIS77	13	ILE76, TYR78, SER36, GLY35, TYR205, ILE142, ARG141, ASN146, SER143, VAL144, ASP75
4c	-8.3	0		13	PRO314, PRO312, SER315, TYR205, ILE142, SER37, VAL144, SER36, ILE76, TYR78, HIS77, GLY79, ILE311
4d	**-9.3**	1	SER315	12	ILE311, PRO312, TYR78, TYR205, HIS77, ILE76, VAL144, SER36, GLY35, ILE142, ARG141, ARG202
4e	-8.1	0		15	PRO209, PRO208, PRO207, ILE206, GLY78, HIS76, TYR77, ILE75, SER35, ASN37, SER36, ILE37, VAL39, TYR100, SER210
4f	-8.5	0		16	ASP66, PRO67, ASP68, VAL162, LYS187, ASP11, LEU186, THR185, GLU182, TYR184, ASP183, ARG65
4g	-7.7	1	SER315	12	PRO314, PRO312, ILE311, HIS77, GLY79, TYR78, ILE76, VAL144, ILE142, SER36, SER37
4h	-7.9	1	SER315	12	PRO314, PRO312, ILE311, HIS77, GLY79,TYR78, ILE76, VAL144, ILE142, SER36, SER37, TYR205

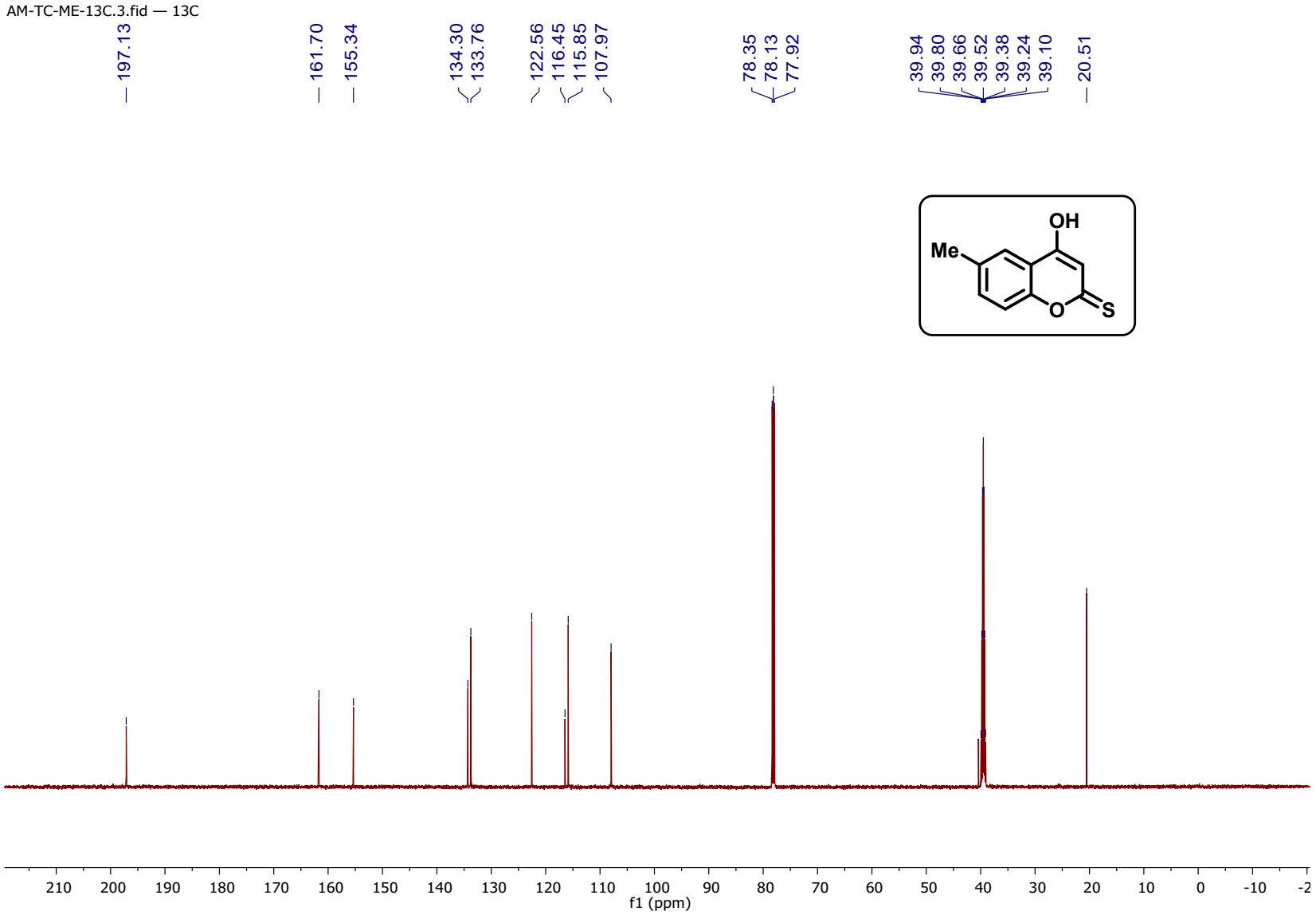
4i	-8.6	1	SER315	11	TYR205, TYR78, GLY35, SER36, HIS77, VAL144, ILE76, ASP75, PRO312, ILE311, ILE142
4j	-8.3	2	GLN121, LYS120	18	ASP50, LEU49, LYS47, ALA175, ASP172, SER82, ASP174, ASP75, SER84, PRO122, LYS120
4k	-8.8	2	SER315	12	ARG202, ARG141, ILE142, VAL144, ILE76, GLY35, TYR78, SER36, HIS77, TYR205, PRO312, ILE311
4l	**-9.8**	1	SER315	14	ILE311, TYR205, PRO312, GLY35, ILE142, SER35, ILE76, HIS77, ASP75, VAL144, PRO314, ARG202, ARG141
4m	-8.0	1	SER315	10	ILE142, ASN38, VAL144, ILE76, SER37, SER36, TYR78, TYR205, PRO314, HIS77
4n	-8.6	0		13	ILE142, GLY79, TYR78, TYR205, ILE311, PRO312, SER315, ARG202, ARG141, HIS77, VAL144, ILE76, SER36
4o	-7.9	2	SER315, TYR205	10	PRO312, ILE311, VAL144, HIS77, TYR78, GLY35, SER36, ILE76, ILE142
4p	**-9.1**	1	LYS293	17	LYS293, VAL268, LEU292, THR291, GLU288, ARG171, ASP289, ASP172, PRO173, ASP174, SER295
4q	-8.1	0		16	GLY78, ILE206, PRO209, PRO207, PRO208, SER210, TYR100, ILE37, SER35, SER36, ASN37, ILE75, VAL39, TYR77, HIS76
4r	-7.7	0		17	VAL39, ILE75, ASN37, SER36, SER35, ILE37, TYR100, SER210, PRO209, PRO207, PRO208, ILE206, GLY78, TYR77, HIS76
4s	-7.8	2	HIS76	14	ILE206, GLY78, TYR100, ILE37, TYR77, ILE75, SER35, SER36, ASN37, VAL39, ASP74, ASN40, PRO207
4t	-7.8	1	SER210	13	PRO209, PRO207, PRO208, ILE206, HIS76, TYR77, GLY78, VAL39, ILE37, ILE75, SER35, SER36, TYR100
4u	-8.0	0		16	VAL39, TYR77, HIS76, GLY78, ILE206, PRO209, PRO207, PRO208, SER210, TYR100, ILE37, SER35, SER36, ASN37, ILE75
4v	-8.0	1	LYS187	17	LEU186, SER189, ASP68, ASP66, PRO67, ARG65, ASP183, GLU182, THR185, VAL162,
4w	-8.5	2	LYS120,SER82	17	LYS47, ASP50, GLN121, LEU49, LYS120, ASP172, SER84,ASP174, PRO122, HIS77, ASP75, ALA175,GLN176
4x	-8.6	1	SER315	11	ASP75, HIS77, VAL144, ILE76, ILE142, SER37, SER36, TYR205, ILE311, PRO312
4y	-8.8	0		15	VAL268, LYS293, ASP174, PRO173, ASP172, ASP289, GLU288, LEU292, THR291, ARG171, ASP12
4z	-8.1	1	SER210	12	PRO209, VAL39, ILE37, ILE75, SER36, SER35, HIS76, TYR77, ILE206, PRO207, PRO208
4aa	-8.3	0		14	PRO312, PRO313, PRO314,SER315, TYR205, ILE142, ILE76, VAL144, SER37, SER36, TYR78, HIS77, GLY79, ILE311
4ab	-7.8	1	TYR205	15	SER37, SER36, ASN38, ILE76, TYR78, VAL144,

					ILE142, HIS77, GLY79, PRO314, ILE311, PRO312, PRO313, SER315
Control - inhibitor(NAG)	-5.6	3	ASN145, THR72	6	ASN146, VAL147, VAL114, LEU87, ASN70

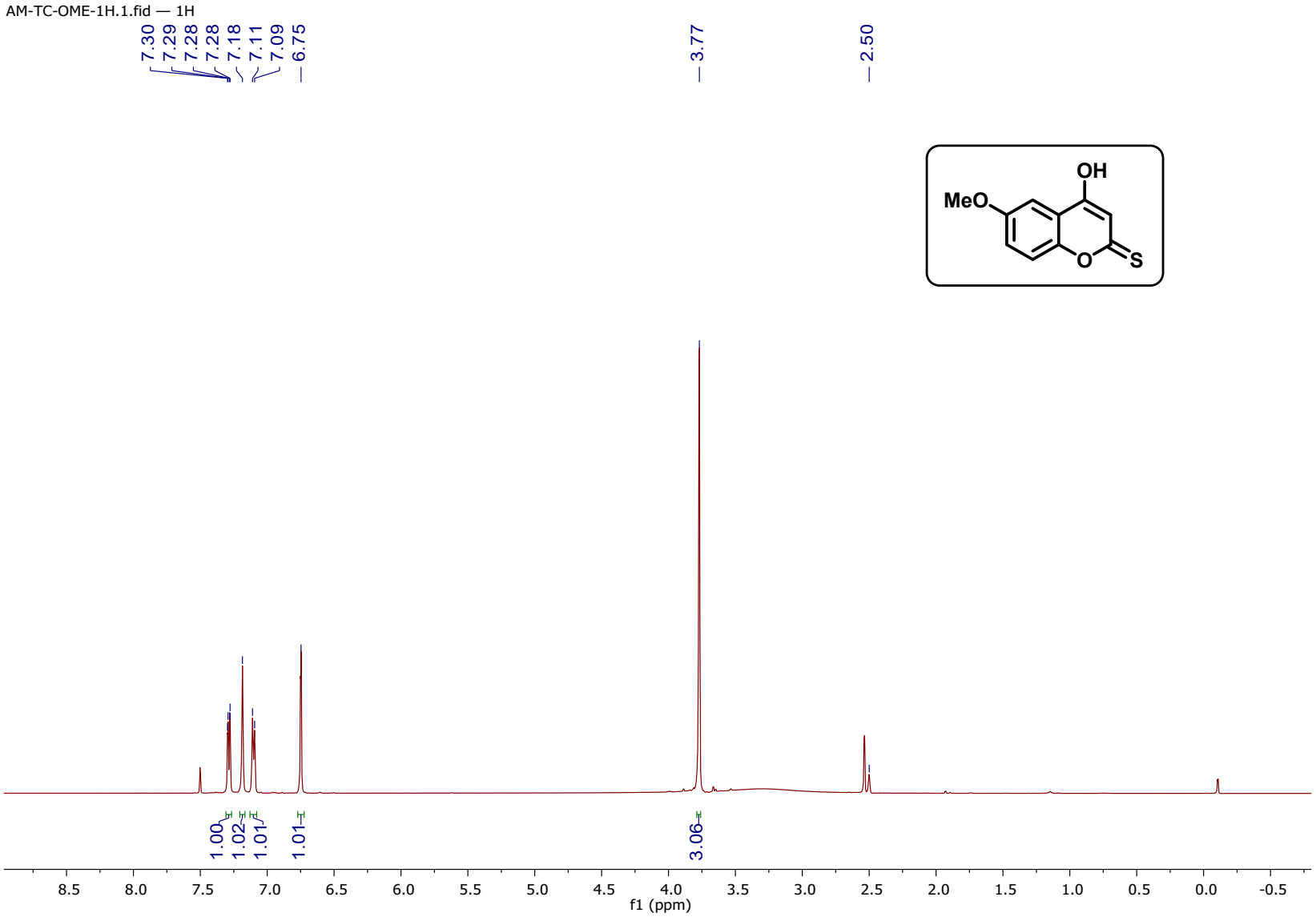
^{1}H NMR (600MHz, 1:1 $CDCl_3$ and D6-DMSO) spectrum of 4-hydroxy-6-methyl-2H-chromene-2-thione



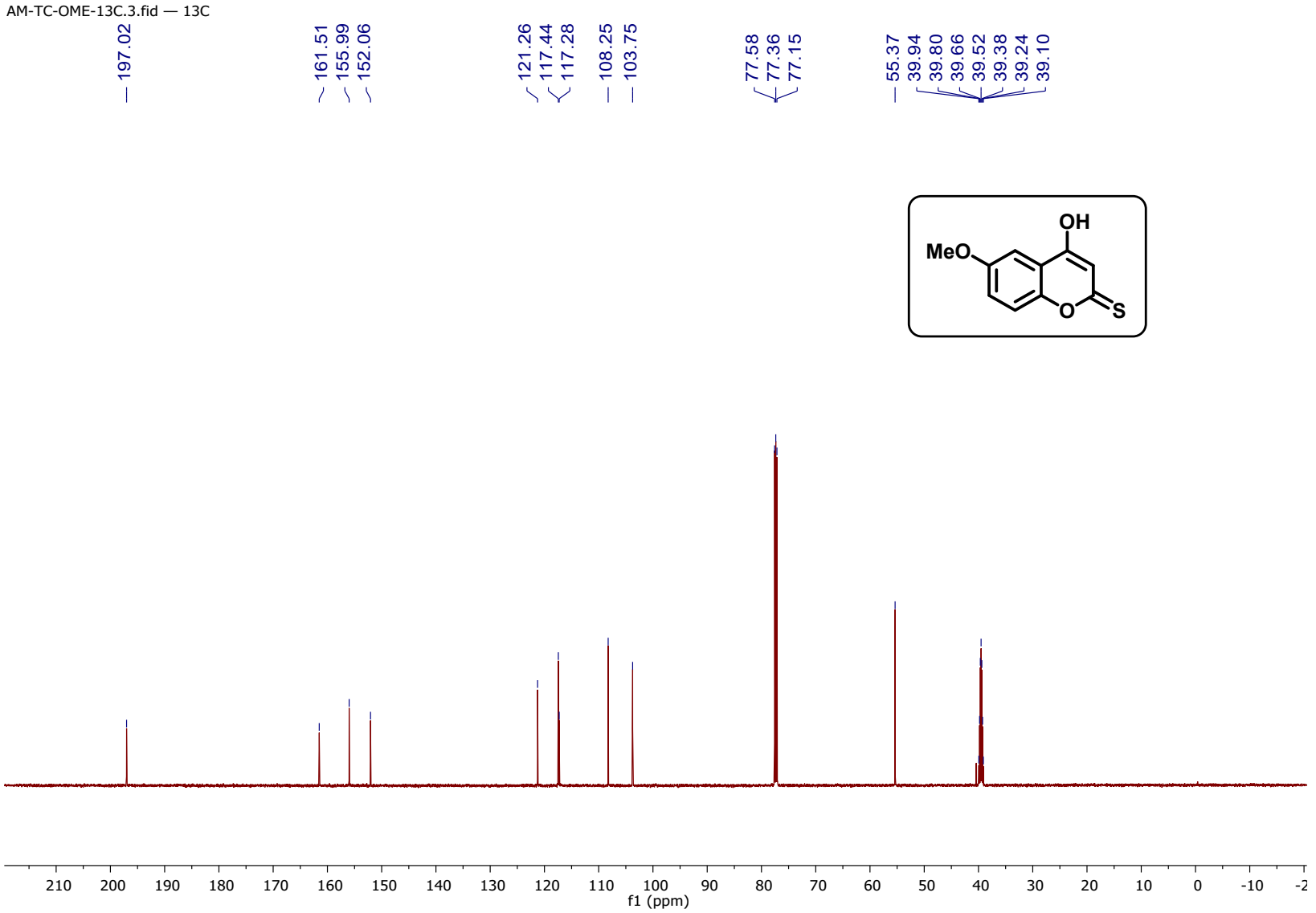
^{13}C {^{1}H} NMR (150MHz, 1:1$CDCl_3$ and D6-DMSO) spectrum of 4-hydroxy-6-methyl-2H-chromene-2-thione



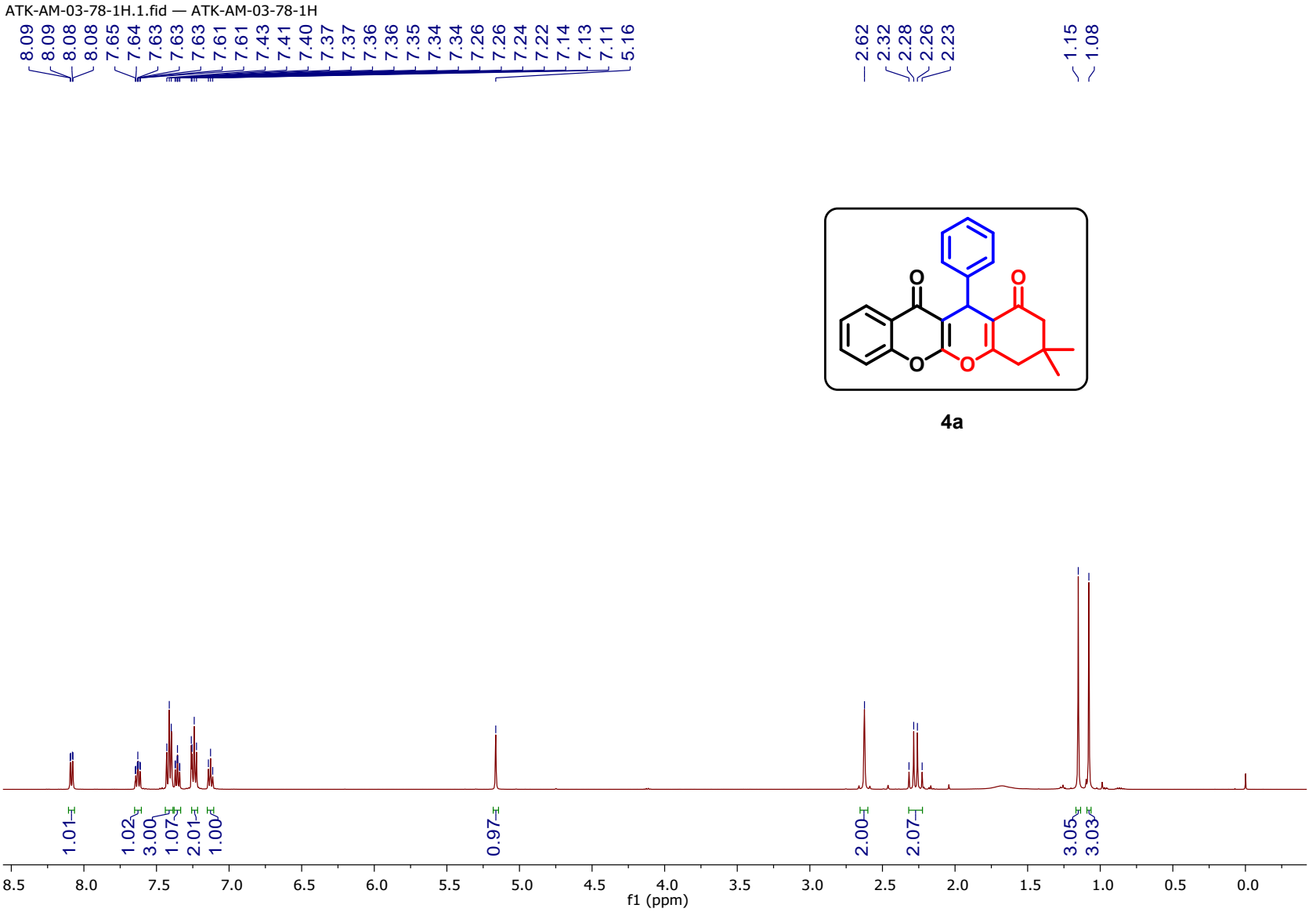
^{1}H NMR (600MHz, 1:1$CDCl_3$ and D6-DMSO) spectrum of 4-hydroxy-6-methoxy-2H-chromene-2-thione



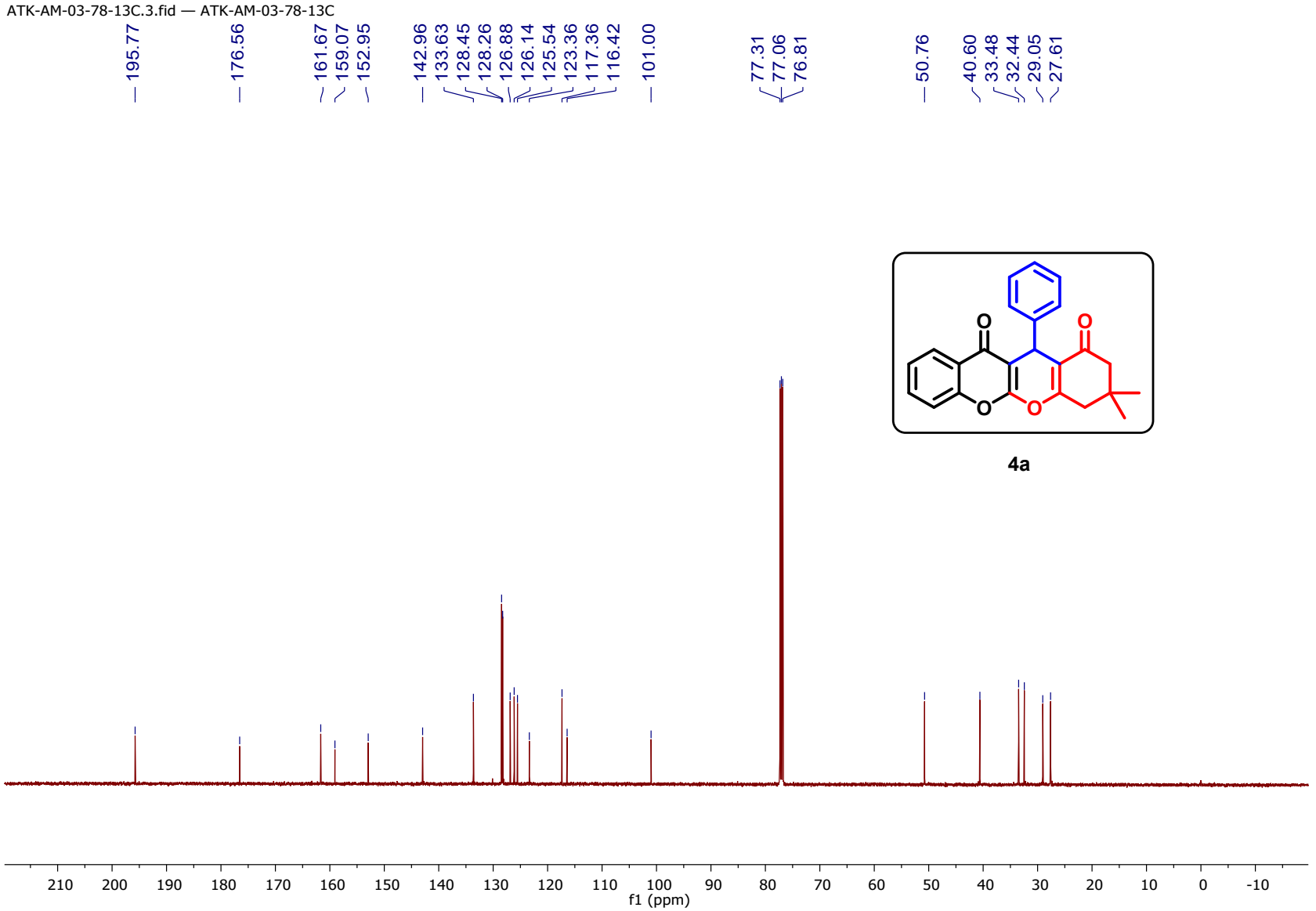
^{13}C {^{1}H} NMR (150MHz, 1:1$CDCl_3$ and D6-DMSO) spectrum of 4-hydroxy-6-methoxy-2H-chromene-2-thione



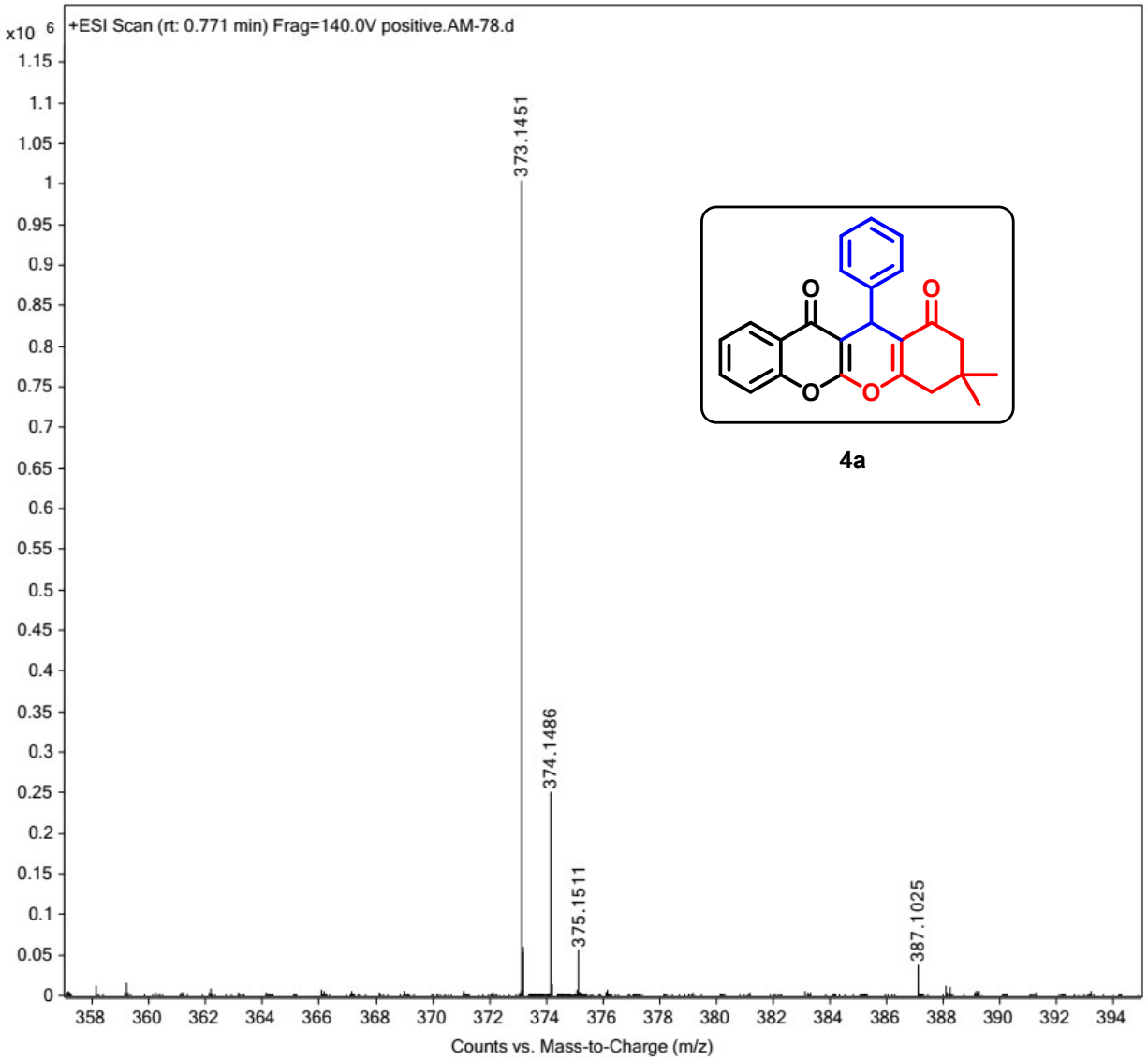
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4a



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4a

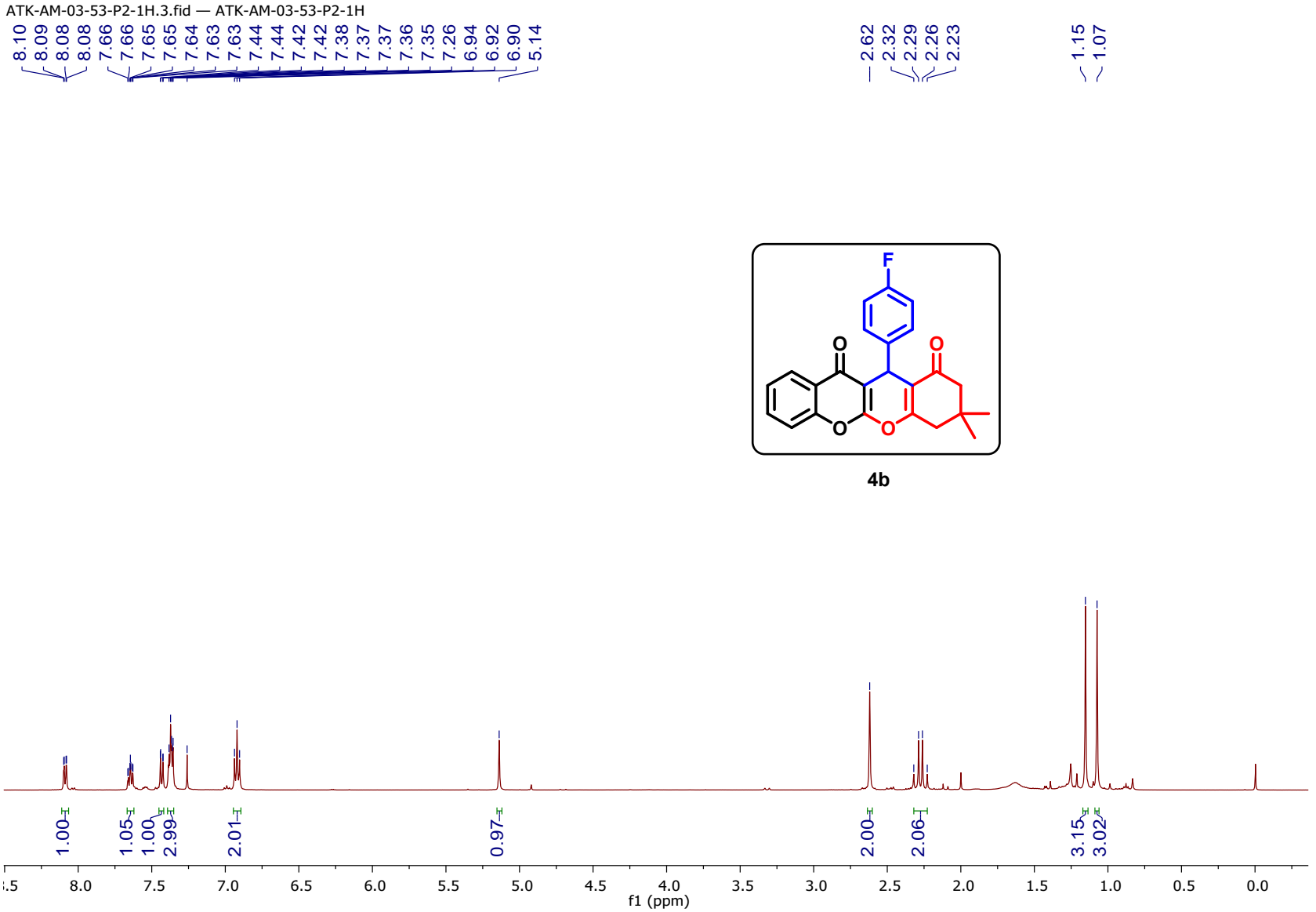


HRMS spectrum of compound 4a

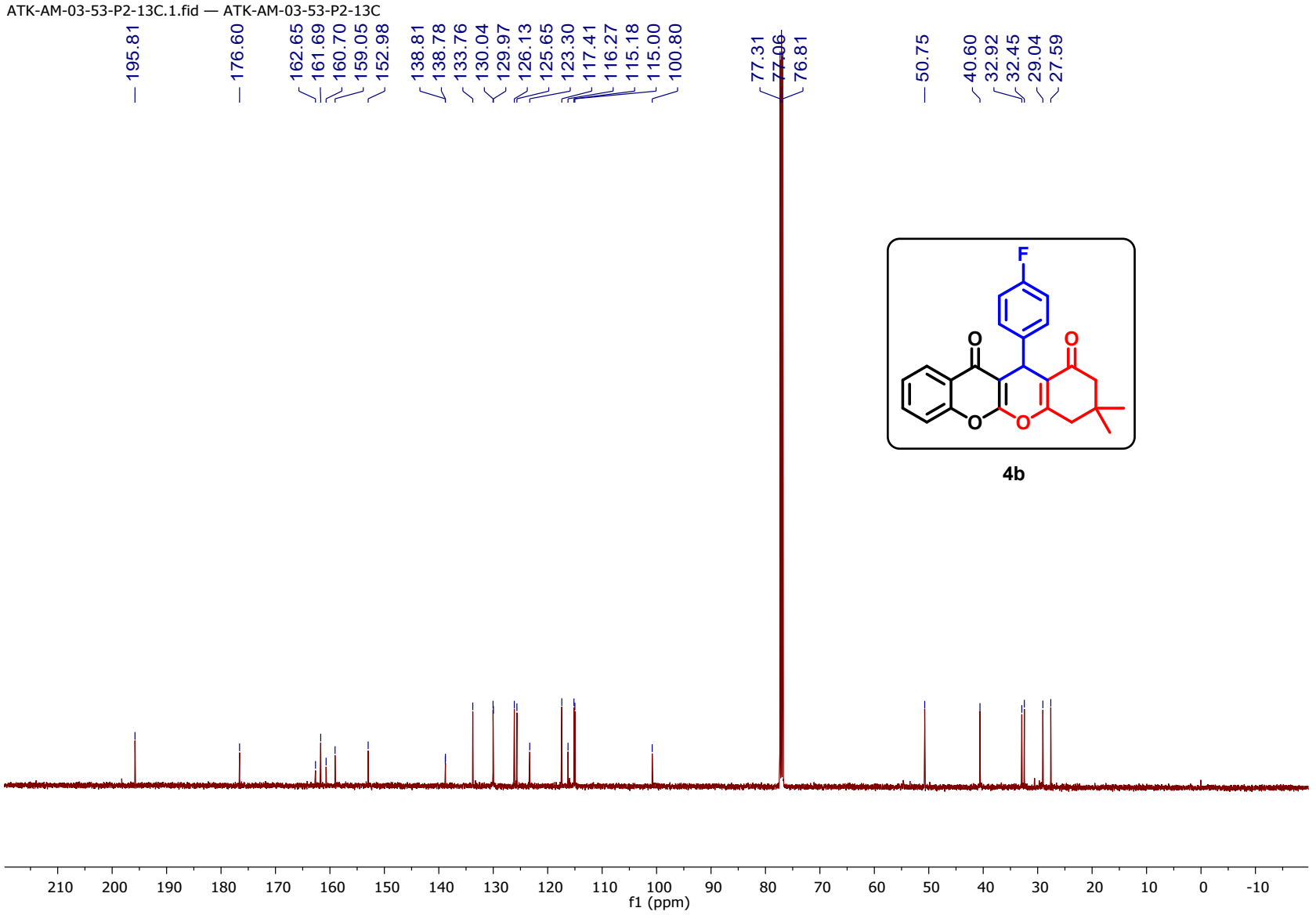


4a

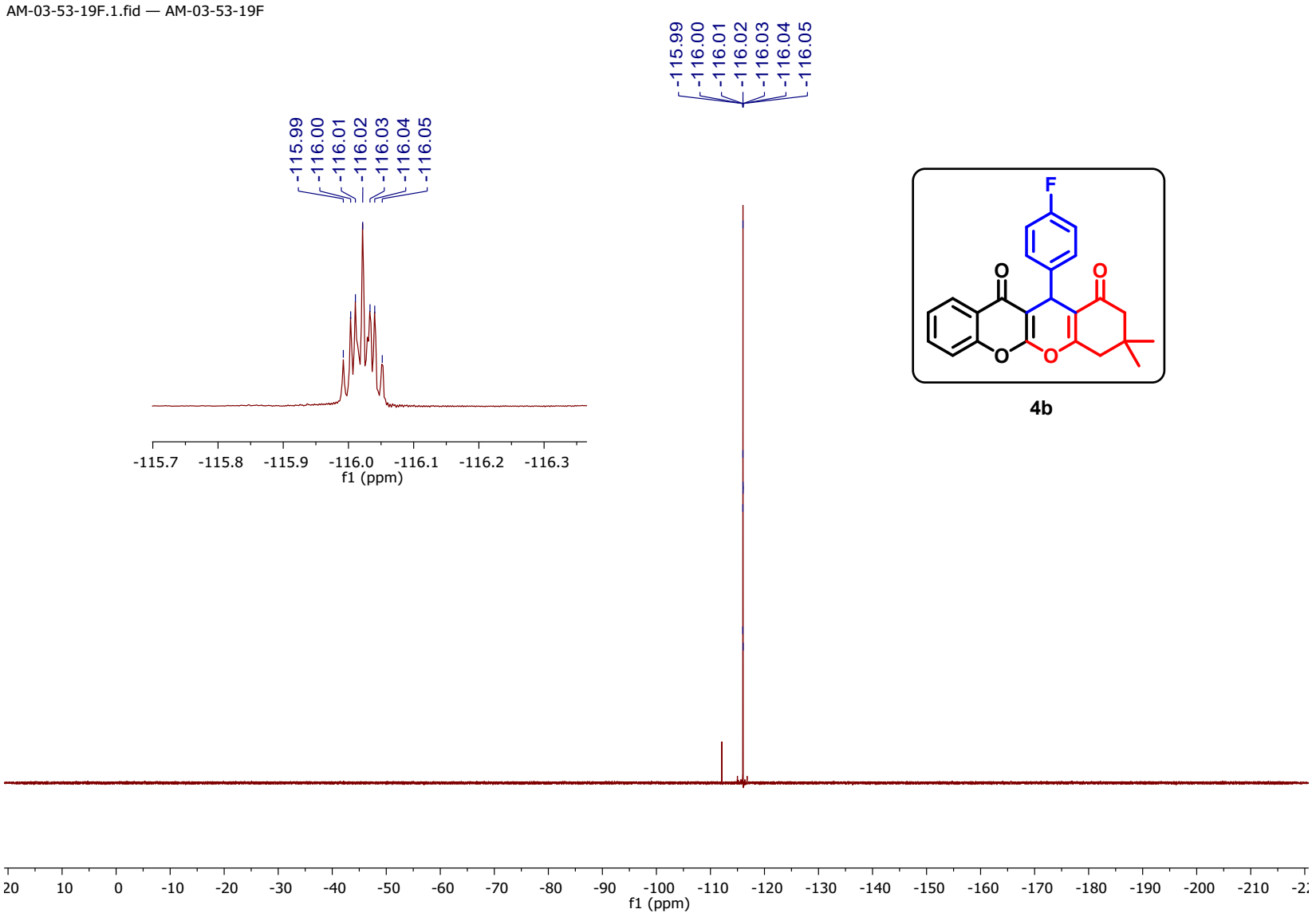
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4b



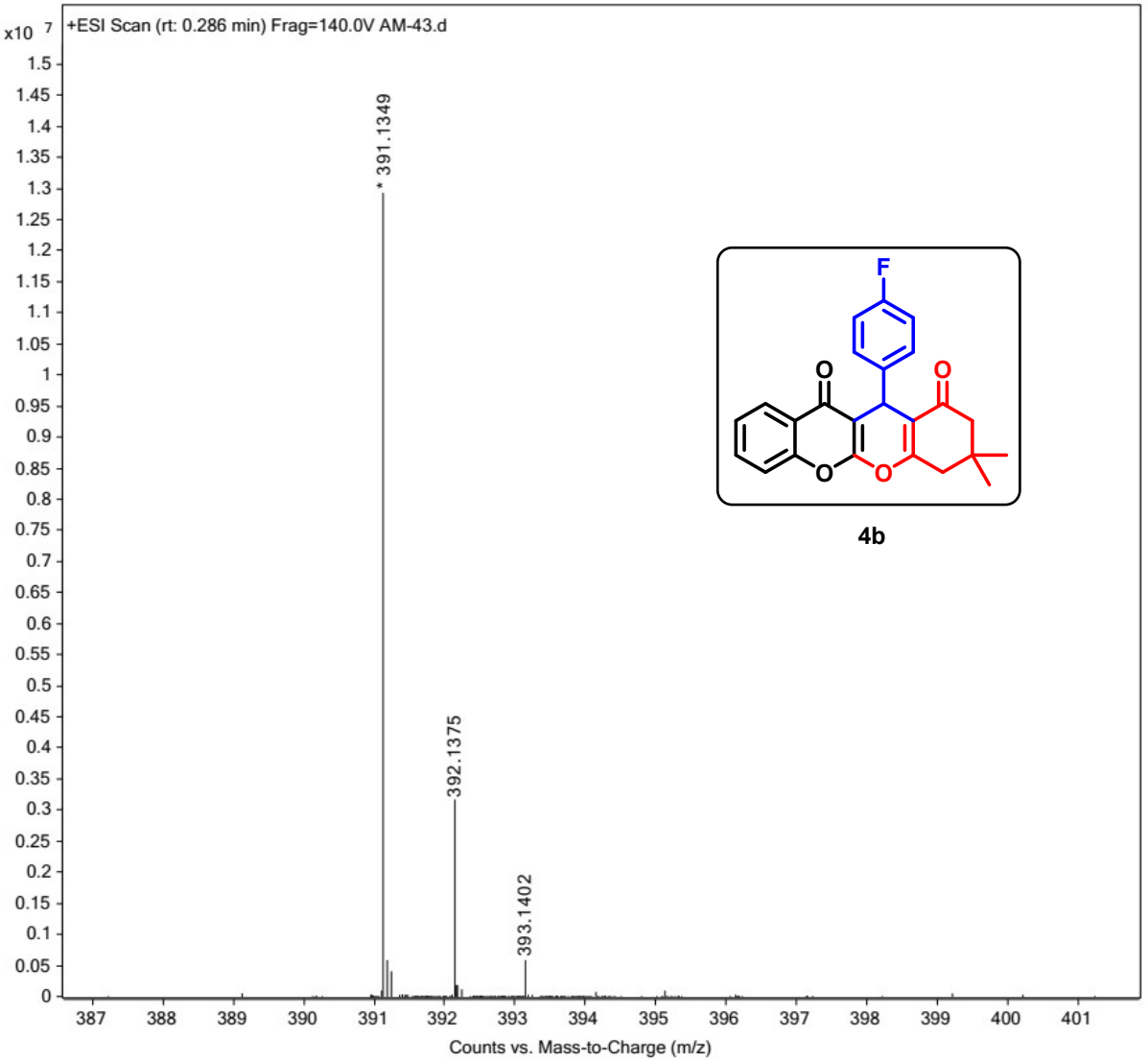
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4b



