Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Discovery of new pyrimidine-5-carbonitrile derivatives as anticancer agents targeting $EGFR^{WT}$ and $EGFR^{T790M}$

Ahmed A. Nasser[a], Ibrahim H. Eissa[a]*, Mohamed R. Oun[a], Mohamed A. El-Zahabi[a]*, Mohammed S. Taghour[a], Amany Belal[b], Abdulrahman M. Saleh[a], Ahmed B.M. Mehany[c], Hendrik Luesch[d,e], Ahmad E. Mostafa[f], Wael M Afifi[f,g], James R. Rocca[d,h], Hazem A. Mahdy[a]*

[a] Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt.

[b] Medicinal Chemistry Department, Faculty of Pharmacy, Beni-Suef University, Beni-Suef 62514, Egypt.

[c] Zoology Department, Faculty of Science, Al-Azhar University, Cairo 11884, Egypt.

[d] Department of Medicinal Chemistry, University of Florida, Gainesville, FL 32610, USA.

[e] Center for Natural Products, Drug Discovery and Development (CNPD3), University of Florida, Gainesville, FL 32610, USA.

[f] Pharmacognosy and Medicinal Plants Dept., Faculty of Pharmacy (Boys), Al-Azhar University, Cairo, 11884, Egypt.

[g] Department of Pharmacognosy, Faculty of Pharmacy, Sinai University, Ismailia, Egypt.

[h] AMRIS Facility, McKnight Brain Institute, University of Florida, Gainesville, FL 32610, USA.

*** Corresponding authors:**

Ibrahim H. Eissa

Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt

Email: Ibrahimeissa@azhar.edu.eg

Mohamed A. El-Zahabi

Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt

Email: malzahaby@yahoo.com

Hazem A. Mahdy

Pharmaceutical Medicinal Chemistry & Drug Design Department, Faculty of Pharmacy (Boys), Al-Azhar University, Cairo 11884, Egypt

Email: Hazem_hady2001@azhar.edu.eg

Content

1	**Reported reactions between nitrile group and cysteine amino acid**
2	**Docking studies**
3	**ADMET studies**
4	**Spectral analysis**

1- Reported reactions between nitrile group and cysteine amino acid

a- **Fig. 1** was reported by Montanari *et al* [1] showing two possible mechanisms for the inhibition reaction. In the first proposed mechanism, the protonation of N1 occurs concertedly with the nucleophilic attack of the thiolate anion. The second proposed mechanism is the nucleophilic addition and proton transfer in a stepwise mechanism, where the thiolate group of Cys25 attacks the carbon atom of the nitrile group first to form a covalent bond.

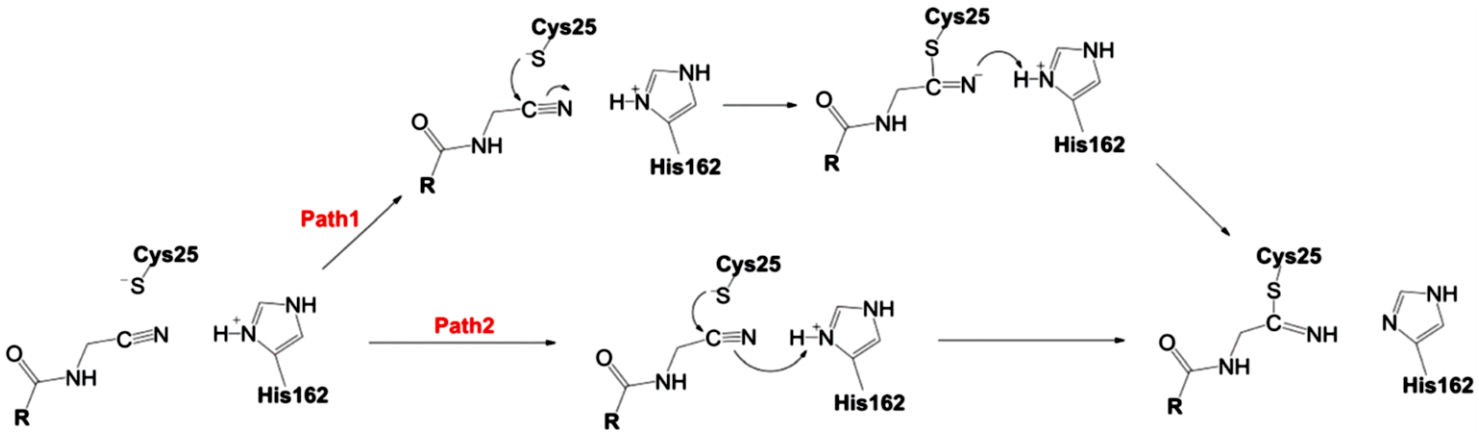


Fig. 1: Possible reaction mechanisms of nitrile derivatives with cysteine amino acid.

b- **Fig. 2** was reported by Liang *et al* [2] showing the condensation reaction between cysteine and 2-cyanobenzothiazole that occurs *in vitro* and in living cells under the control of pH.

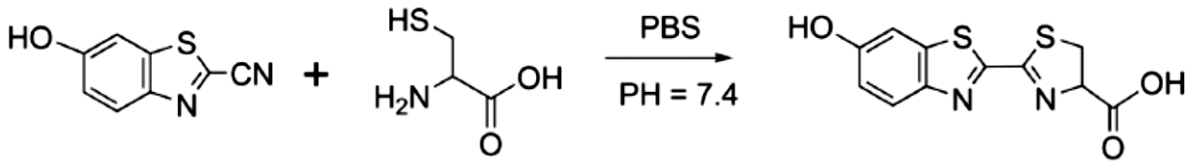


Fig. 2: the condensation reaction between cysteine and 2-cyanobenzothiazole pH = 7.4.

c- **Fig. 3** was reported by Zheng *et al* [3] showing the mechanism of condensation reaction between L-cysteine and the cyano group of 2-cyano-6-aminobenzothiazole to yield Aminoluciferin

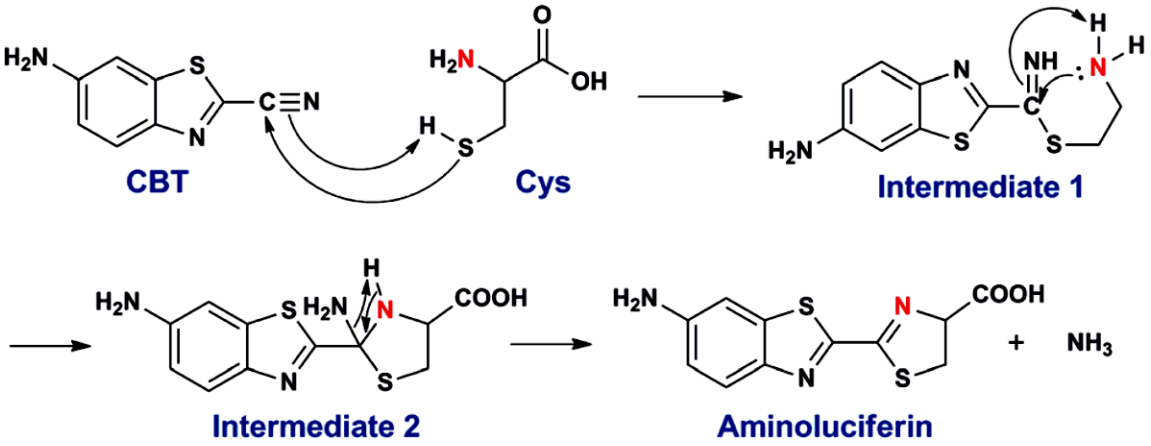


Fig. 3: Proposed reaction mechanism between 2-cyano-6-aminobenzothiazole and L-cysteine.

2. In silico studies

A) Docking studies

- **The binding mode of compound 16_b against $EGFR^{WT}$**

The binding mode of compound **16_b** was similar to that of the reference ligand, erlotinib. It gave affinity value of -9.55 kcal/mol. The pyrimidine-5-carbonitrile moiety was embedded in the adenine pocket of the $EGFR^{WT}$ forming a hydrophobic interaction with Leu694. The sulfamethazine moiety occupied the hydrophobic pocket I forming a hydrophobic bond with Met742, Leu834, Leu837, and Cys751. Also, it formed two hydrogen bonds with Thr830 and Thr766 with a distance of 2.10 and 1.92 A˚, respectively. Moreover, the phenyl group at position-4 occupied the hydrophobic region II. (**Fig 1**).

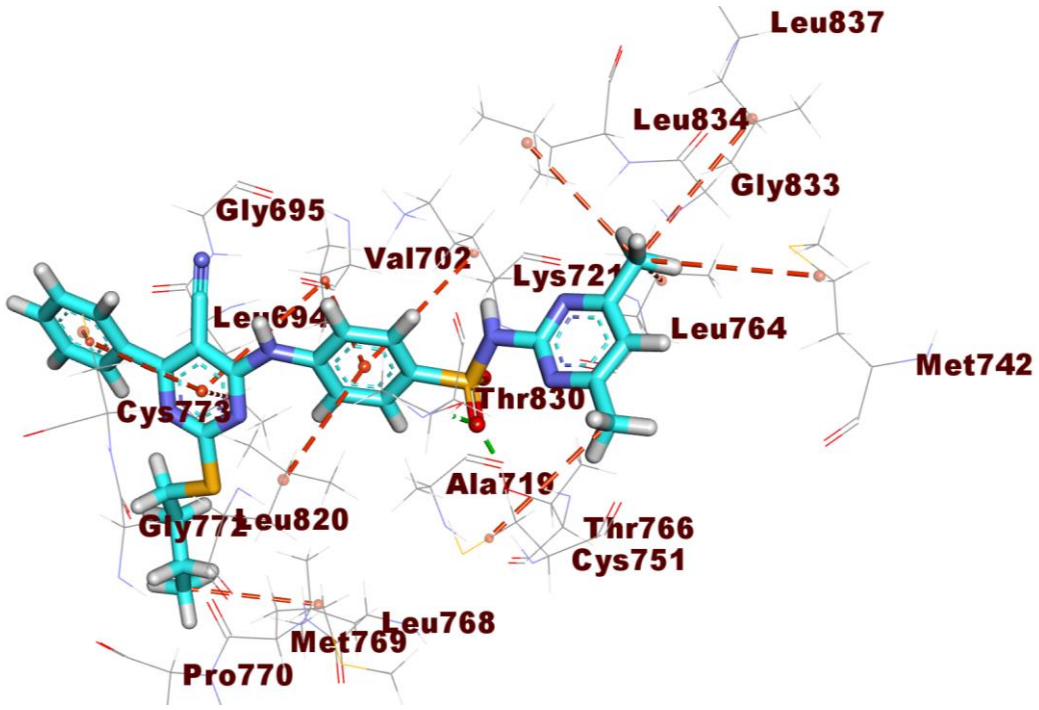


Fig. 1: Compound **16b** docked into the active site of $EGFR^{WT}$ formed two hydrogen bonds (green) with Thr830 and Thr766 residues. The pi interactions are represented in orange lines.

- **The binding mode of compound 16b against $EGFR^{T790M}$**

The binding mode of compound **16b** was similar to that of the reference ligand, TAK-285. It gave affinity value of -6.74 kcal/mol. The pyrimidine-5-carbonitrile moiety was embedded in the adenine pocket. The sulfur atom of thiobutyl moiety formed a hydrogen bond with Thr854. The sulfamethazine moiety occupied the hydrophobic pocket I forming a hydrophobic bond with Lys728, Leu718, and Leu844. Moreover, the phenyl group at position-4 occupied the hydrophobic region II. (**Fig. 2**).

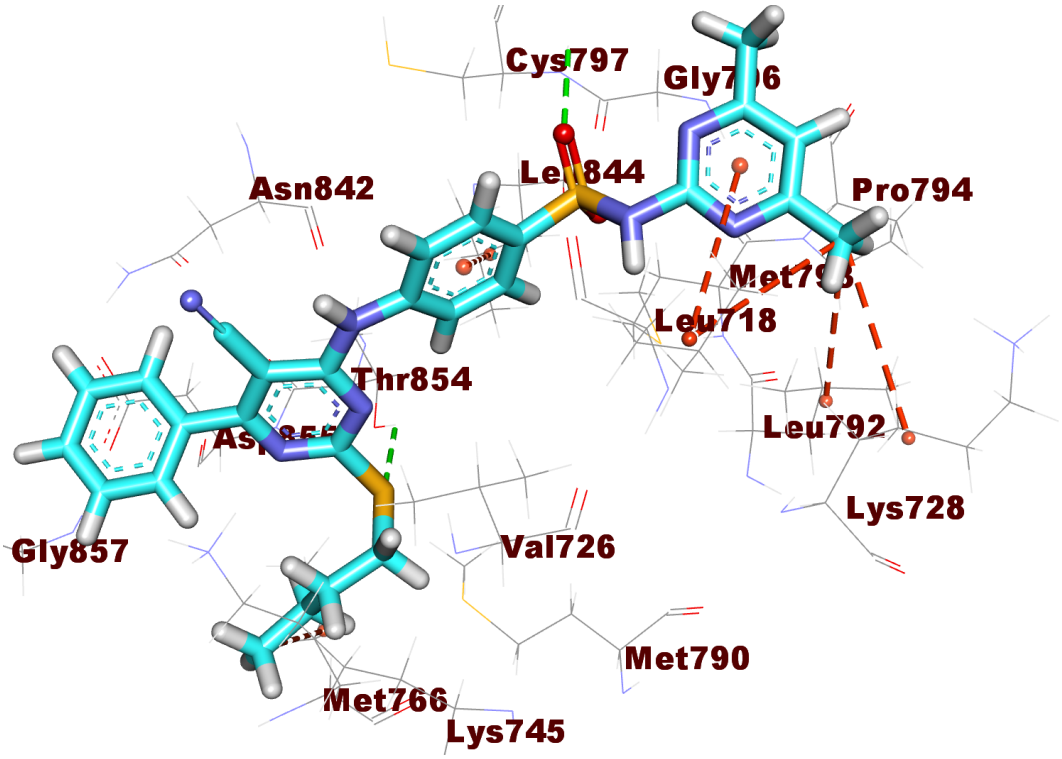


Fig. 2: Binding of compound **16b** with $EGFR^{T790M}$, the hydrogen bonds are represented in green dashed lines and the pi interactions are represented in orange dashed lines.