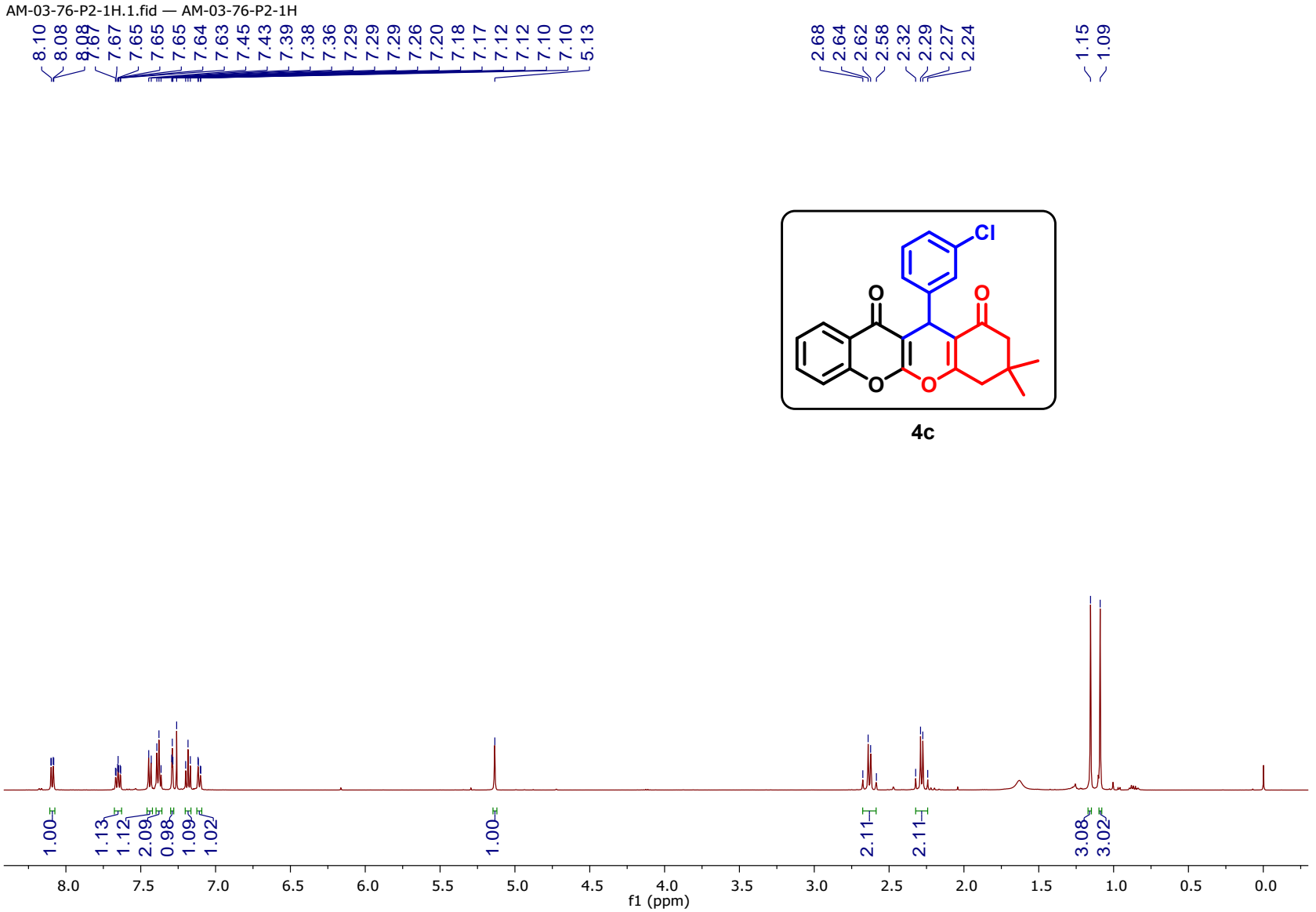
^{19}F NMR (471 MHz, $CDCl_3$) spectrum of compound 4b



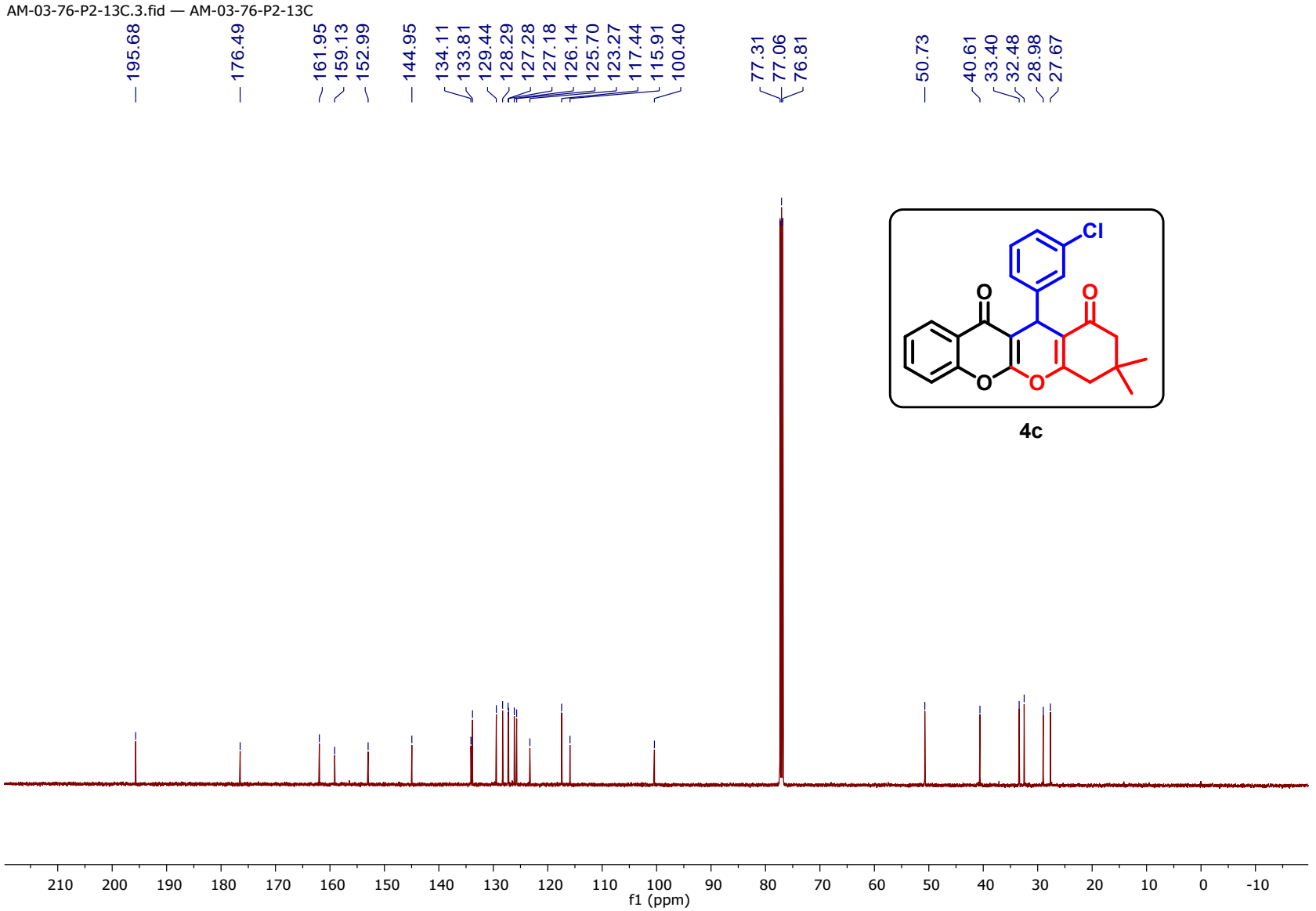
HRMS spectrum of compound 4b



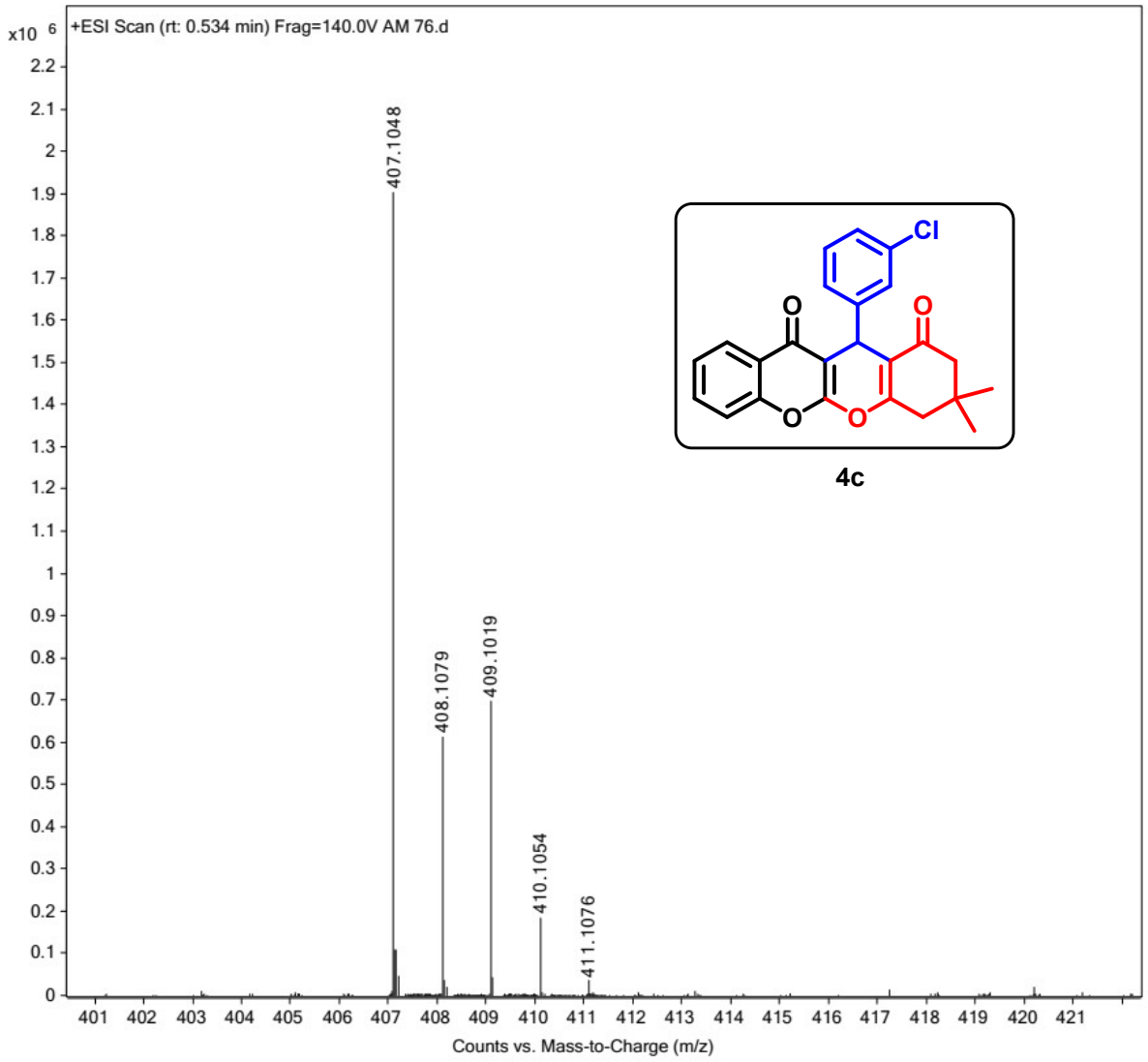
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4c



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4c

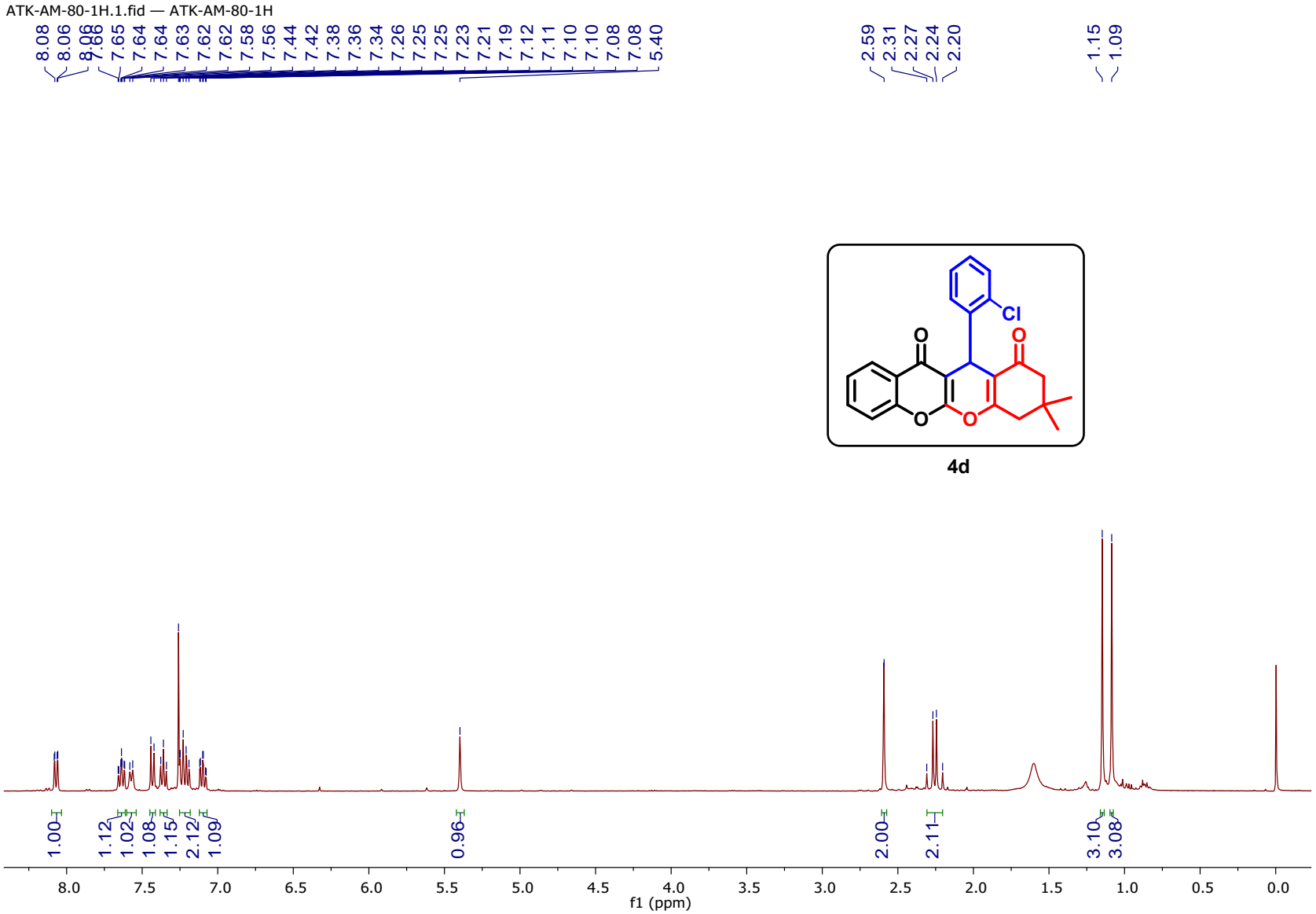


HRMS spectrum of compound 4c

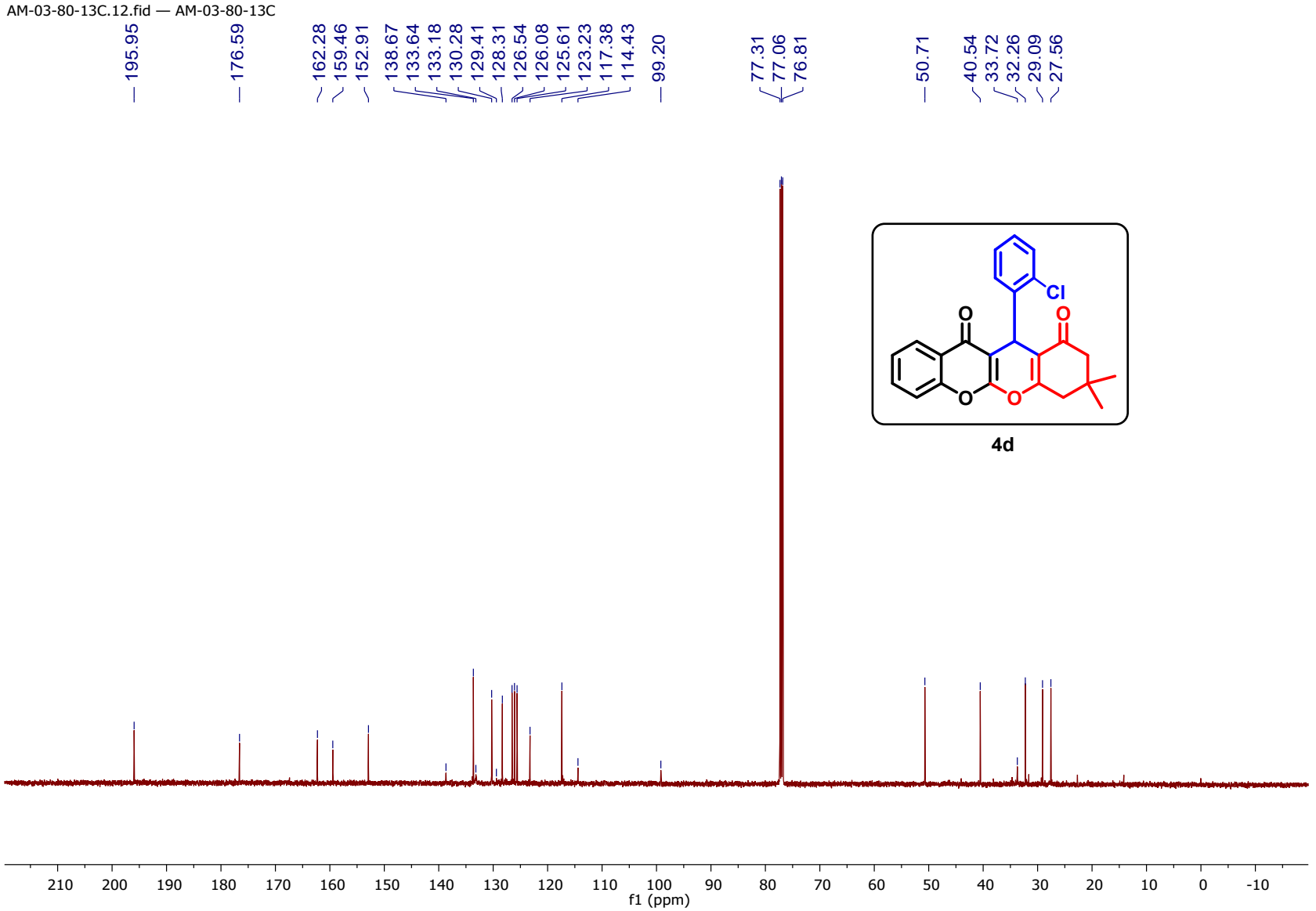


4c

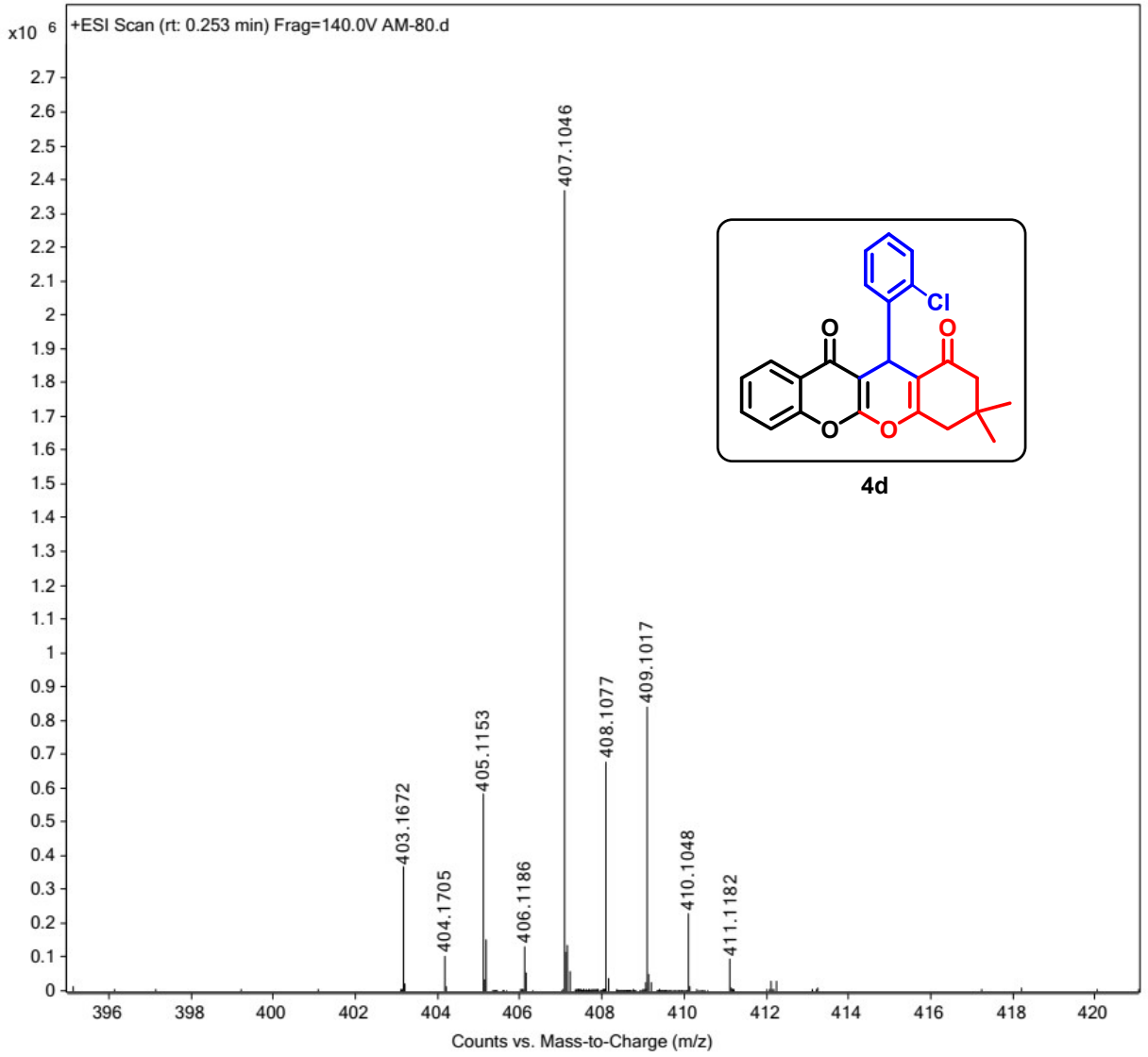
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4d



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4d

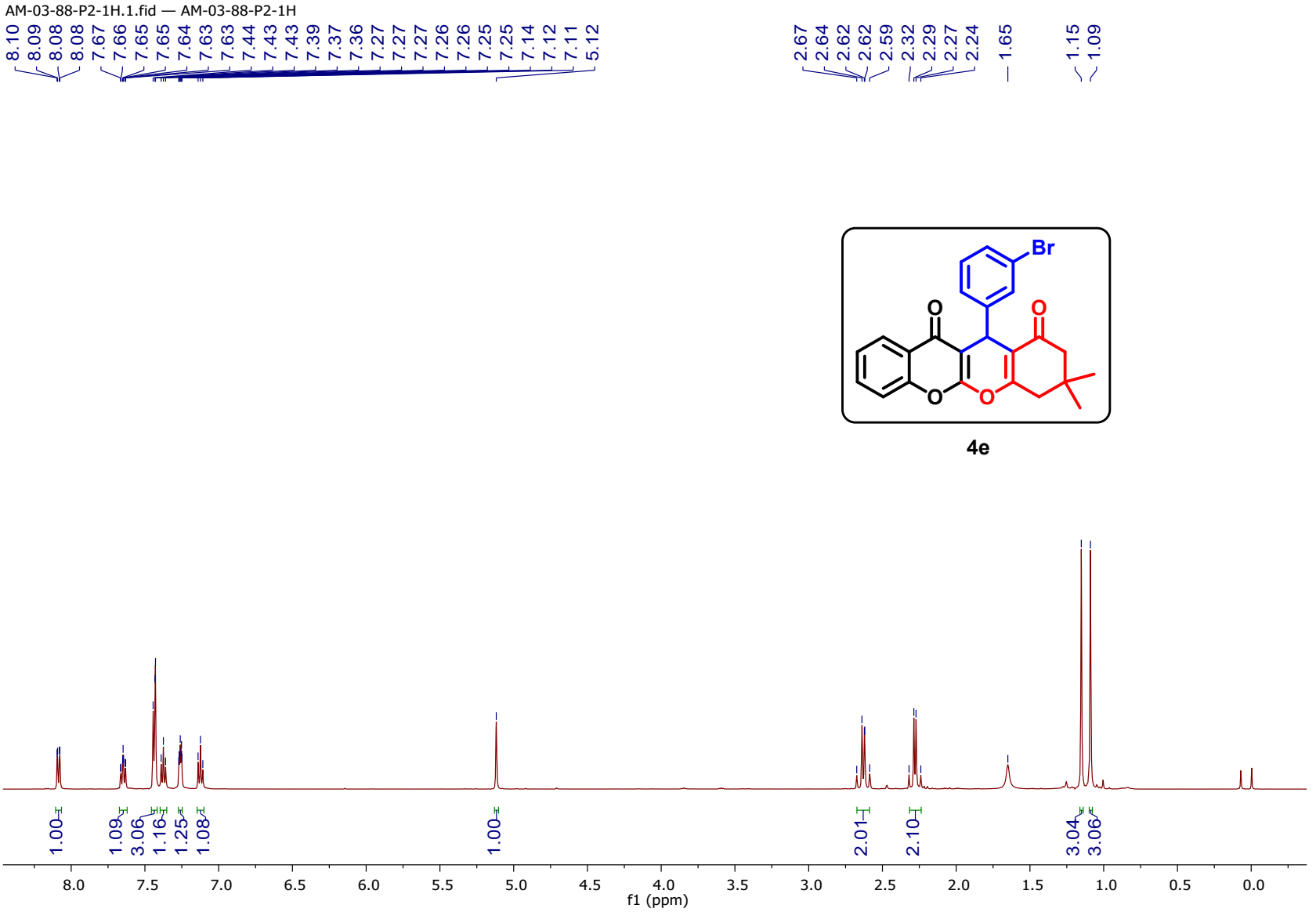


HRMS spectrum of compound 4d



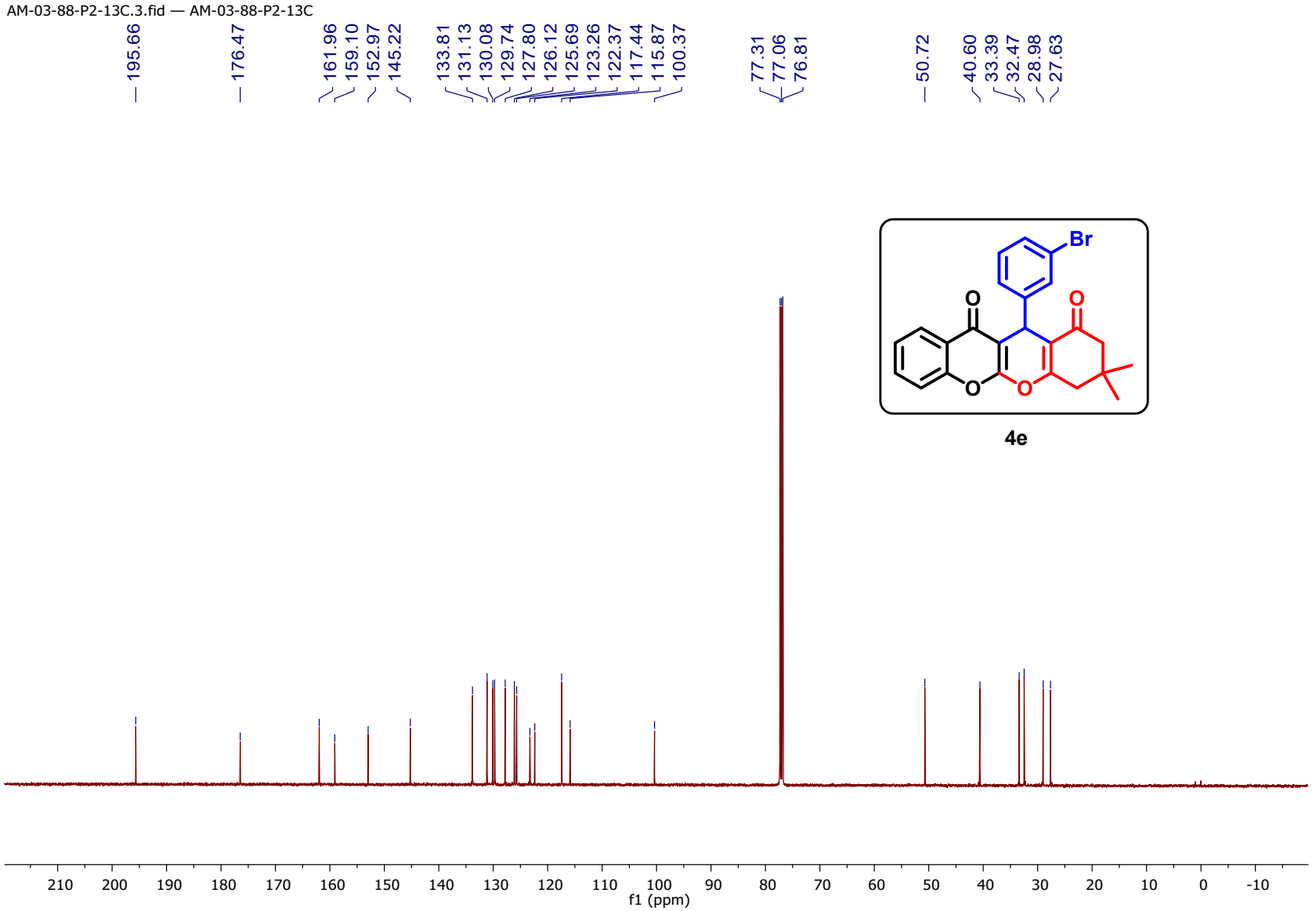
4d

^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4e

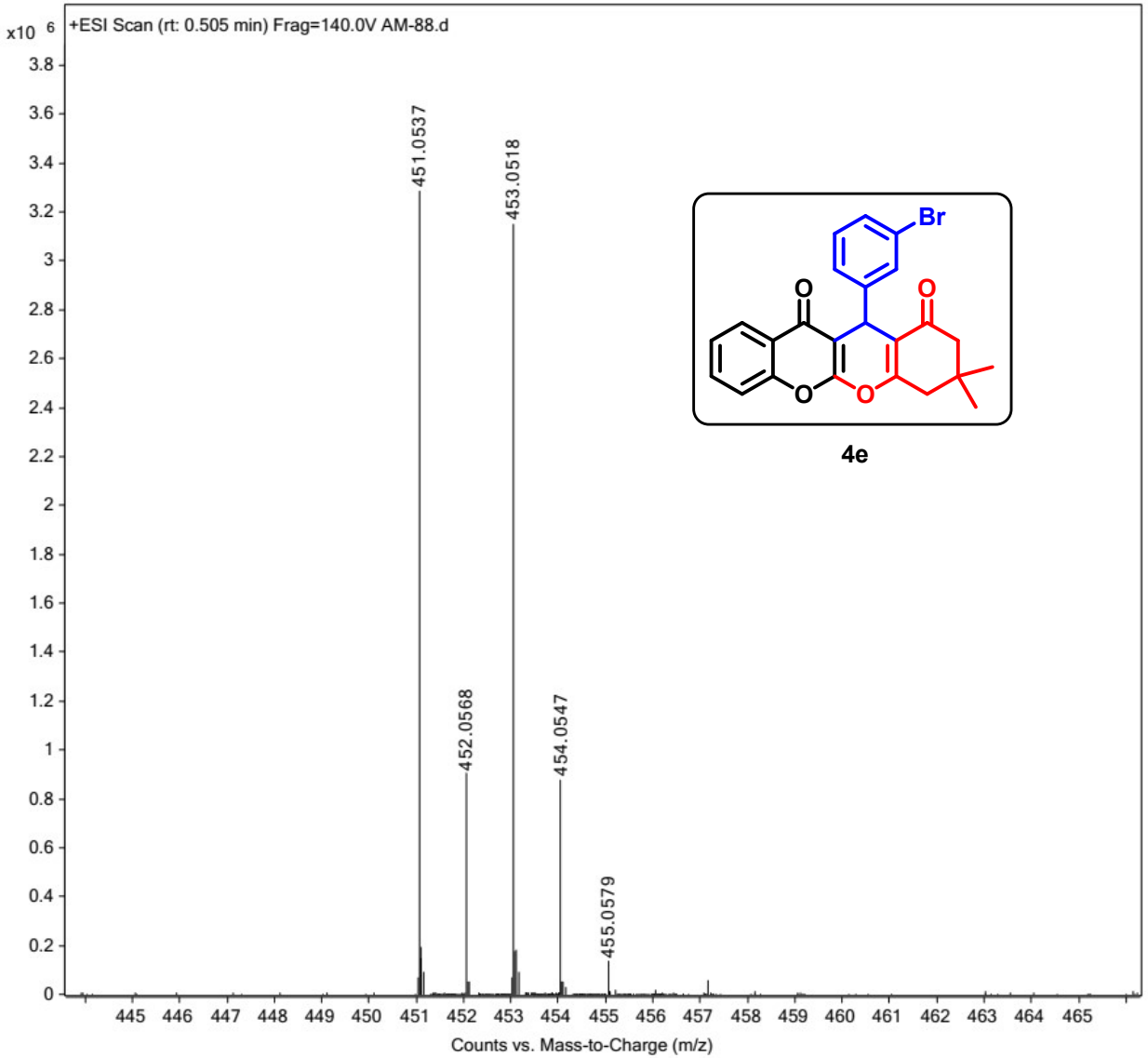


4e

^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4e

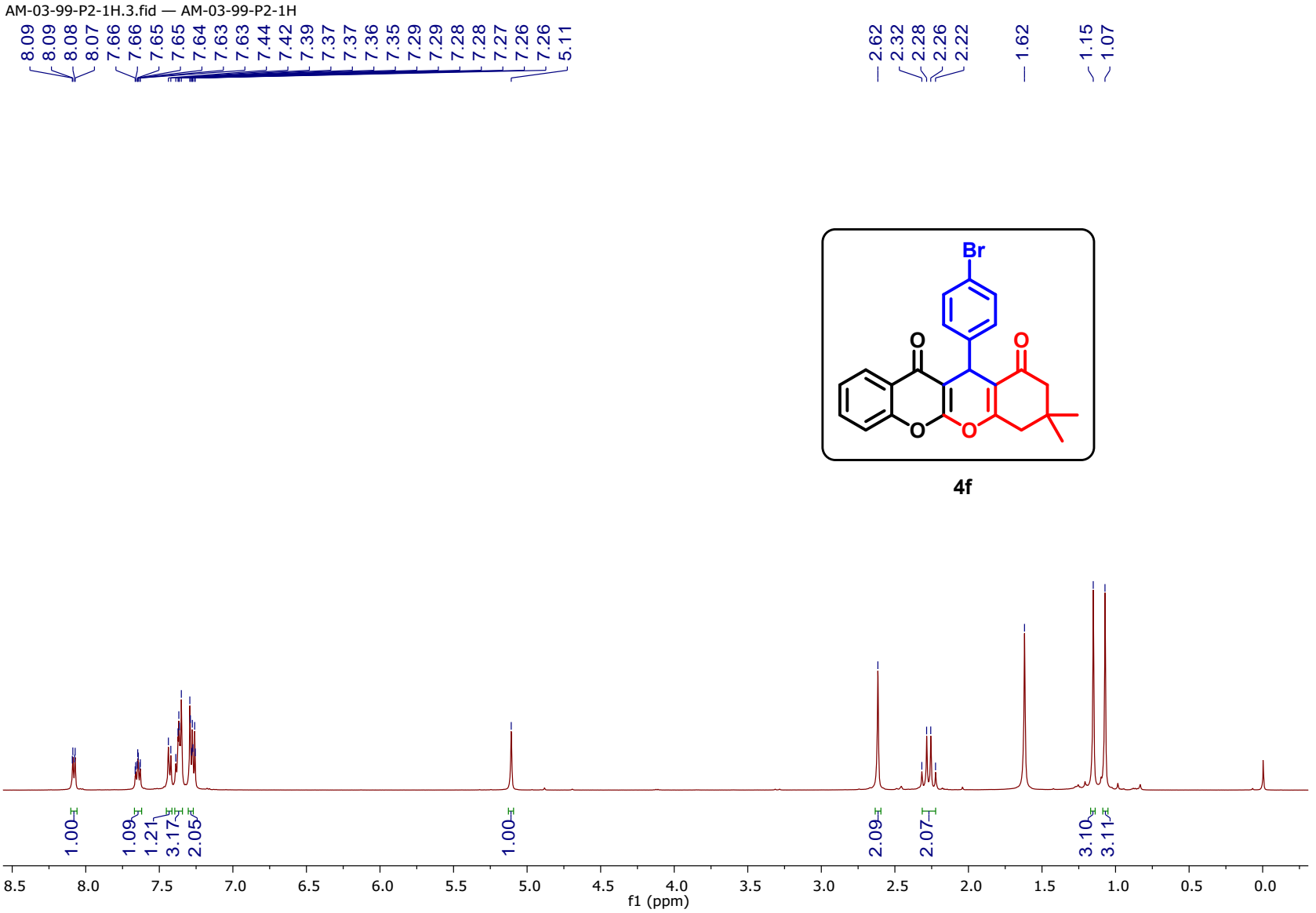


HRMS spectrum of compound 4e

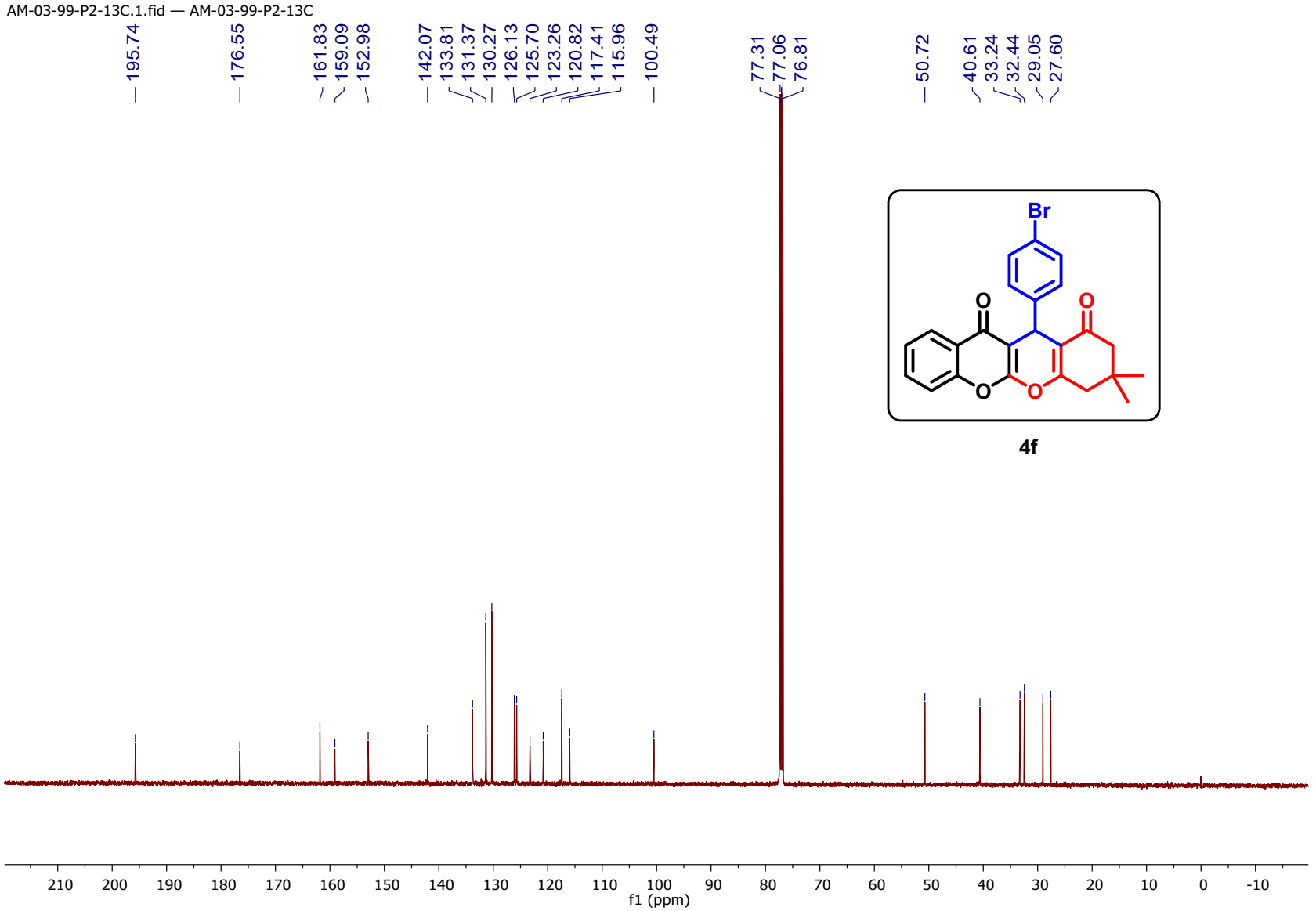


4e

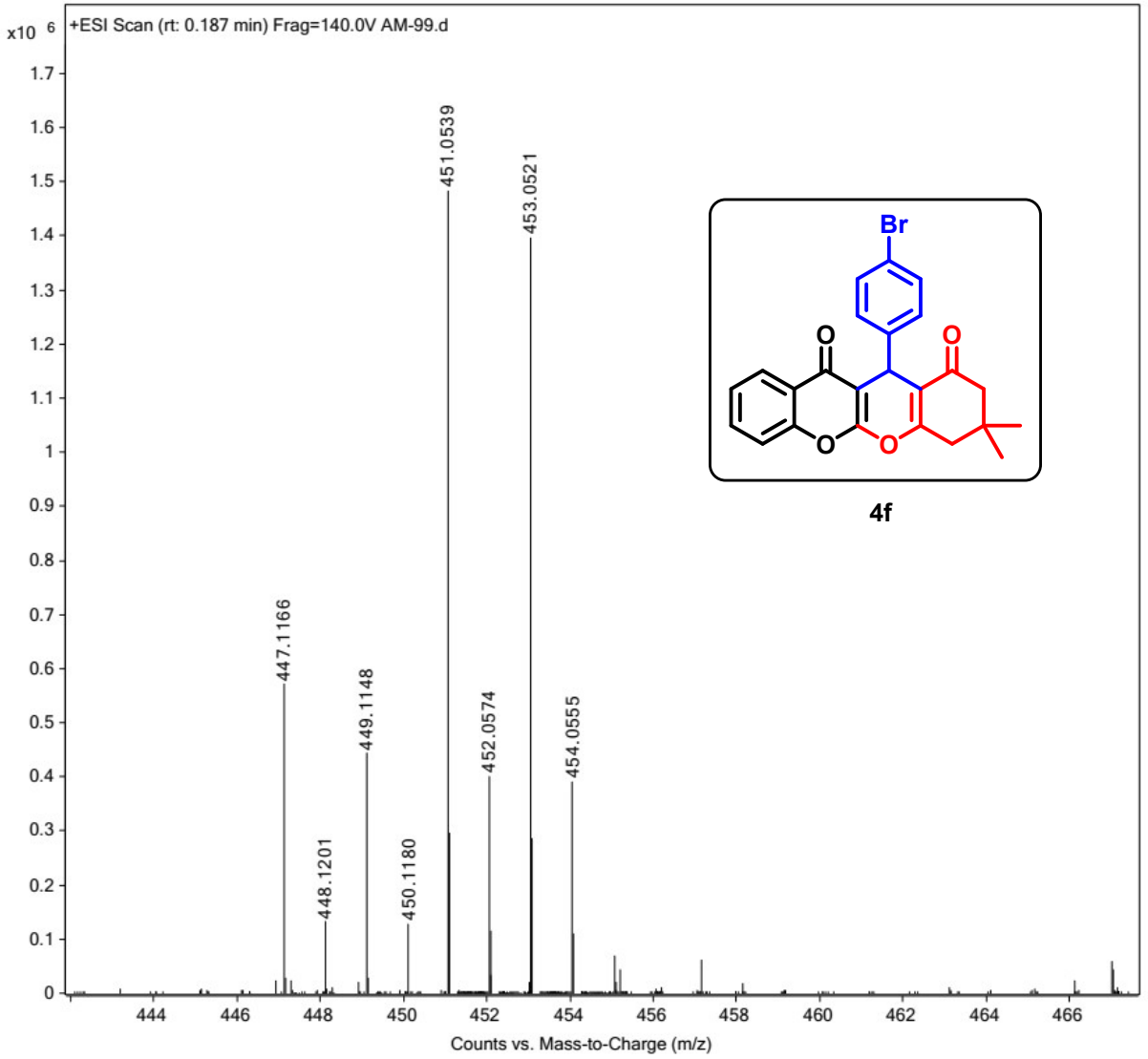
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4f



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4f

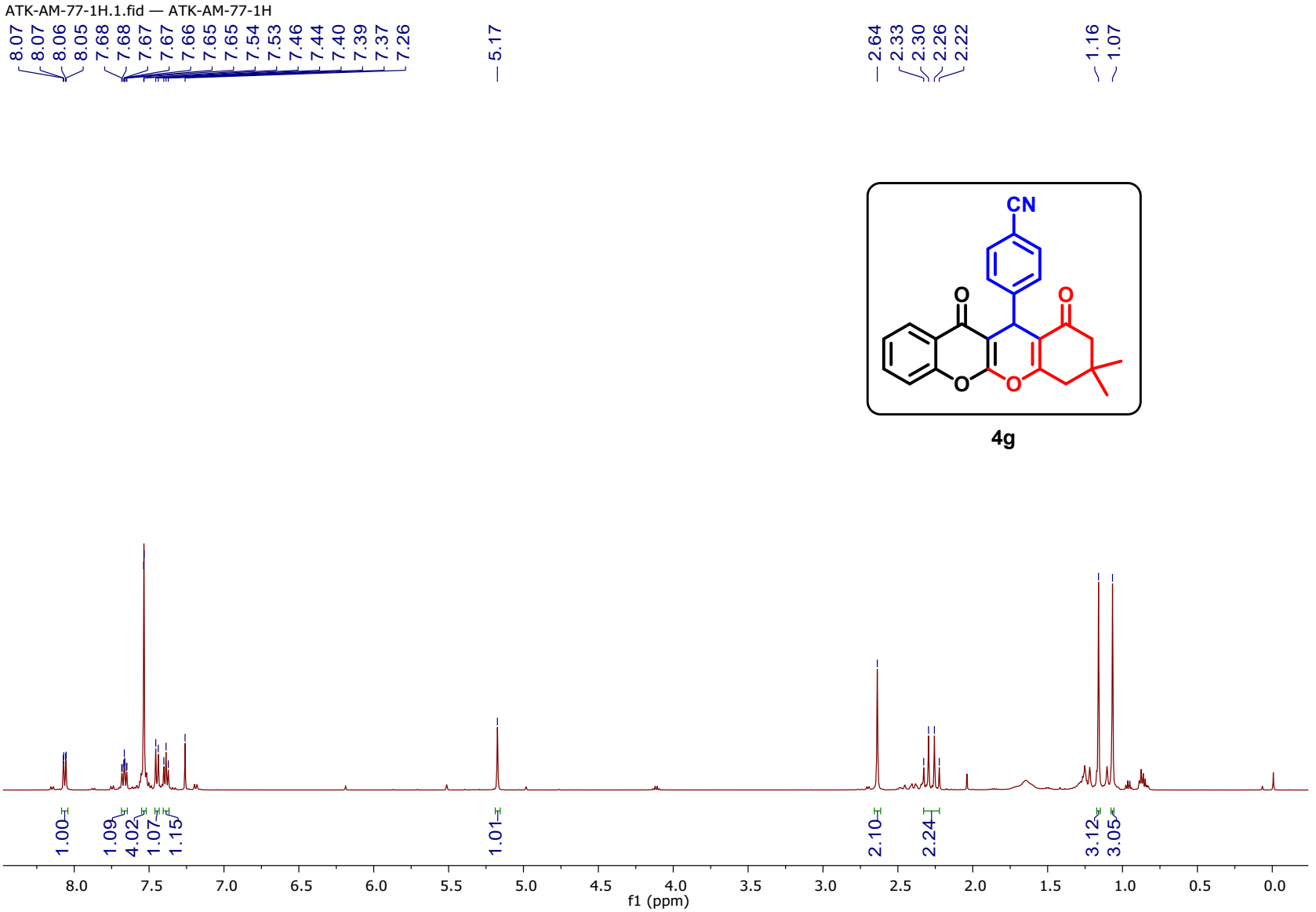


HRMS spectrum of compound 4f

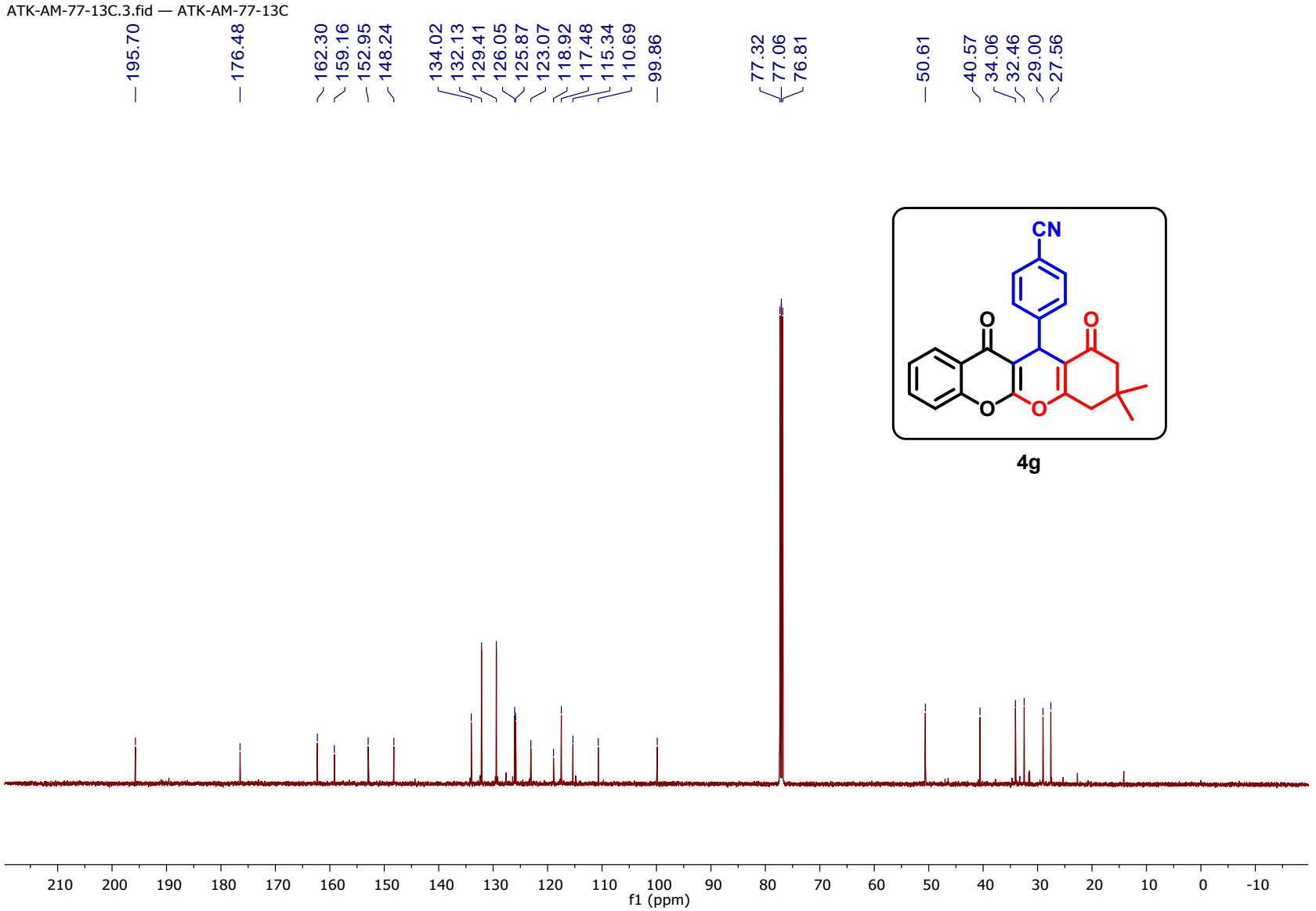


4f

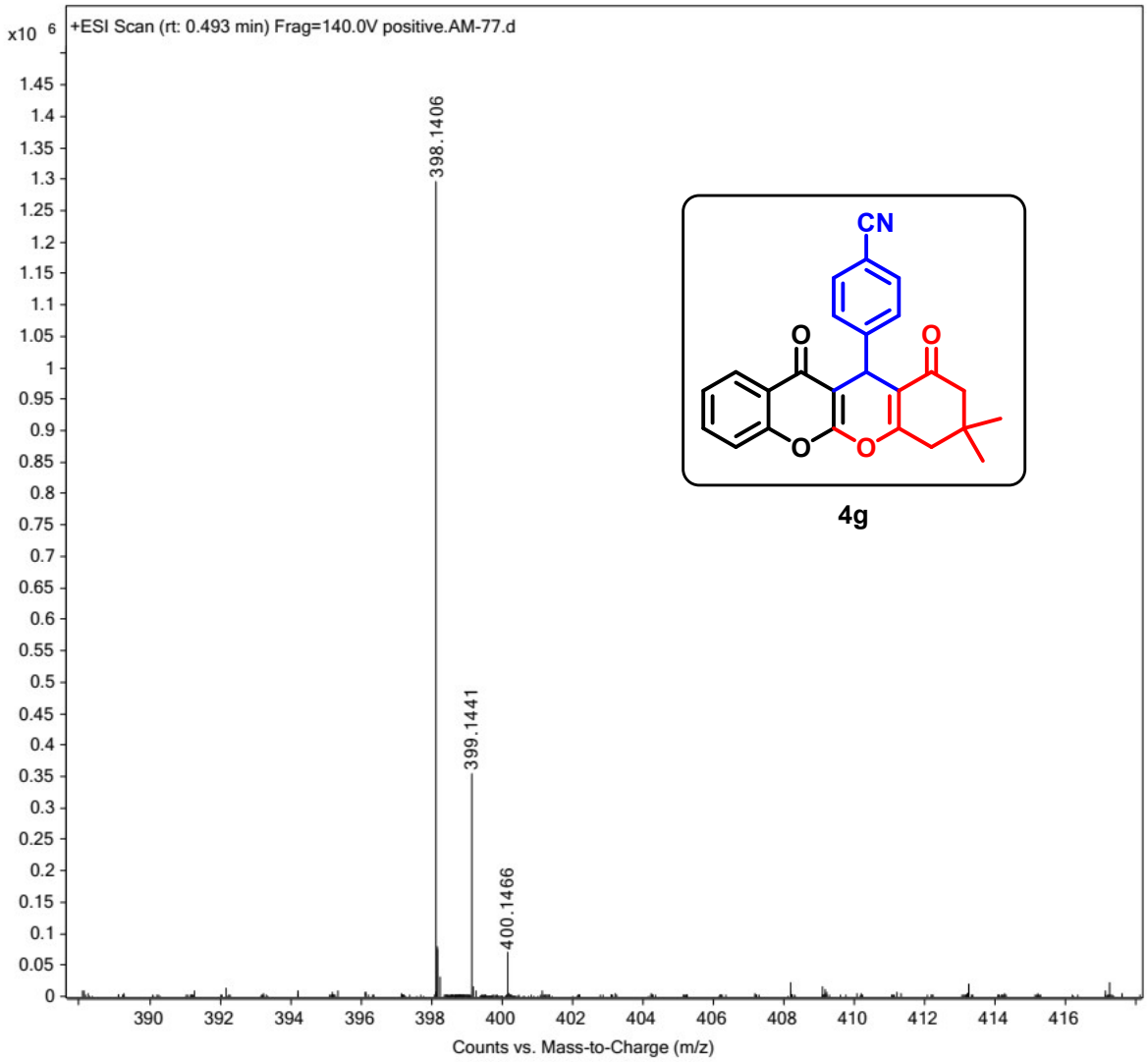
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4g



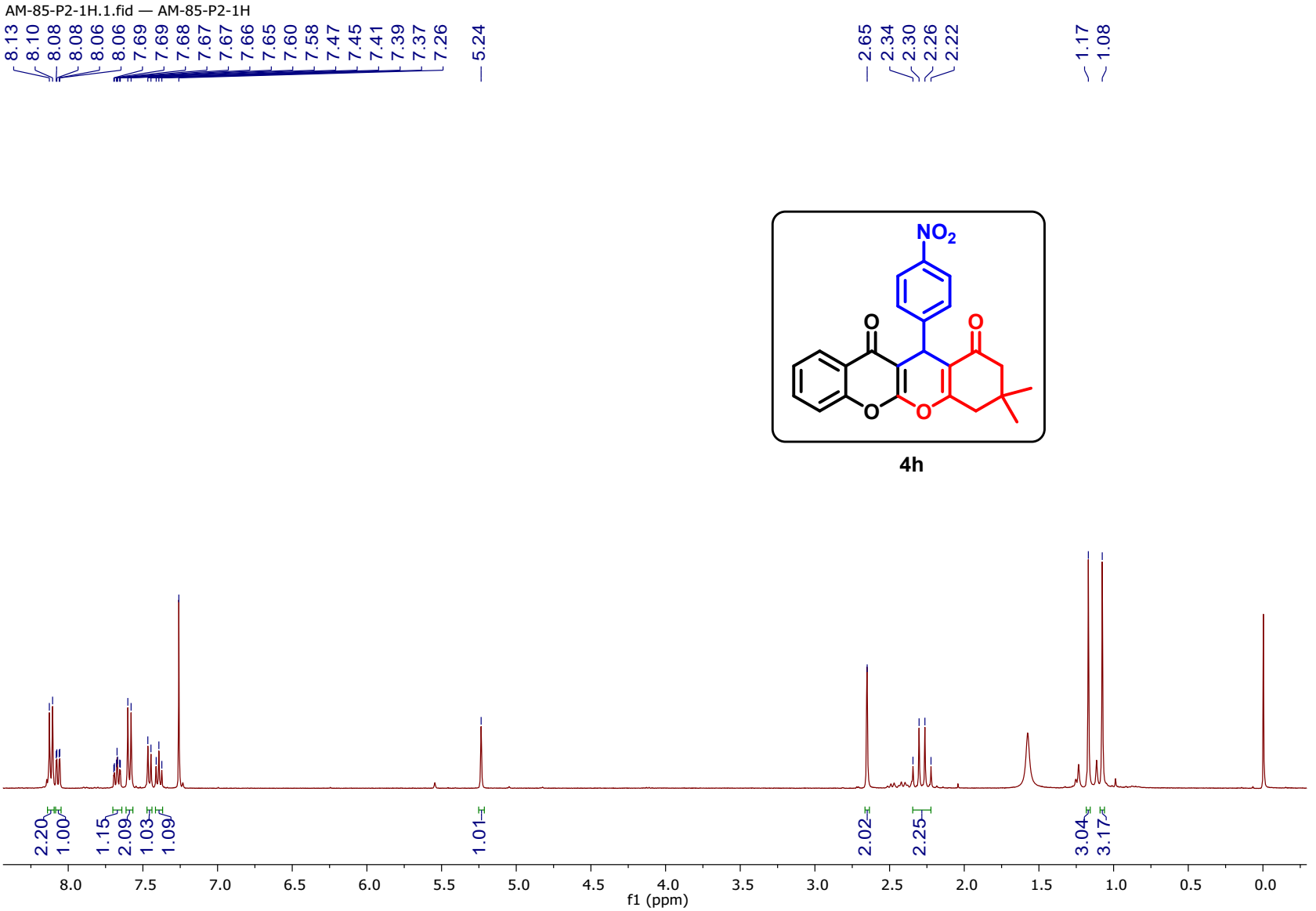
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4g



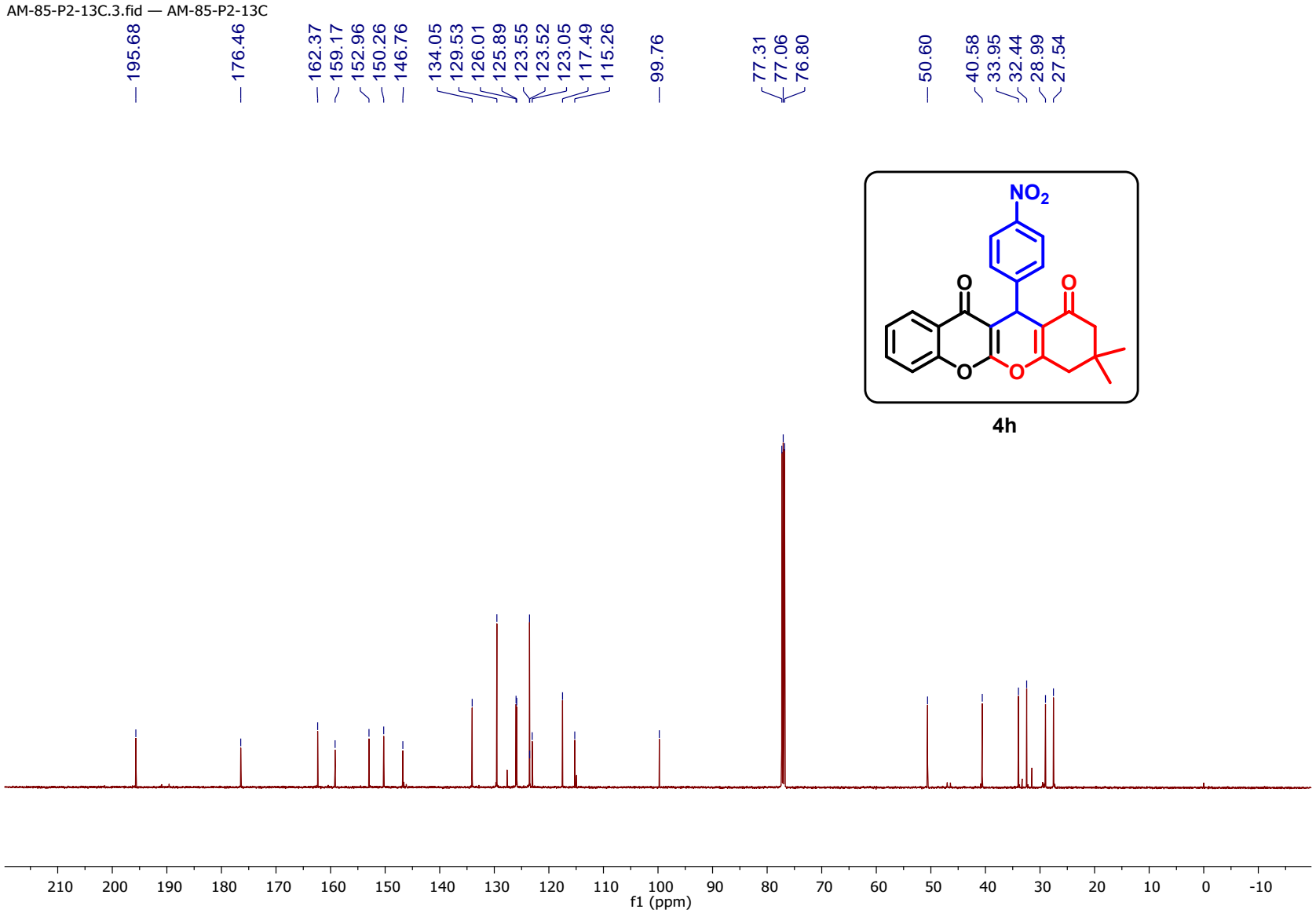
HRMS spectrum of compound 4g



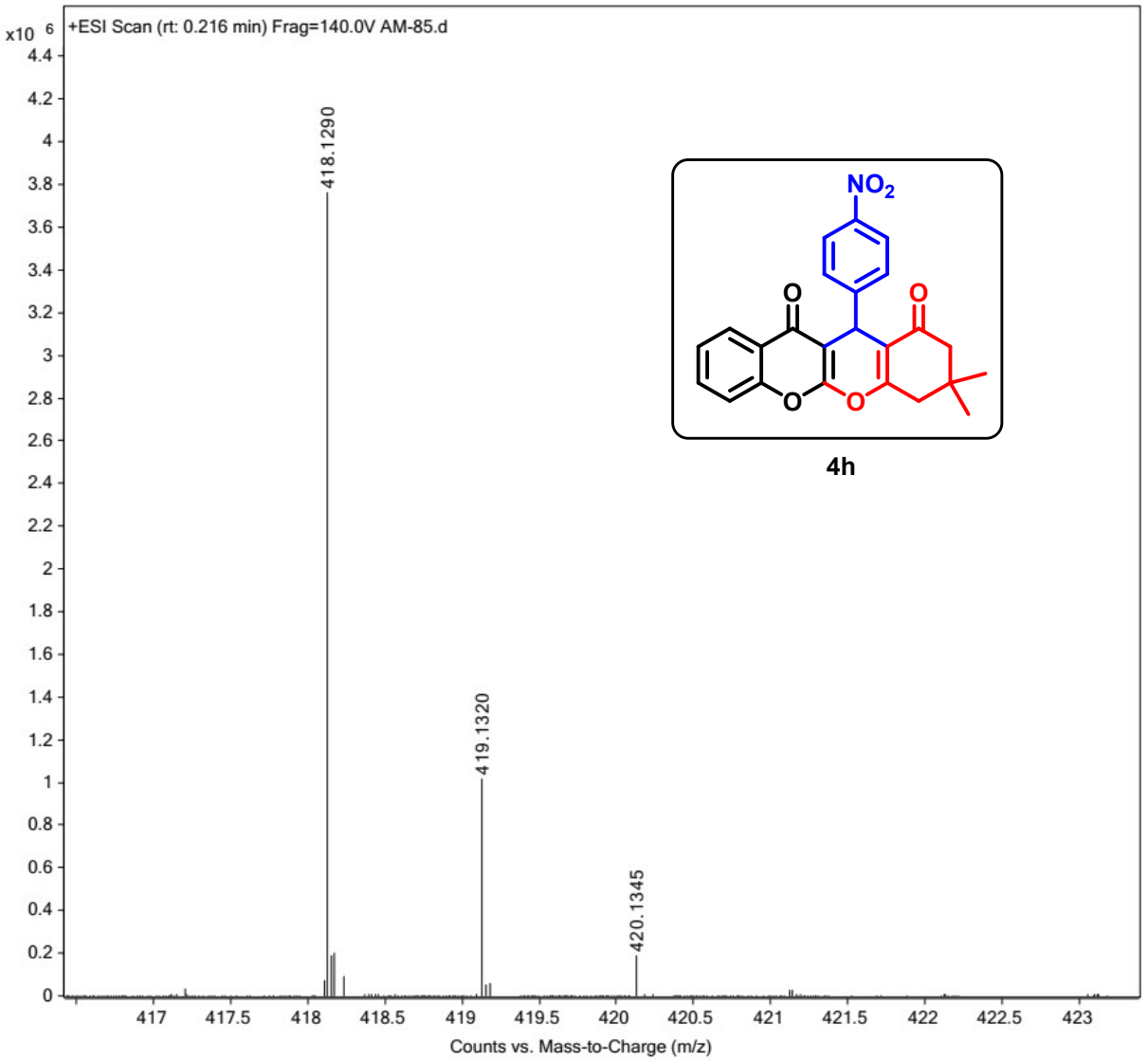
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4h



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4h

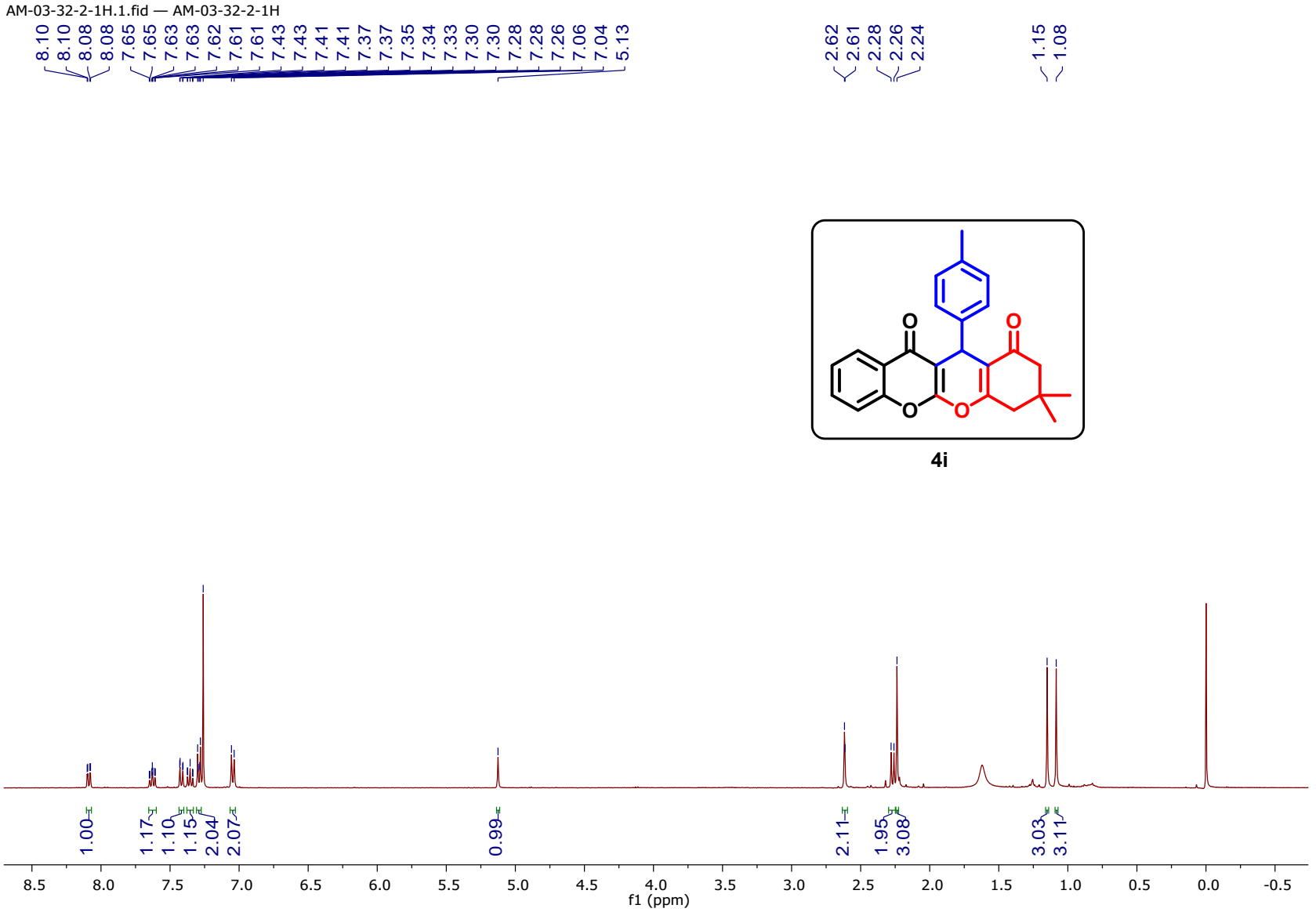


HRMS spectrum of compound 4h



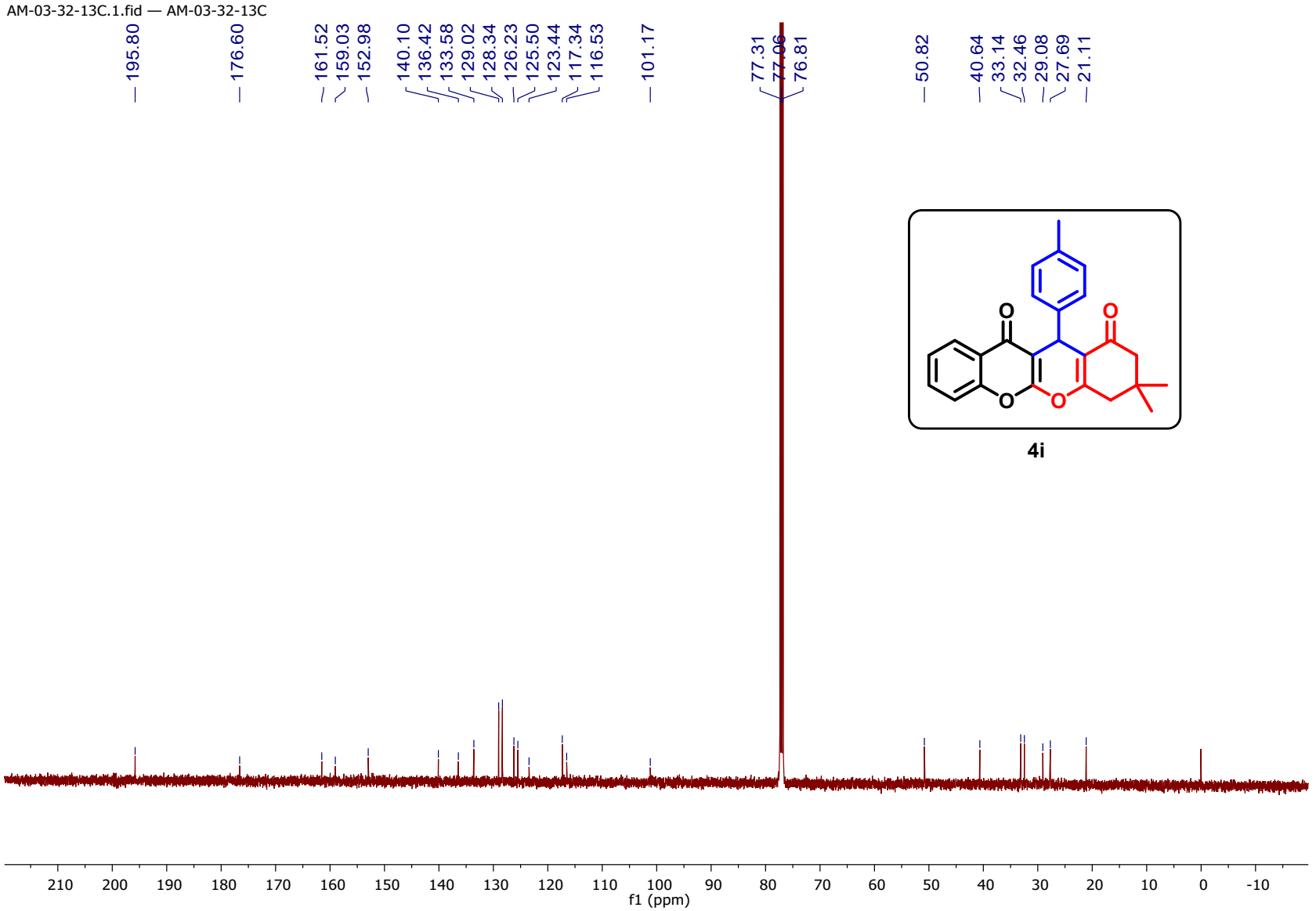
4h

^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4i

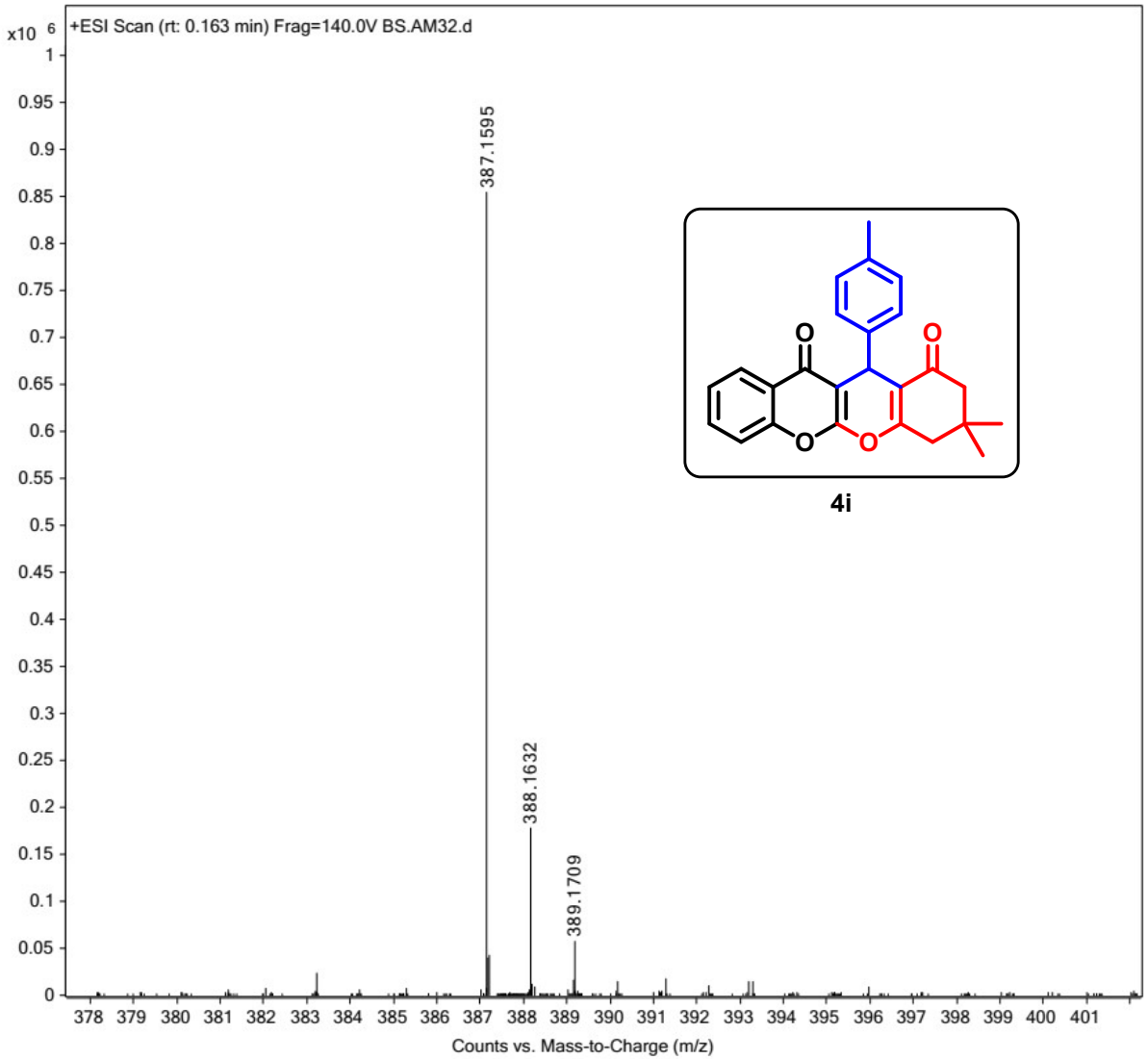


4i

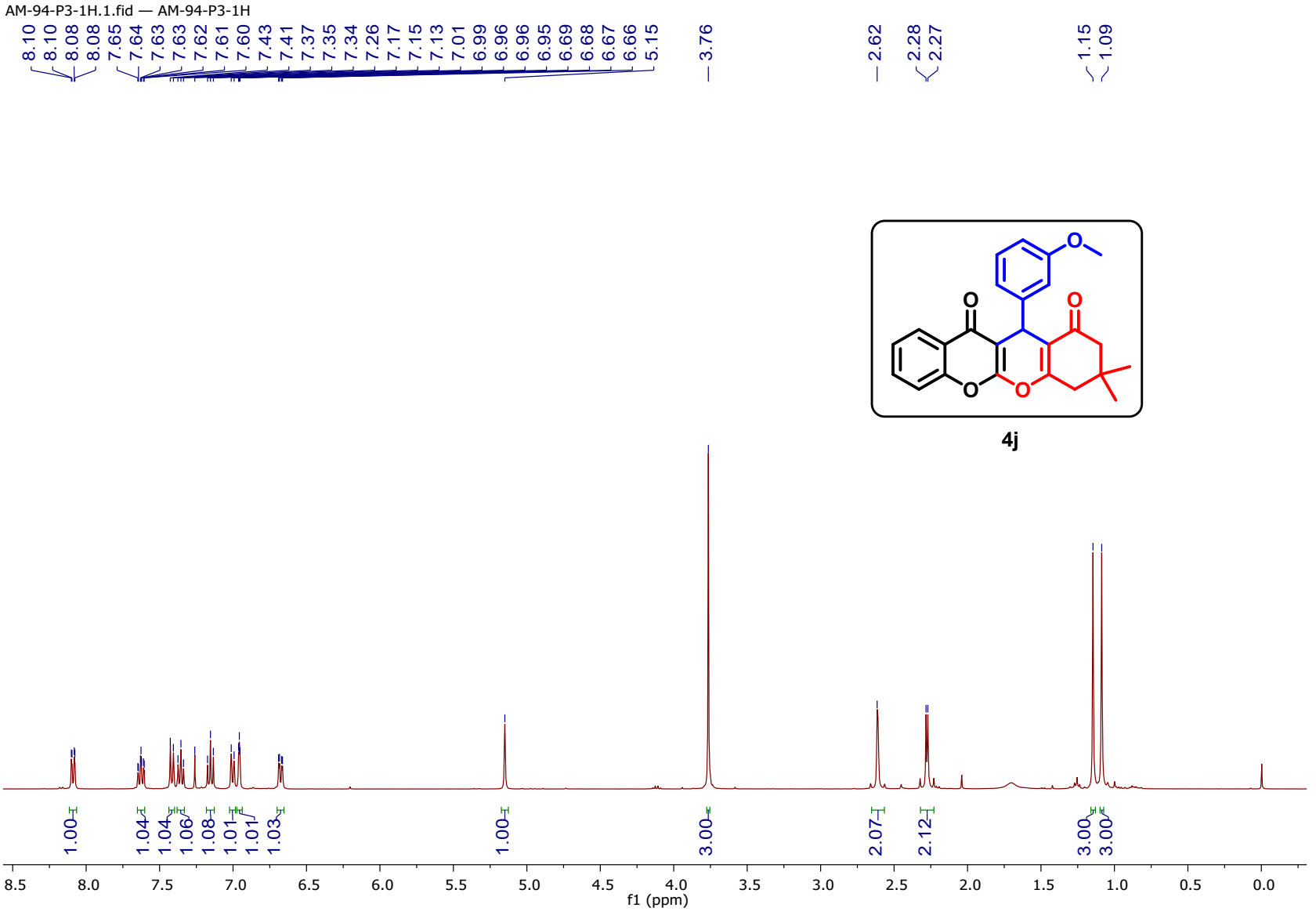
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4i



HRMS spectrum of compound 4i

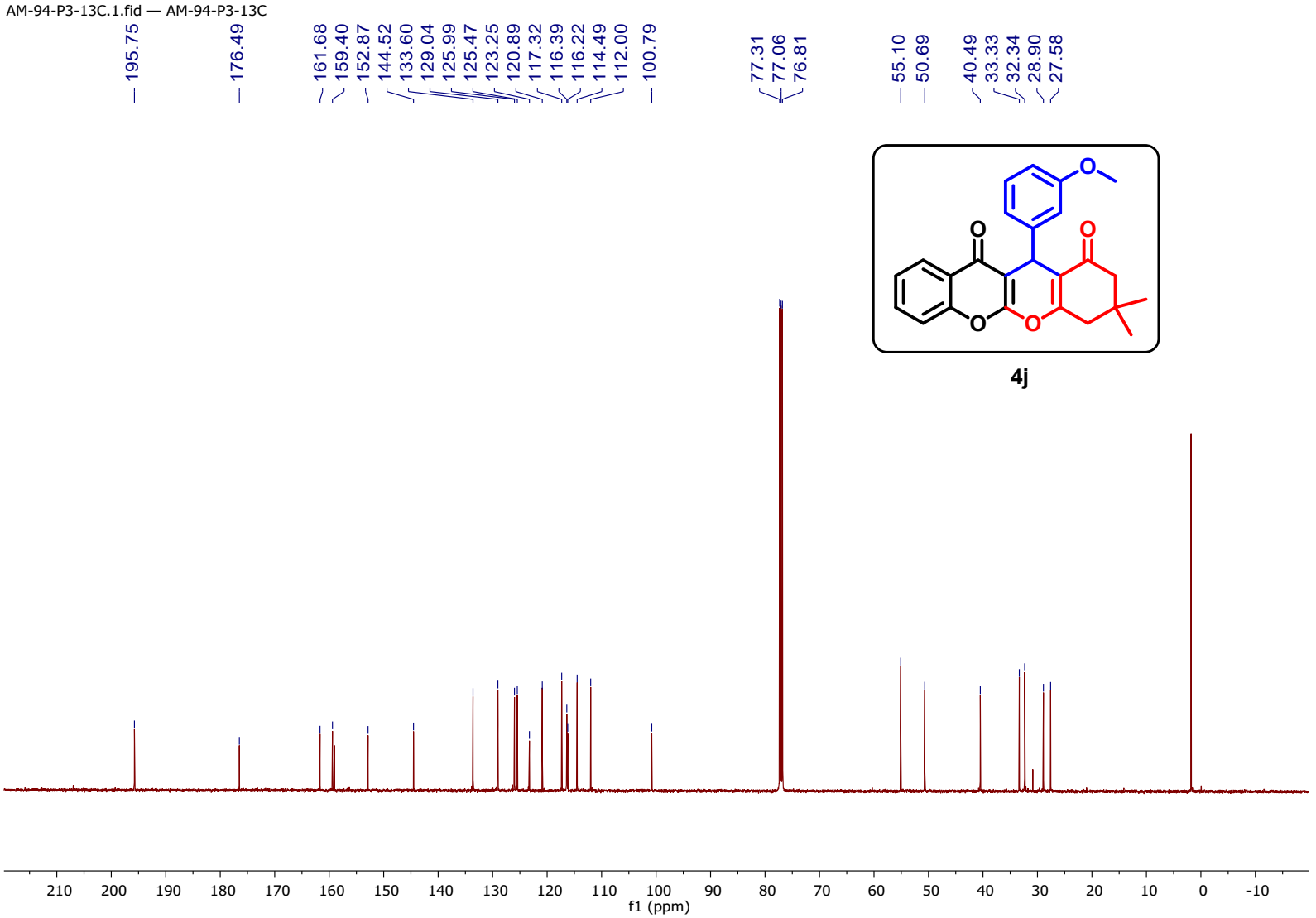


^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4j

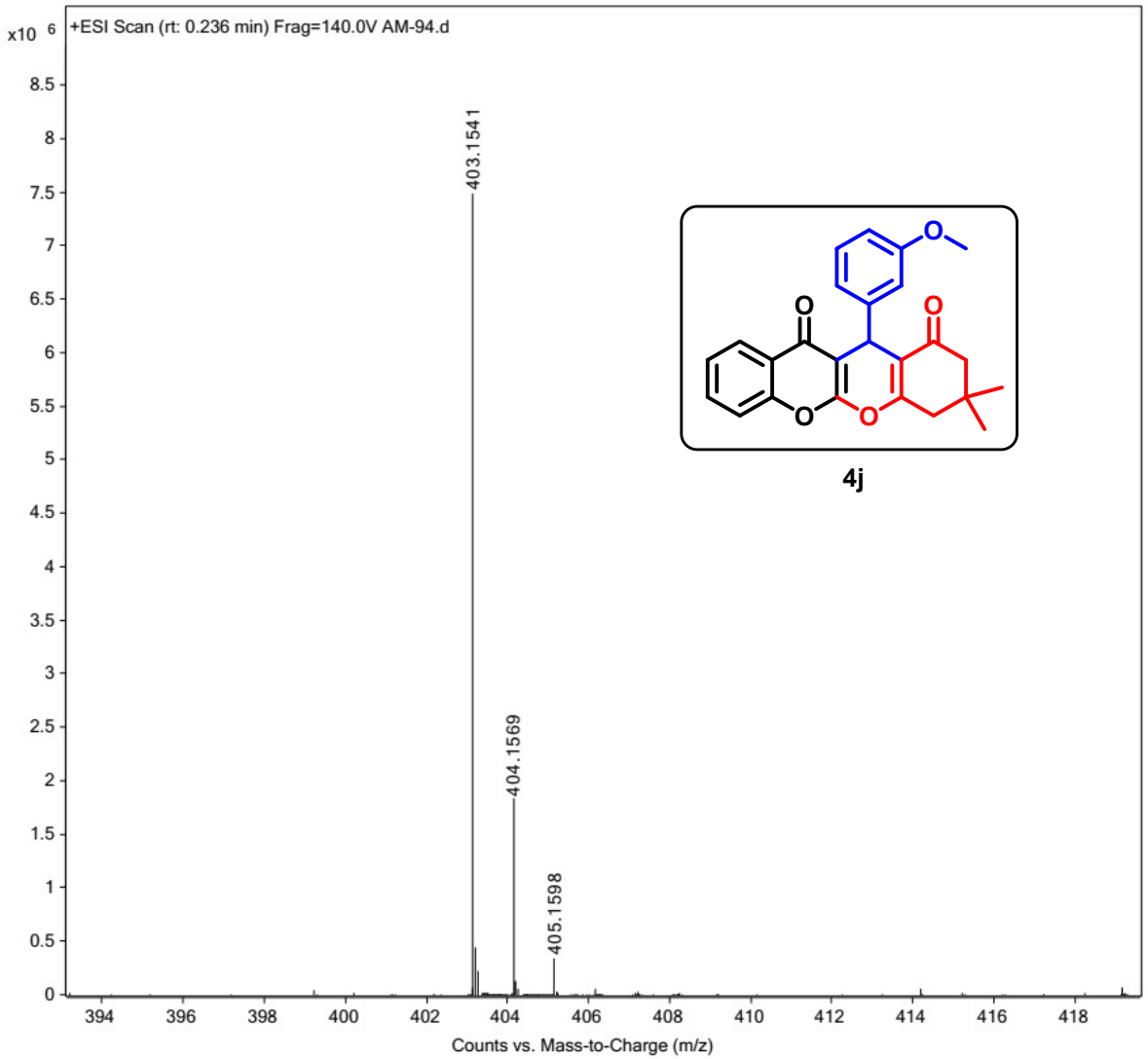


4j

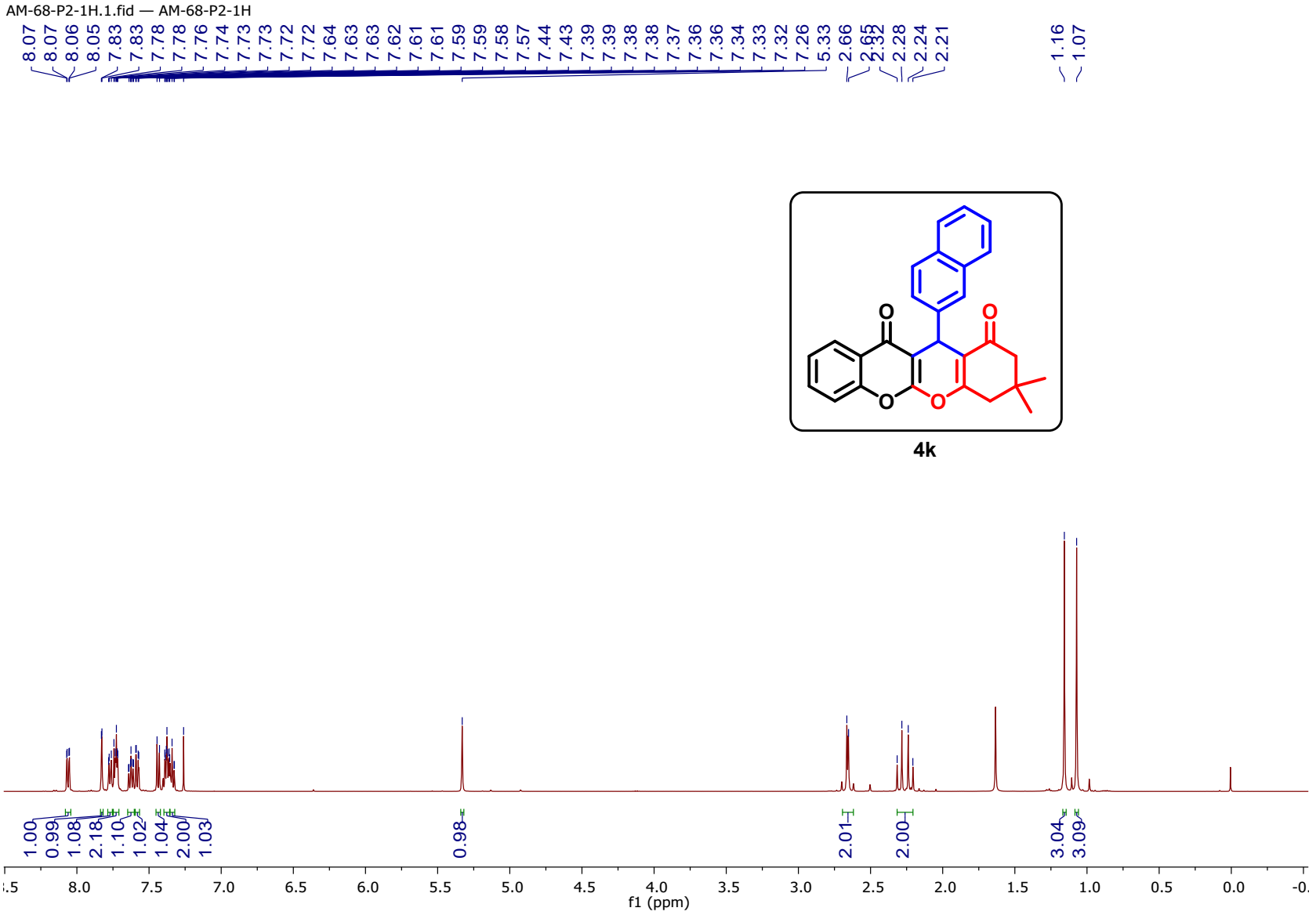
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4j