- **The binding mode of compound 10_a against $EGFR^{WT}$**

The binding mode of compound **10_a** exhibited binding affinity value of -6.93 kcal/mol. The pyrimidine-5-carbonitrile moiety was embedded in the adenine pocket of the $EGFR^{WT}$ forming one hydrogen bonding interaction with Lys721. In addition, it formed two hydrophobic interactions withVal702 and Leu820. The phenyl moiety occupied the hydrophobic pocket I forming a hydrophobic bond with Lys721. Moreover, the phenyl group at position-4 occupied the hydrophobic region II forming three hydrophobic interactions with Ala719, Leu694, and Leu820 (**Fig 1**).

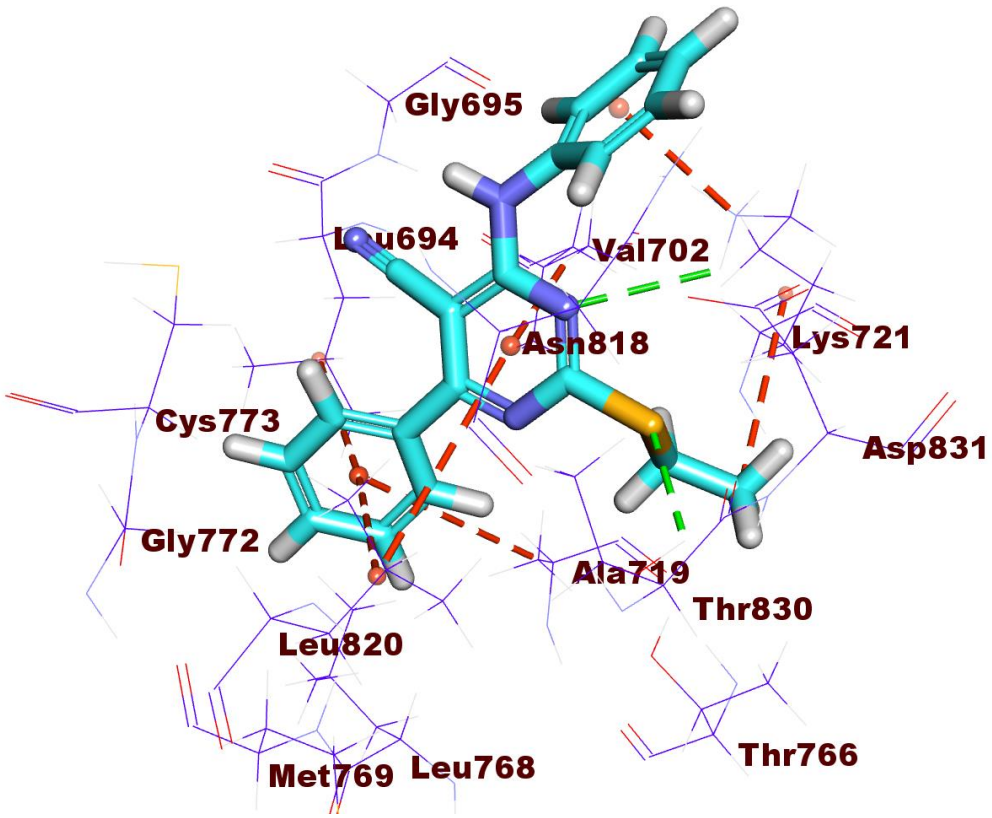


Fig. 3: Compound **10_a** docked into the active site of $EGFR^{WT}$. The hydrogen bonds were represented in green dashed lines and the pi interactions are represented in orange dashed lines.

B) **ADMET Descriptors (*These data were retrieved from Discovery Studio 4.0 Help*)**

Table 1. Predicted ADMET for the designed compounds and reference drugs

Comp.	BBB level	Solubility level	ALog p98	PSA2D	CYP2D6 prediction	PPB prediction
10c	0.736	-6.692	5.864	58.267	-2.615	4.656
10d	< 0.3	-7.346	6.844	58.267	-0.954	5.360
10f	< 0.3	-7.336	6.844	58.267	-0.893	7.828
11a	0.219	-5.752	4.766	69.528	-3.030	2.909
11b	< 0.3	-6.416	5.746	69.528	-1.428	3.614
12b	< 0.3	-6.968	6.474	58.267	-1.826	1.263
13a	0.481	-5.891	5.038	58.267	-3.460	0.959
13b	< 0.3	-6.566	6.018	58.267	-2.033	0.902
15b	< 0.3	-6.695	6.097	75.568	-2.086	6.087
16a	< 0.3	-6.525	5.173	128.201	-14.154	-6.464
17a	0.537	-5.954	5.22	58.267	-1.009	2.812
17b	< 0.3	-6.605	6.199	58.267	0.662	3.327
Erlotinib	0.054	-5.11	4.309	71.052	-2.881	0.620
Gefitinib	-0.371	-4.485	2.642	65.326	-0.801	-0.126

a- Key to human intestinal absorption (ADMET Absorption Level)

Level	Value	Description
0	ADMET_Absorption_T2_2D < 6.1261 (inside 95%)	Good absorption
1	6.1261 ≤ ADMET_Absorption_T2_2D < 9.6026 (inside 99%)	Moderate absorption
2	9.6026 < ADMET_Absorption_T2_2D (outside 99%)	Low absorption
3	ADMET_PSA_2D ≥ 150.0 or ADMET_AlogP98 ≤ -2.0 or ADMET_AlogP98 ≥ 7.0	Very low absorption

- **Note.** ADMET_Absorption_T2_2D is the Mahalanobis distance for the compound in the ADMET_PSA_2D, ADMET_AlogP98 plane. It is referenced from the center of the region of chemical space defined by well absorbed compounds

b- Key to aqueous solubility levels (ADMET Solubility Level)

Level	Value	Drug-likeness
0	log(Sw) < -8.0	Extremely low
1	-8.0 < log(Sw) < -6.0	No, very low, but possible
2	-6.0 < log(Sw) < -4.0	Yes, low
3	-4.0 < log(Sw) < -2.0	Yes, good
4	-2.0 < log(Sw) 0.0=""	Yes, optimal
5	0.0 < log(Sw)	No, too soluble
6	-1000	Warning: molecules with one or more unknown AlogP98 types

c- Key to BBB penetration levels (ADMET BBB)

Level	Value	Description
0	Very High	Brain-Blood ratio greater than 5:1
1	High	Brain-Blood ratio between 1:1 and 5:1
2	Medium	Brain-Blood ratio between 0.3:1 and 1:1
3	Low	Brain-Blood ratio less than 0.3:1
4	Undefined	Outside 99% confidence ellipse
5	AlogP98	Warning: molecules with one or more unknown AlogP98 types

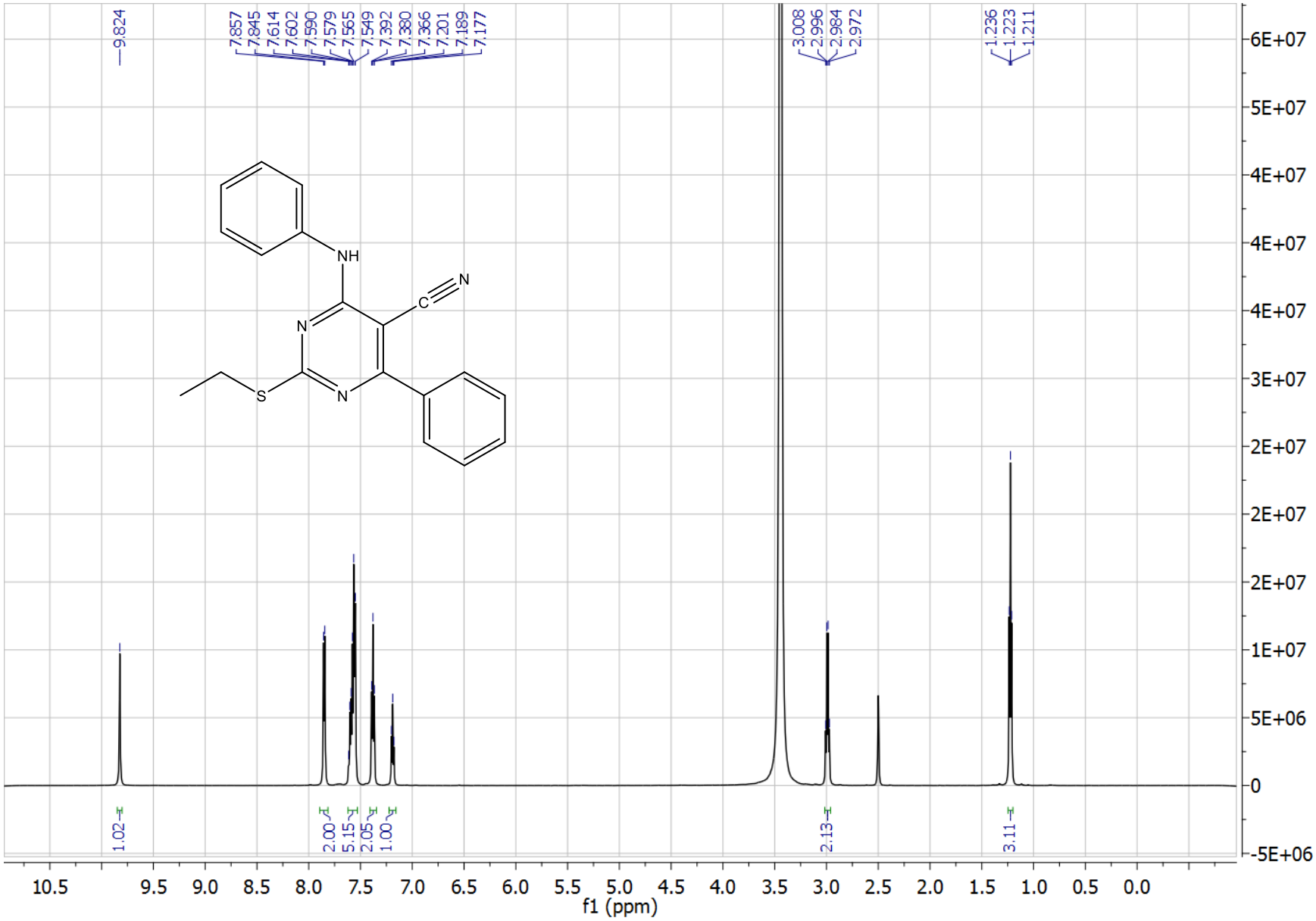
d- ADMET - Cytochrome P450 2D6 (CYP2D6)

ADMET CYP2D6 Prediction	The classification whether a compound is an CYP2D6 inhibitor using the cutoff Bayesian score of 0.161 (obtained by minimizing the total number of false positives and false negatives)

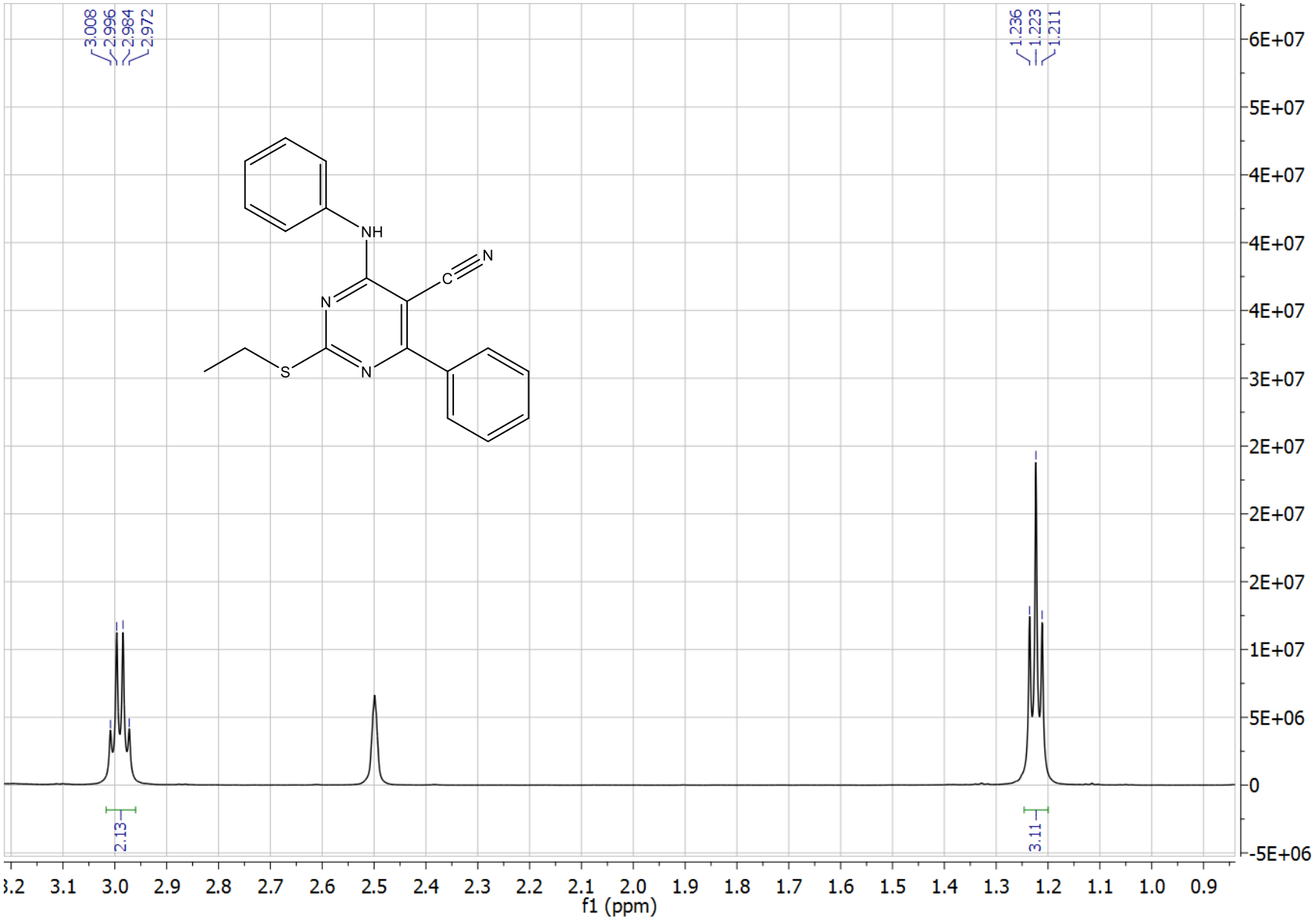
e- ADMET - Plasma Protein Binding

ADMET PPB Prediction	The classification whether a compound is highly bounded (>= 90% bound) to plasma proteins using the cutoff Bayesian score of -2.209 (obtained by minimizing the total number of false positives and false negatives)