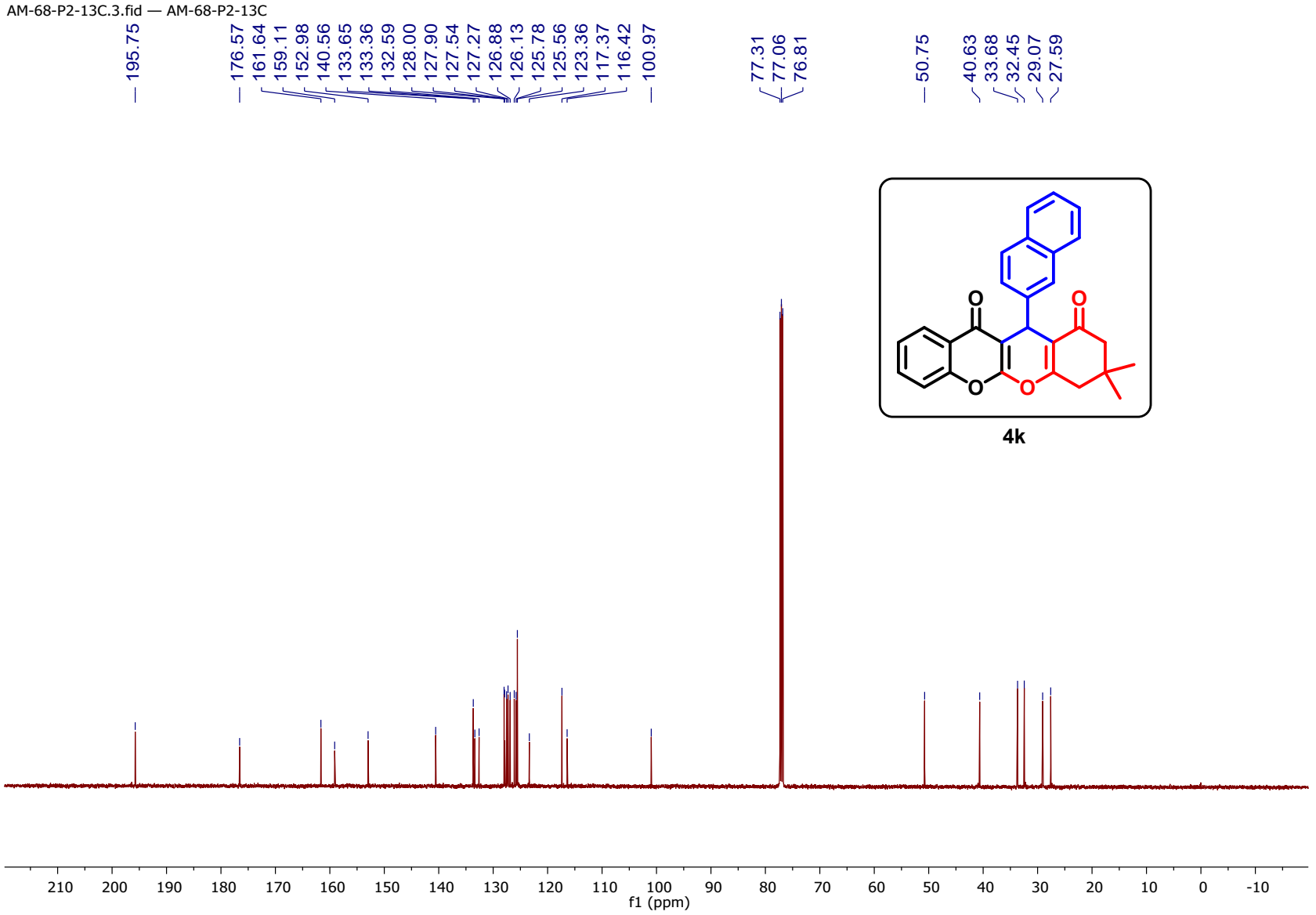
HRMS spectrum of compound 4j



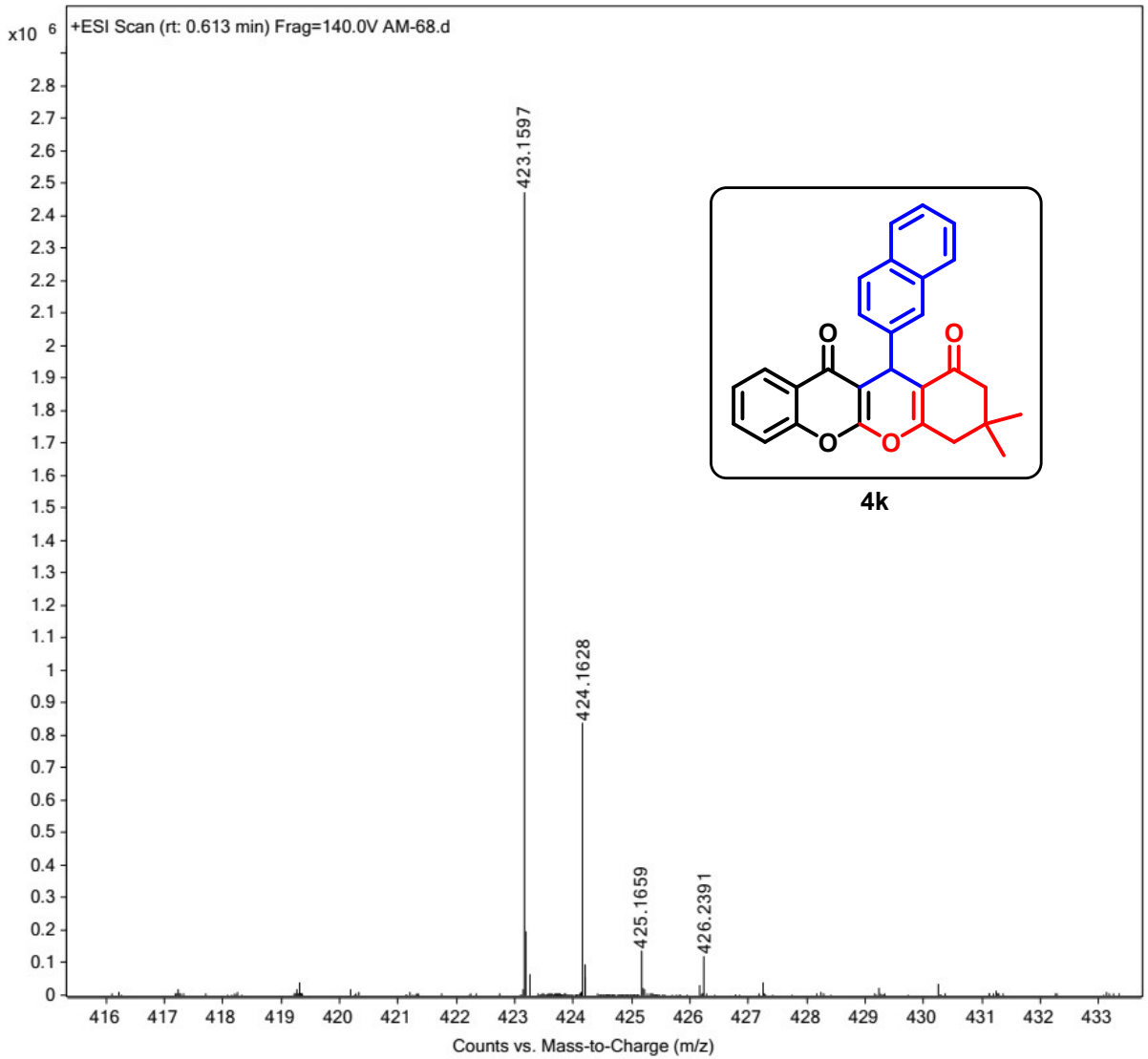
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4k



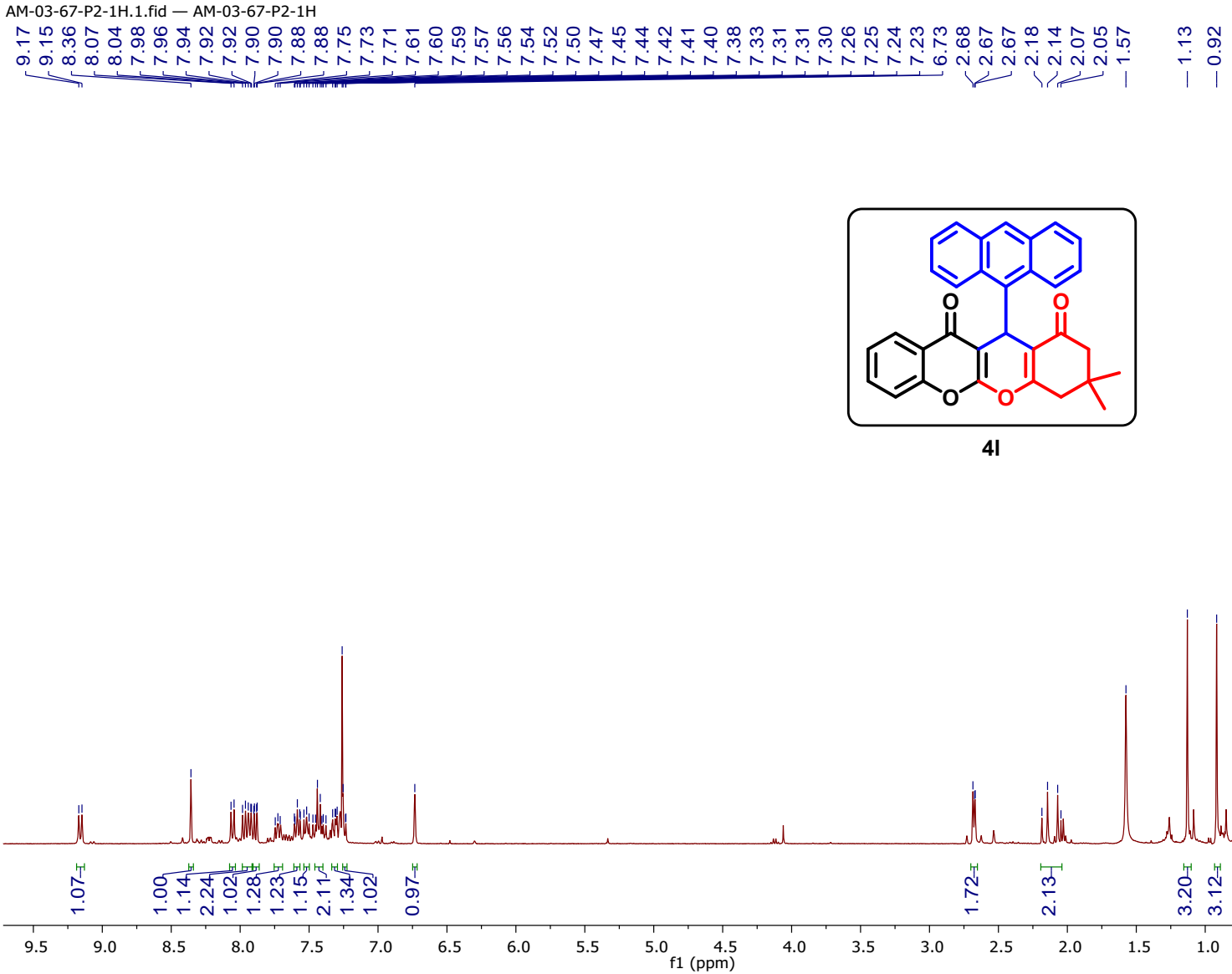
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4k



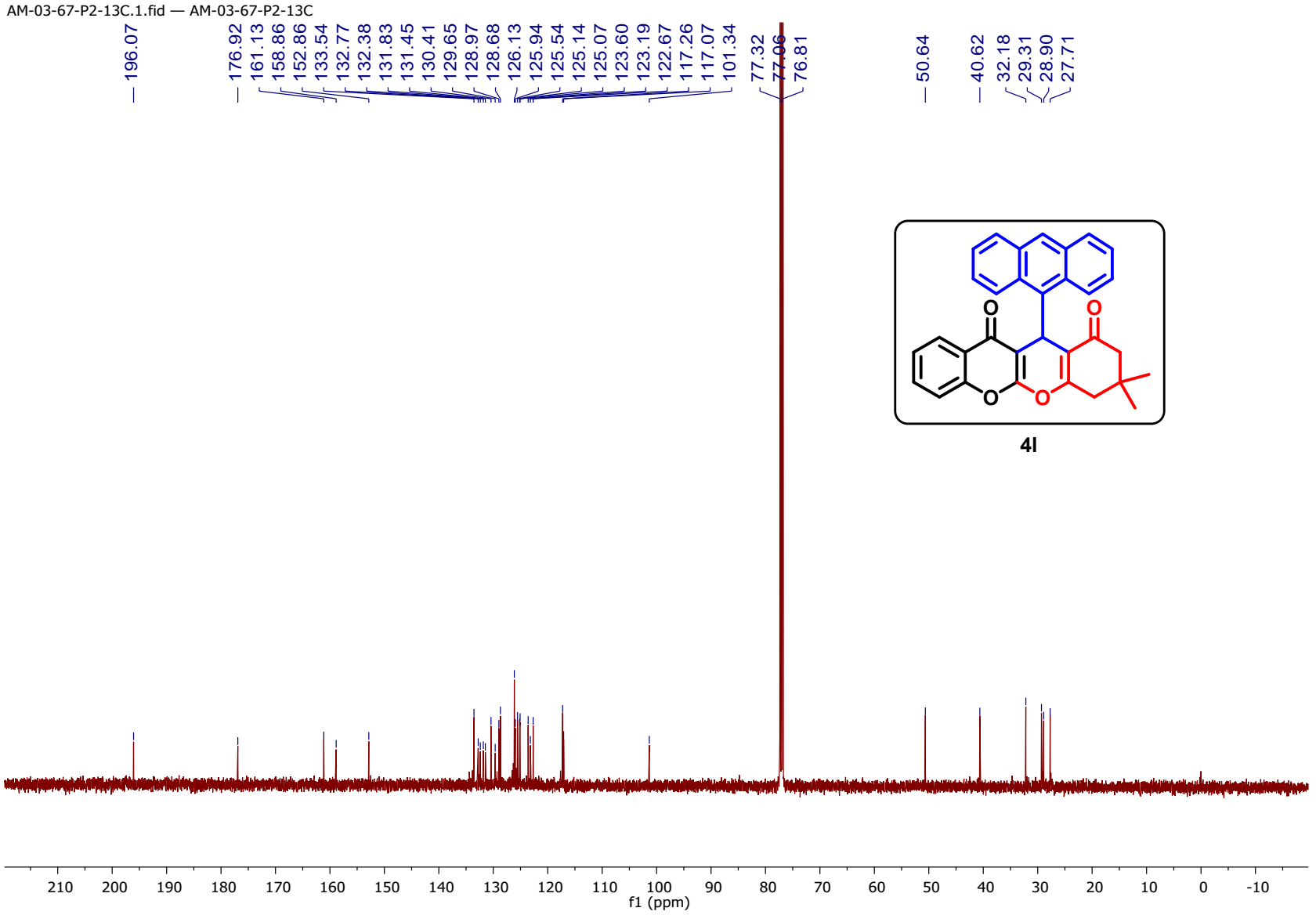
HRMS spectrum of compound 4k



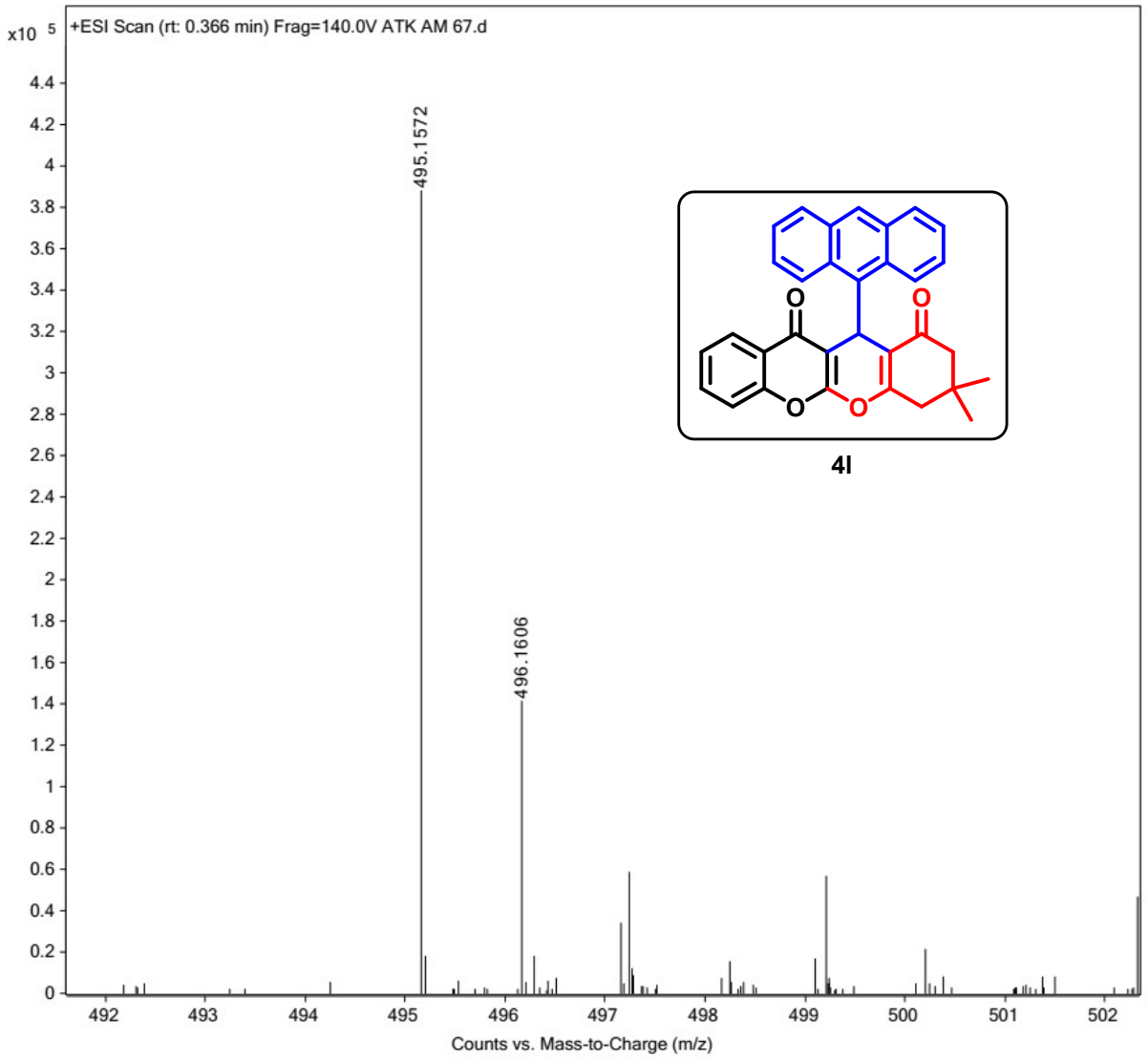
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4l



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4l

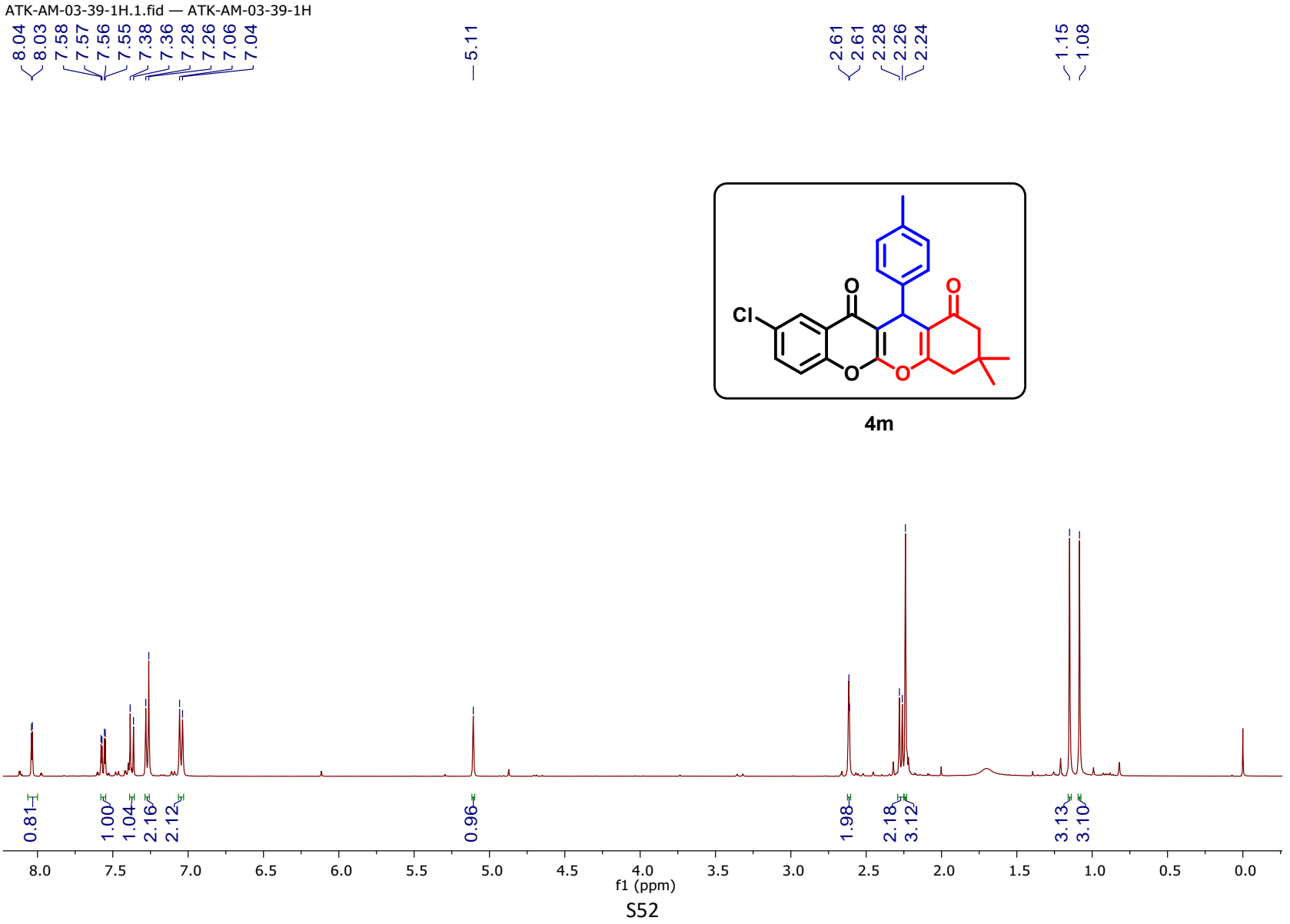


HRMS spectrum of compound 4l

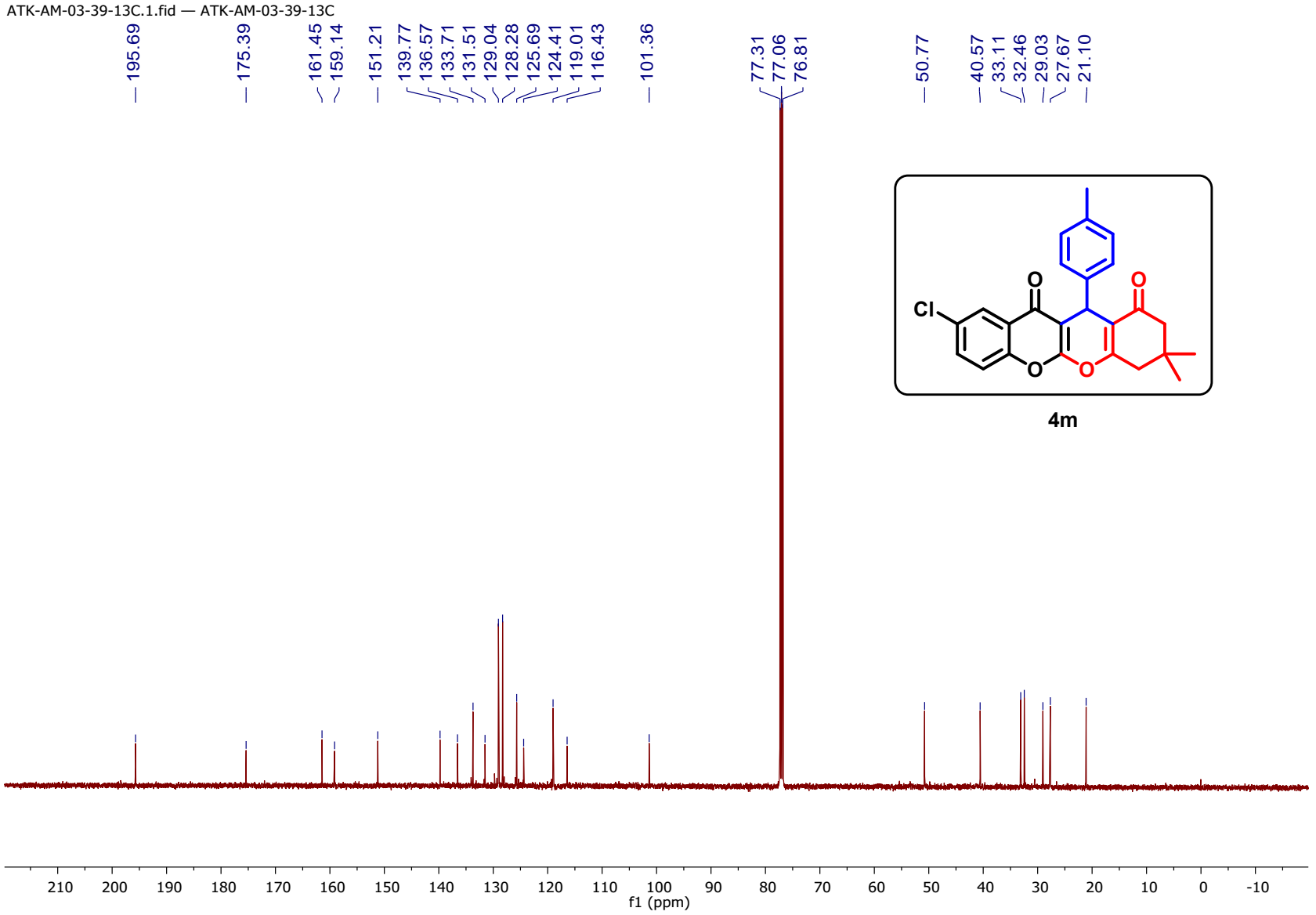


4l

^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4m

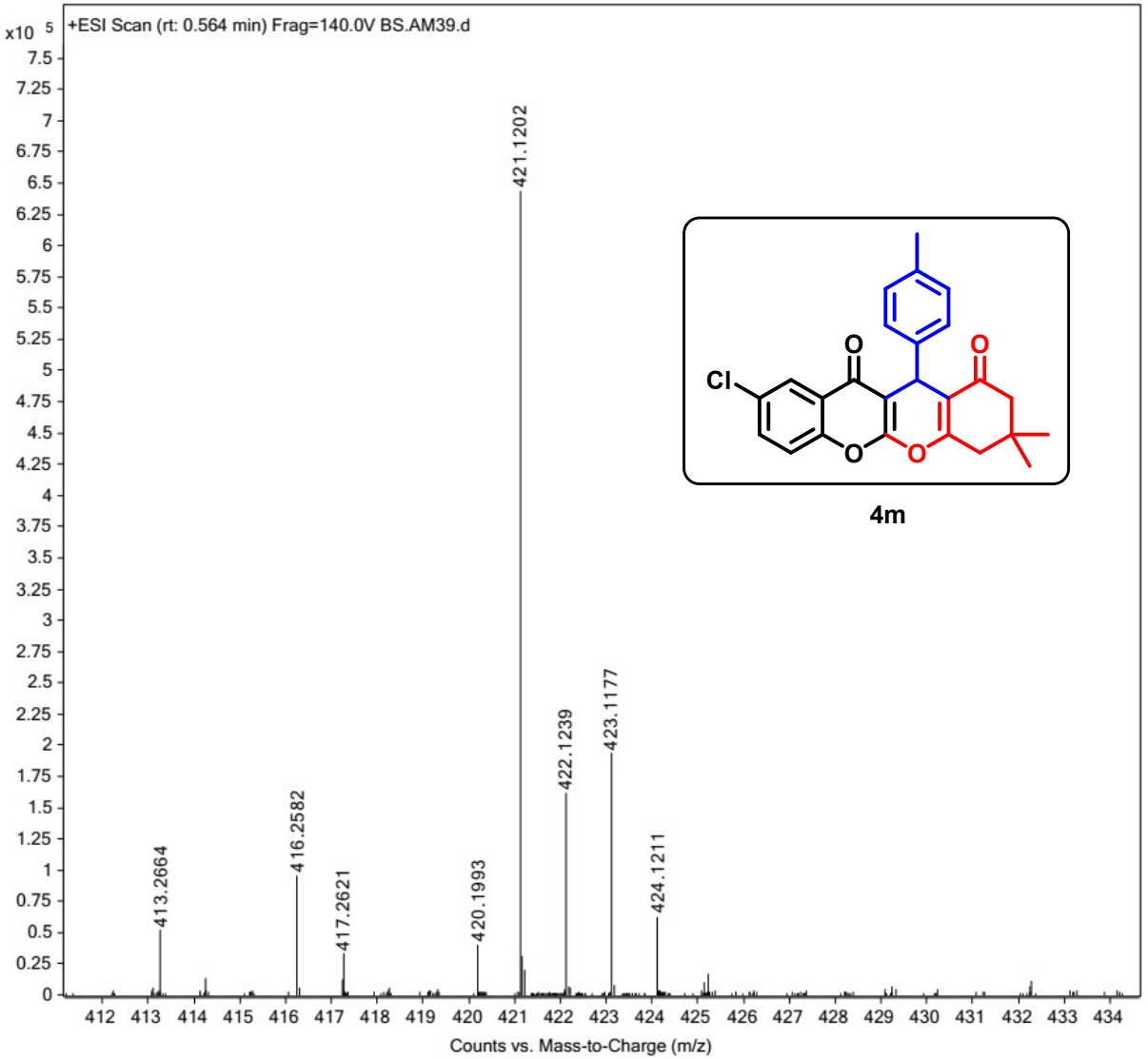


^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4m

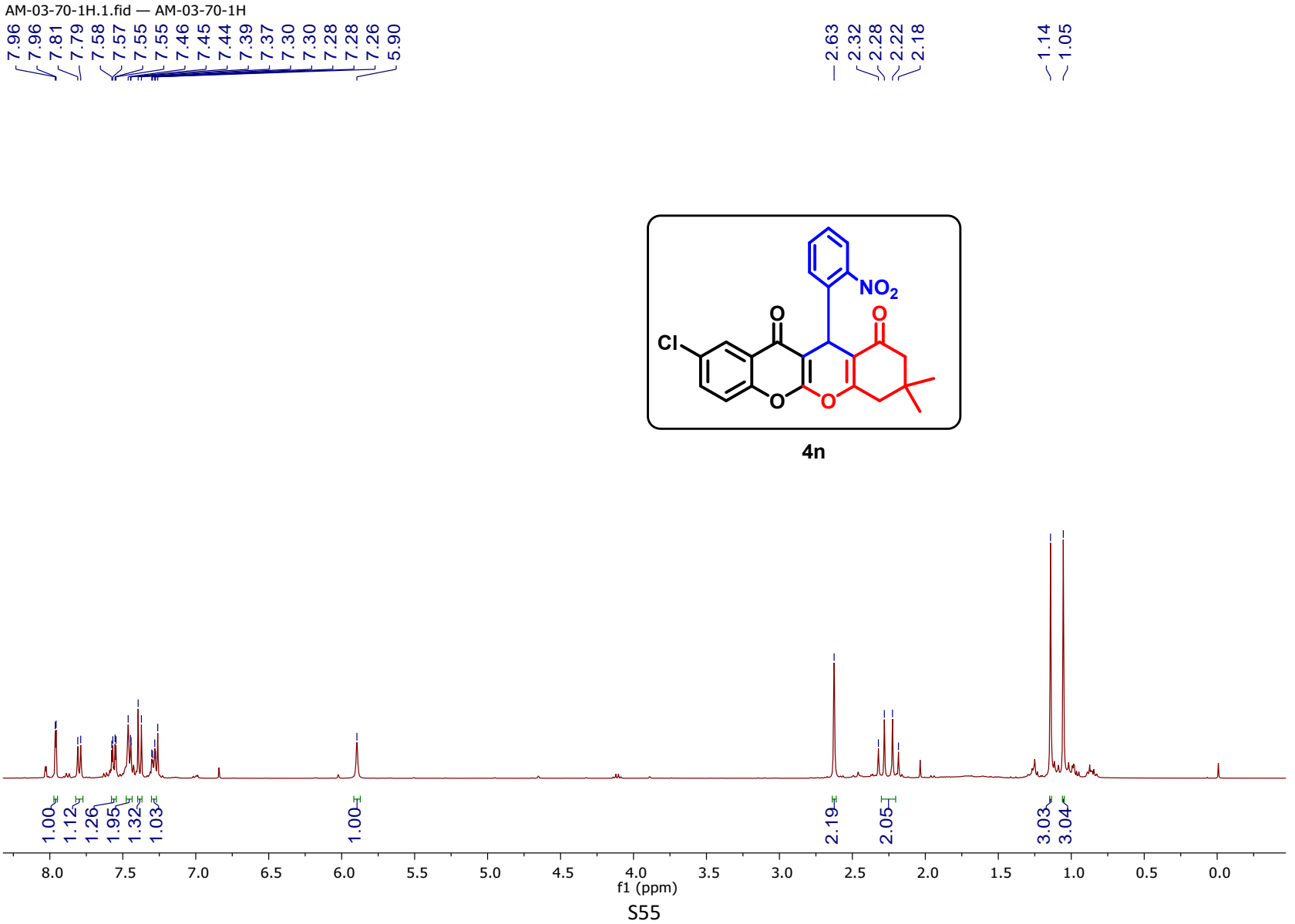


4m

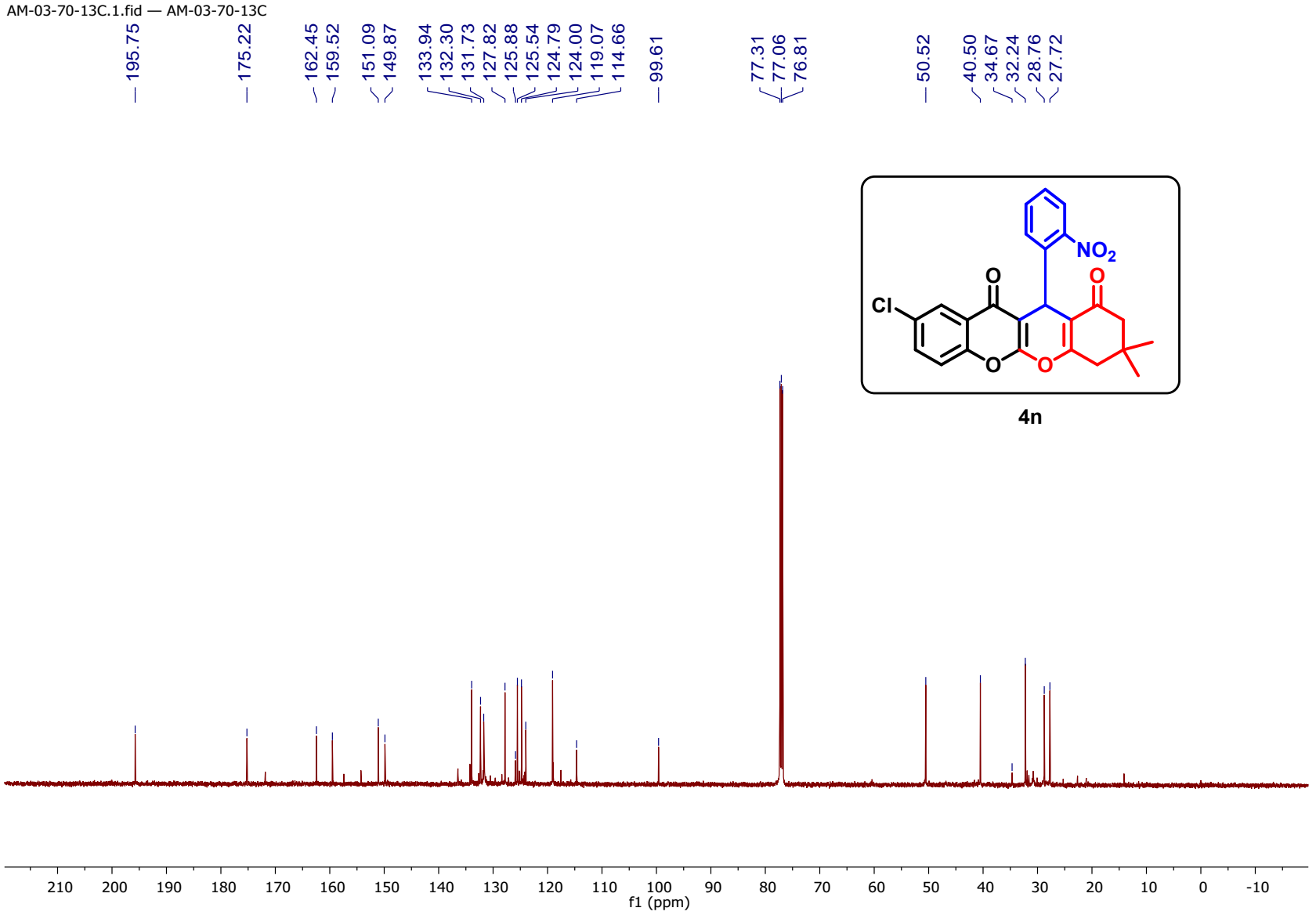
HRMS spectrum of compound 4m



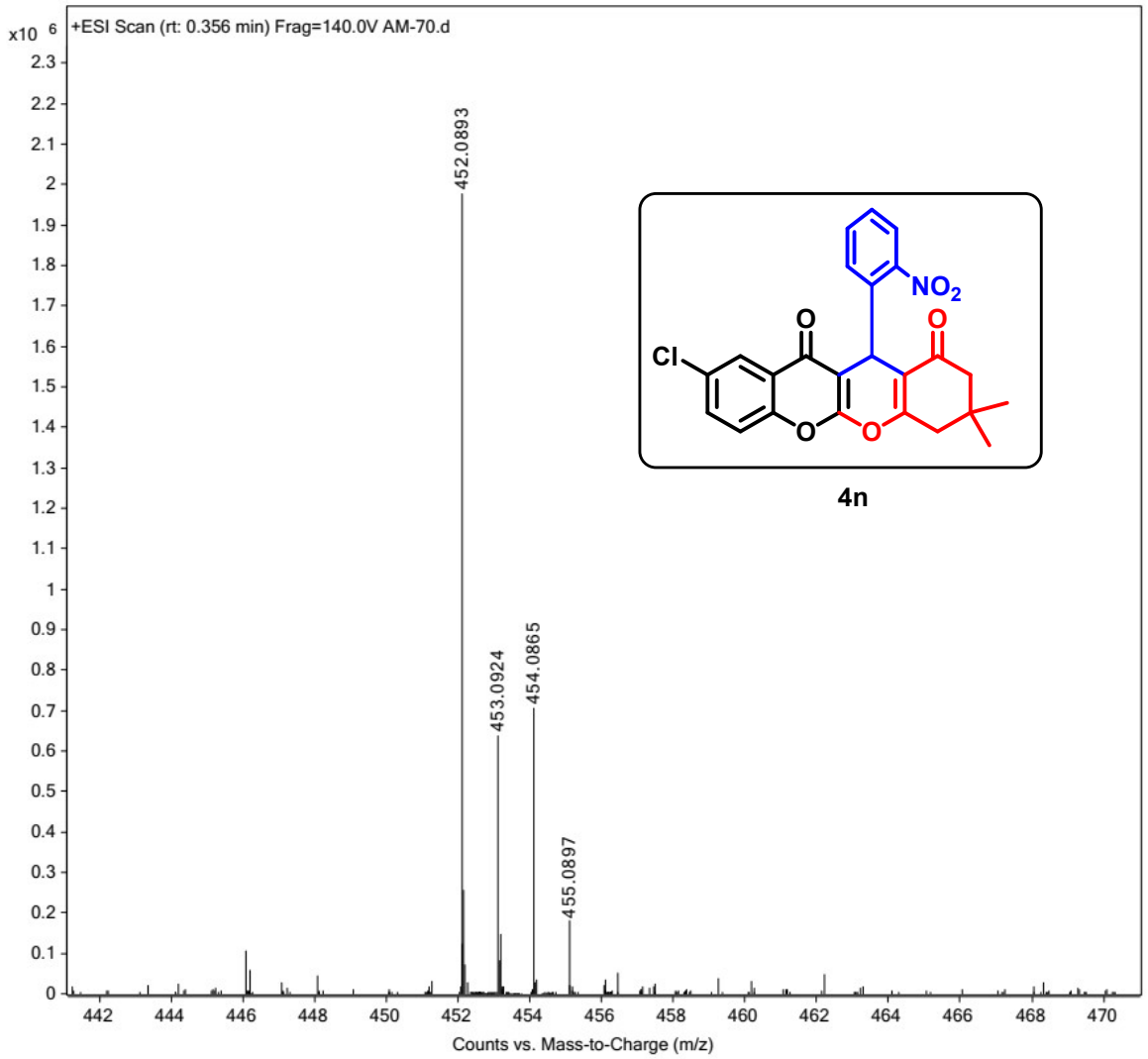
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4n



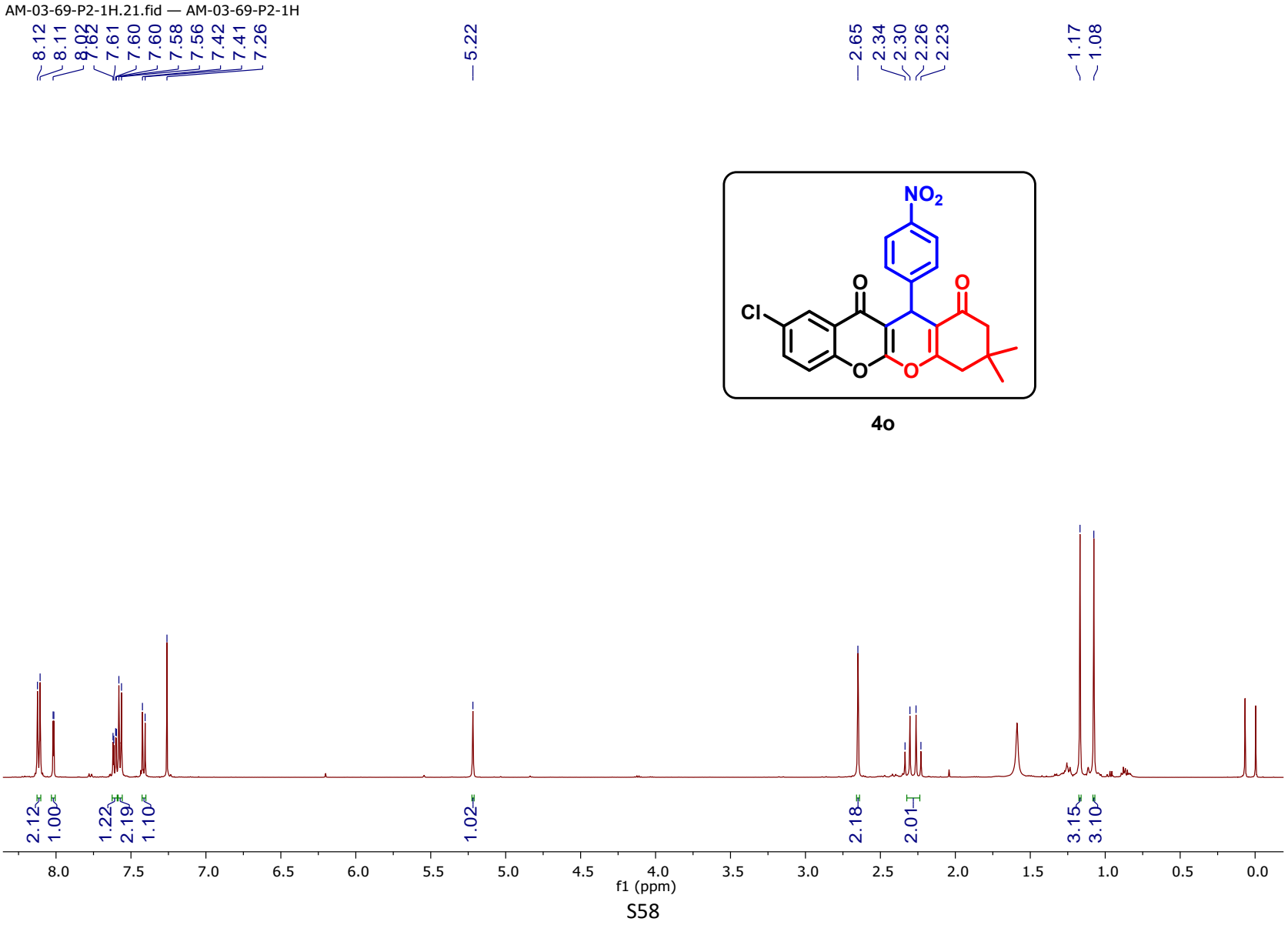
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4n



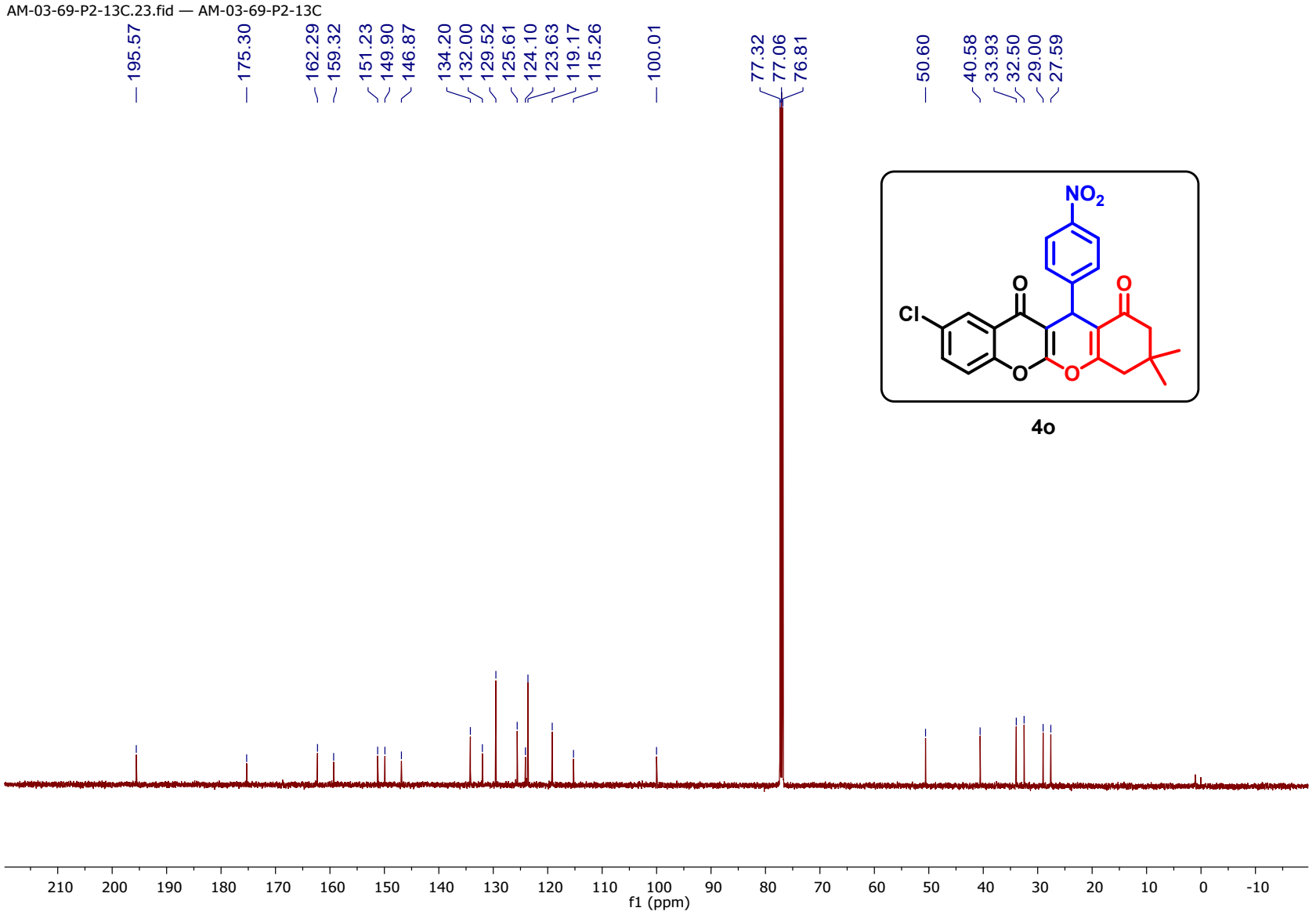
HRMS spectrum of compound 4n



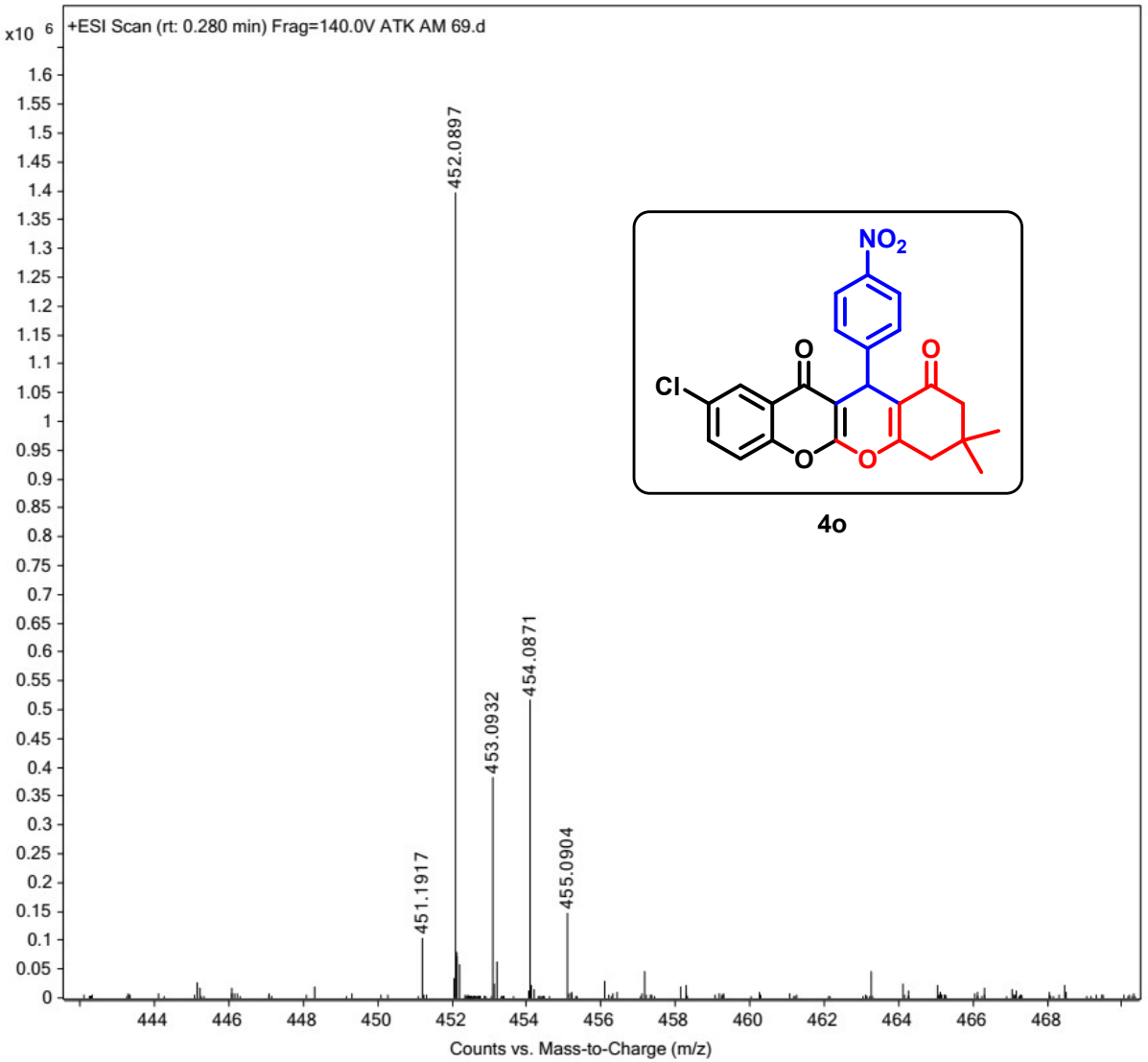
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4o



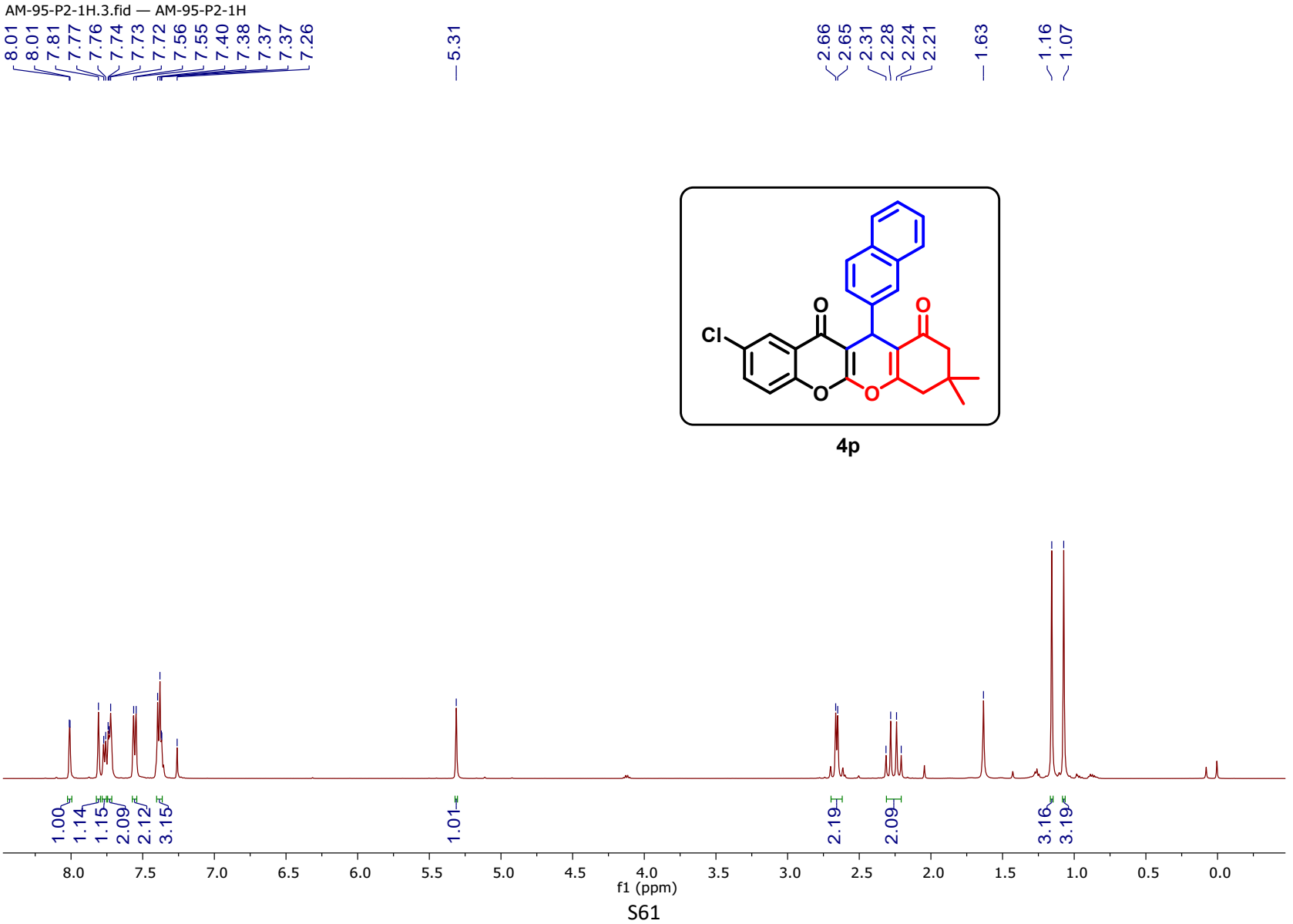
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4o



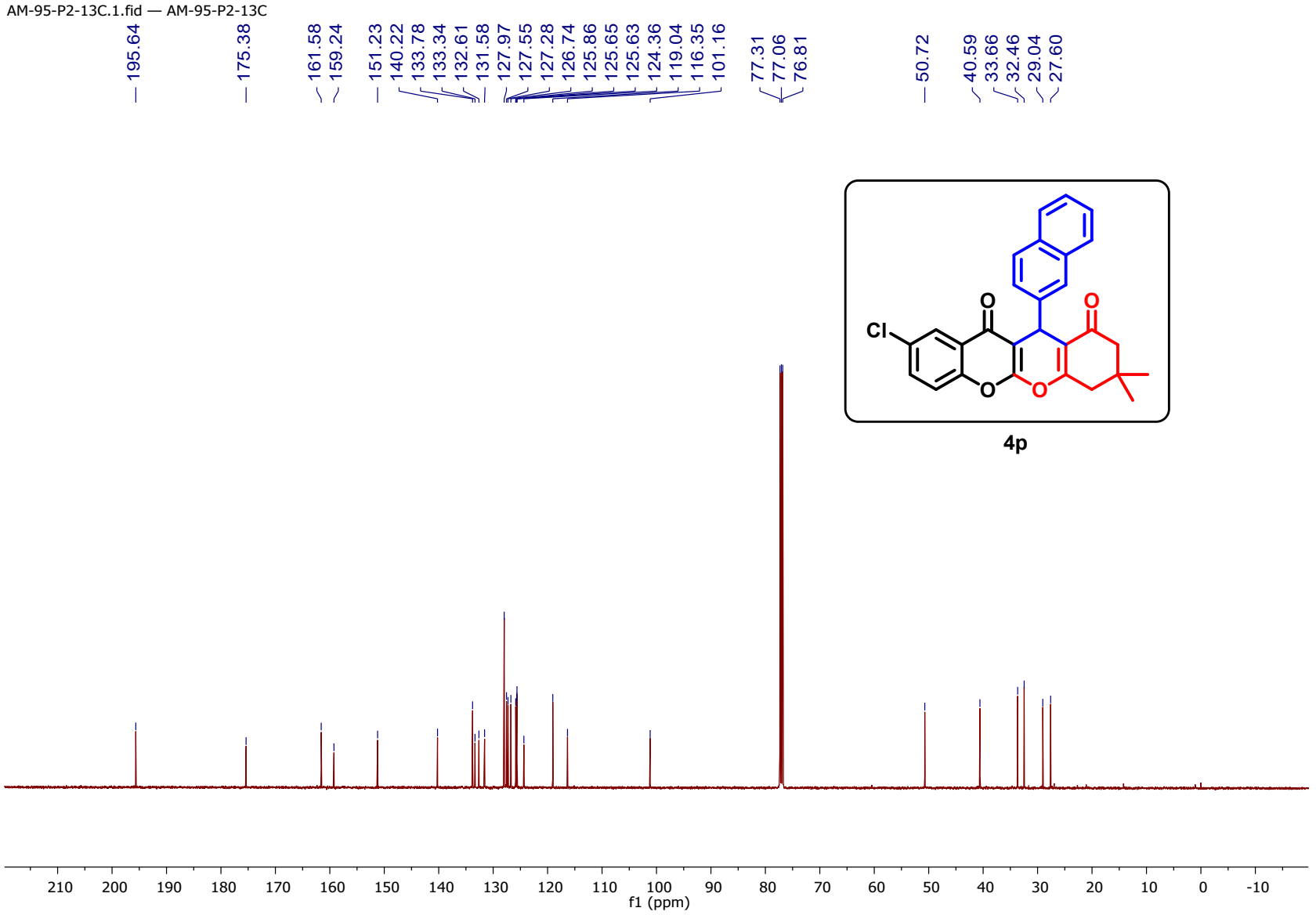
HRMS spectrum of compound 4o



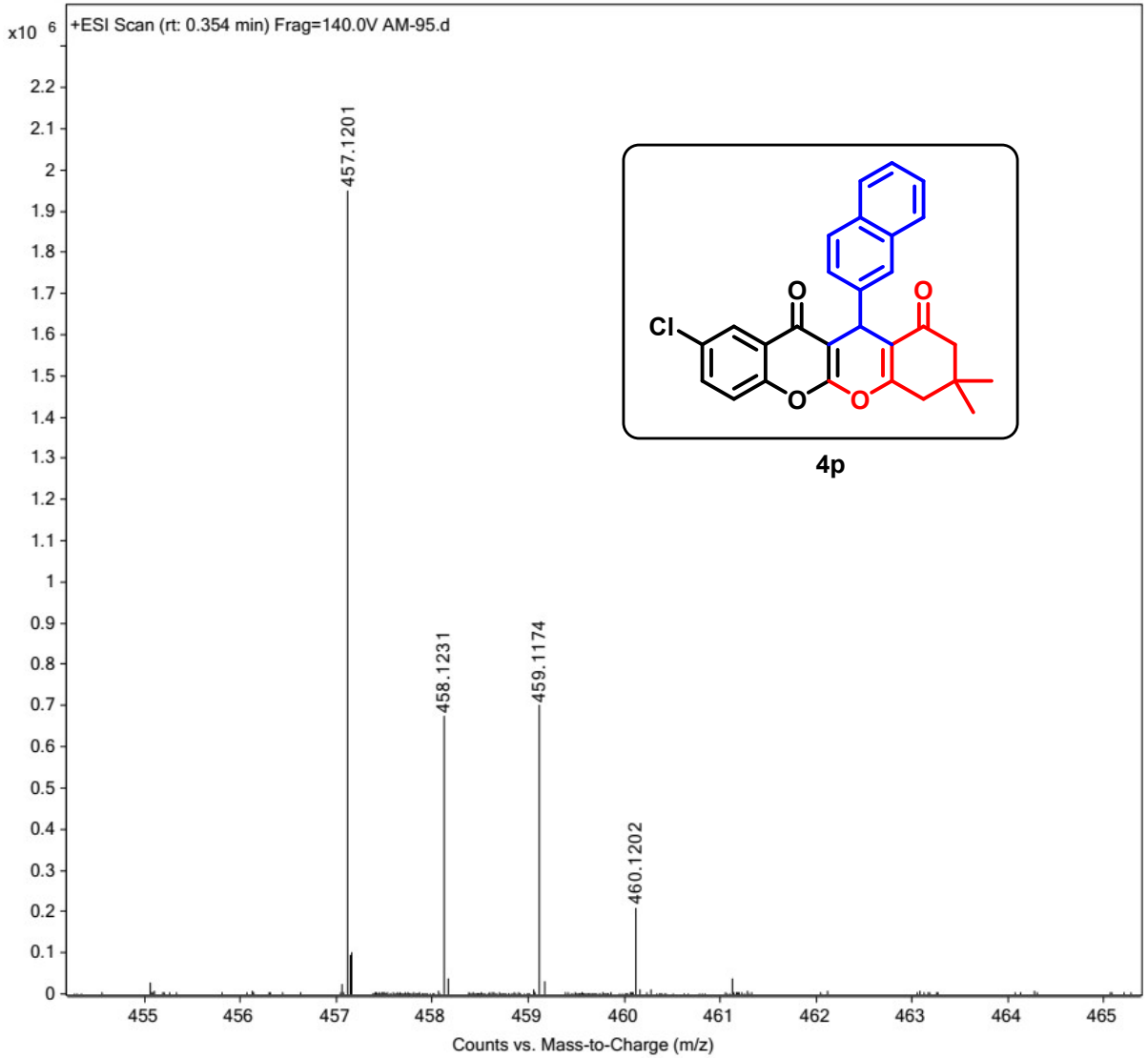
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4p



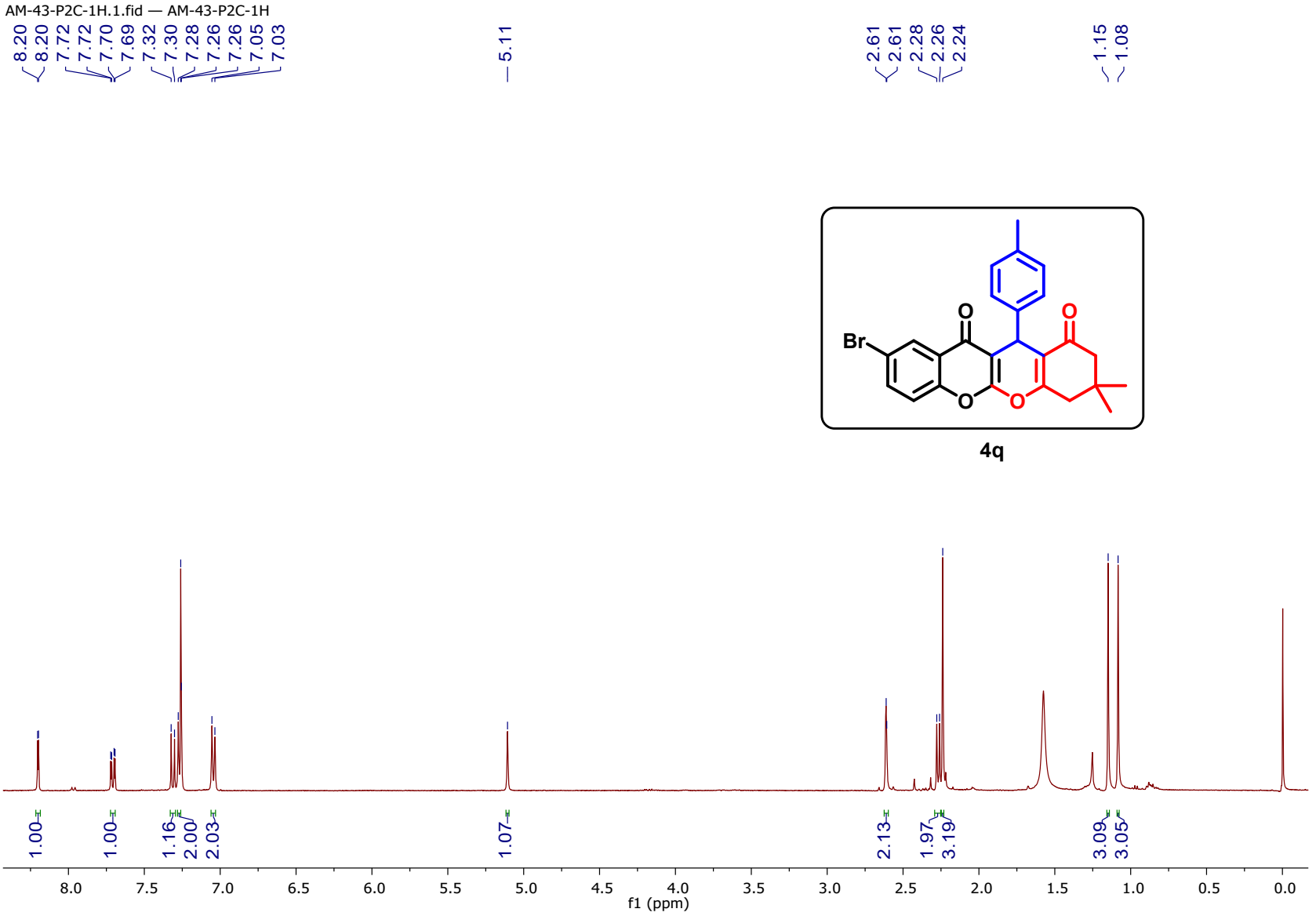
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4p



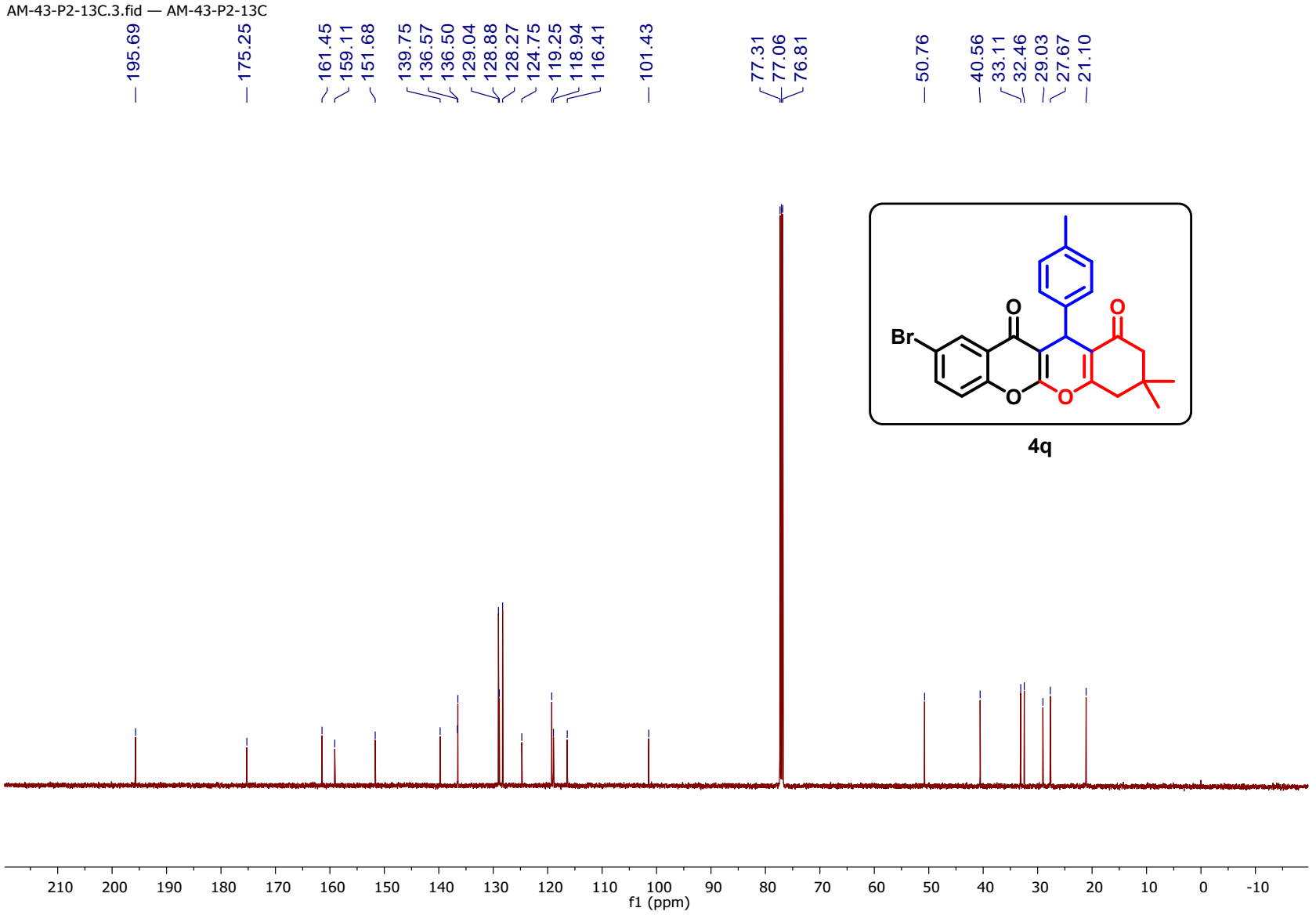
HRMS spectrum of compound 4p



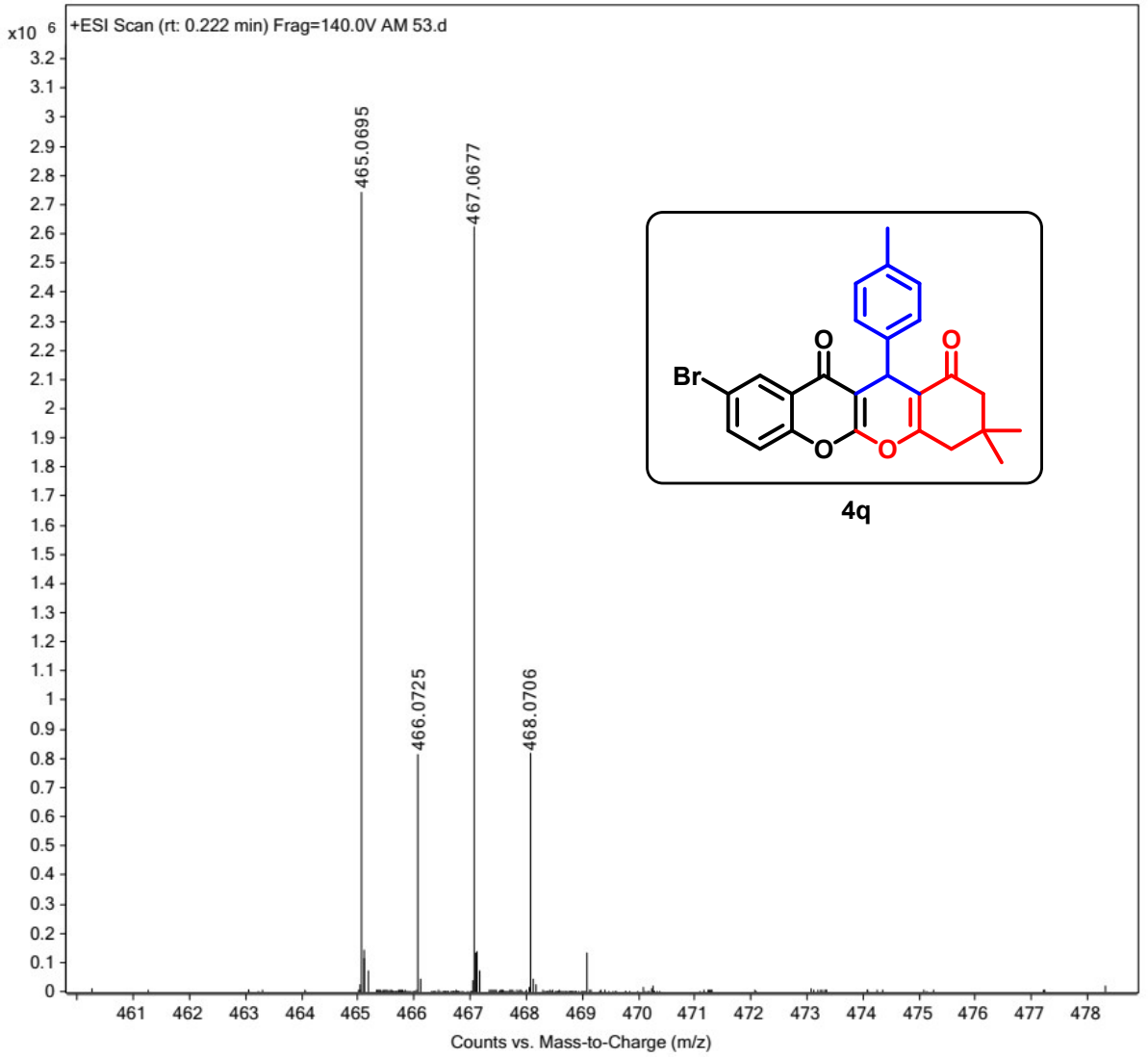
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4q



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4q

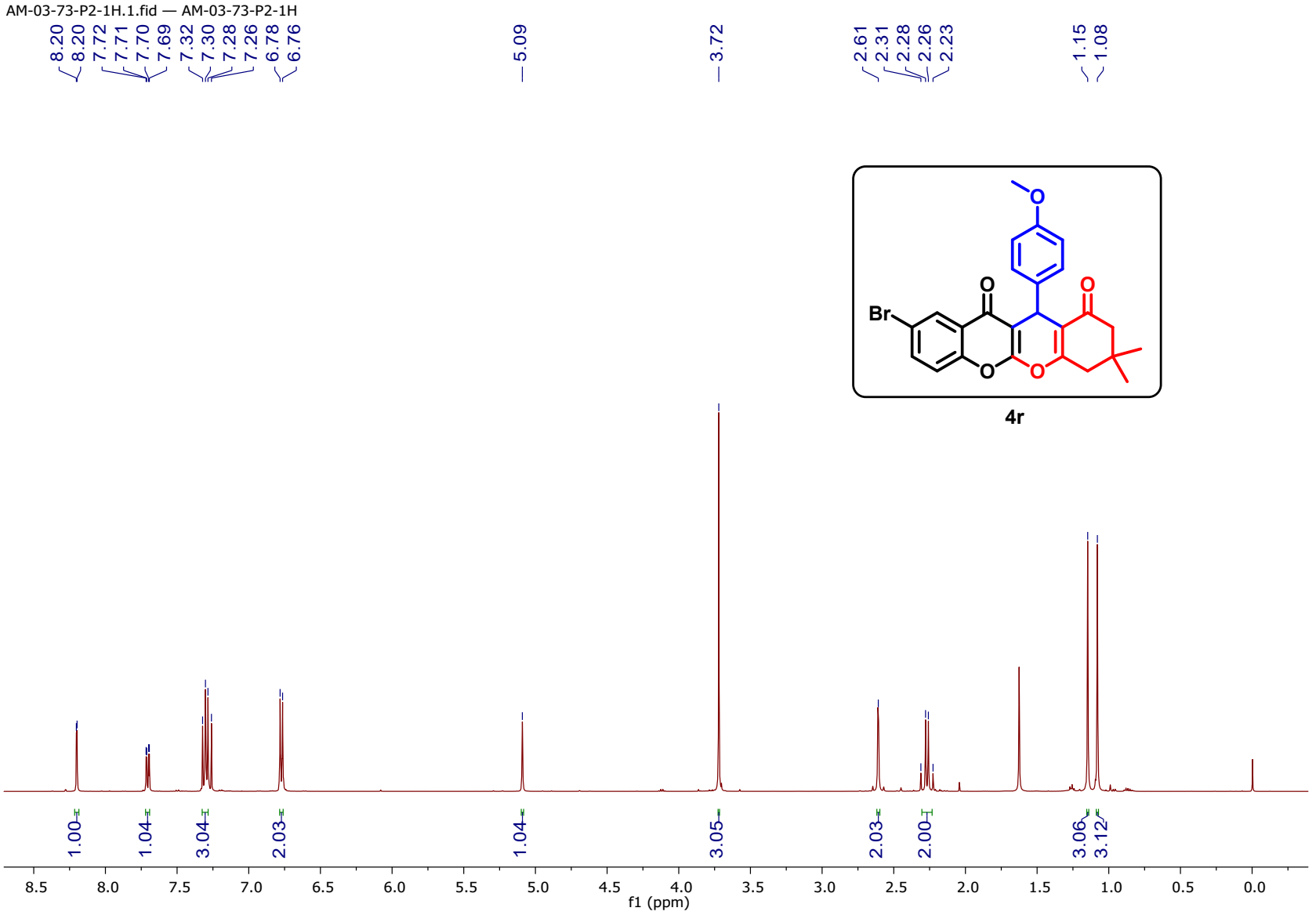


HRMS spectrum of compound 4q

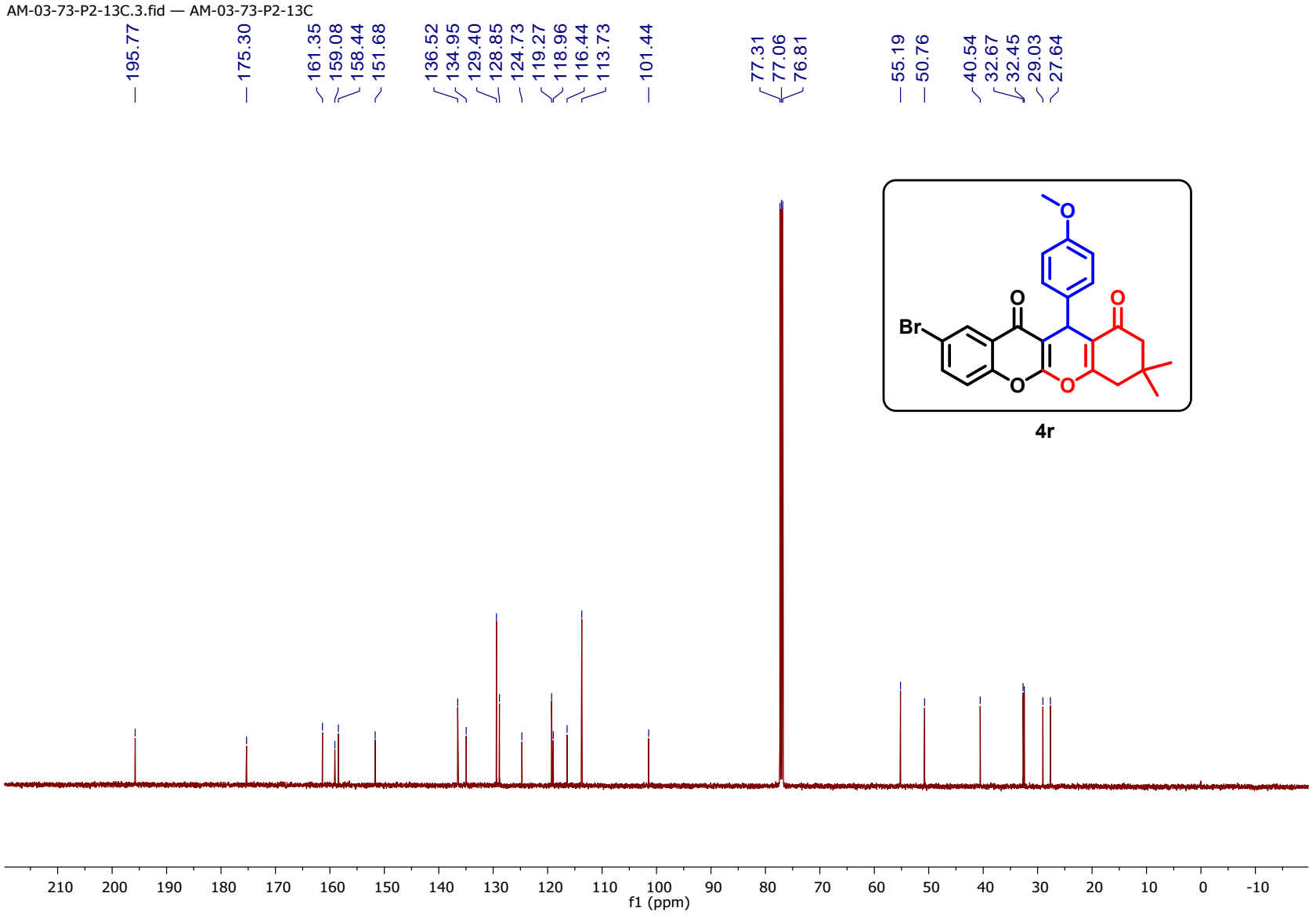


4q

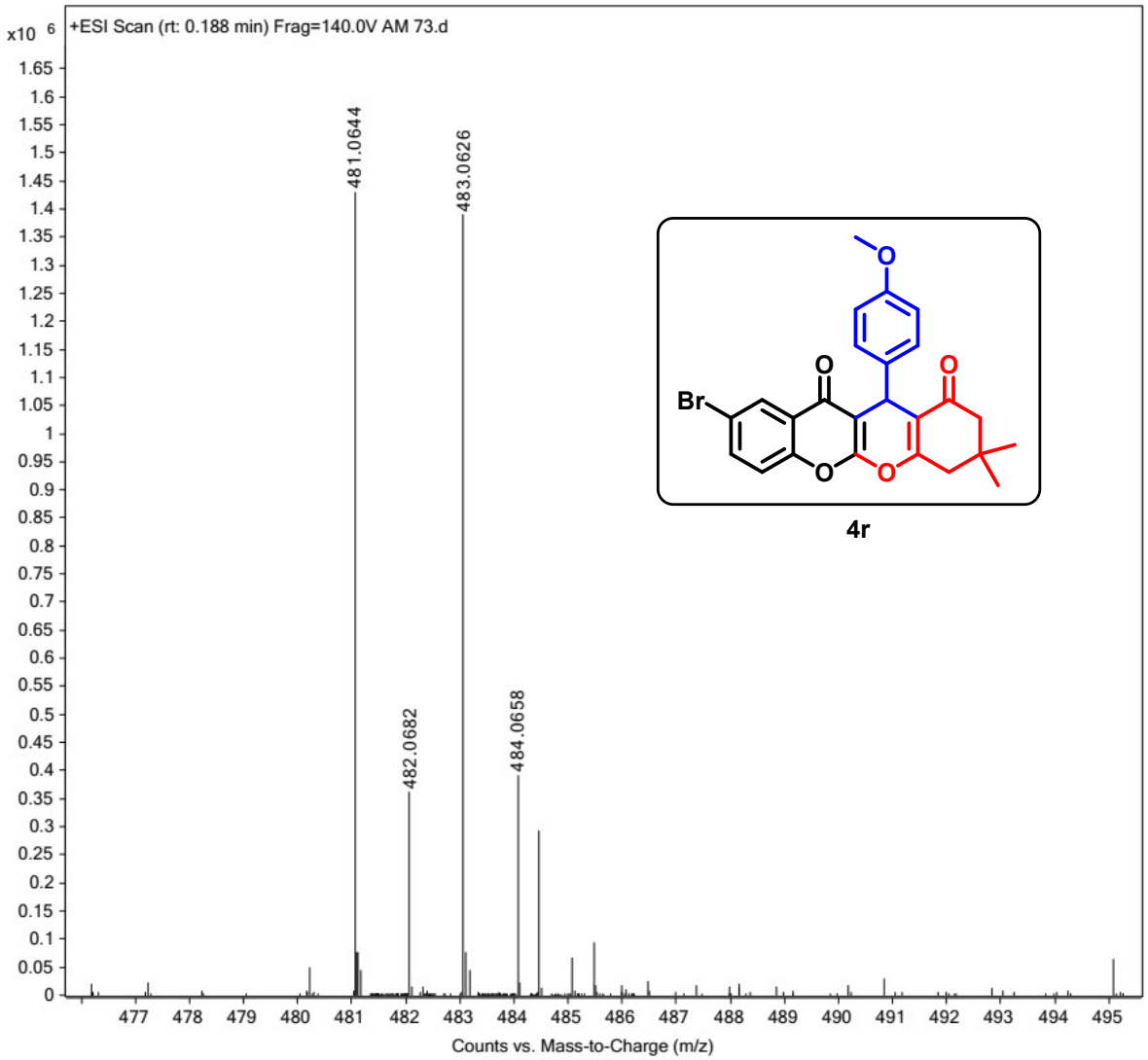
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4r