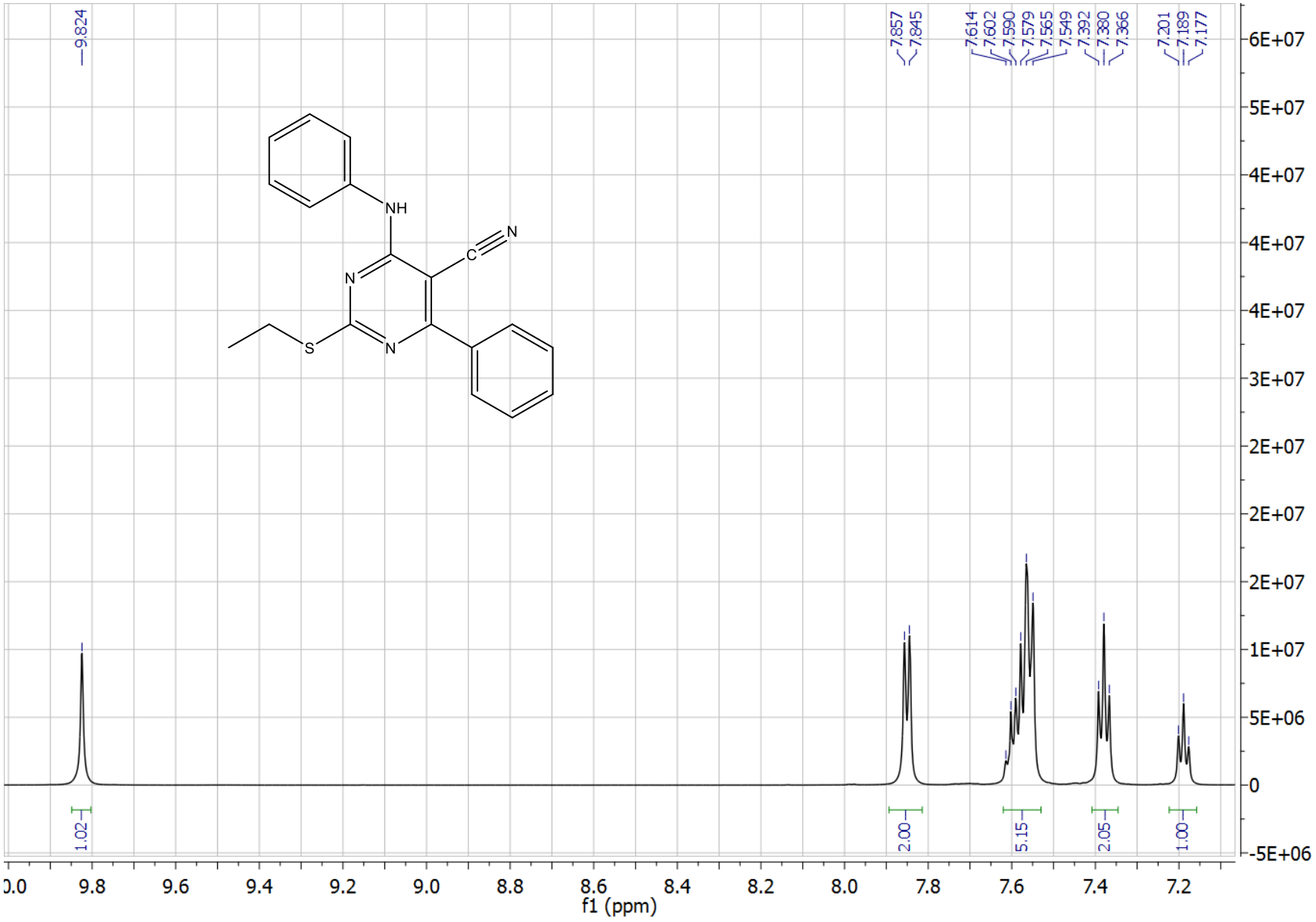
^{1}H NMR of compound 10a



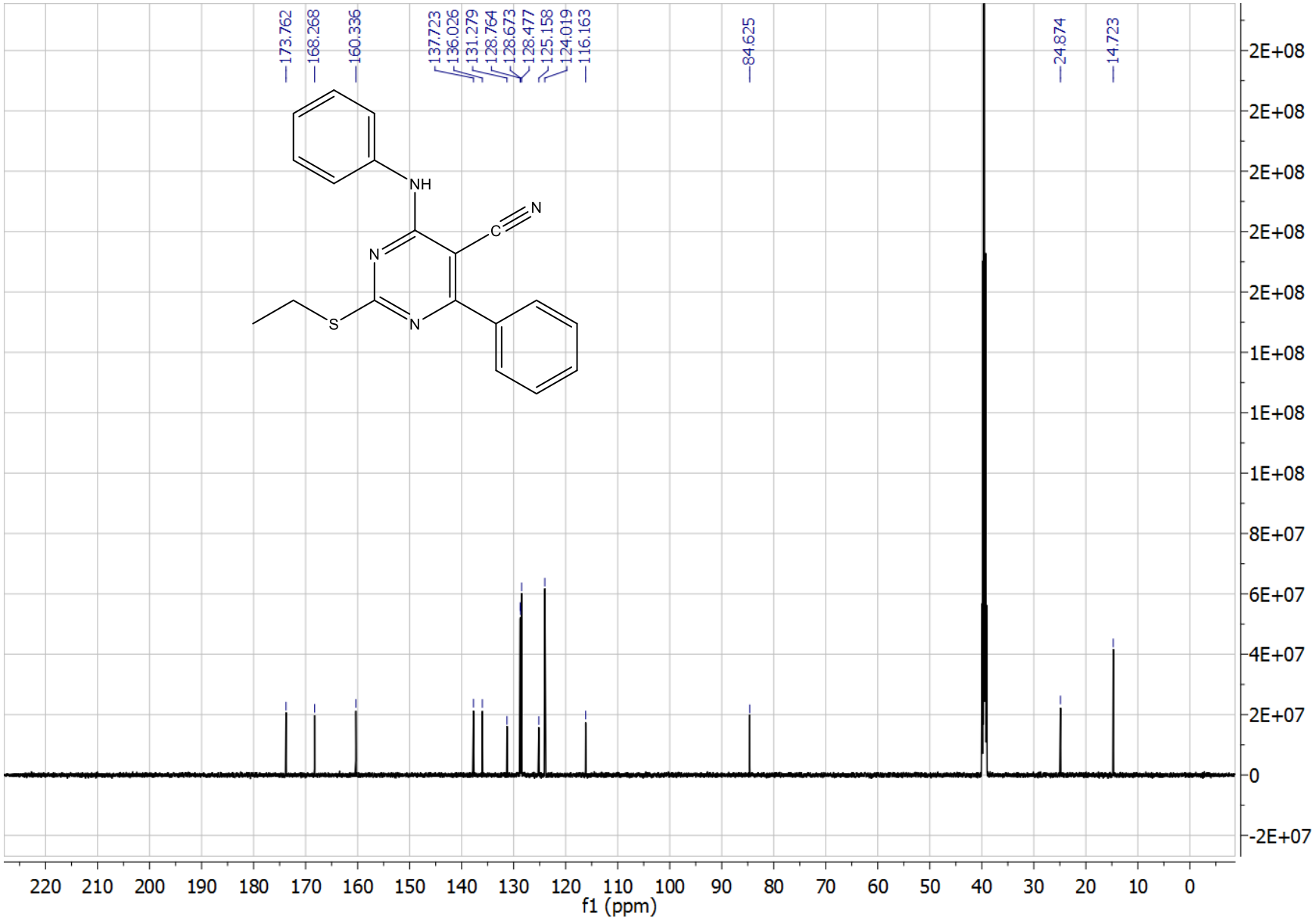
^{1}H NMR of compound 10a



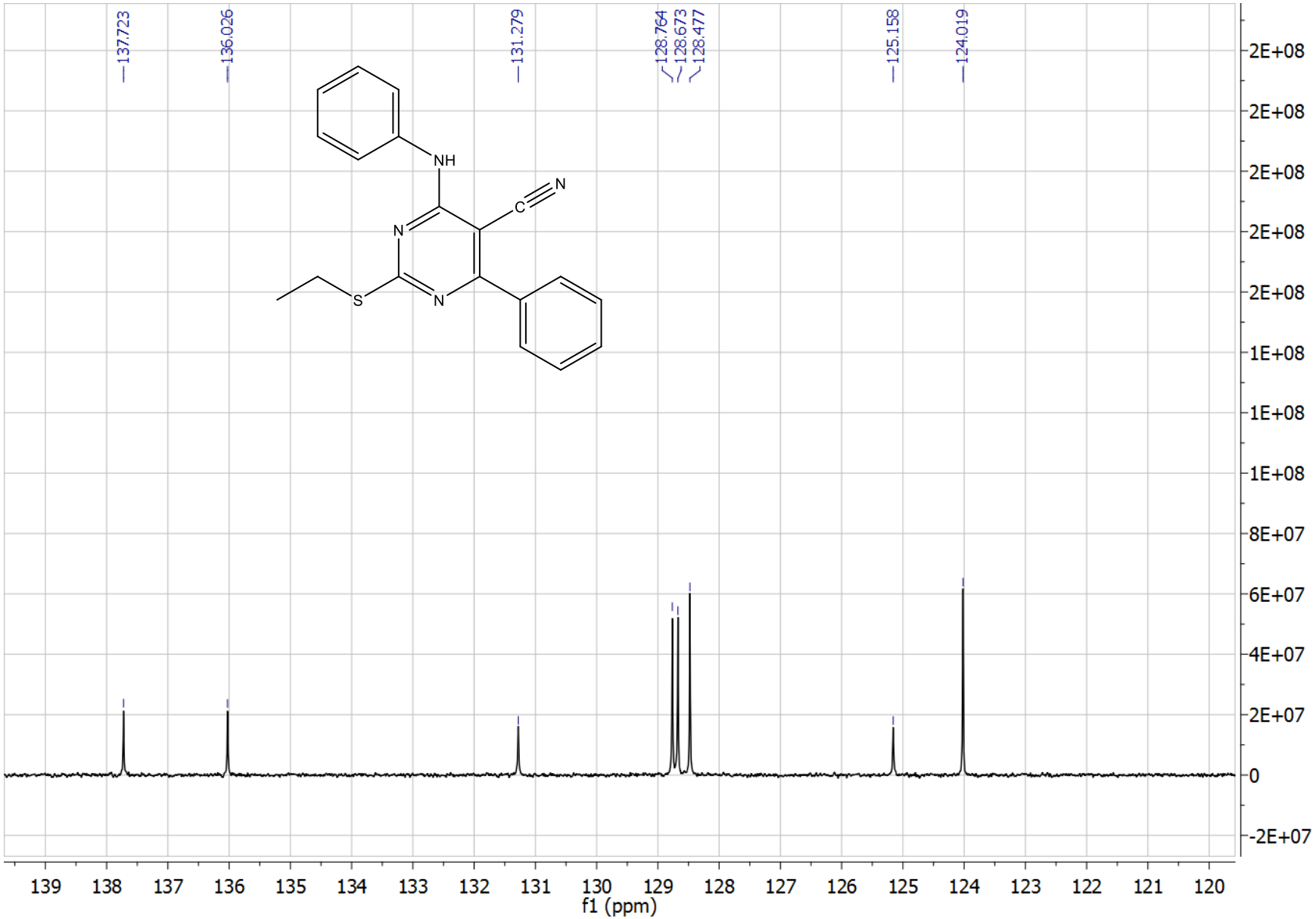
^{1}H NMR of compound 10a



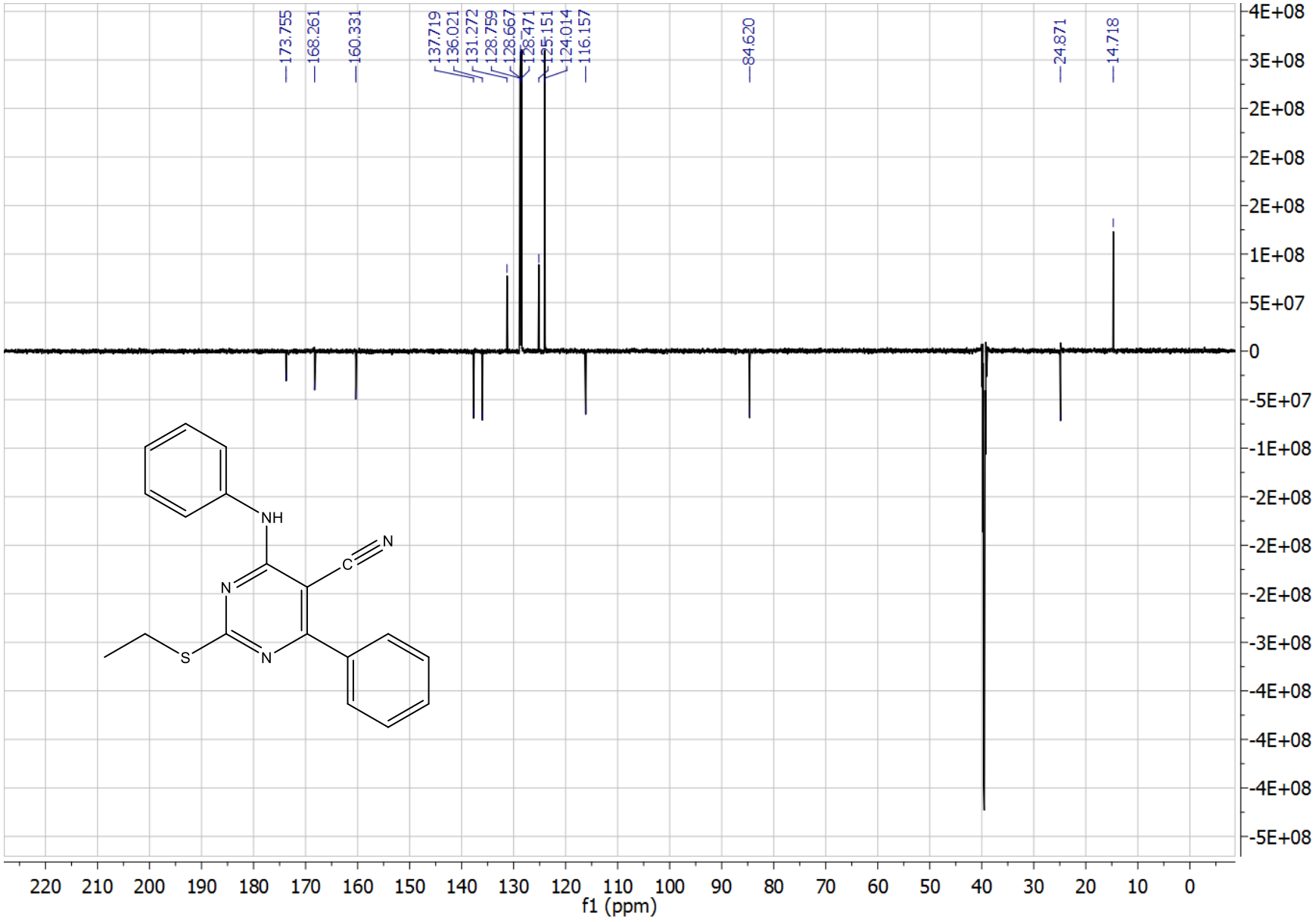