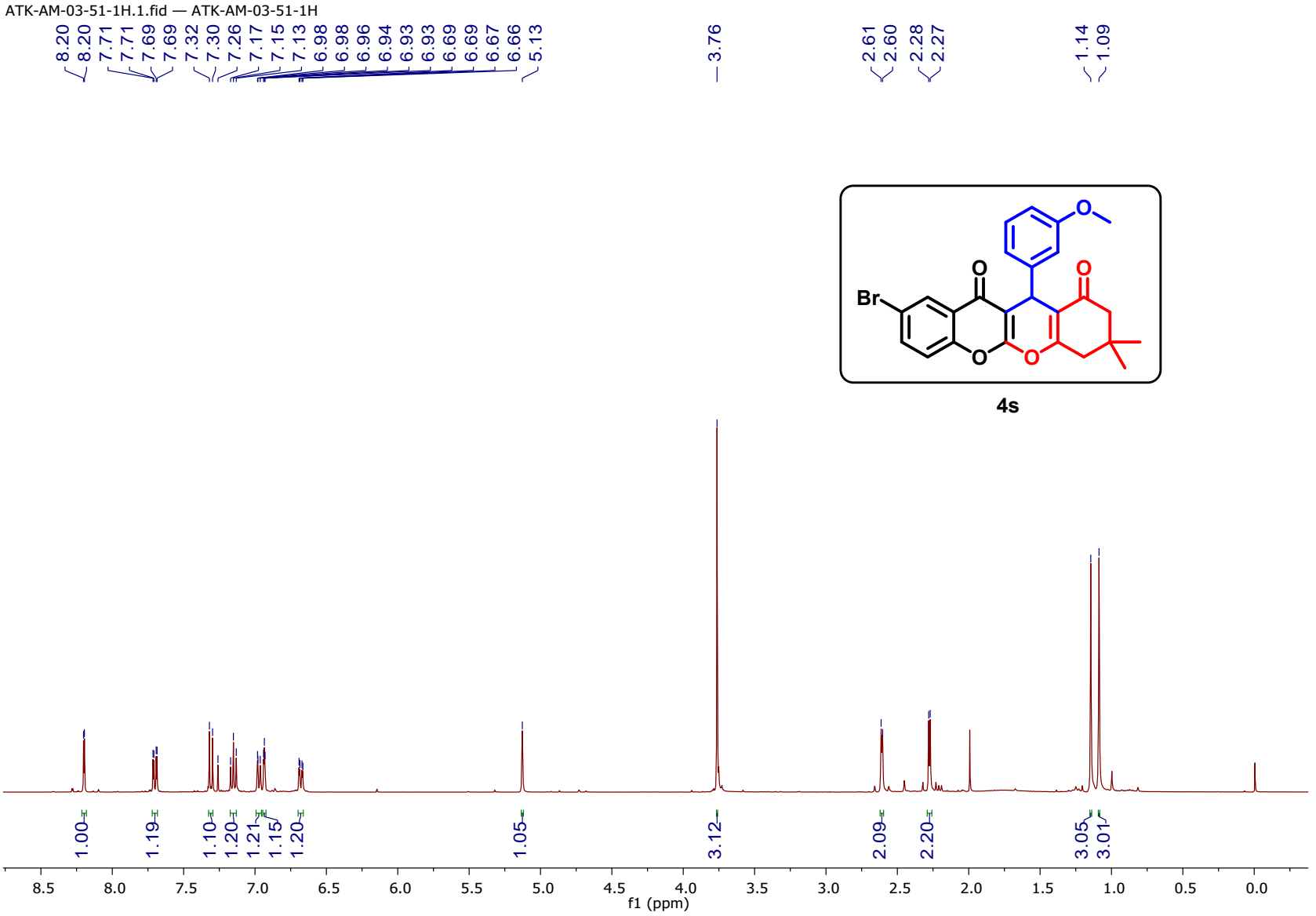
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4r



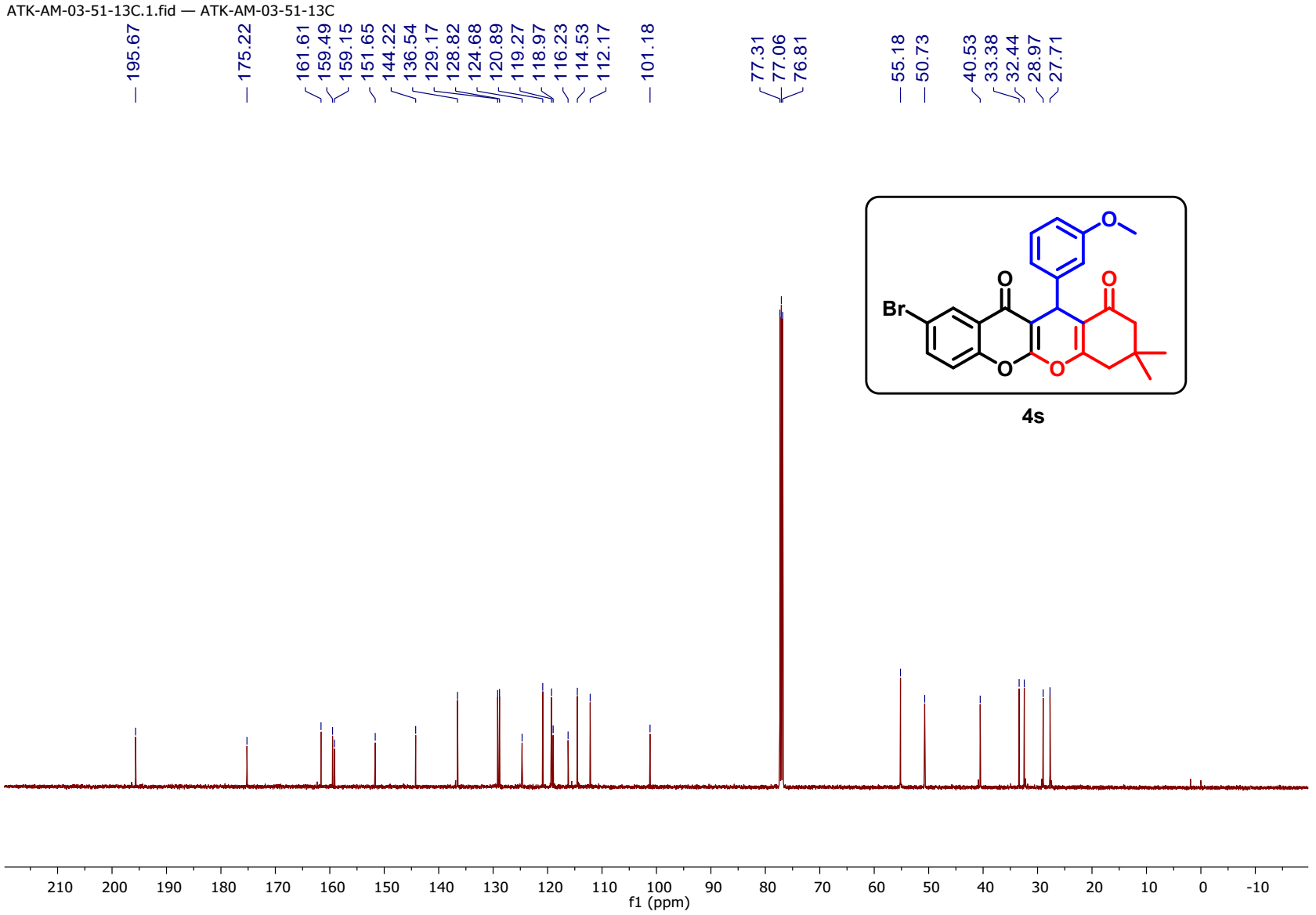
HRMS spectrum of compound 4r



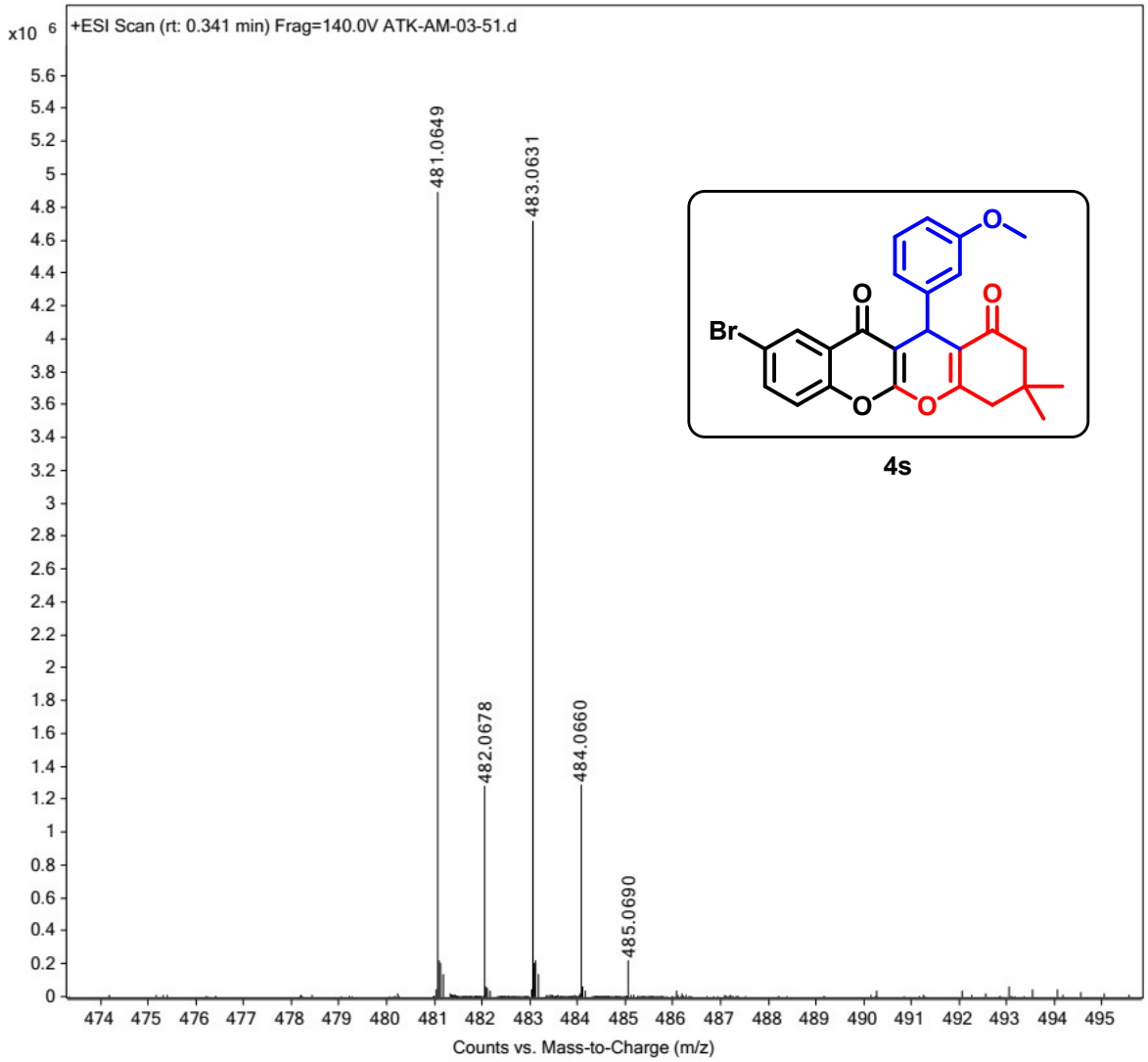
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4s



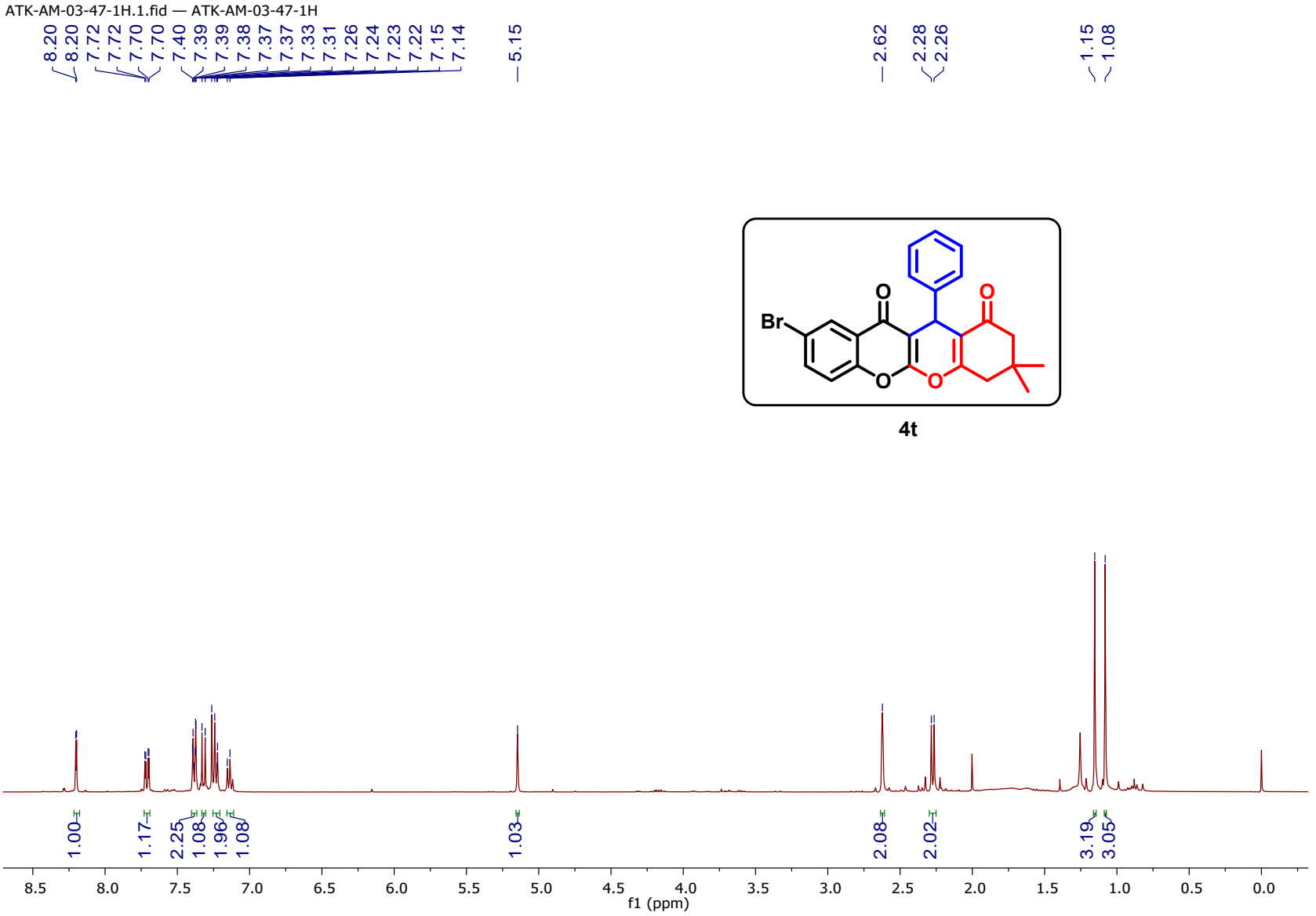
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4s



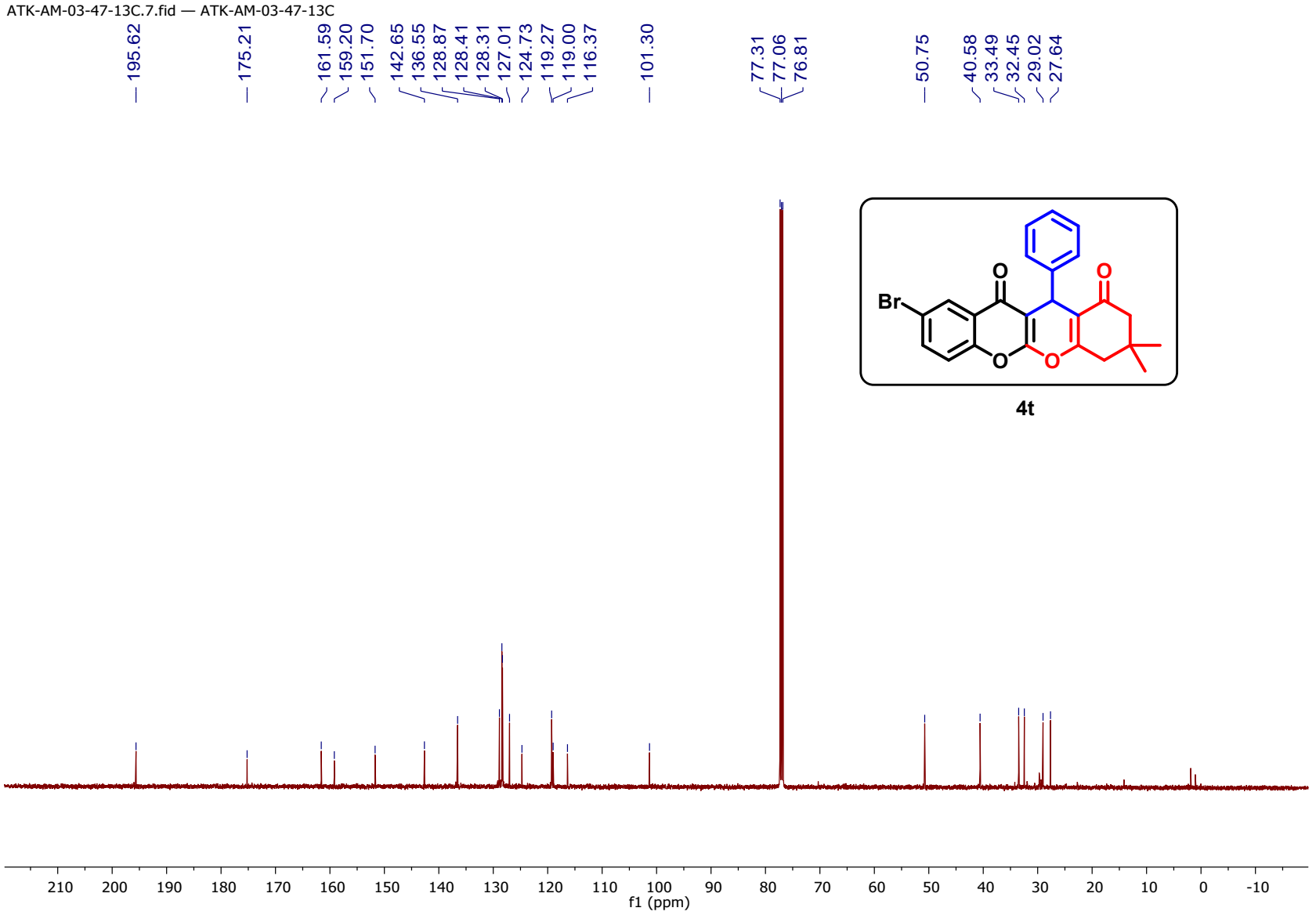
HRMS spectrum of compound 4s



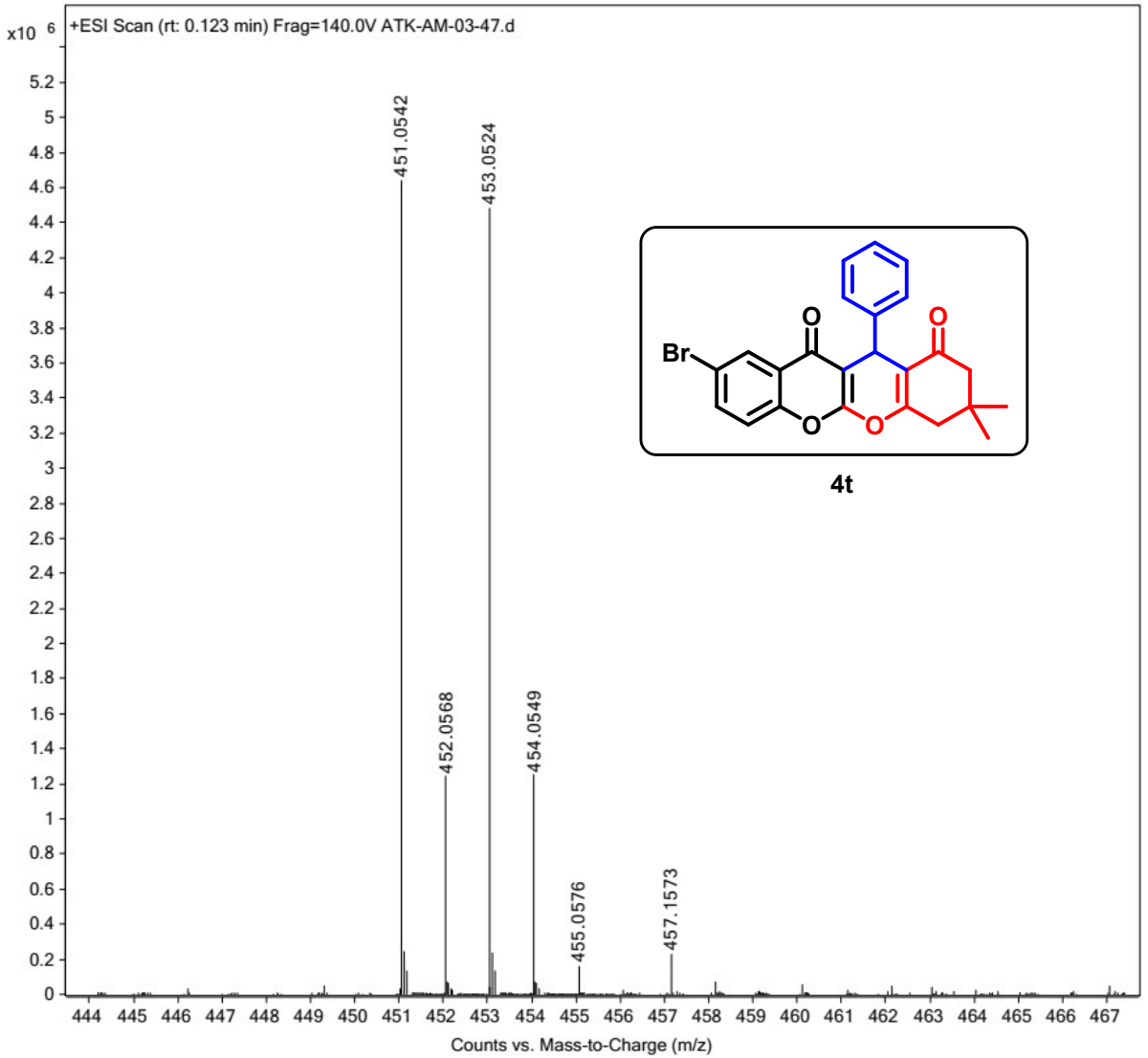
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4t



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4t

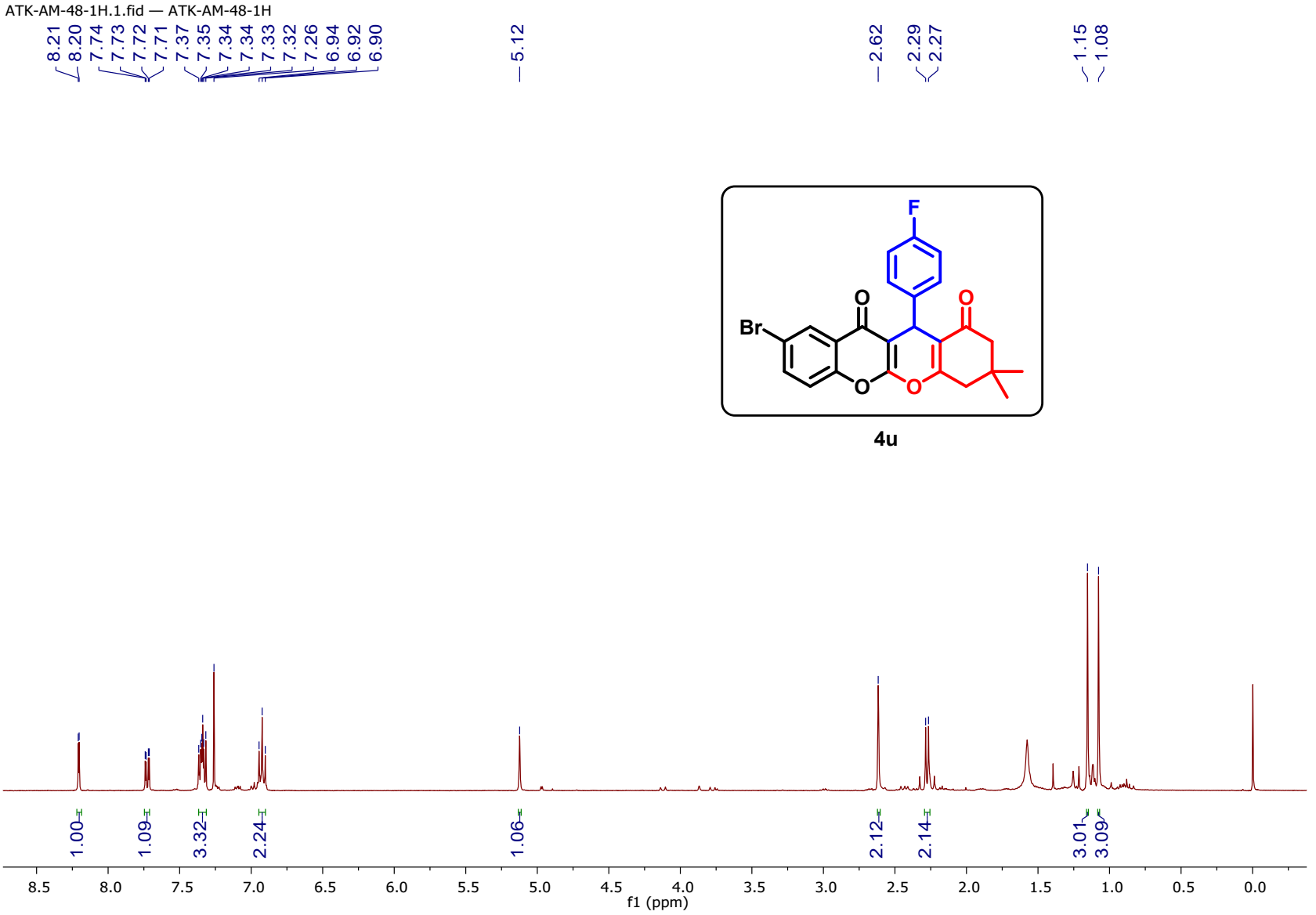


HRMS spectrum of compound 4t



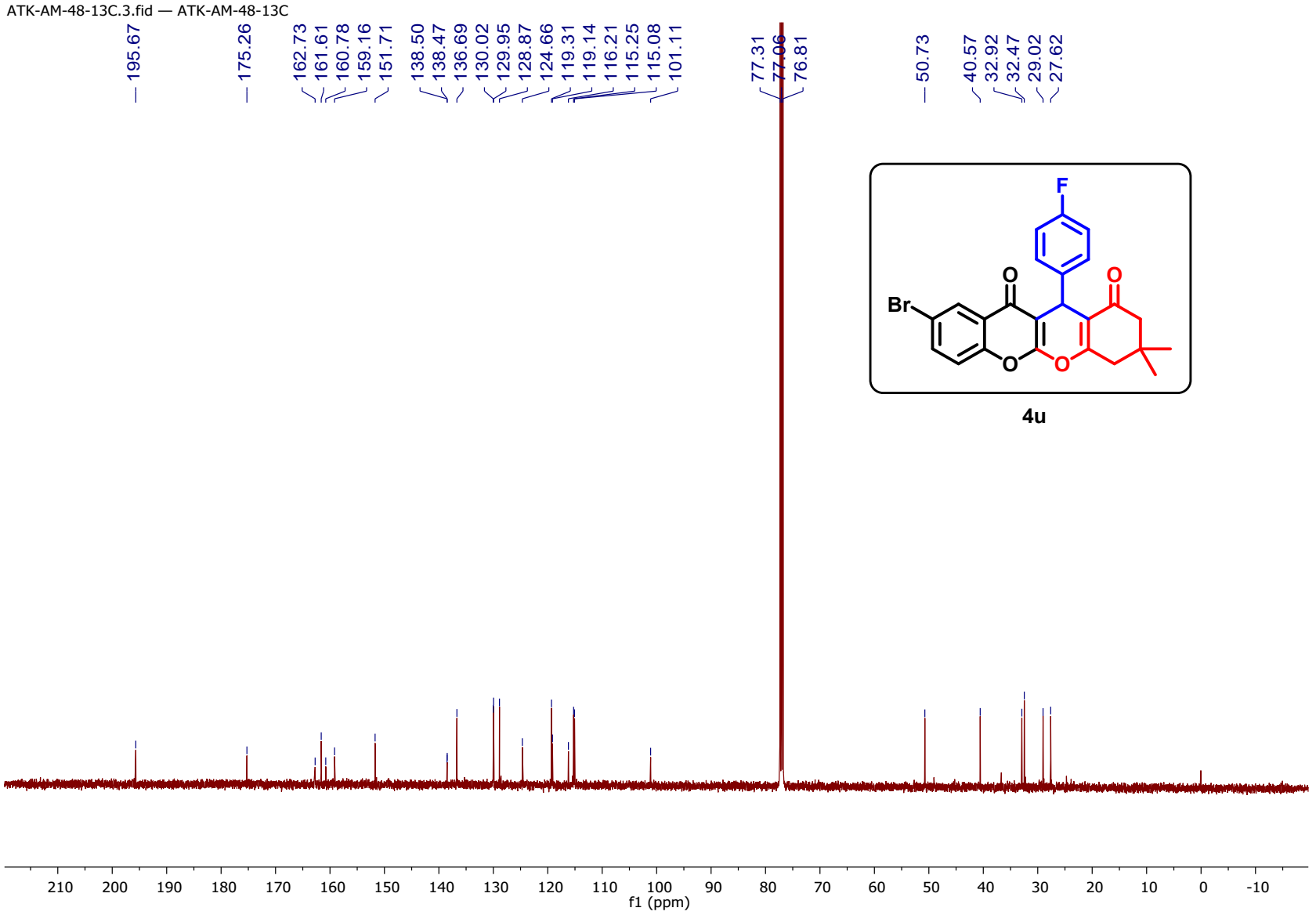
4t

^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4u

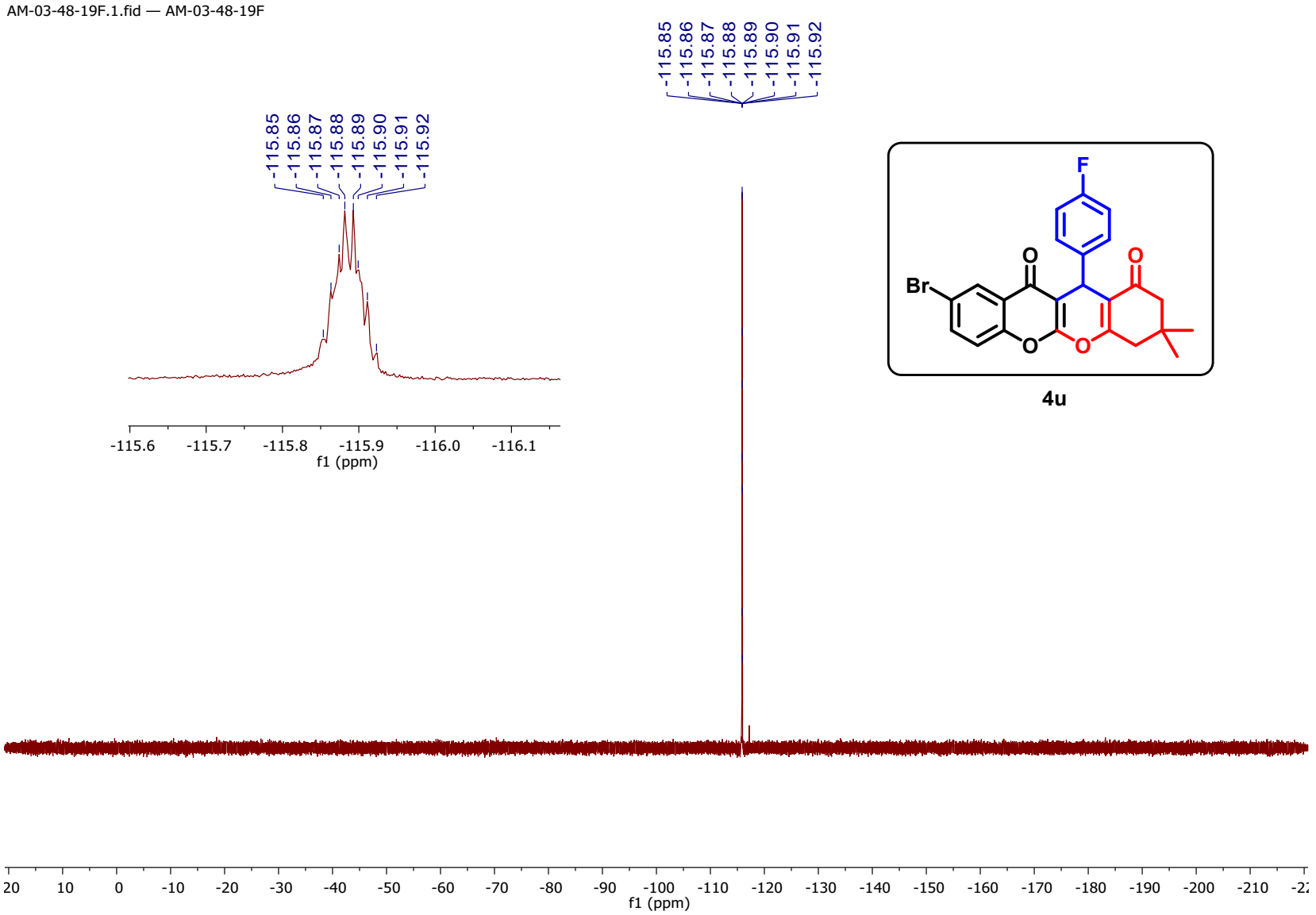


4u

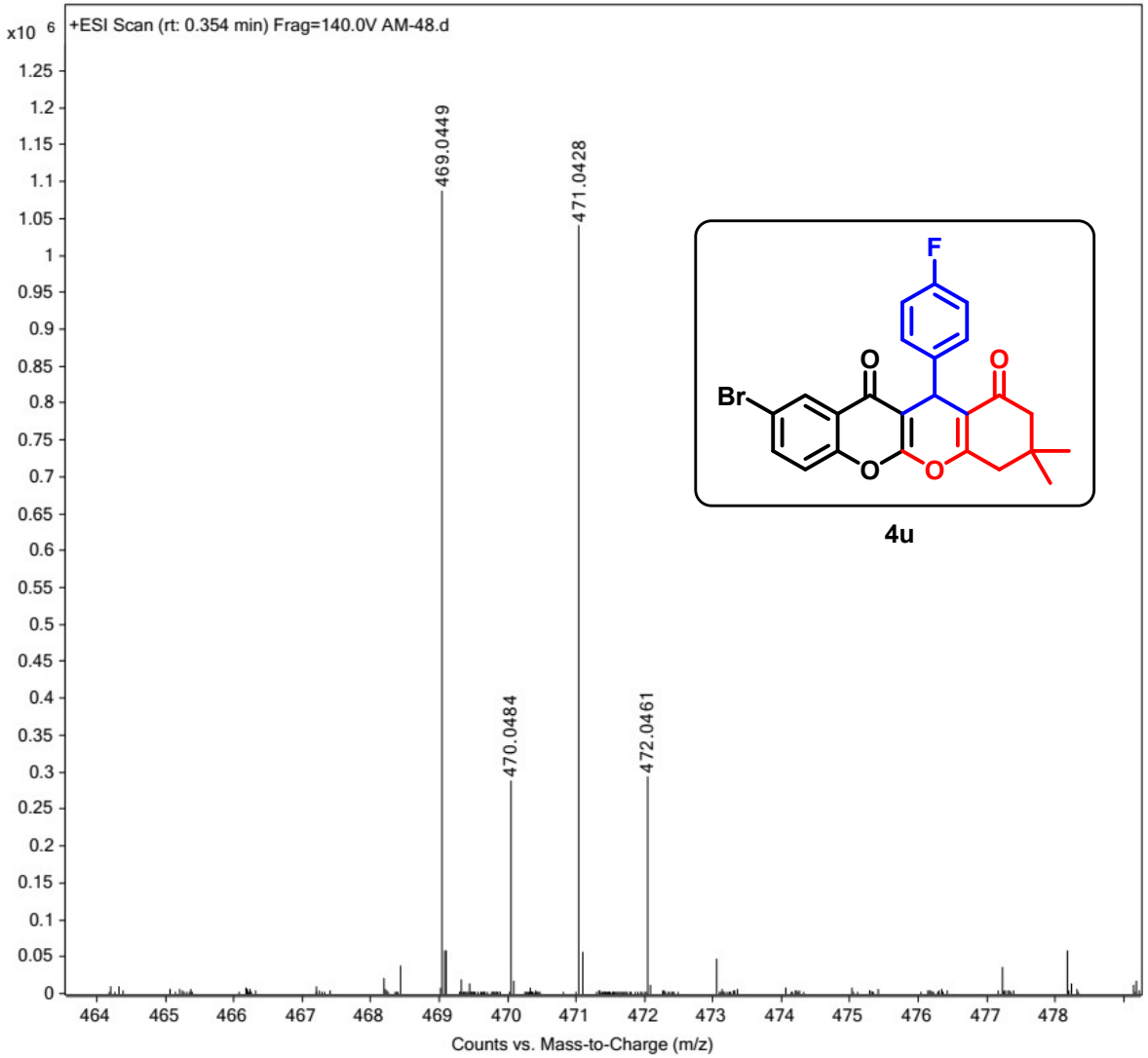
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4u



^{19}F NMR (471 MHz, $CDCl_3$) spectrum of compound 4u

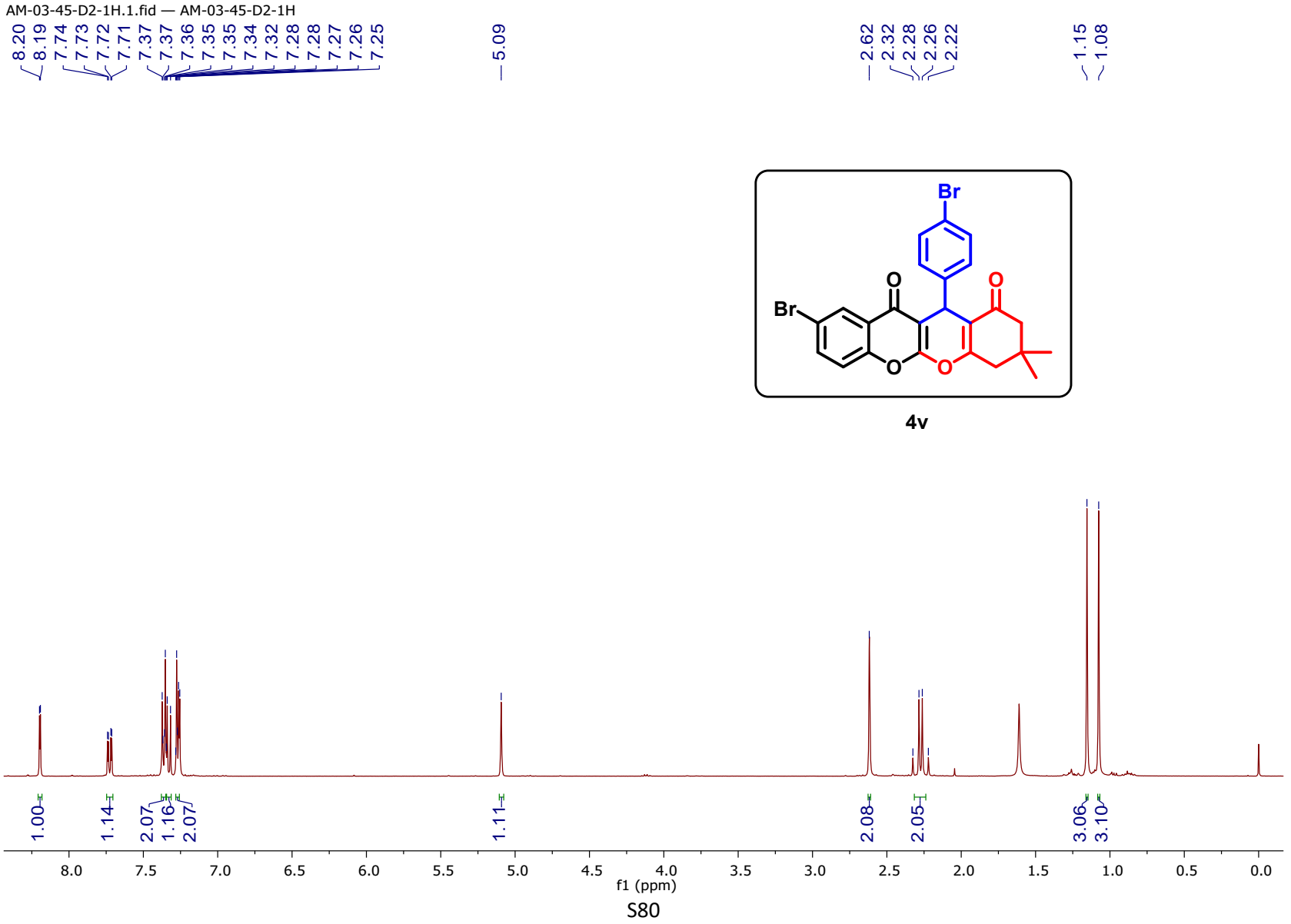


HRMS spectrum of compound 4u

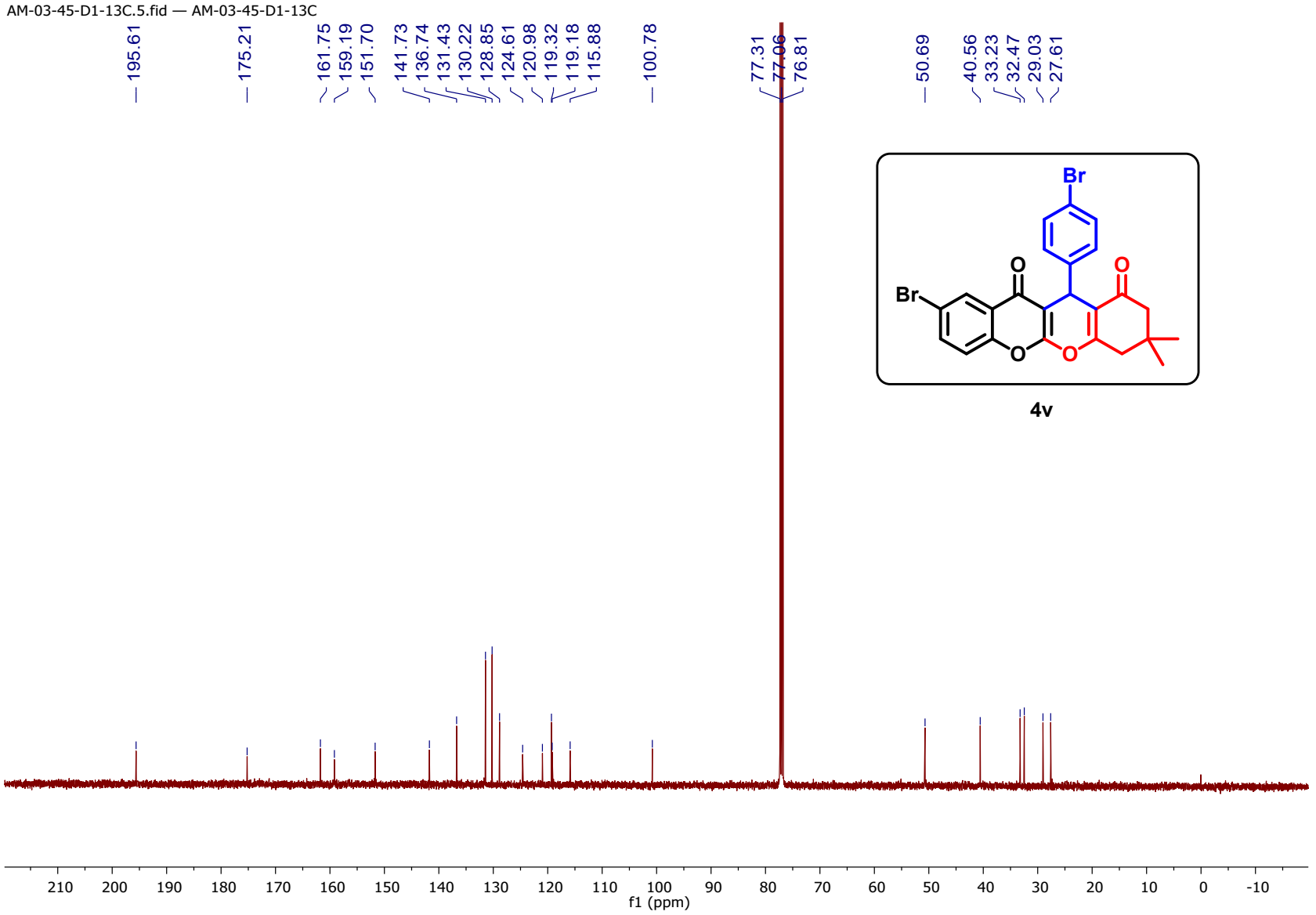


4u

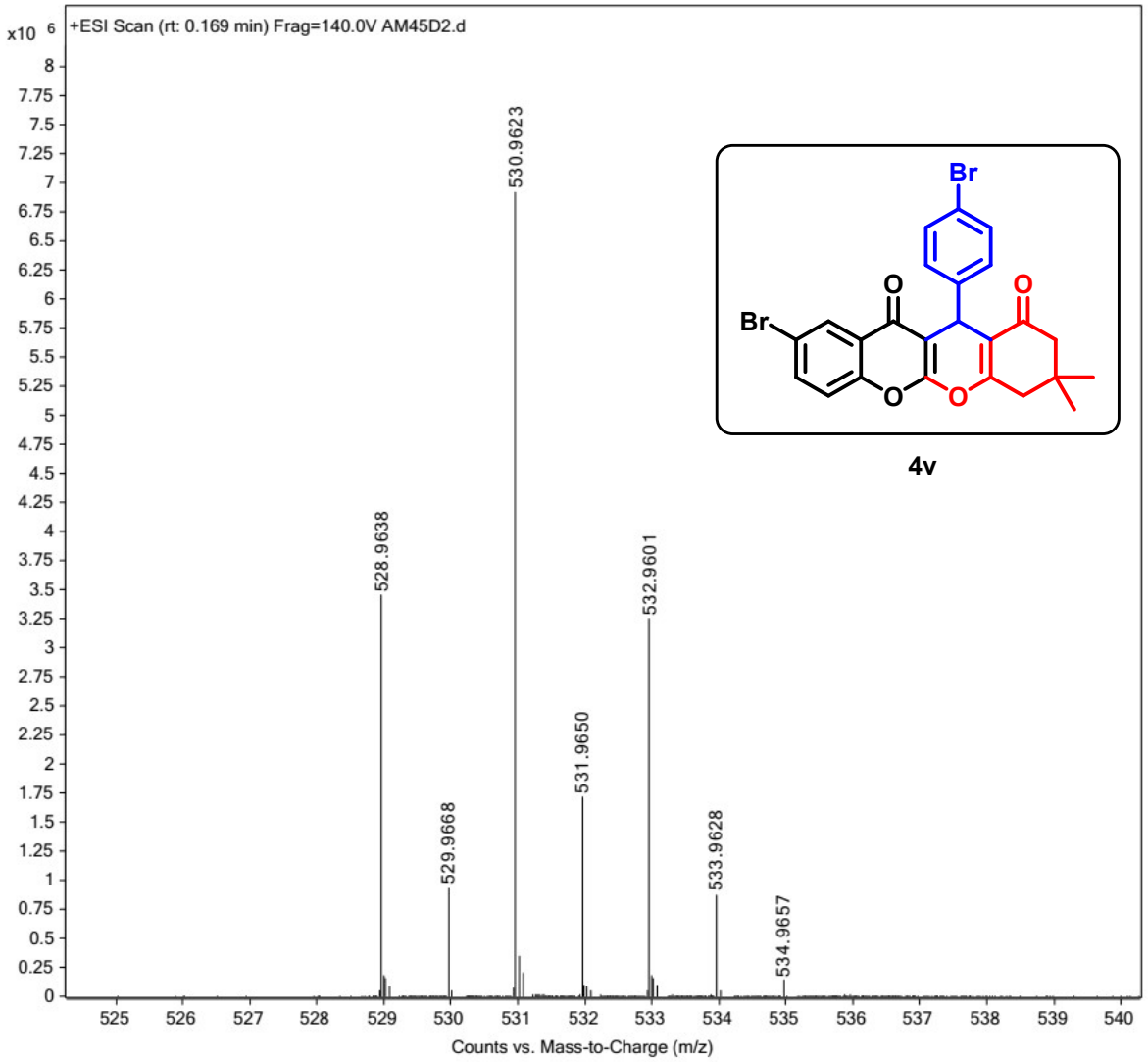
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4v