^{13}C NMR of compound 10a



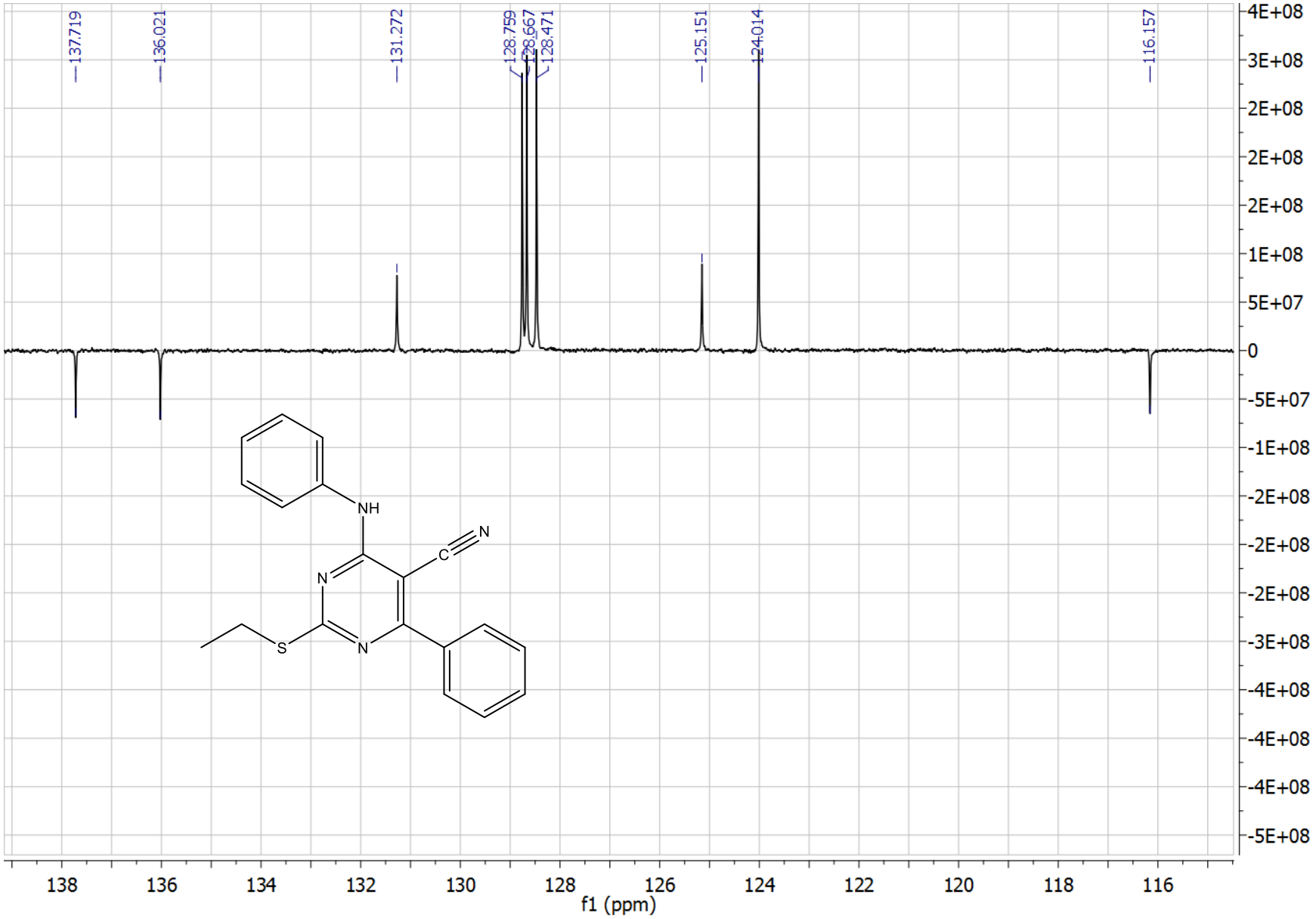
^{13}C NMR of compound 10a



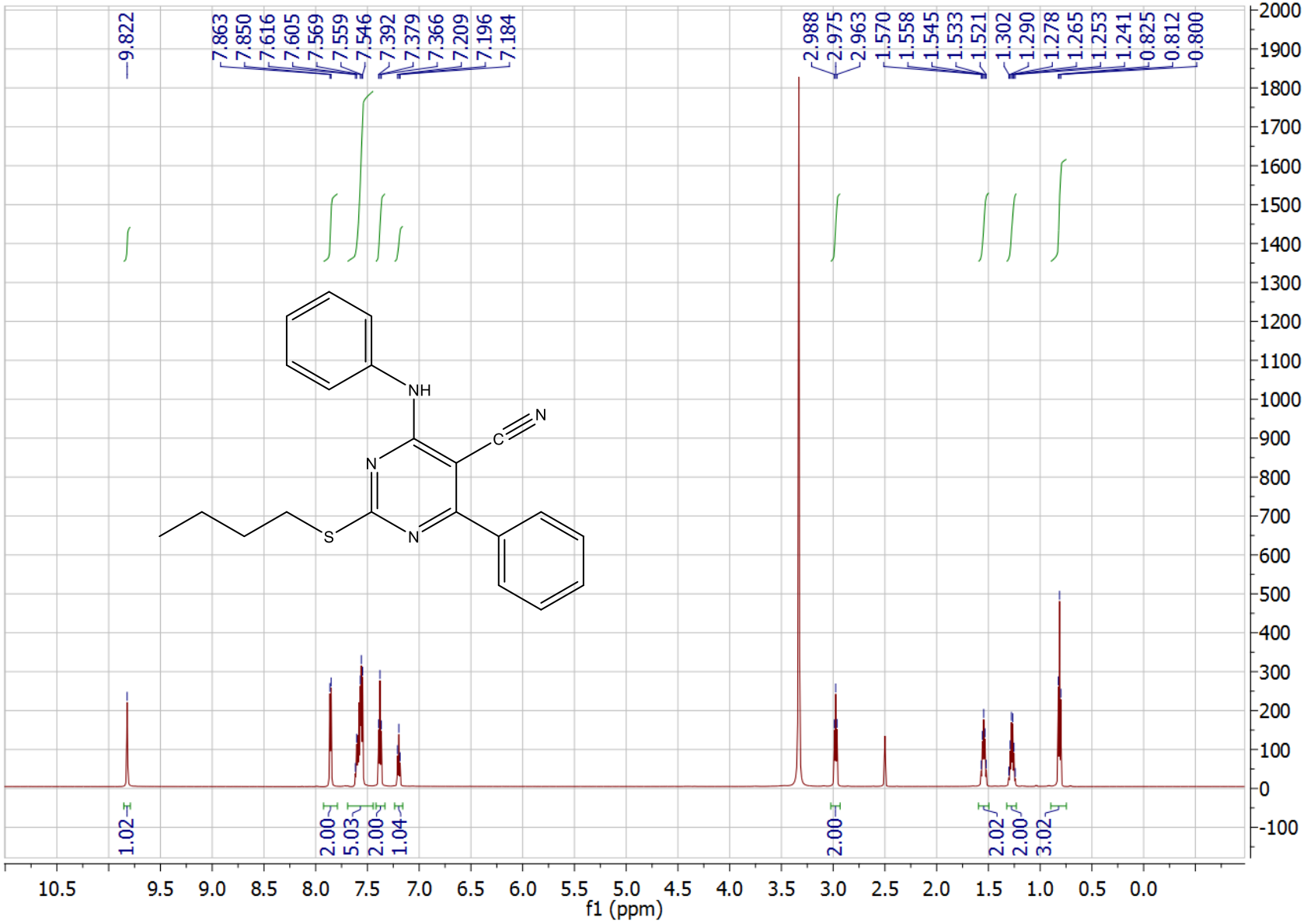
APT of compound 10a



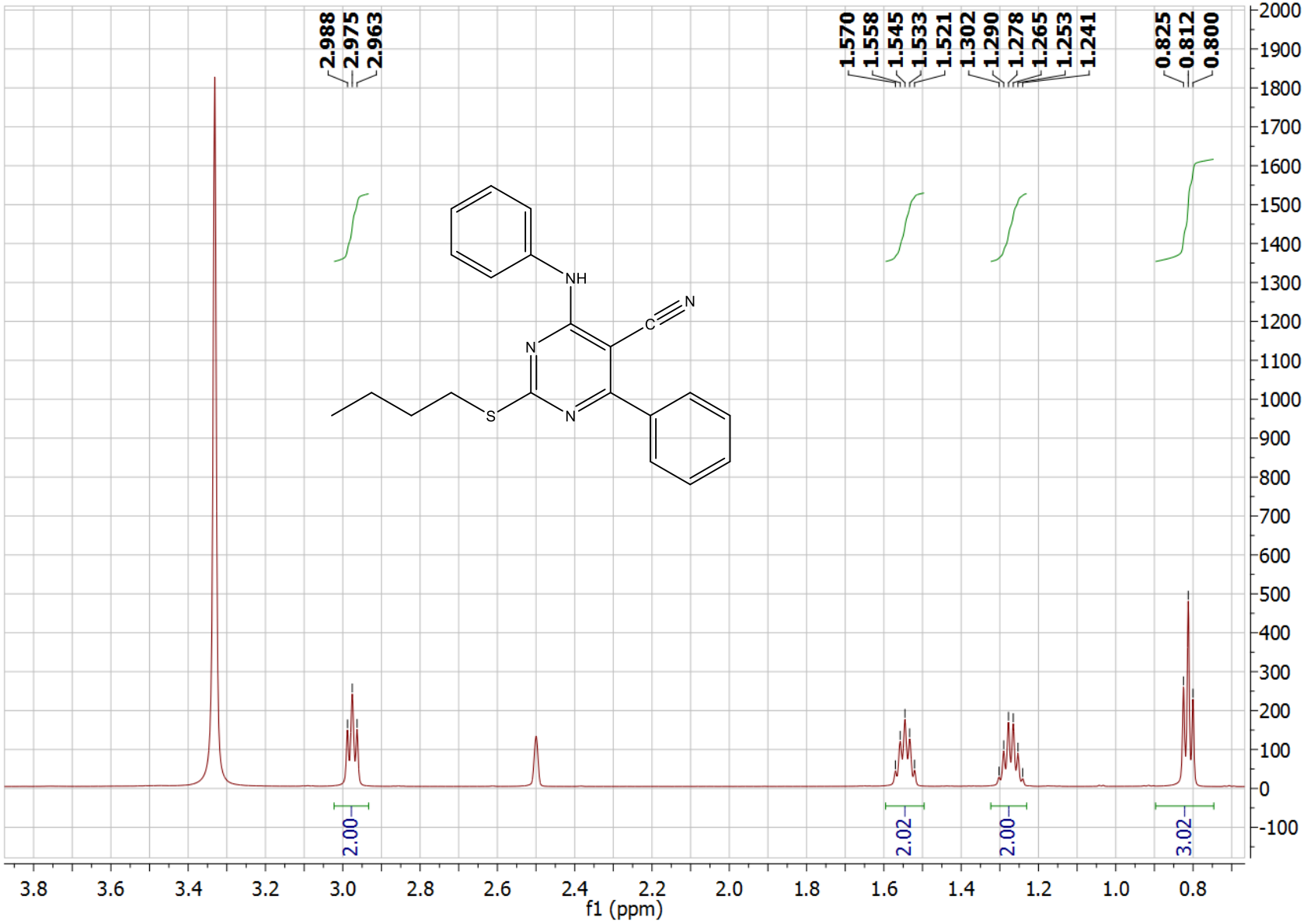
APT of compound 10a



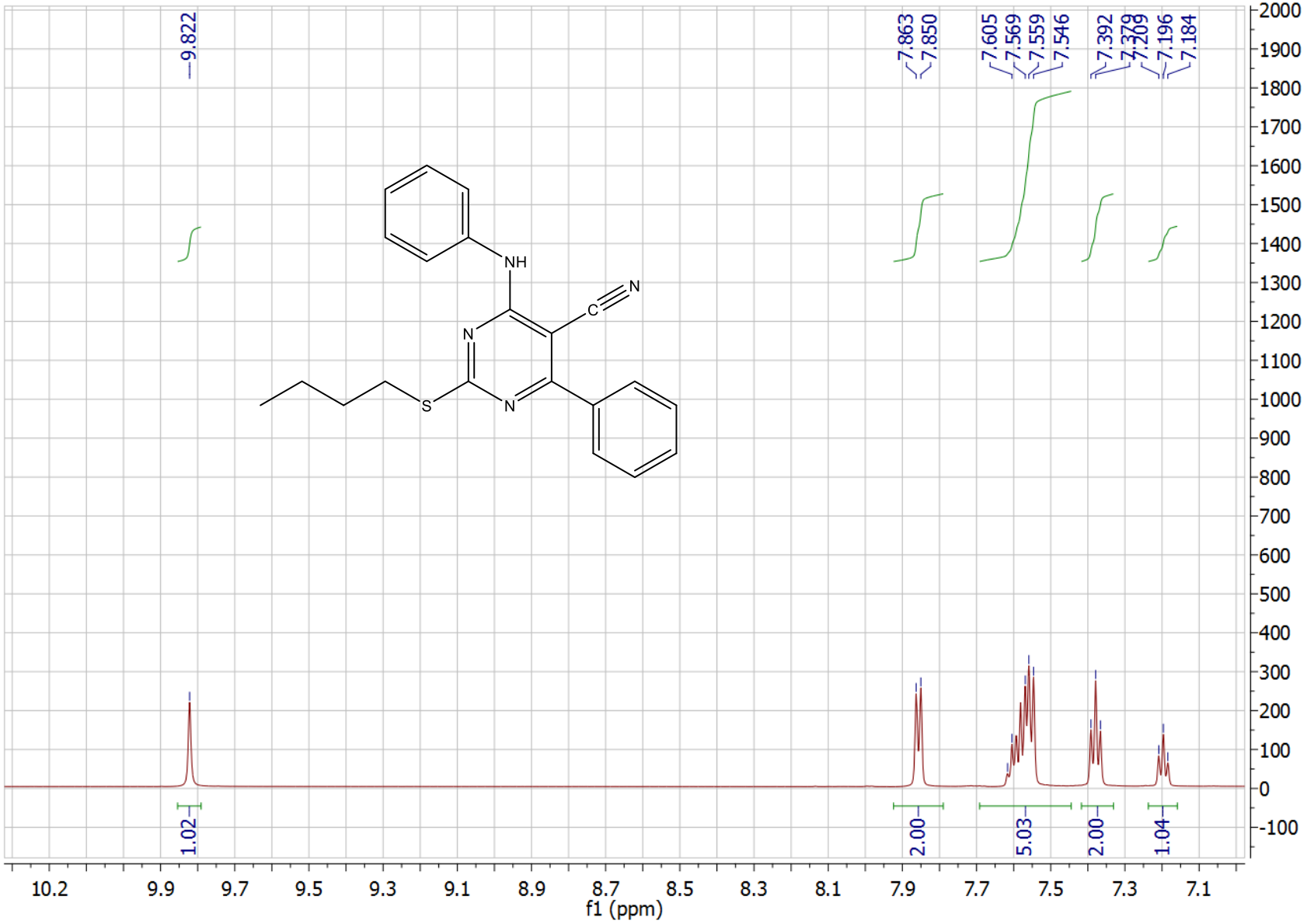
^{1}H NMR of compound 10_b



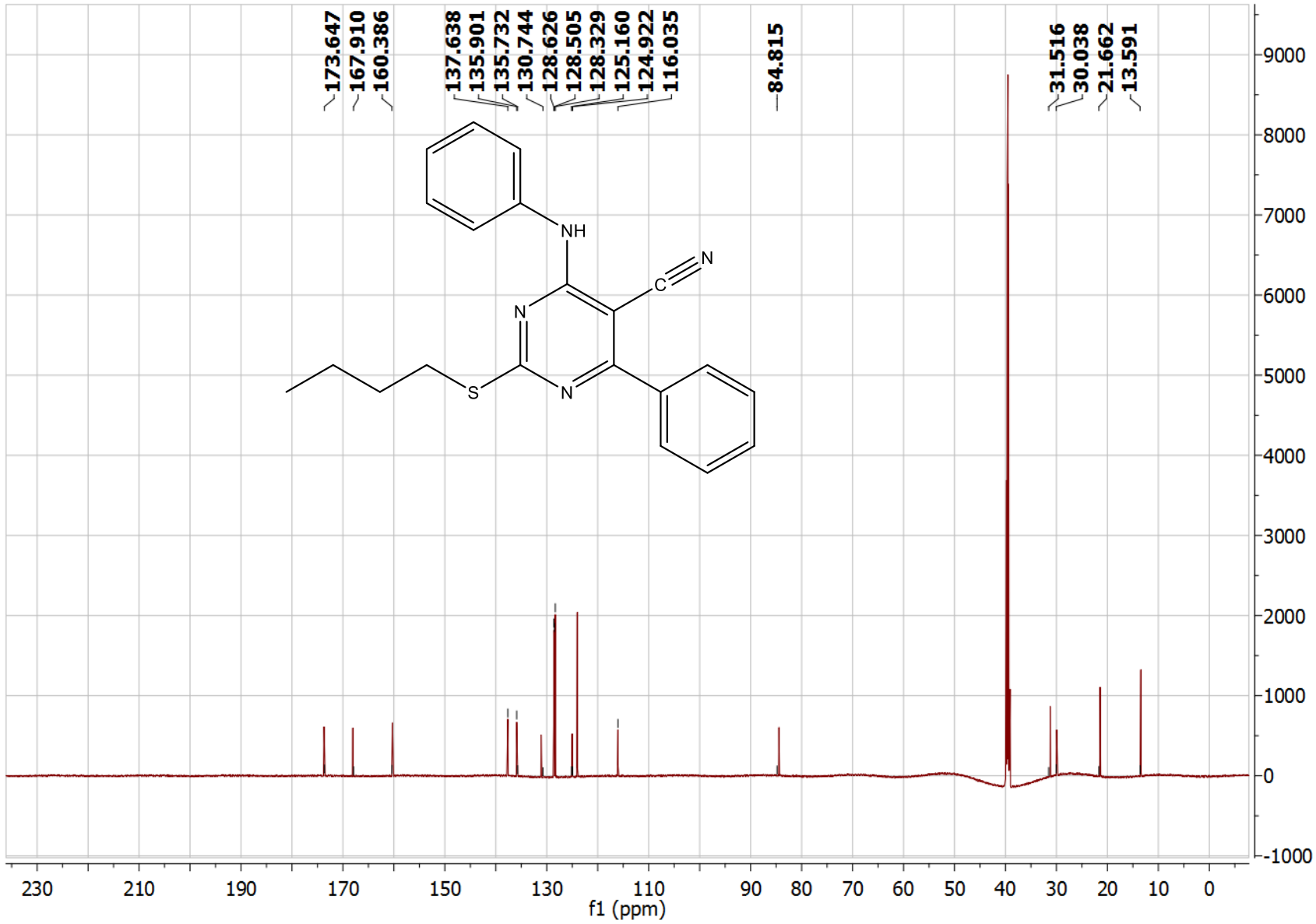
^{1}H NMR of compound 10_b



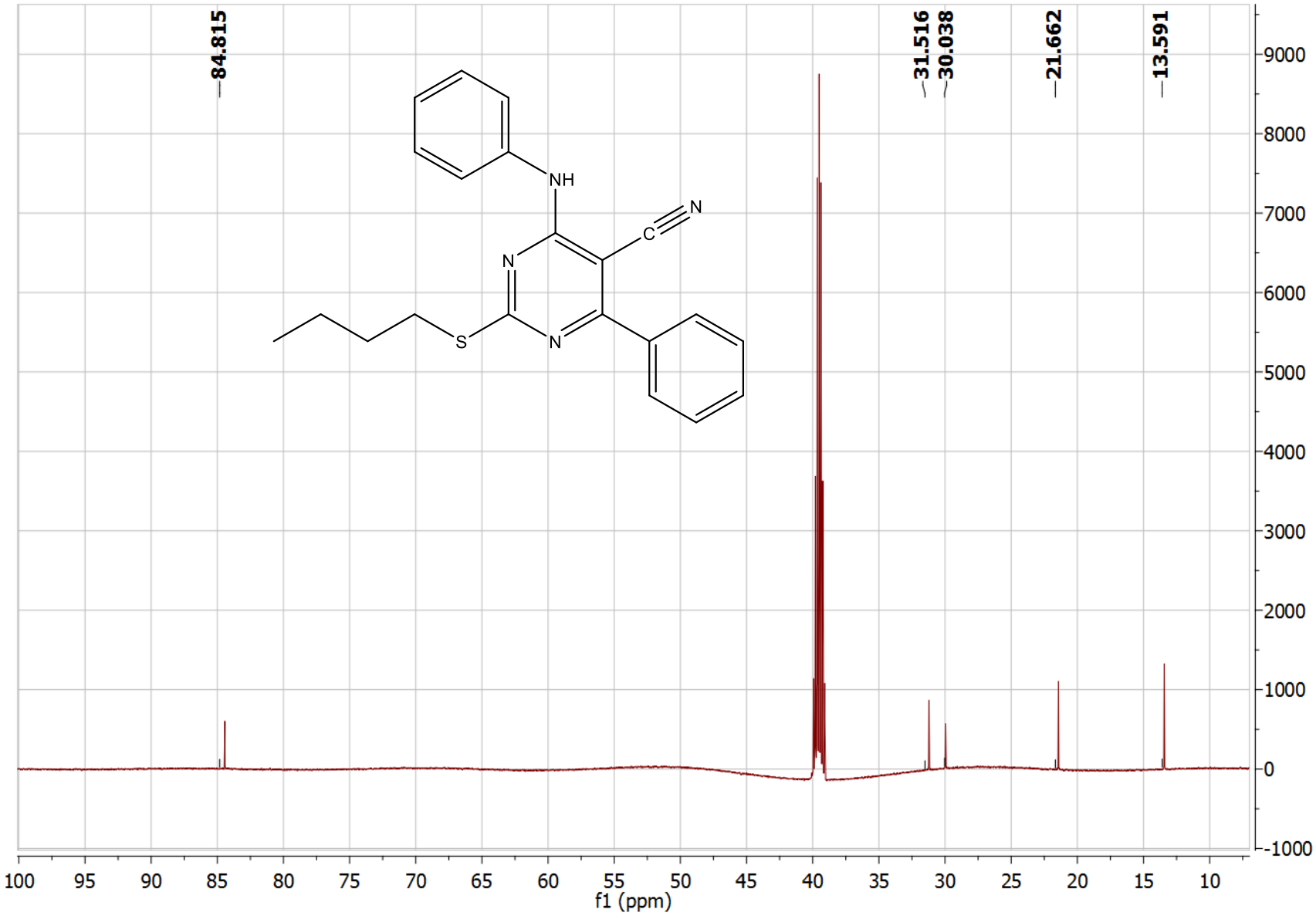
^{1}H NMR of compound 10_b



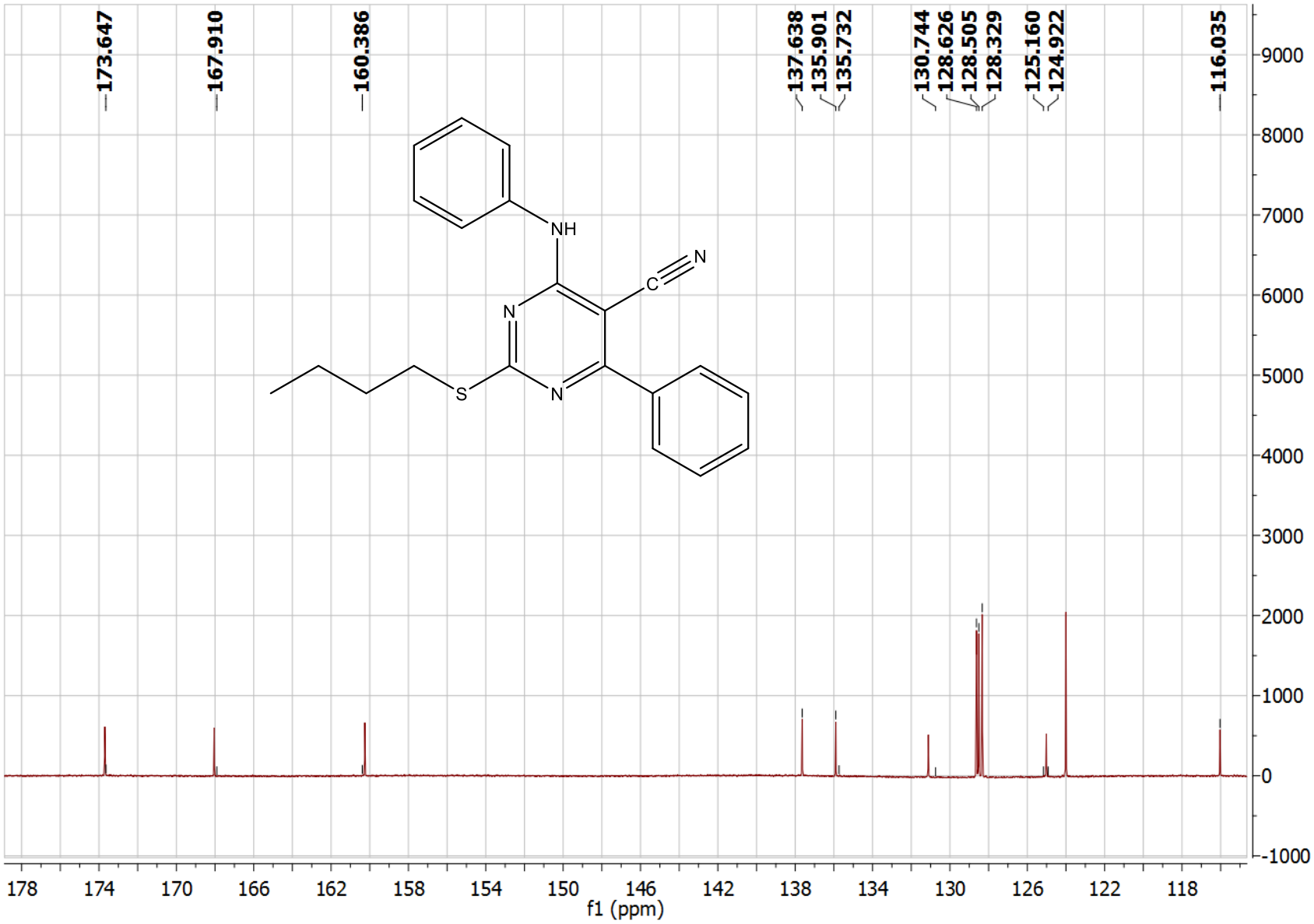
^{13}C NMR of compound 10_b