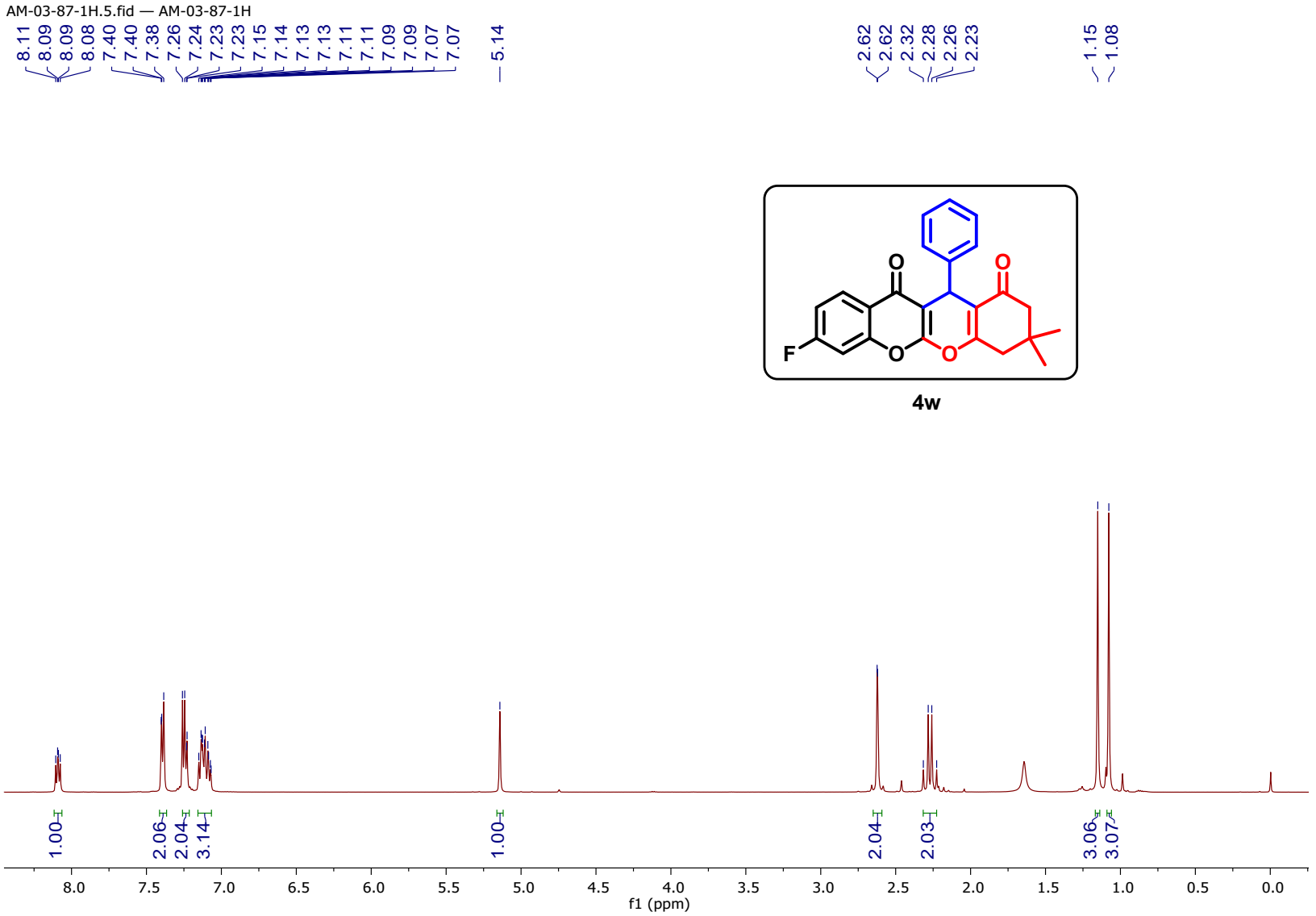
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4v



HRMS spectrum of compound 4v

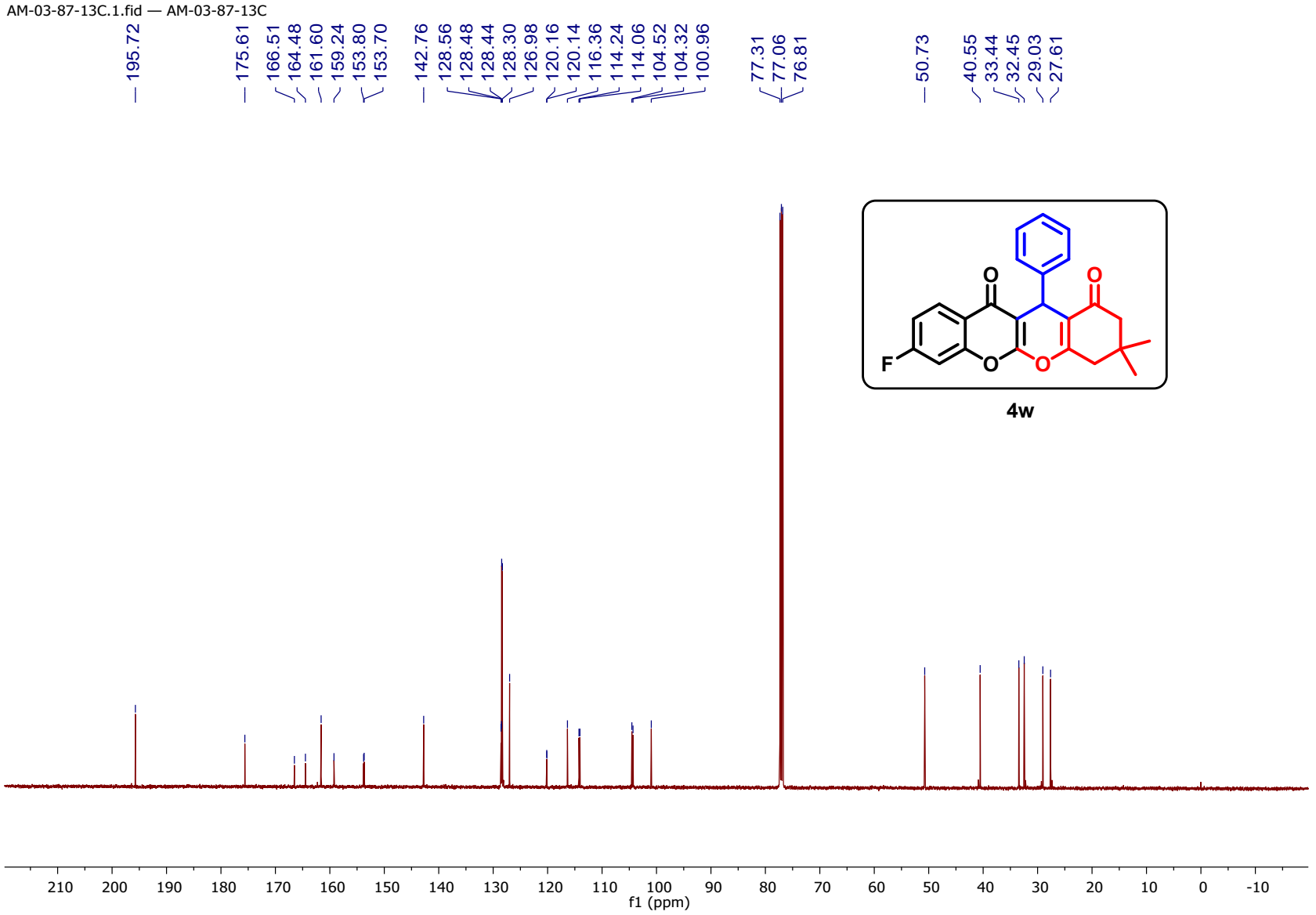


^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4w

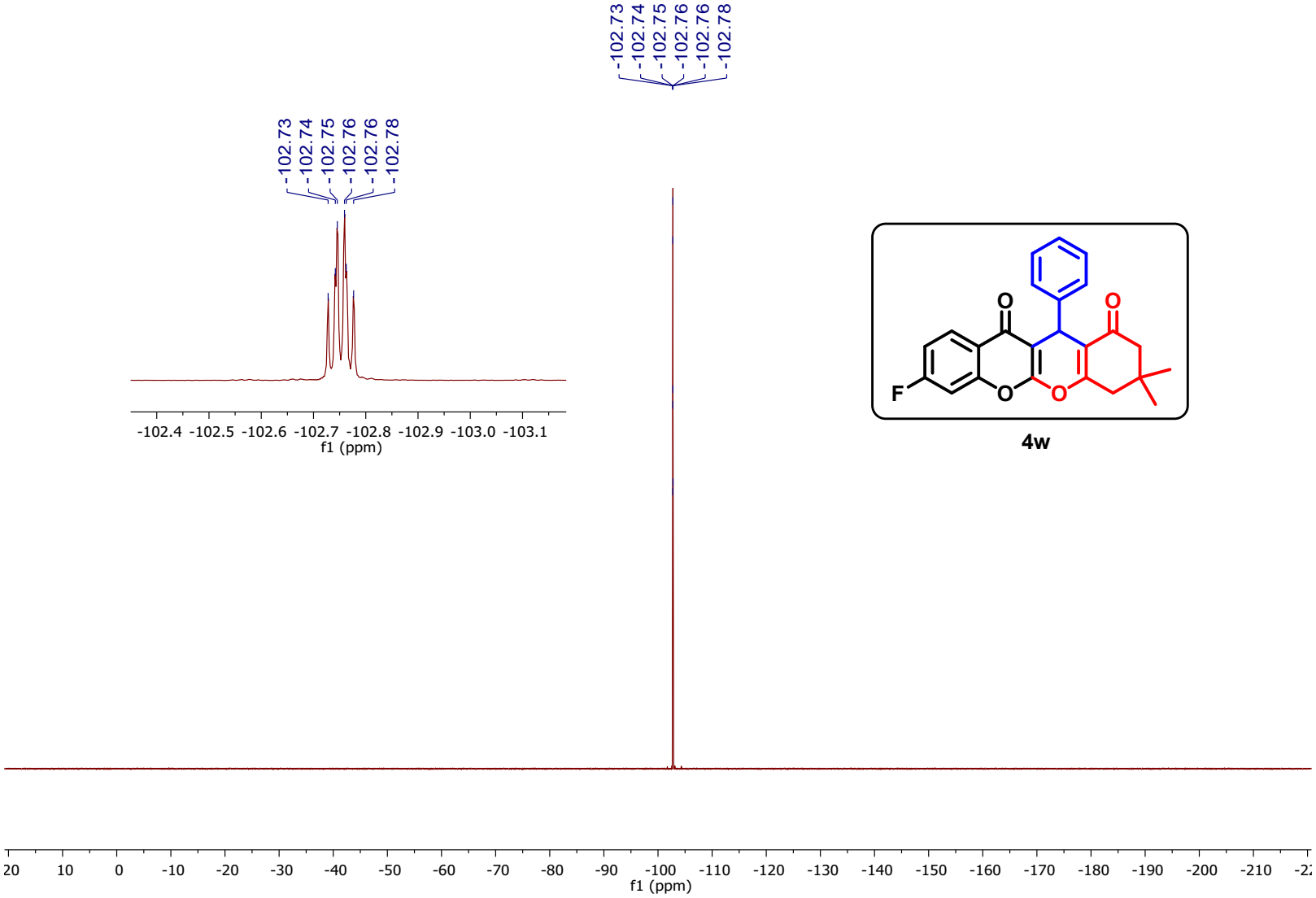


4w

^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4w

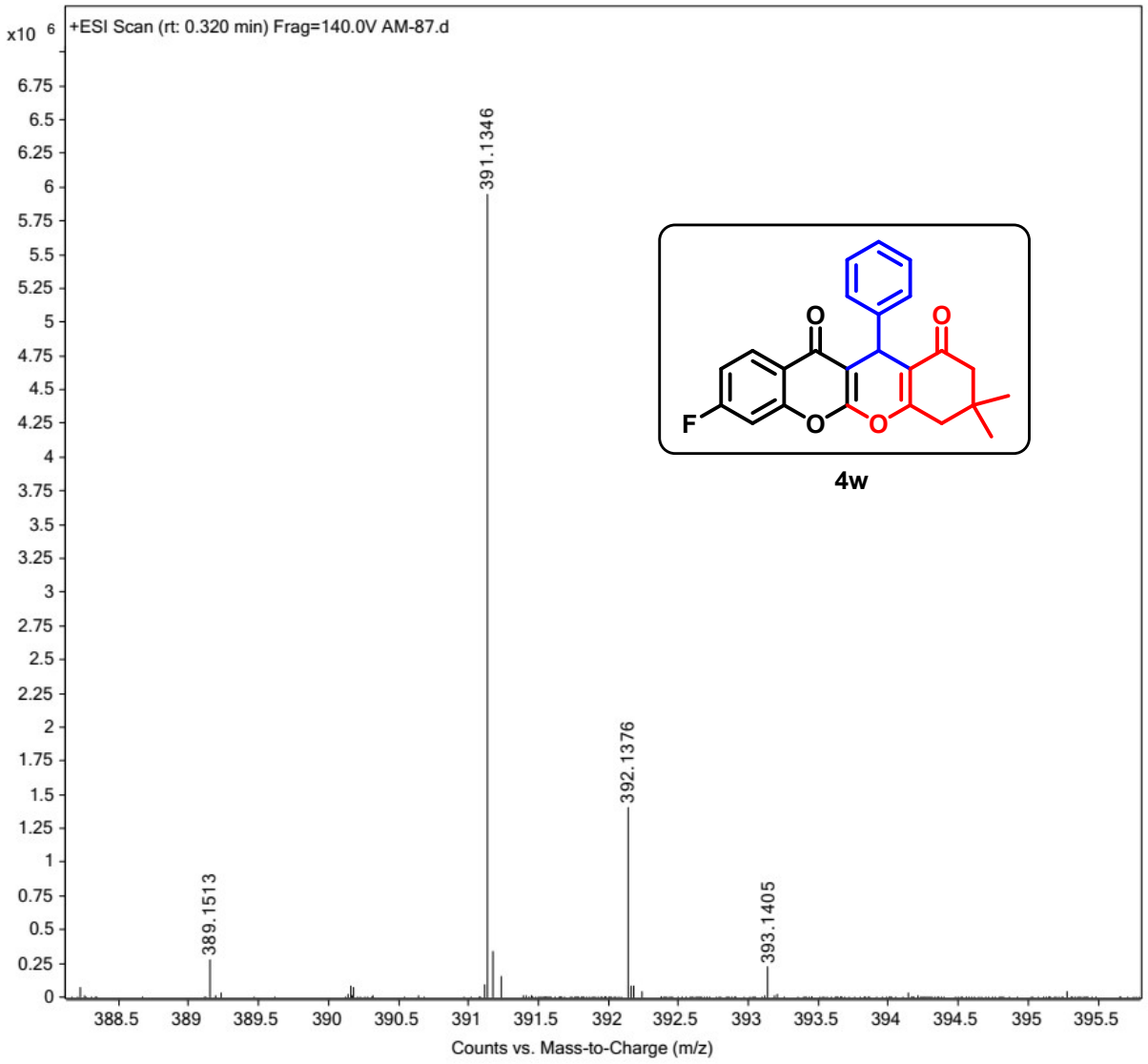


^{19}F NMR (471 MHz, $CDCl_3$) spectrum of compound 4w



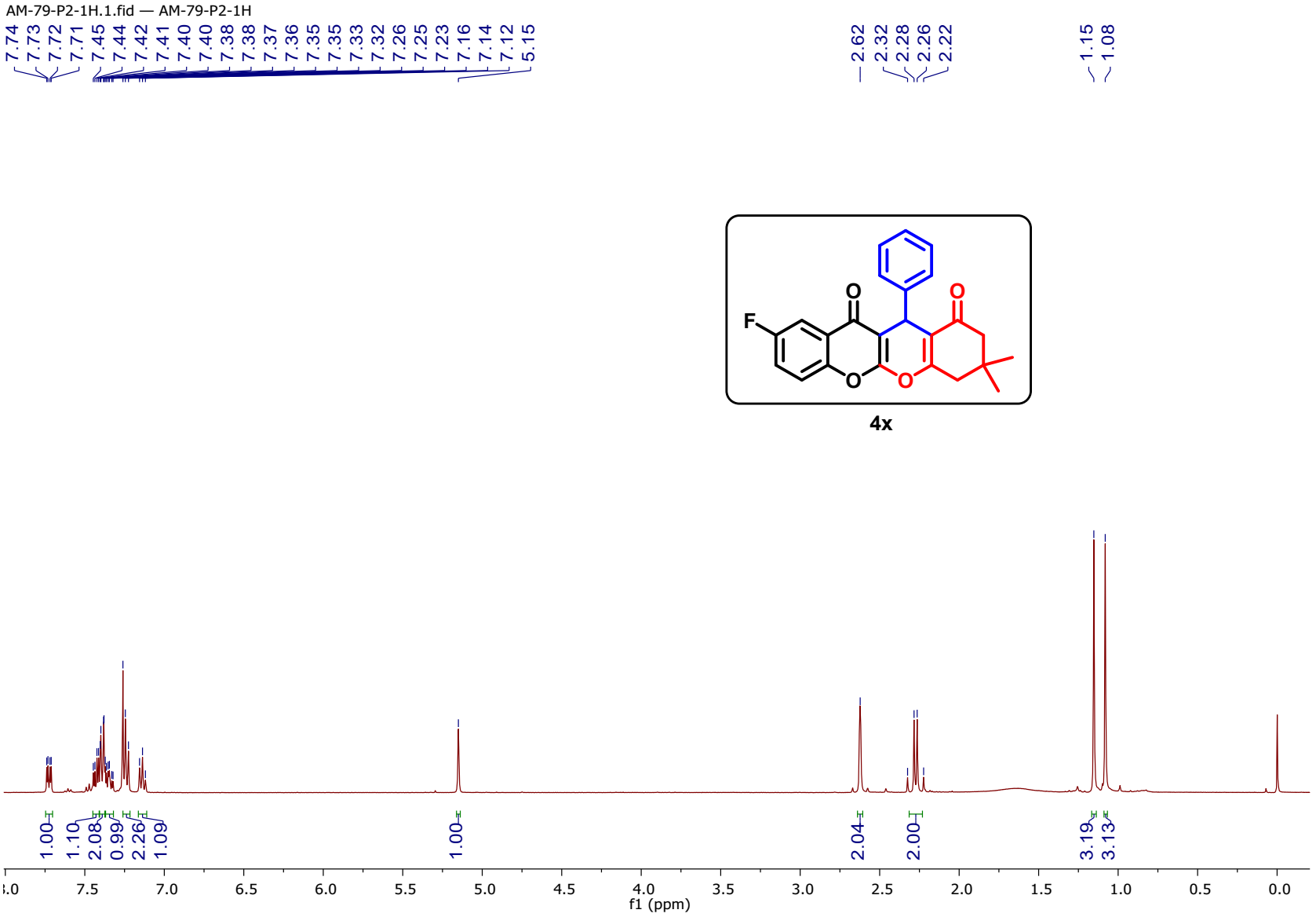
4w

HRMS spectrum of compound 4w



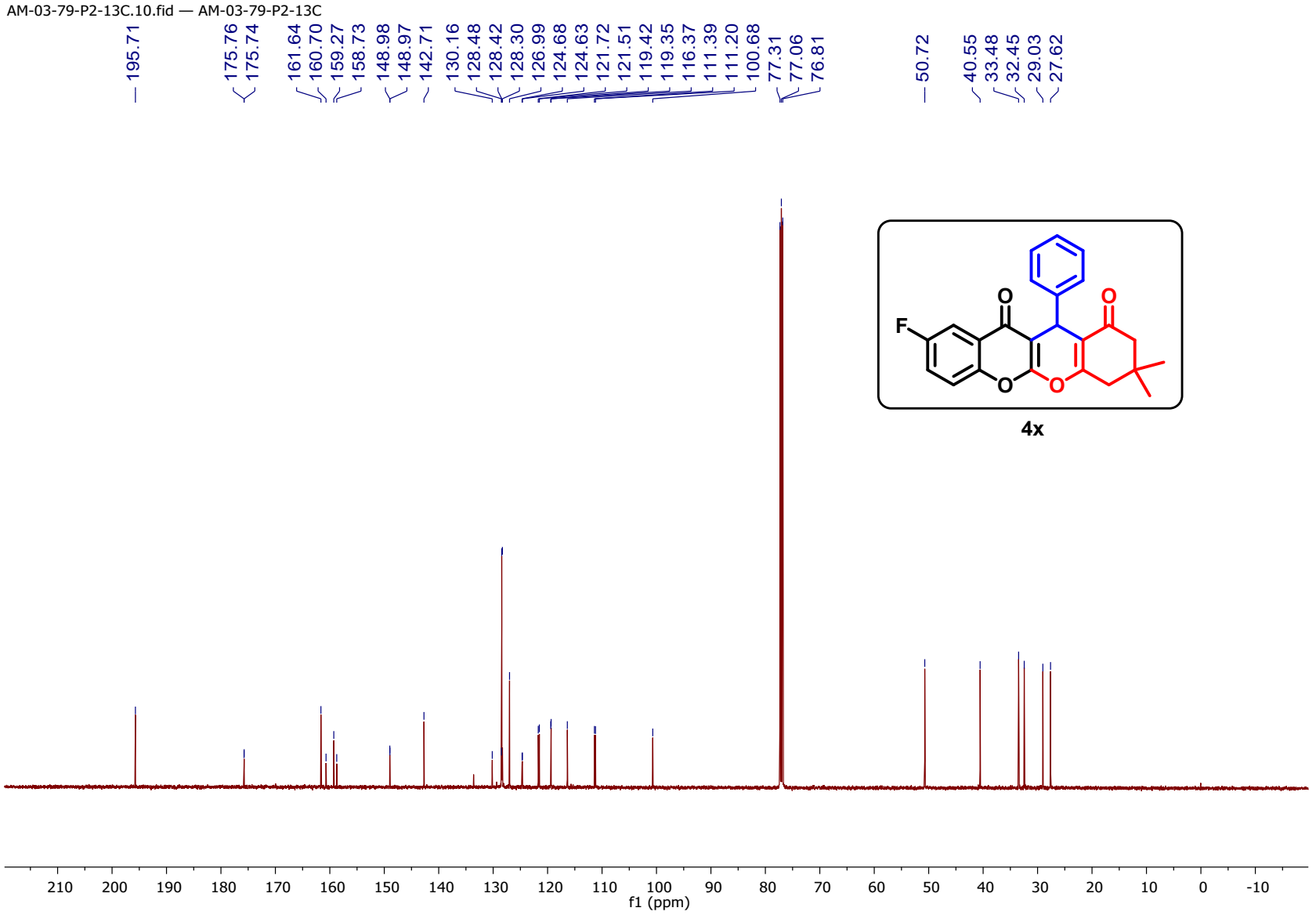
4w

^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4x

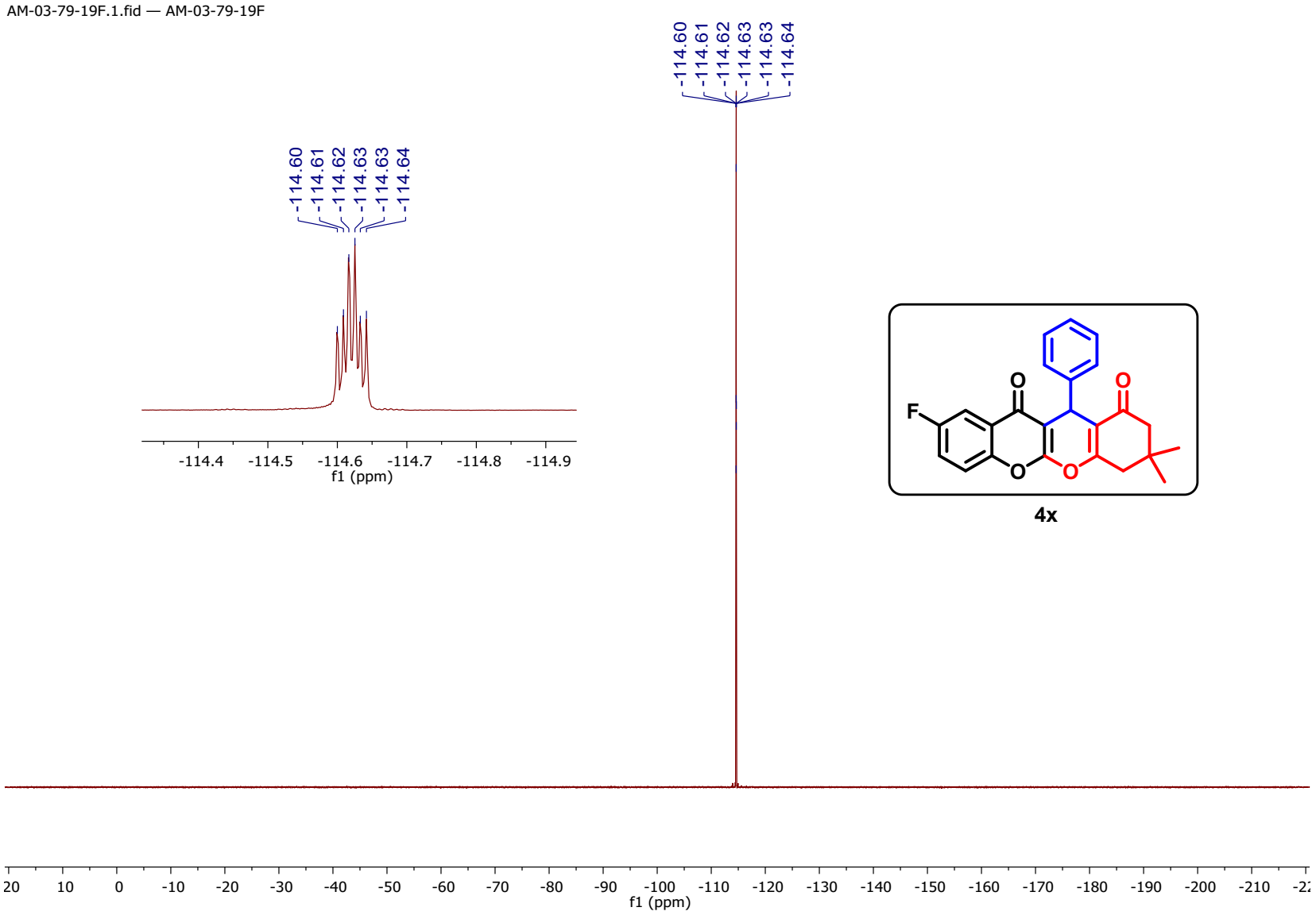


4x

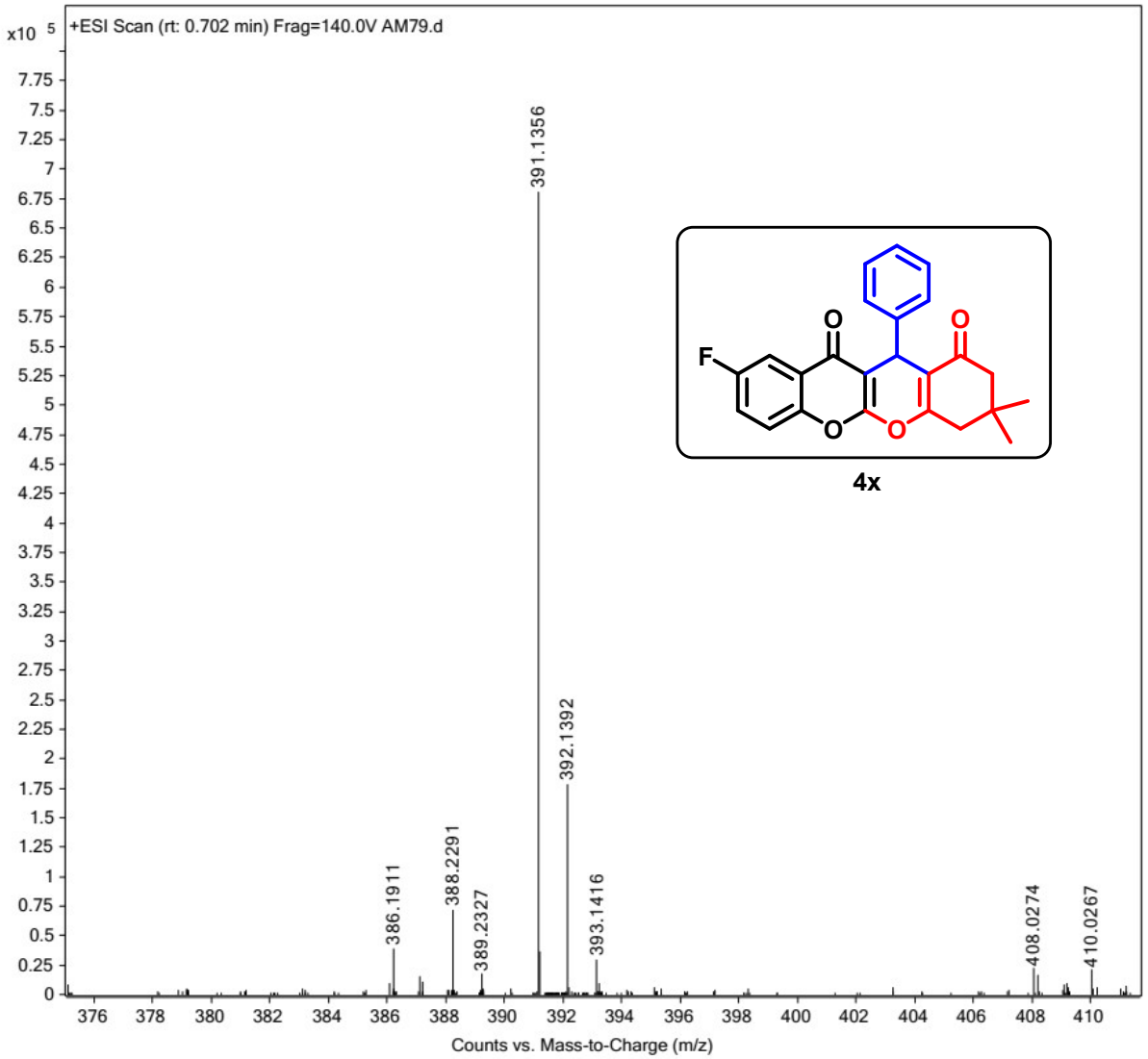
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4x



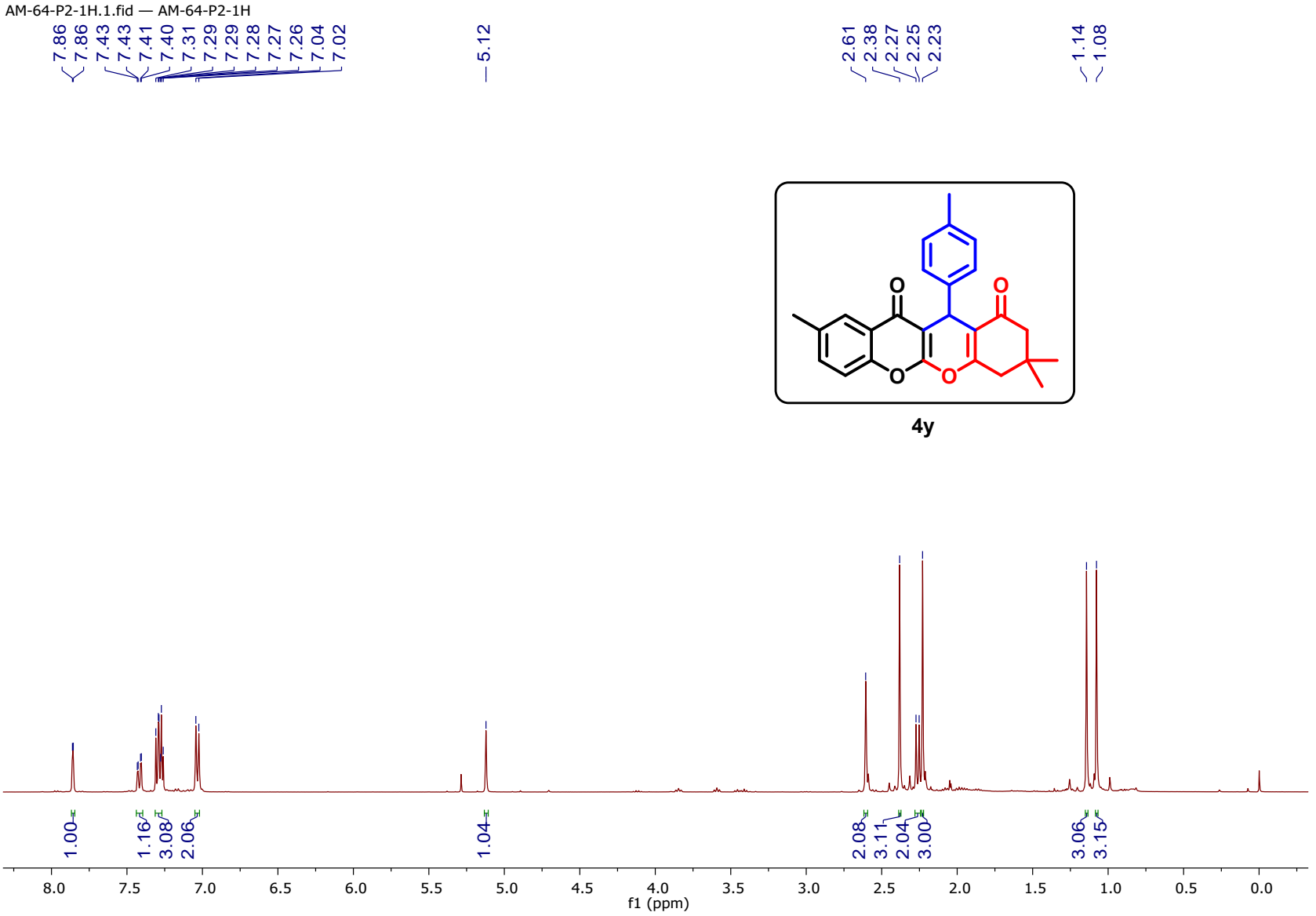
^{19}F NMR (471 MHz, $CDCl_3$) spectrum of compound 4x



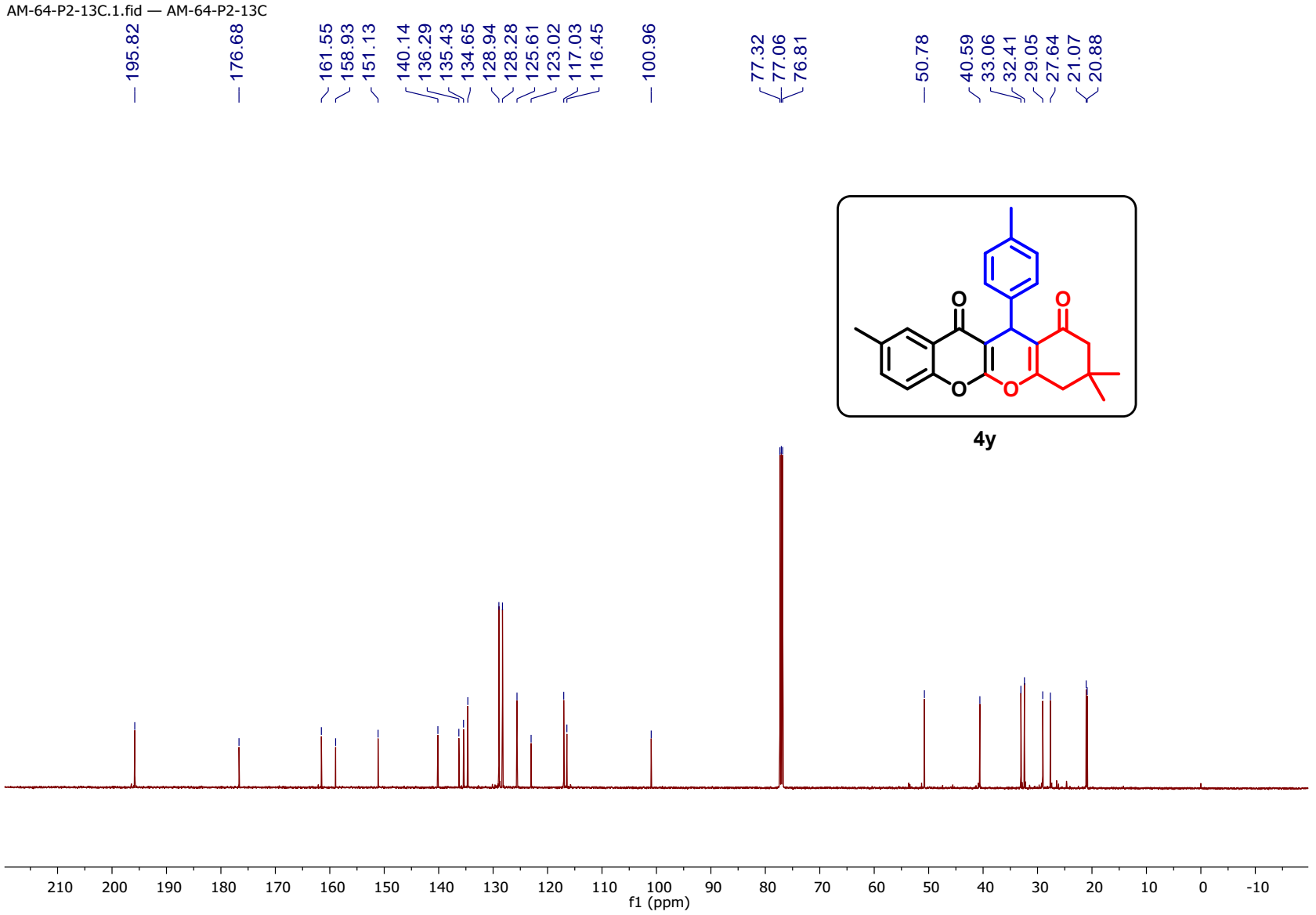
HRMS spectrum of compound 4x



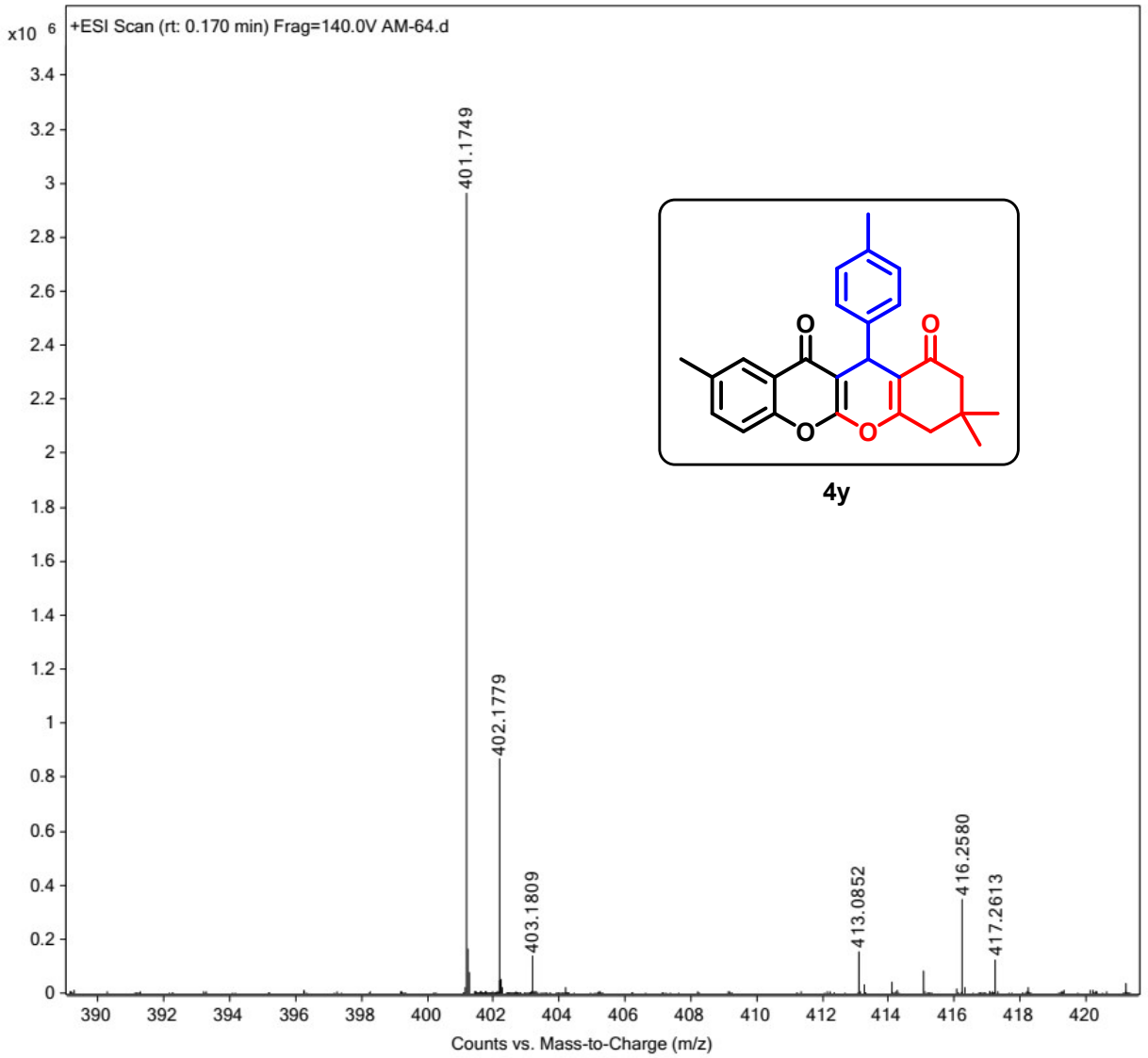
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4y