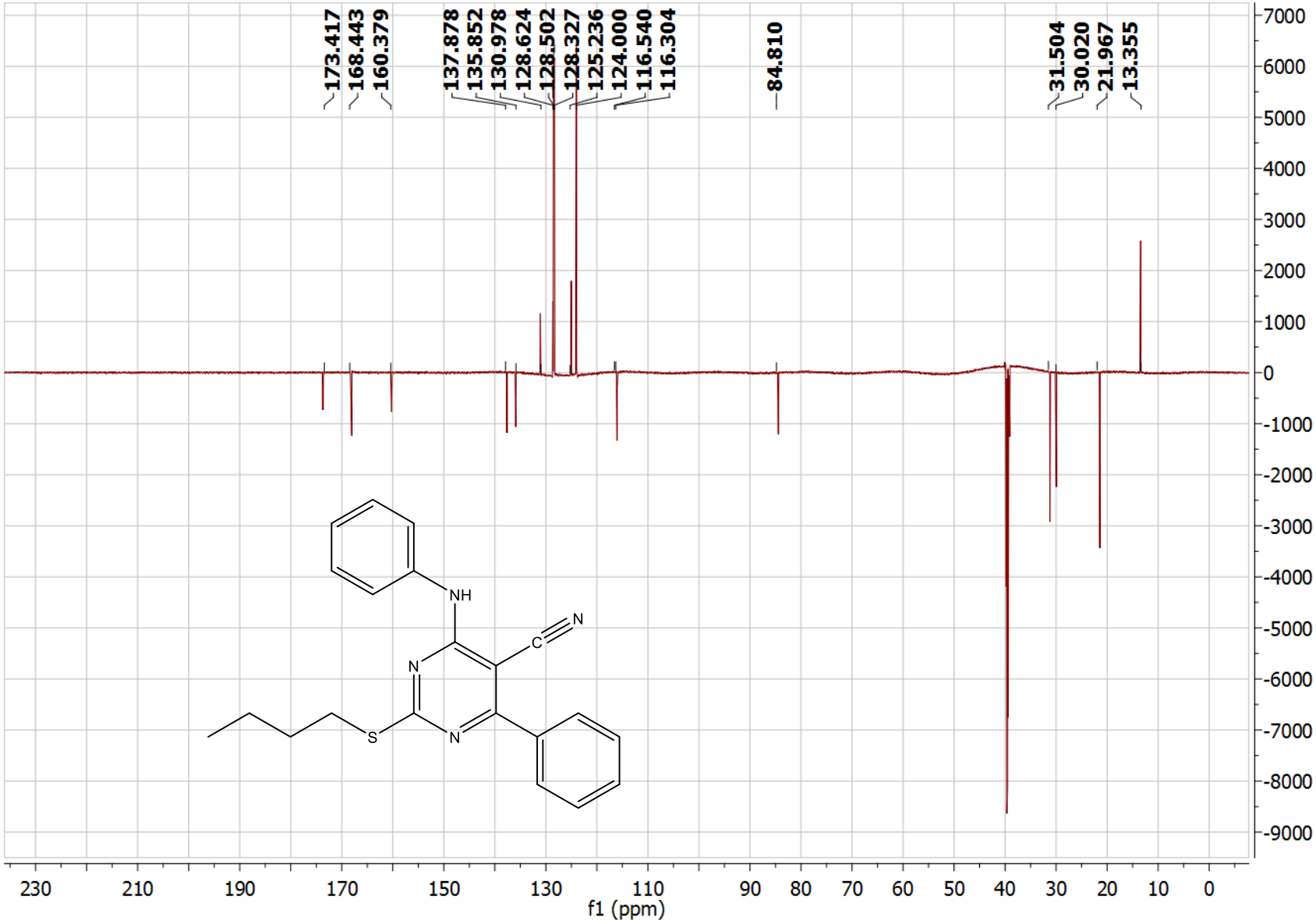
^{13}C NMR of compound 10_b



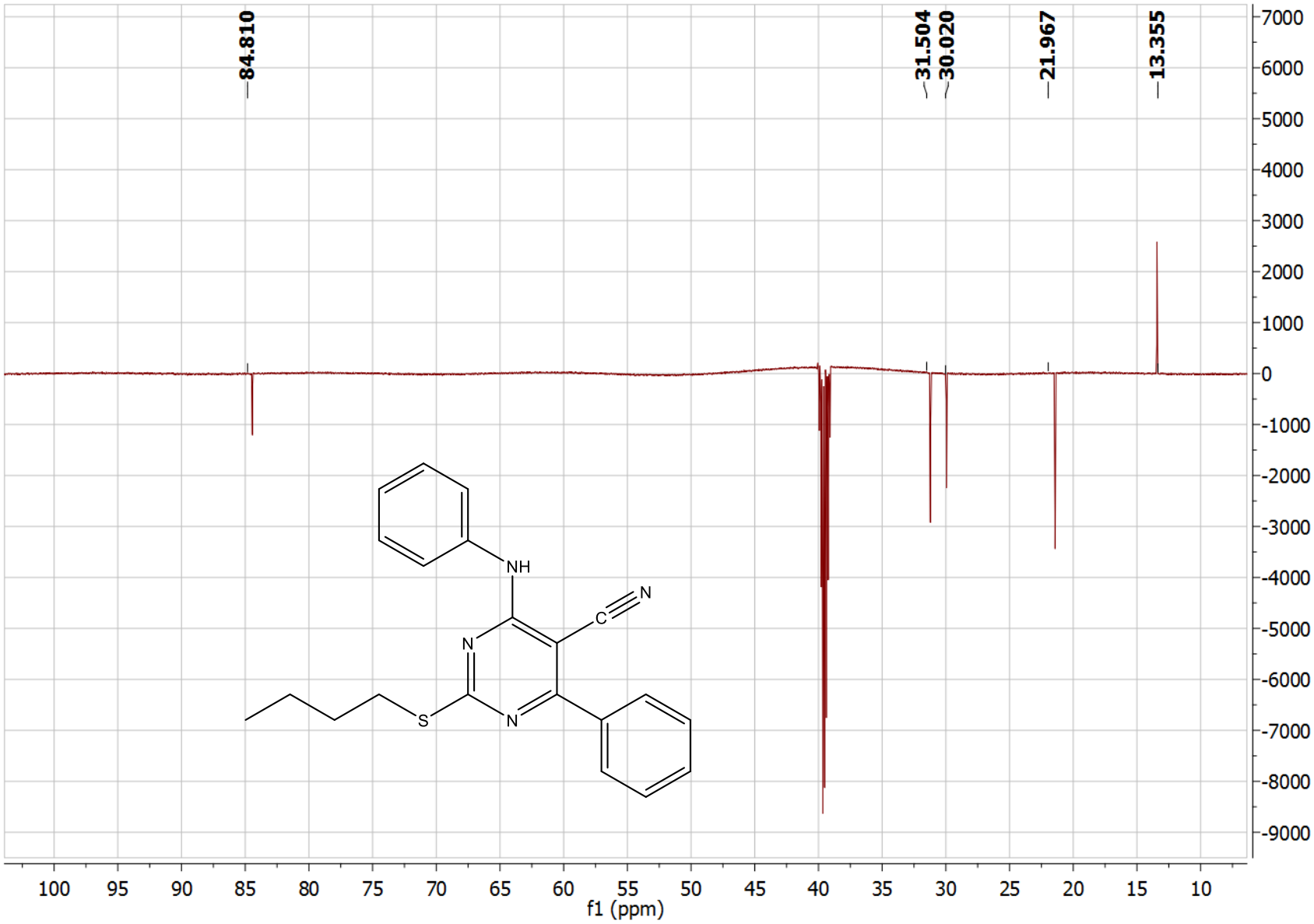
^{13}C NMR of compound 10_b



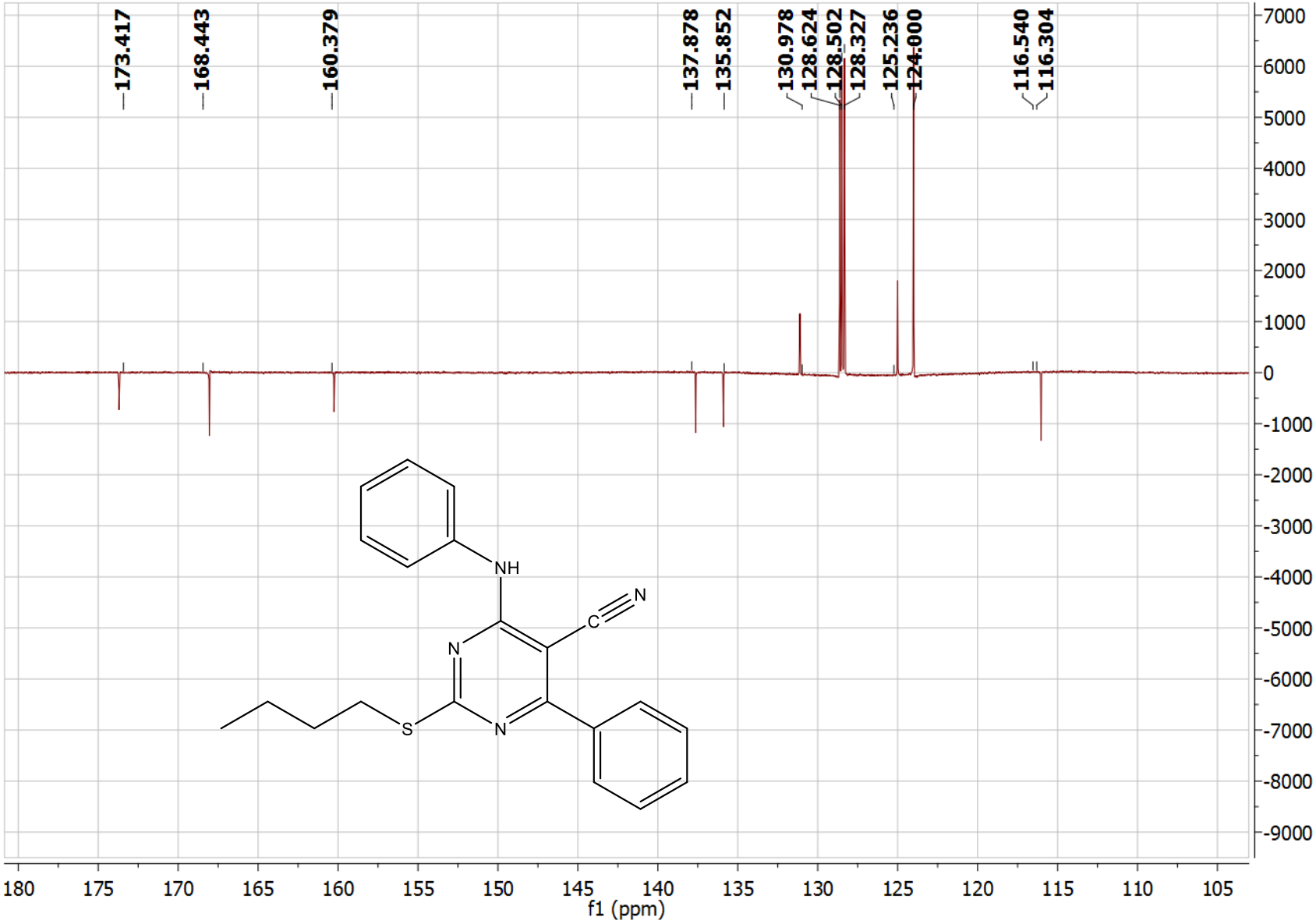
APT of compound 10_b



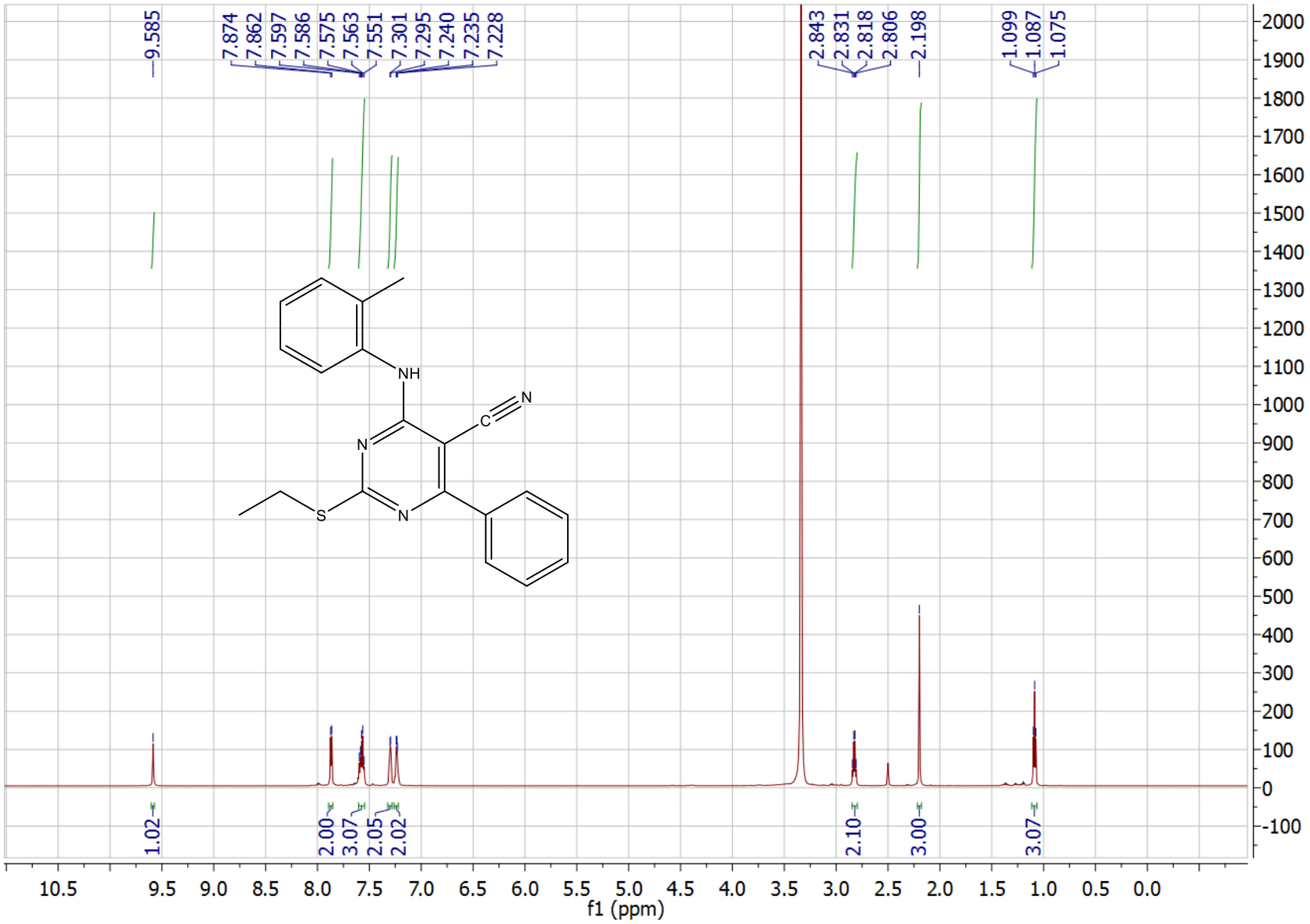
APT of compound 10_b



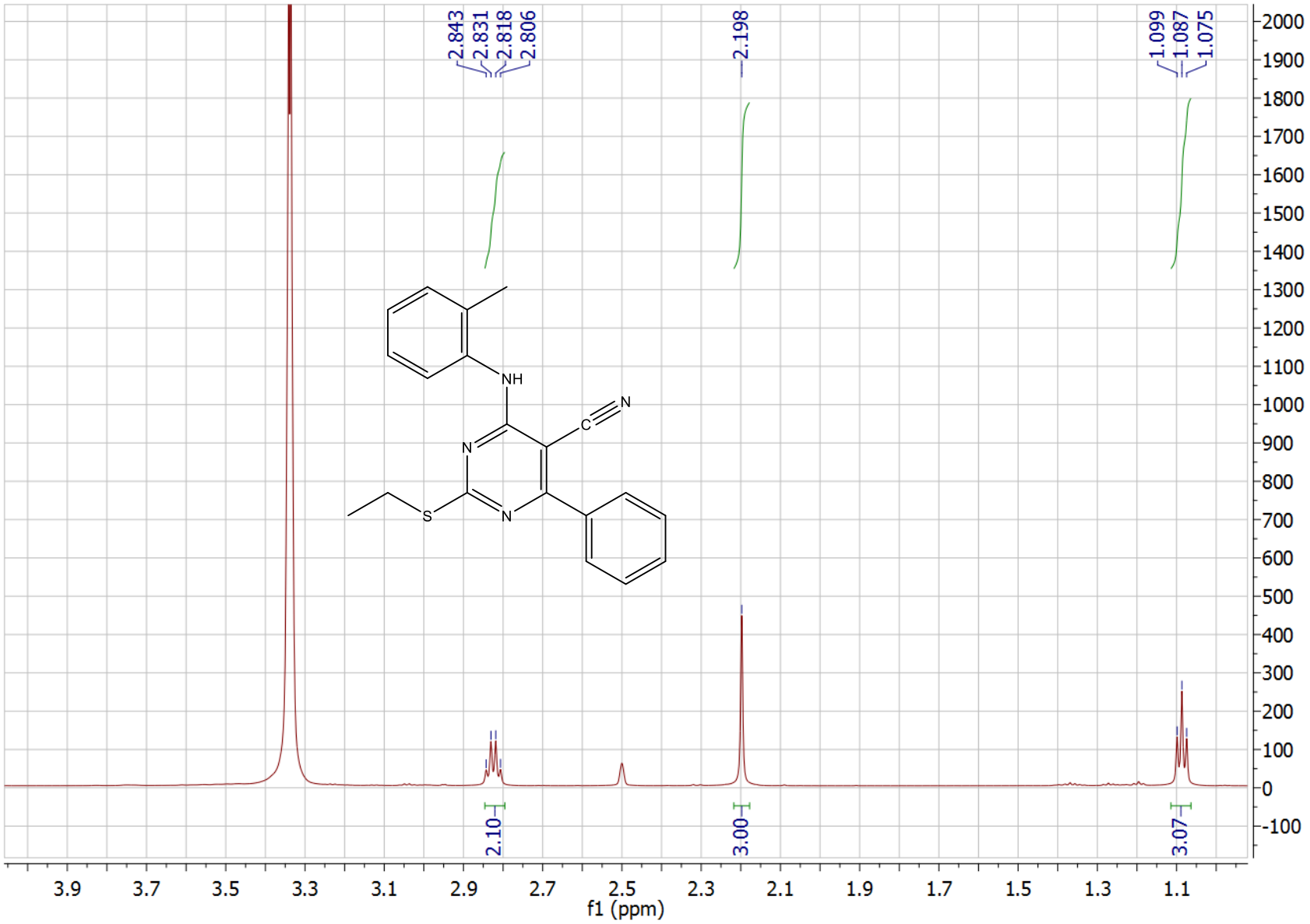
APT of compound 10_b



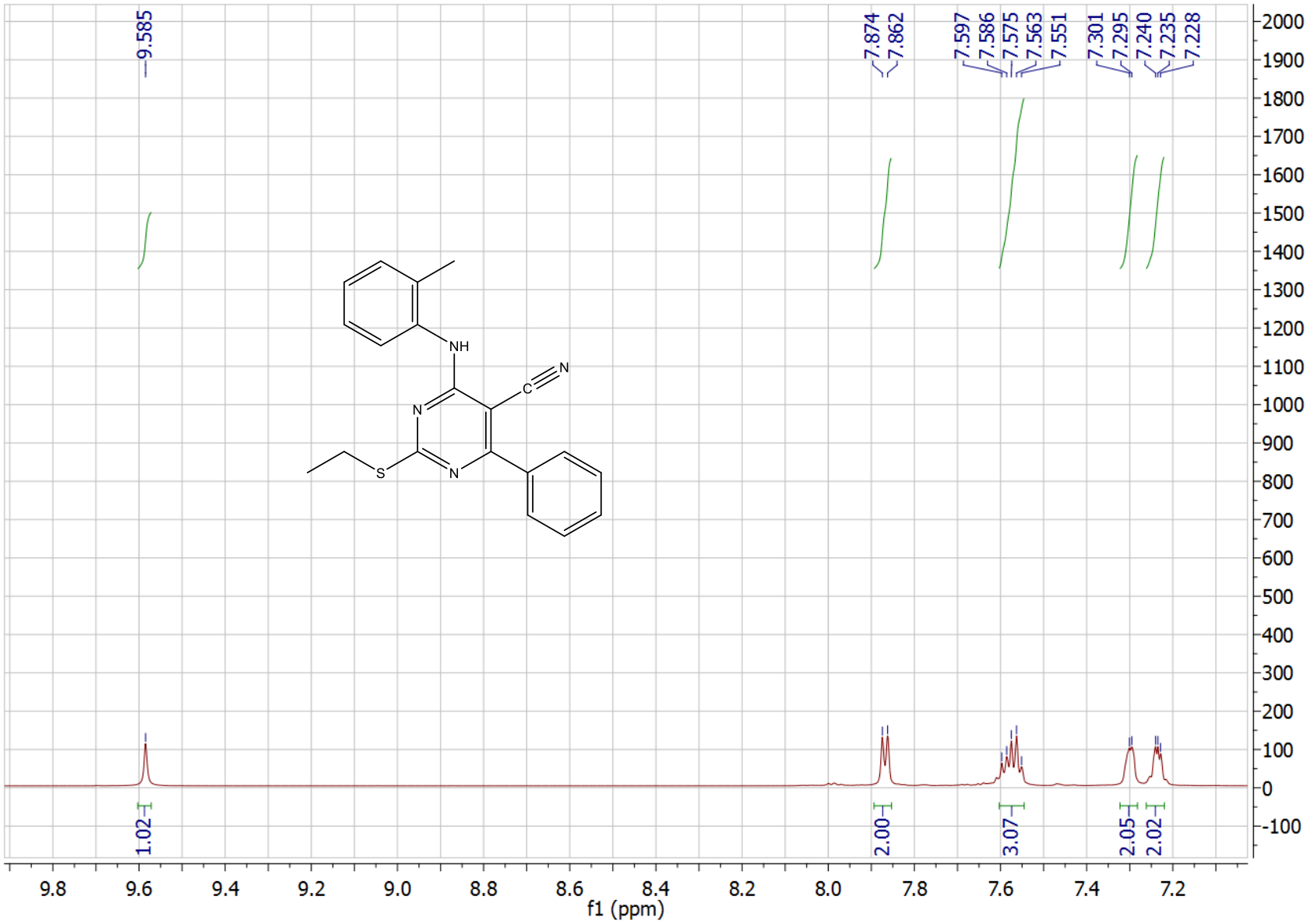
^{1}H NMR of compound 10_c



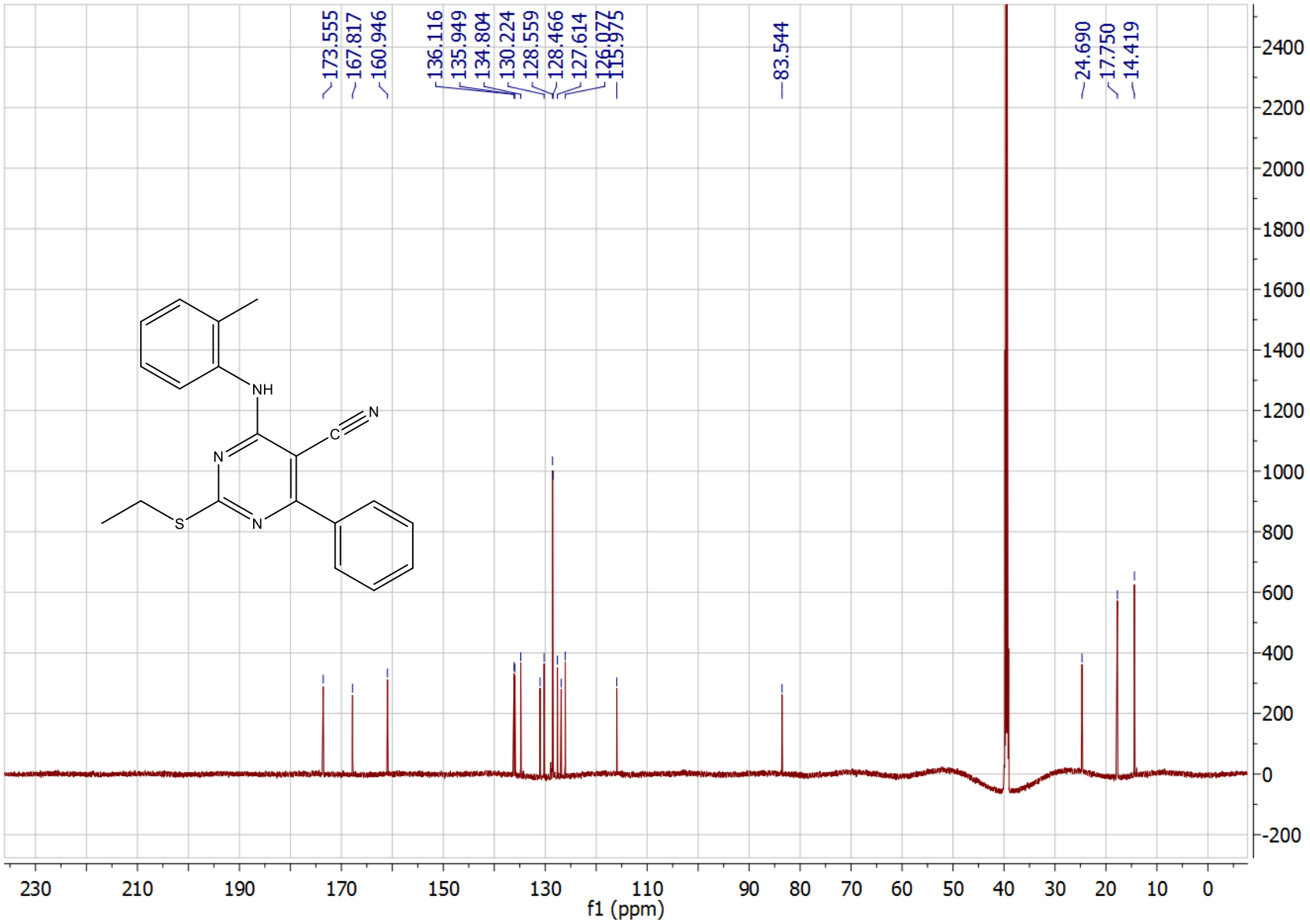
^{1}H NMR of compound 10_c



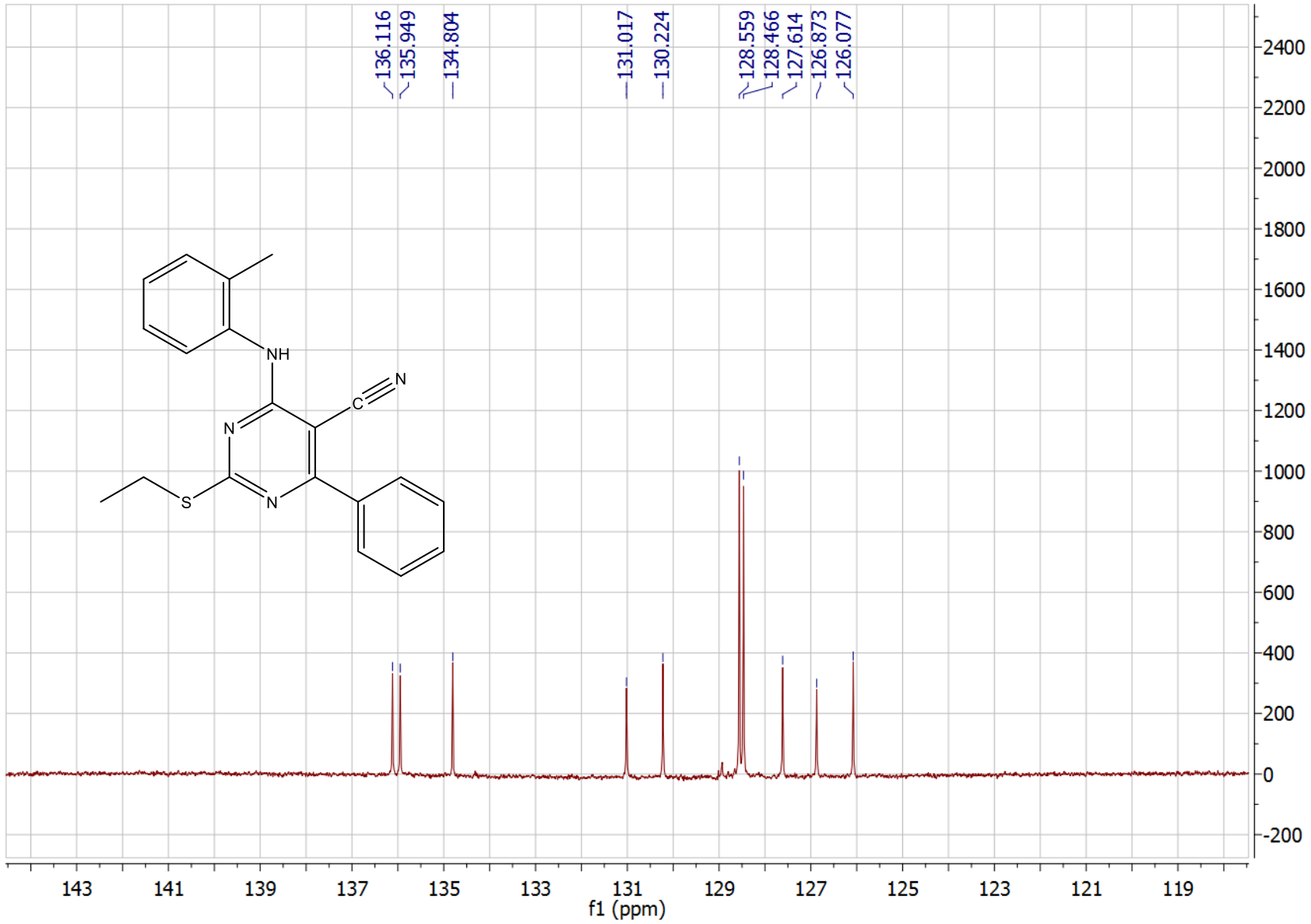