^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4y

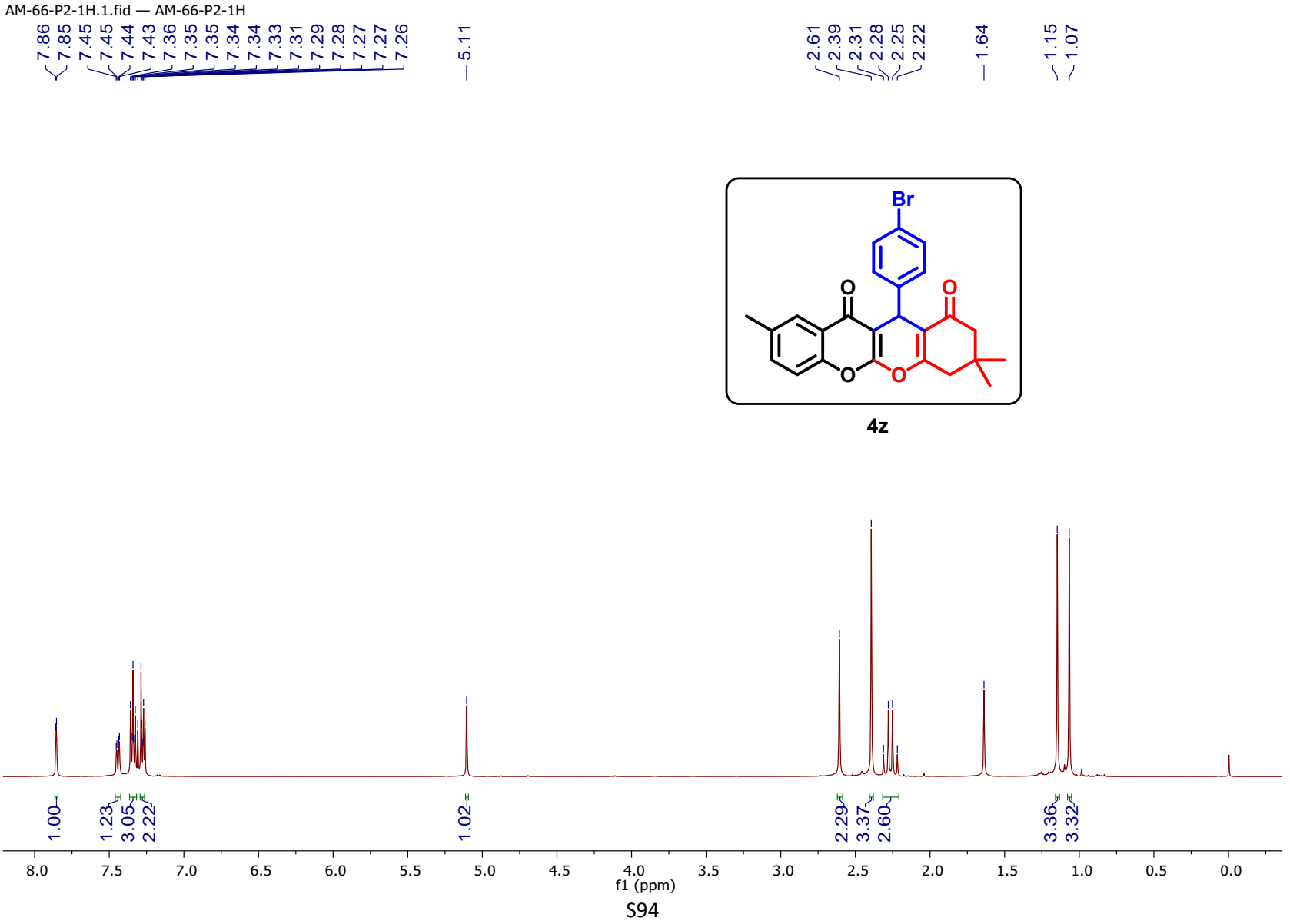


HRMS spectrum of compound 4y

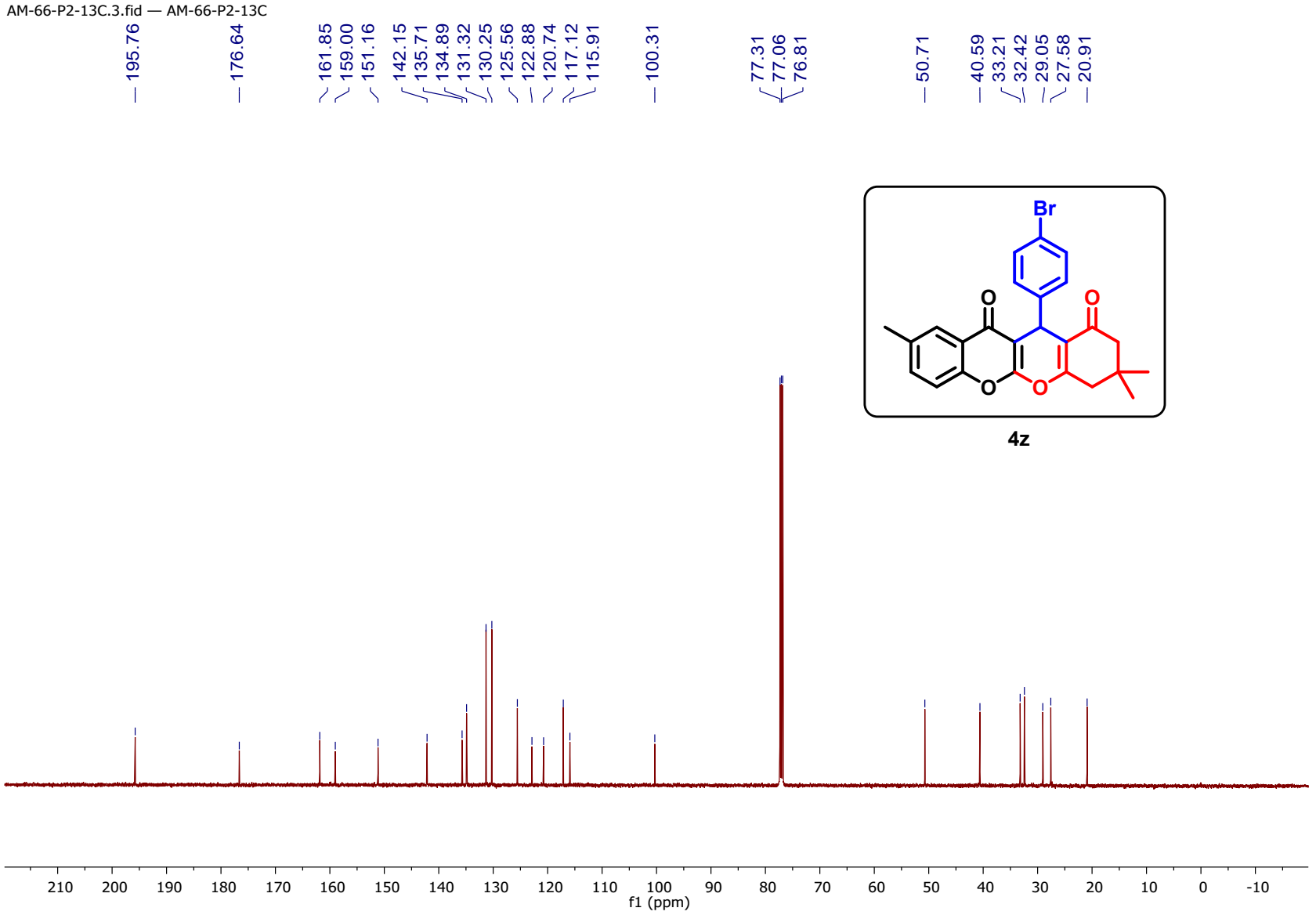


4y

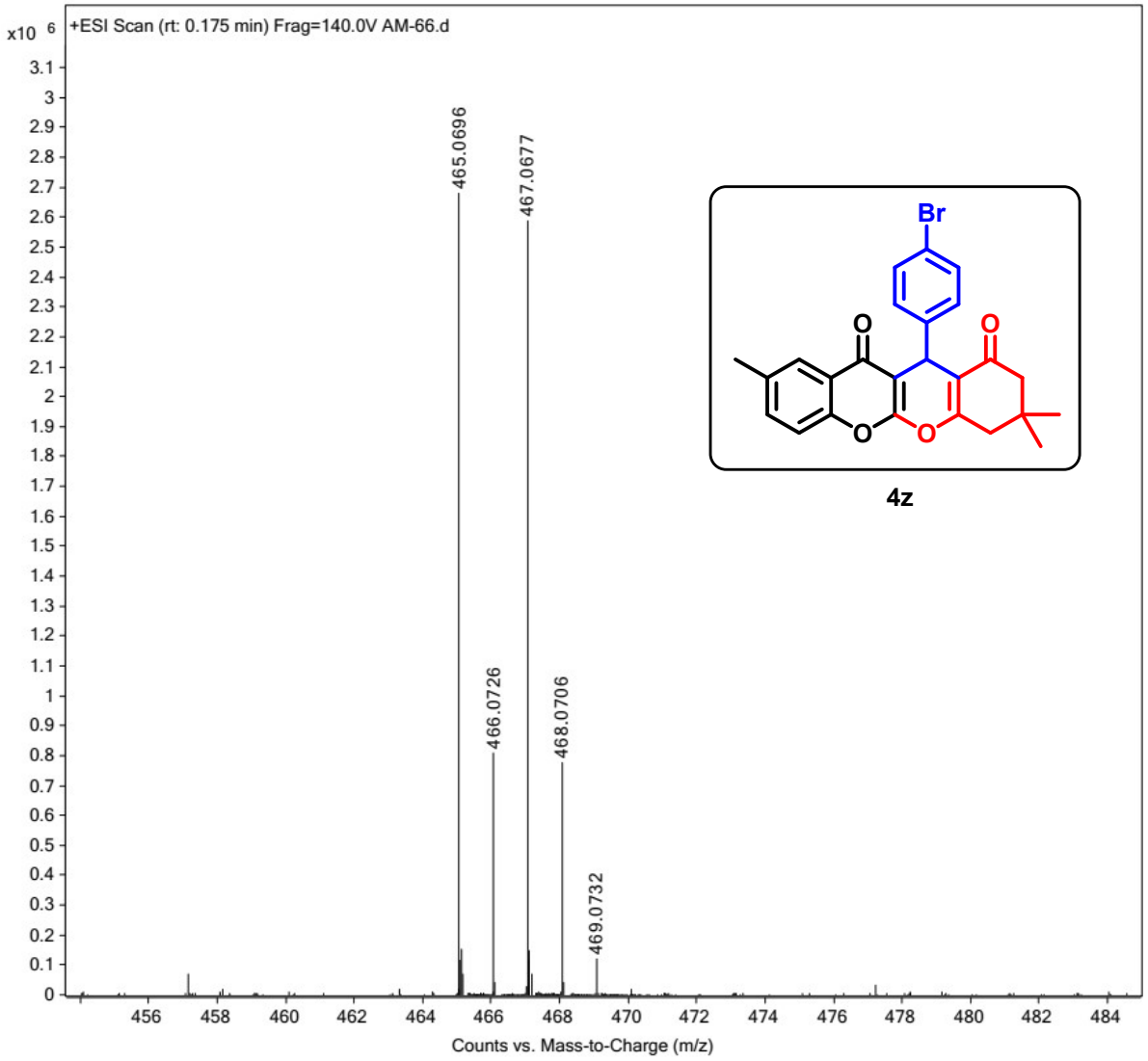
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4z



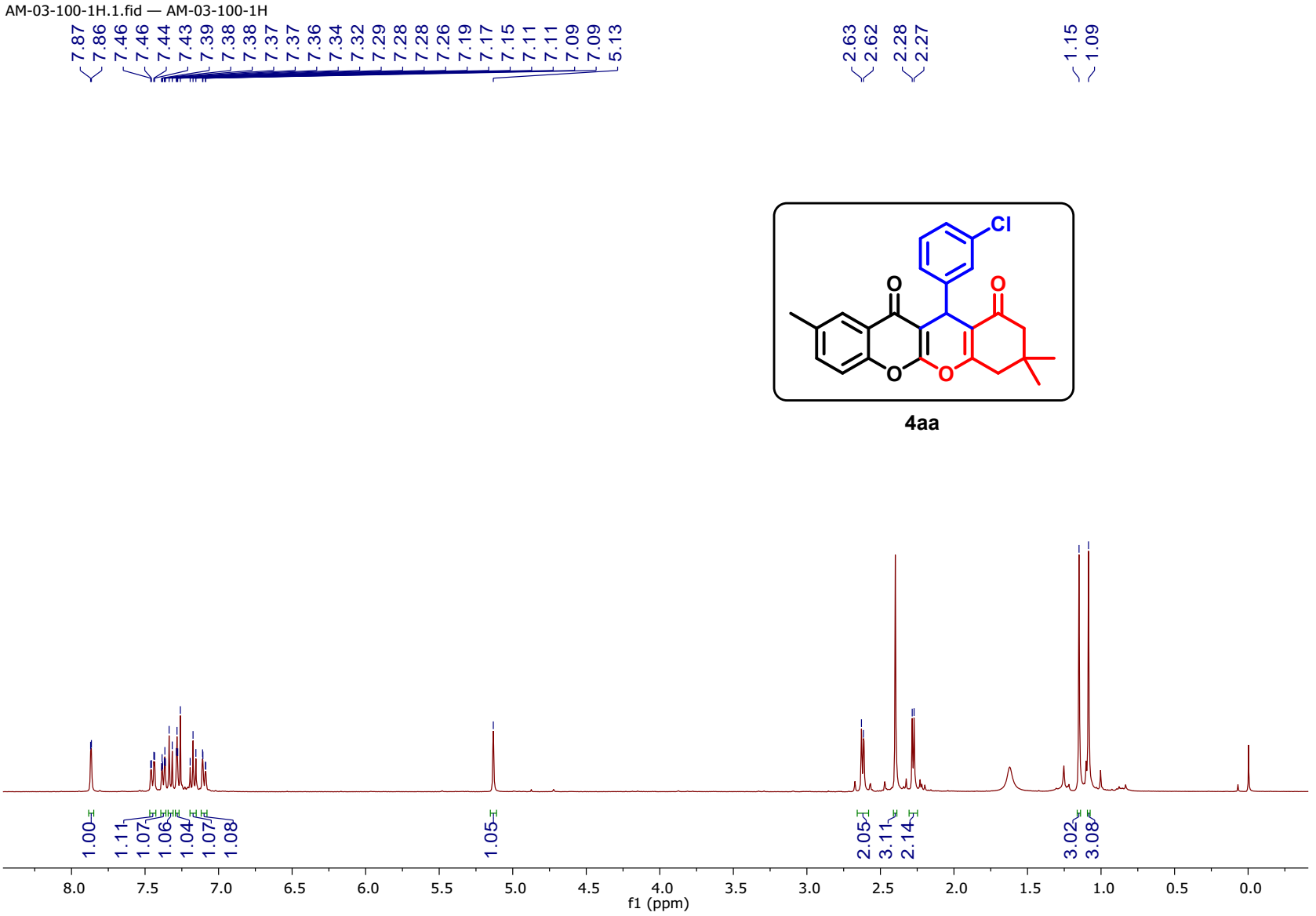
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4z



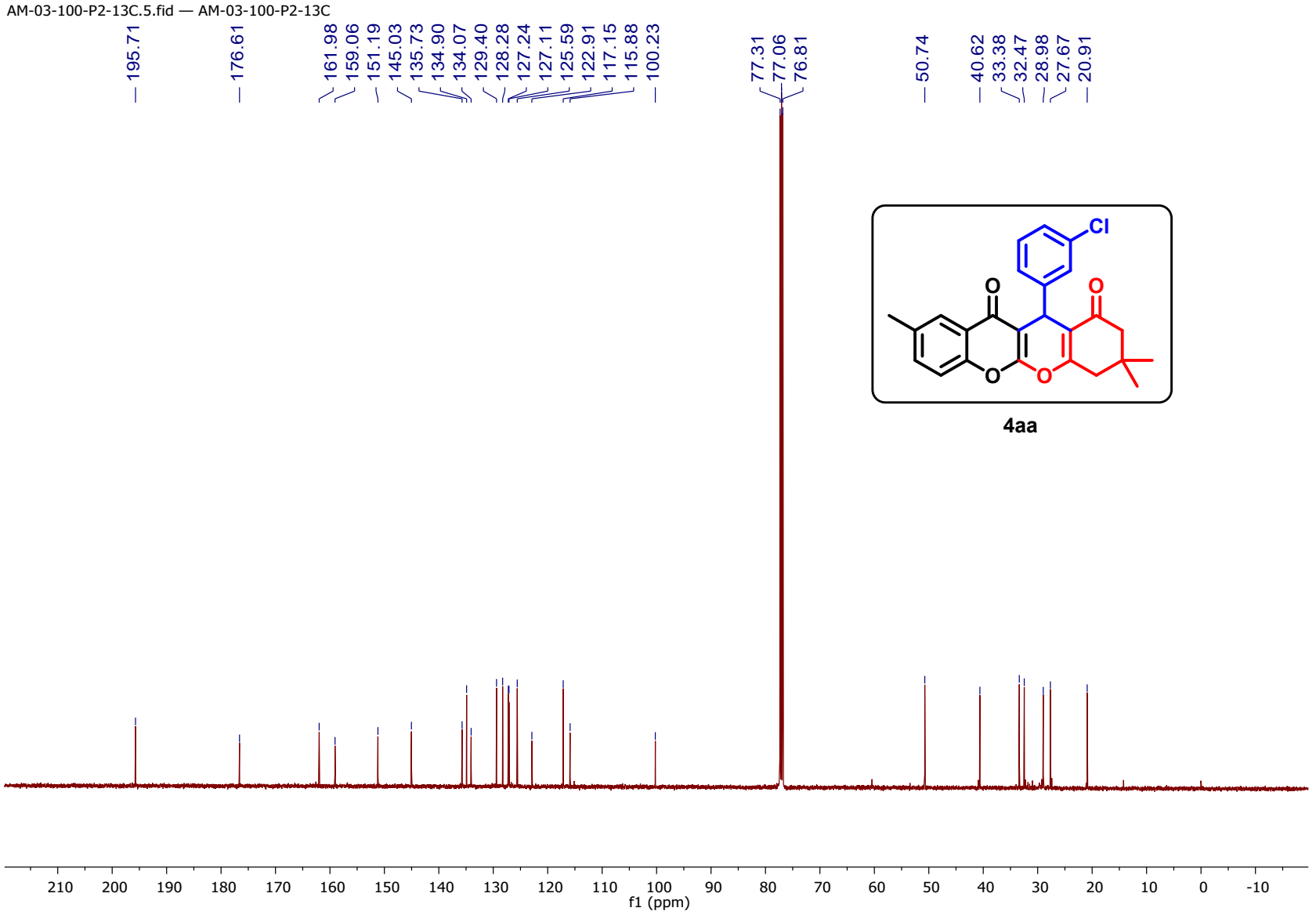
HRMS spectrum of compound 4z



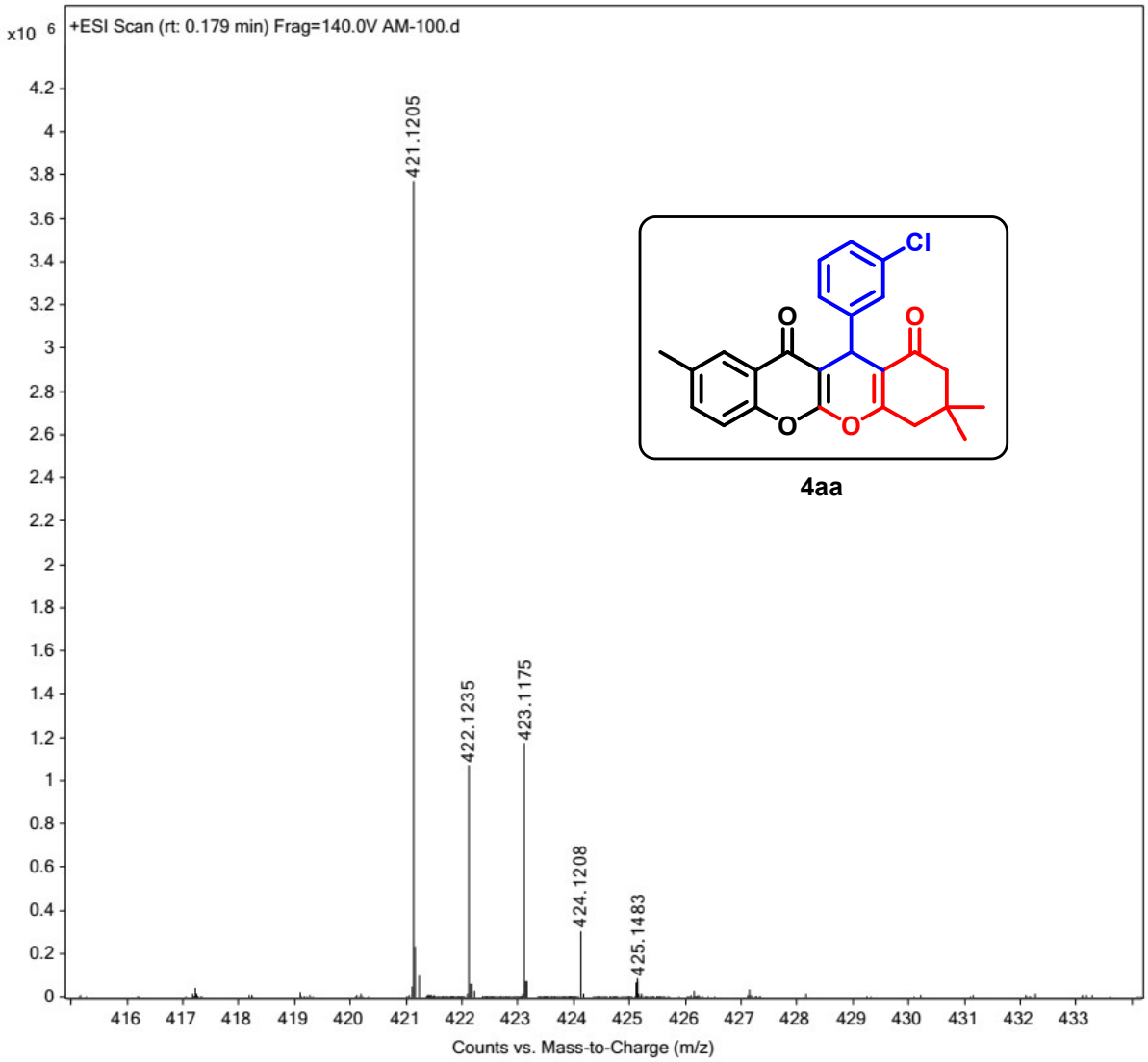
^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4aa



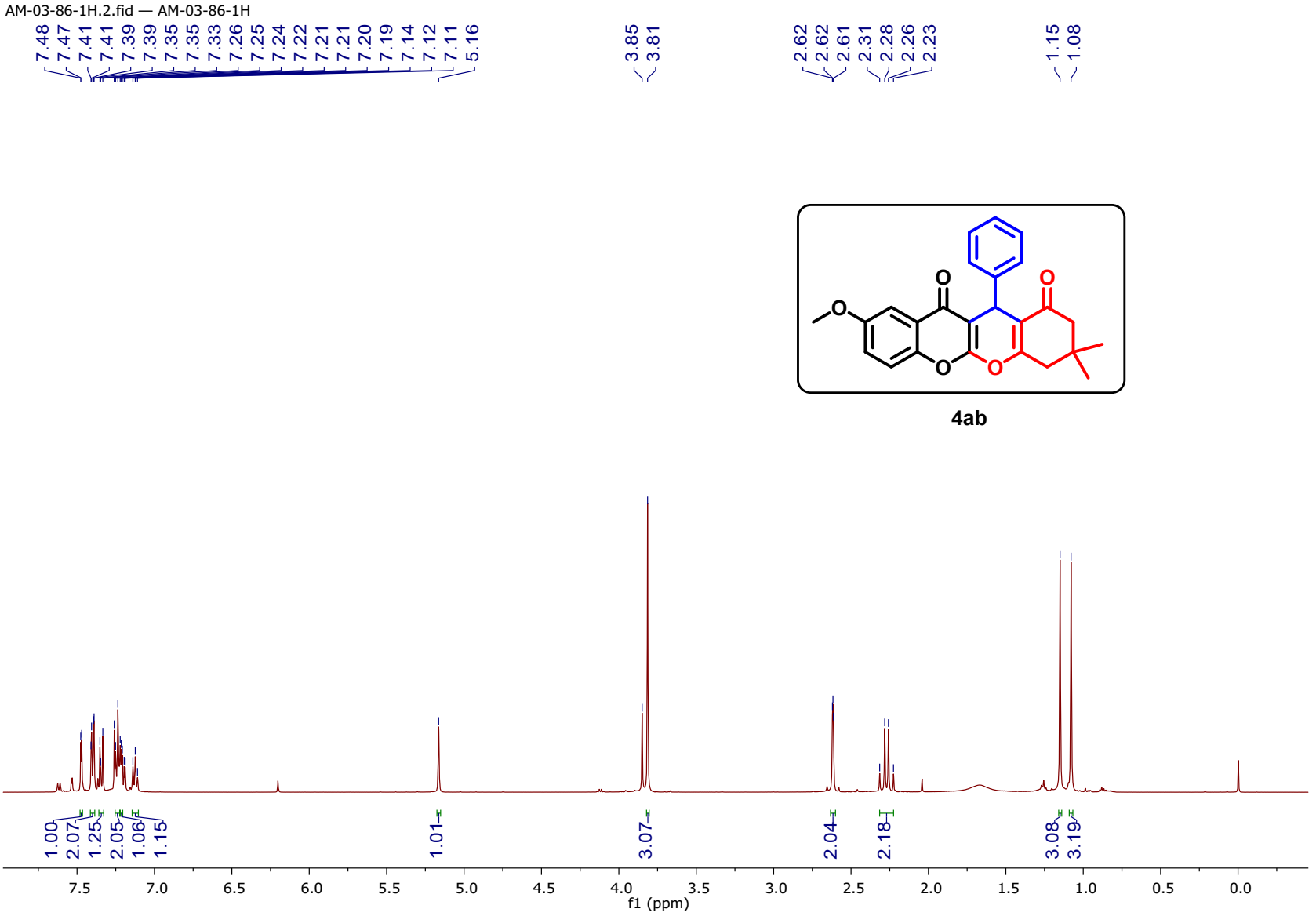
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4aa



HRMS spectrum of compound 4aa

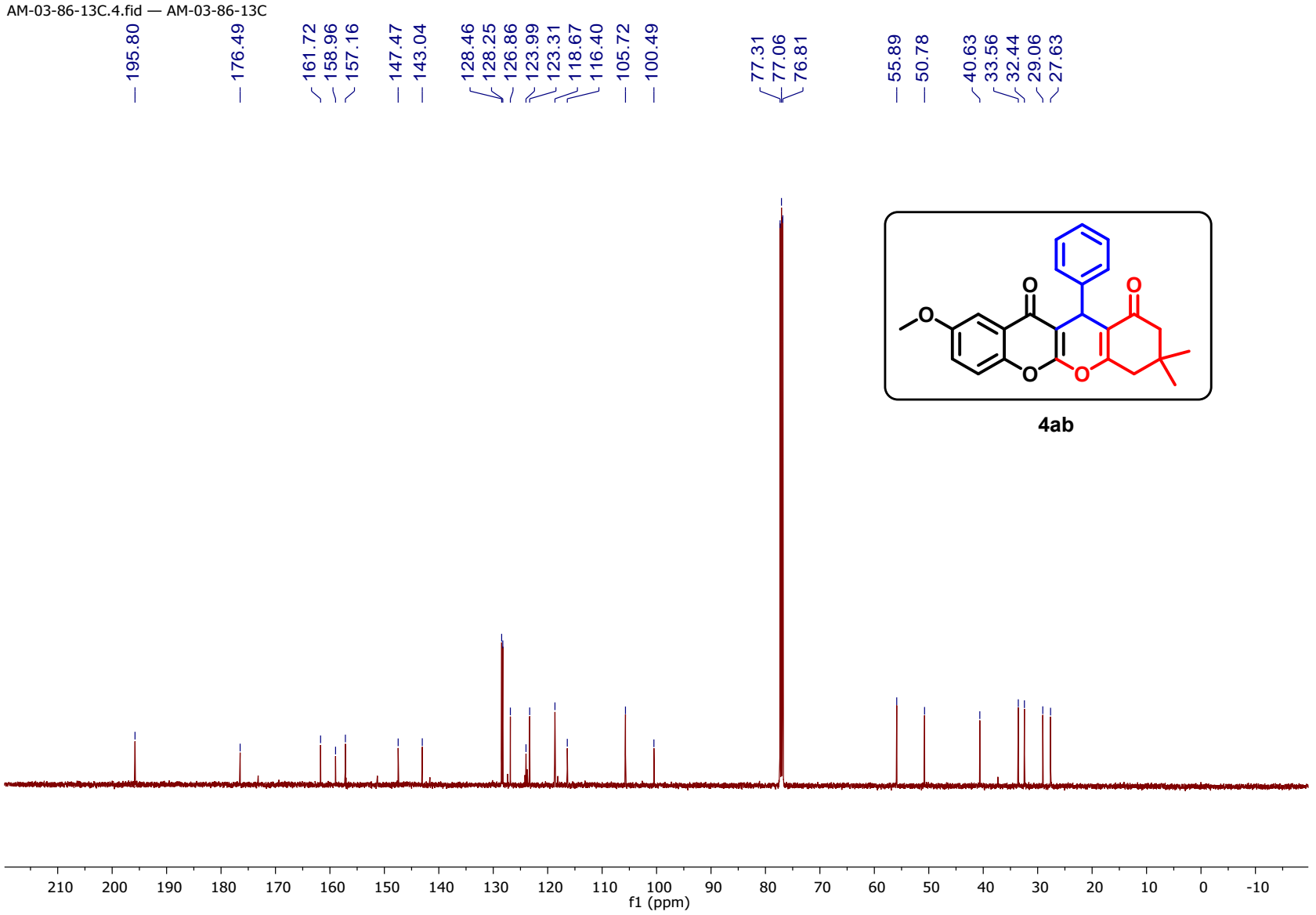


^{1}H NMR (500MHz, $CDCl_3$) spectrum of compound 4ab

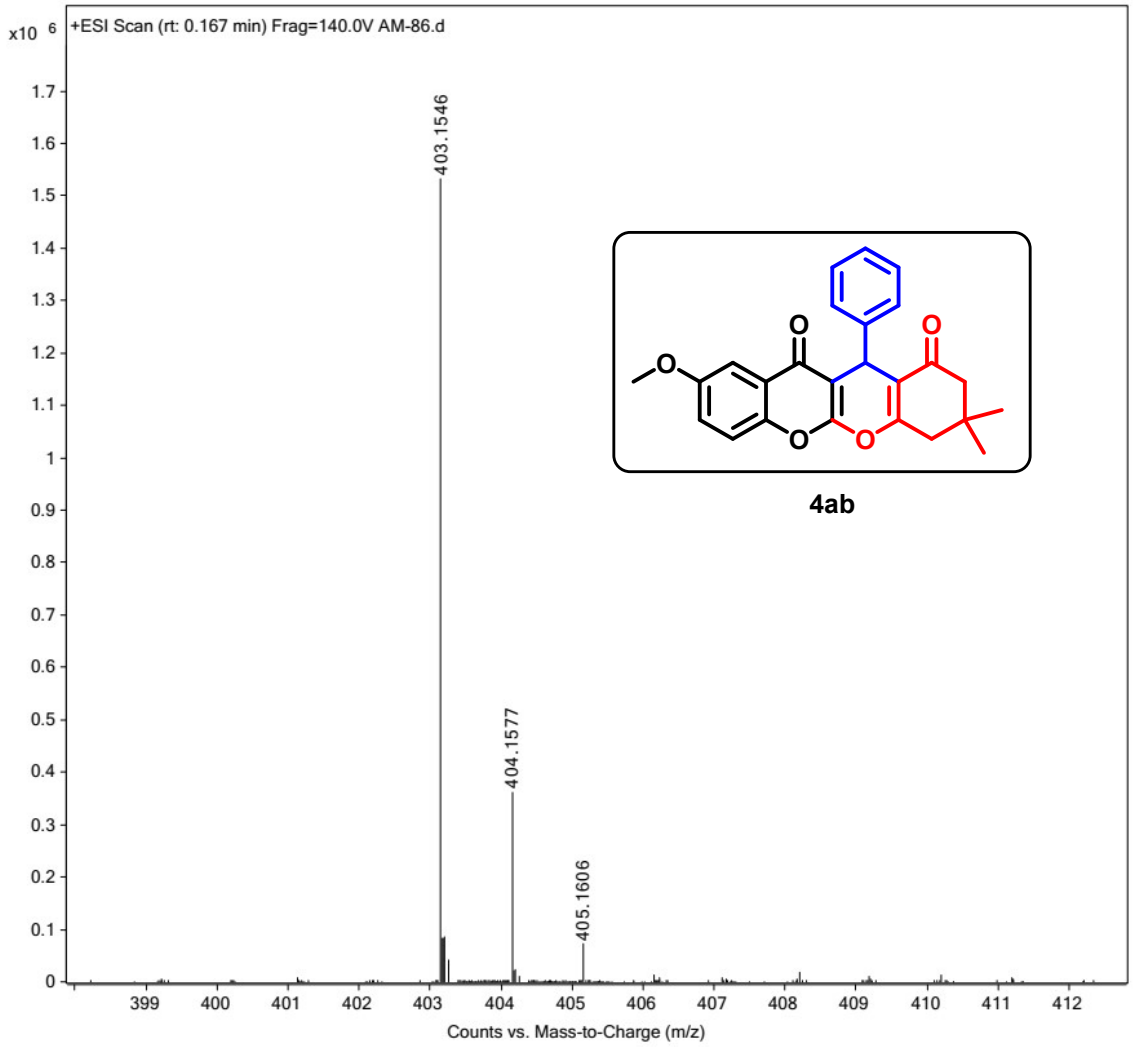


4ab

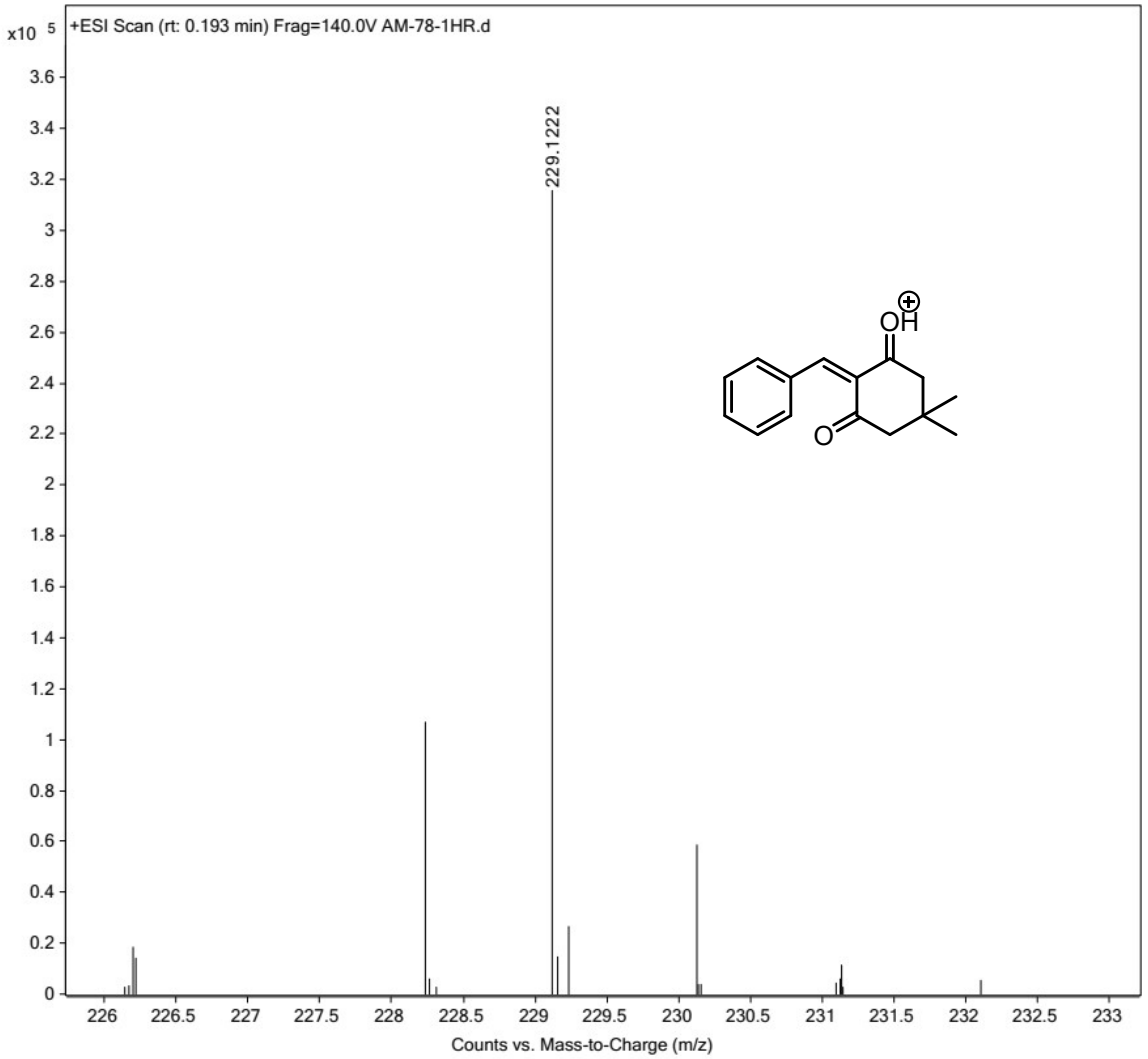
^{13}C {^{1}H} NMR (125 MHz, $CDCl_3$) spectrum of compound 4ab



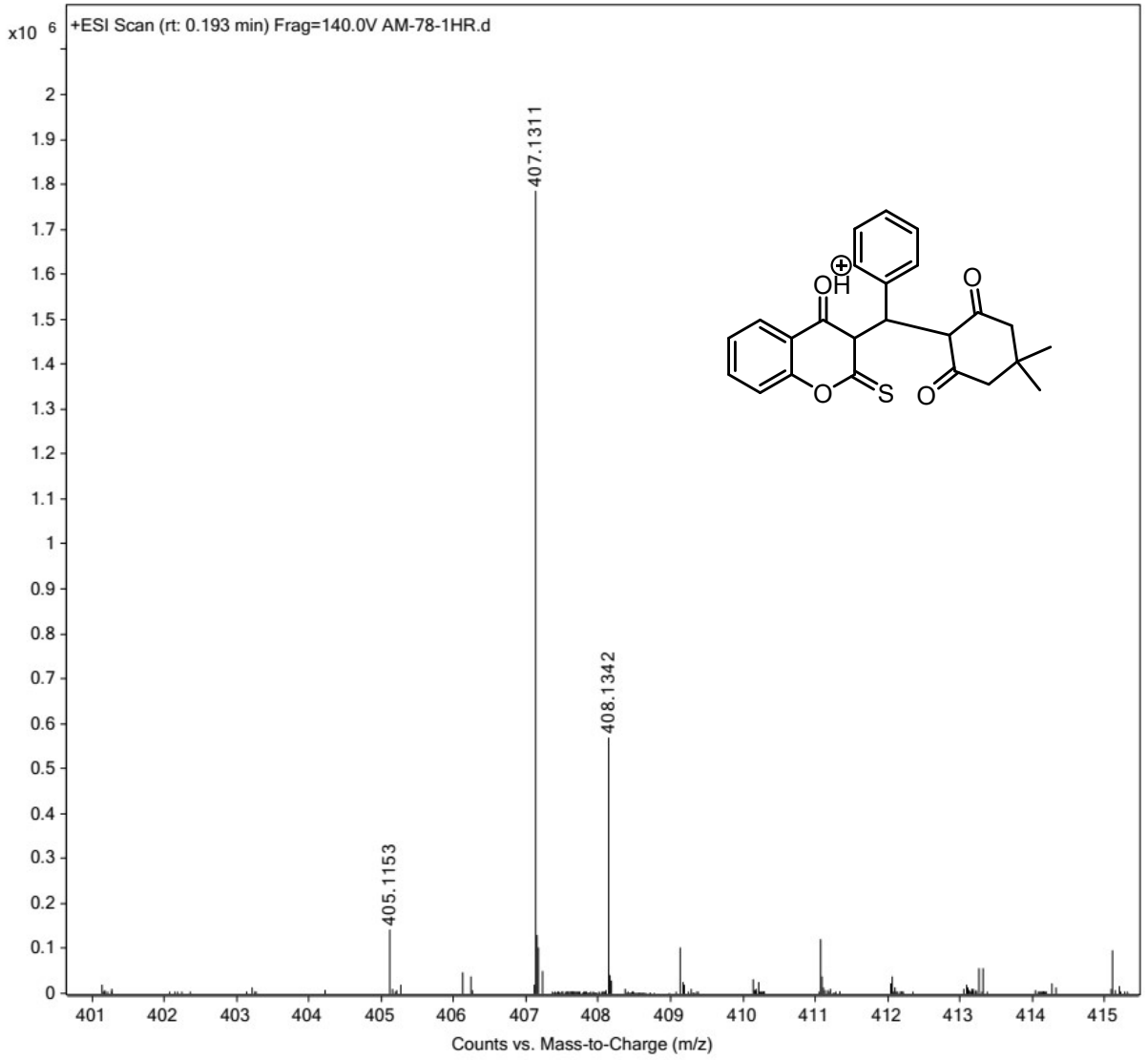
HRMS spectrum of compound 4ab



HRMS spectrum of intermediate B



HRMS spectrum of intermediate D



References

1 D. S. Chandrashekar, S. K. Karthikeyan, P. K. Korla, H. Patel, A. R. Shovon, M. Athar, G. J. Netto, Z. S. Qin, S. Kumar, U. Manne, C. J. Creighton and S. Varambally, *Neoplasia*, 2022, **25**, 18–27.

2 O. Trott and A. J. Olson, *J Comput Chem*, 2010, **31**, 455–461.

3 T. Mosmann, *Journal of Immunological Methods*, 1983, **65**, 55–63.

4 D. Wu and P. Yotnda, *JoVE*, 2011, 3357.