^{1}H NMR of compound 10_c



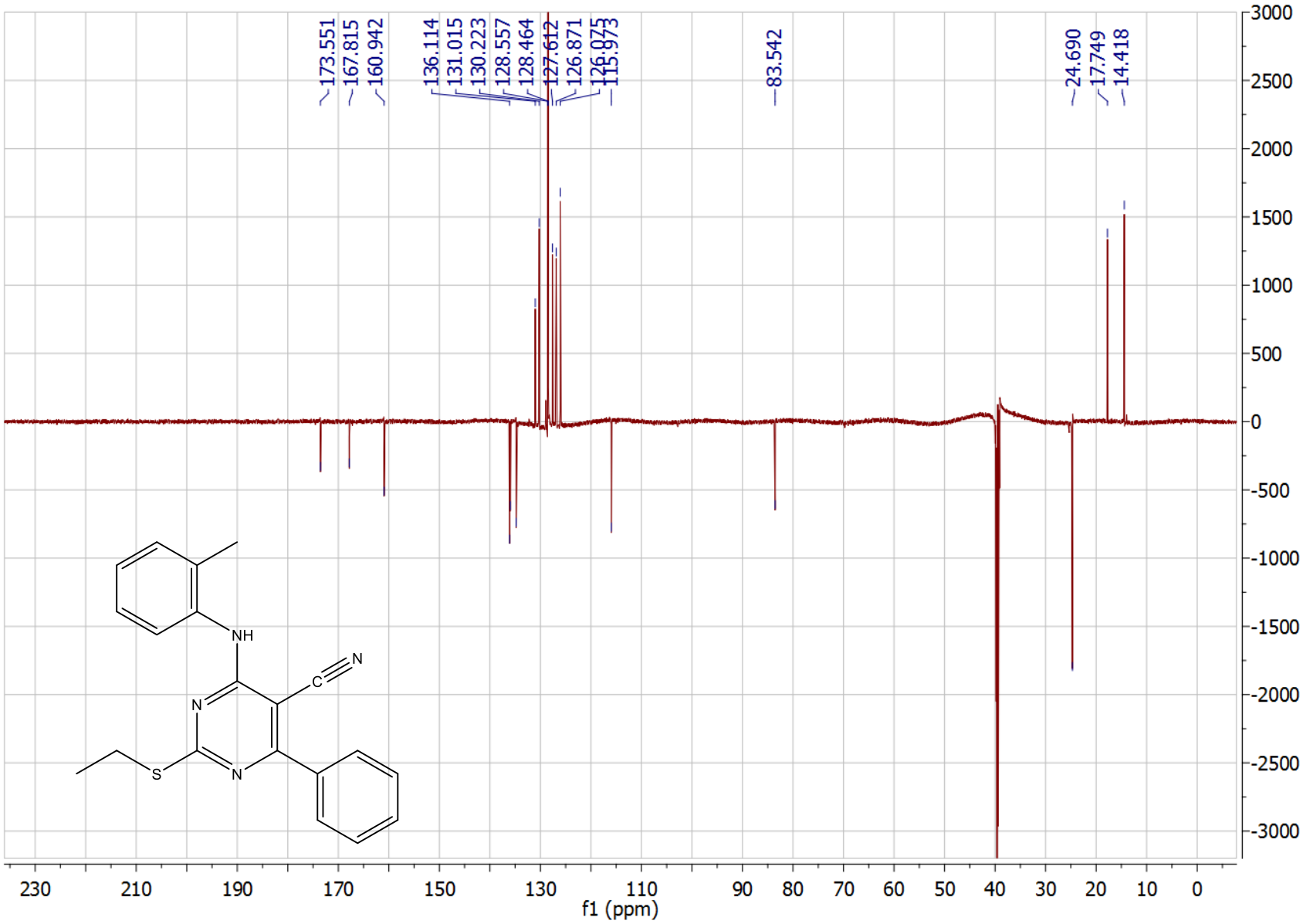
^{13}C NMR of compound 10_c



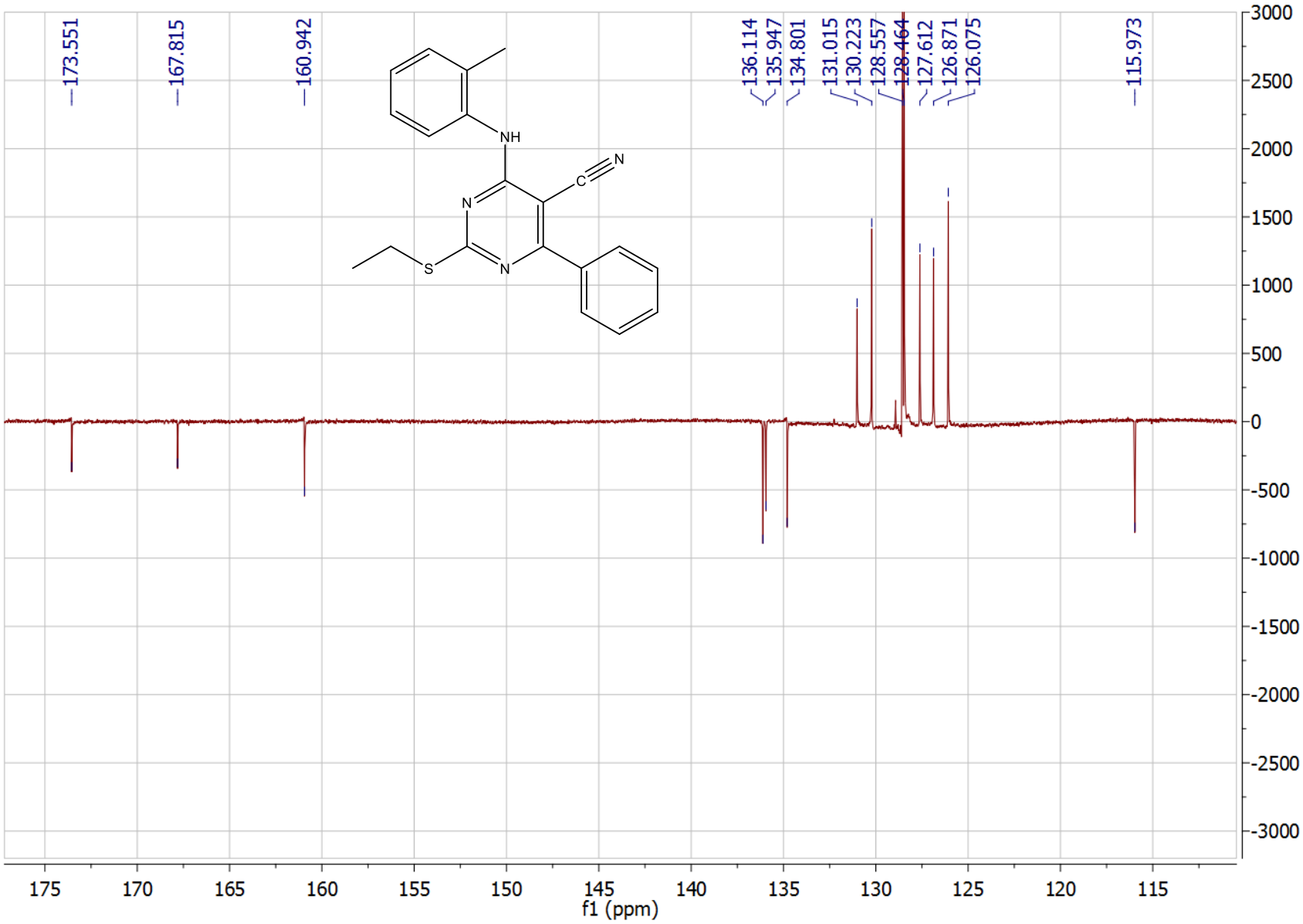
^{13}C NMR of compound 10_c



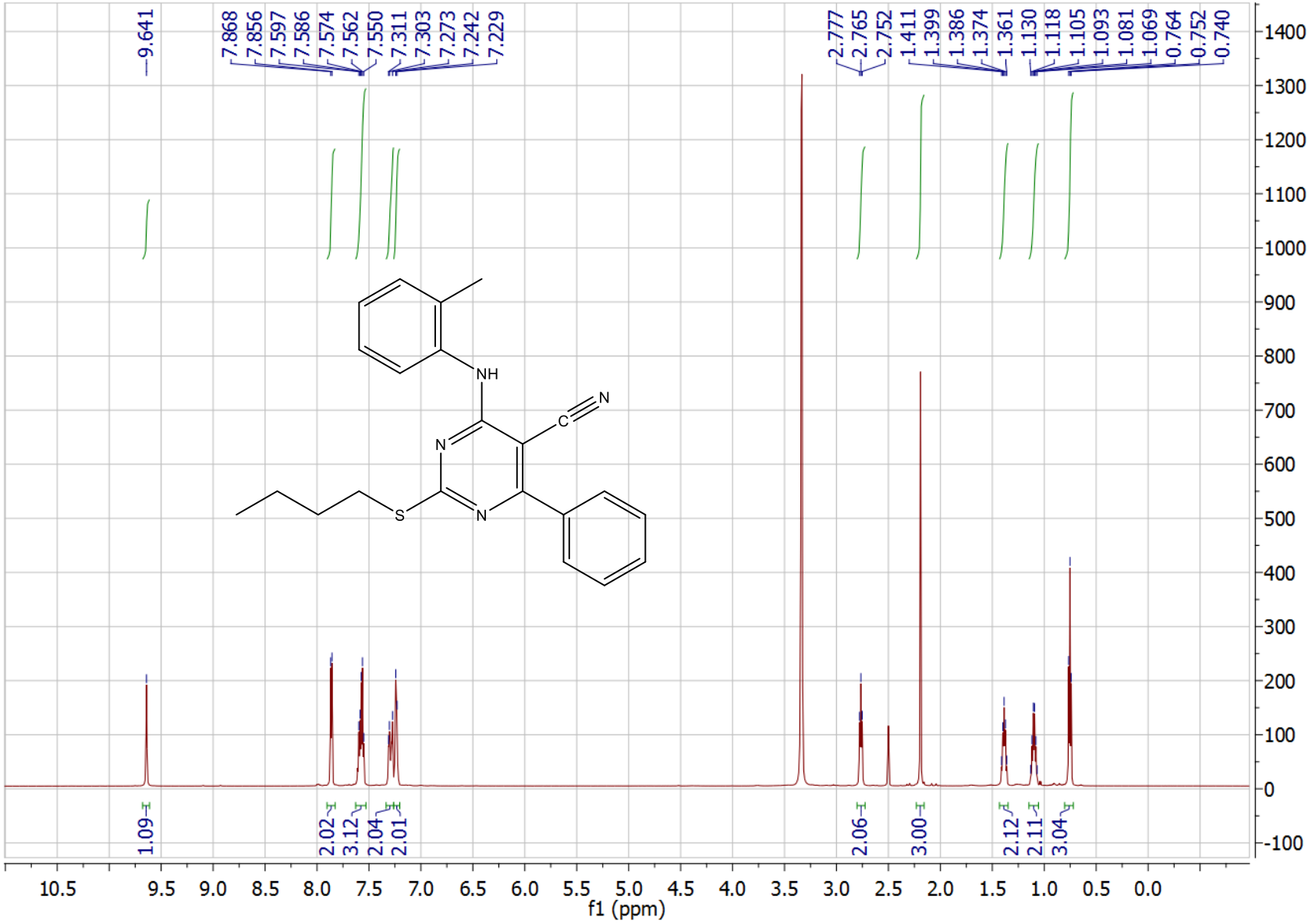
APT of compound 10_C



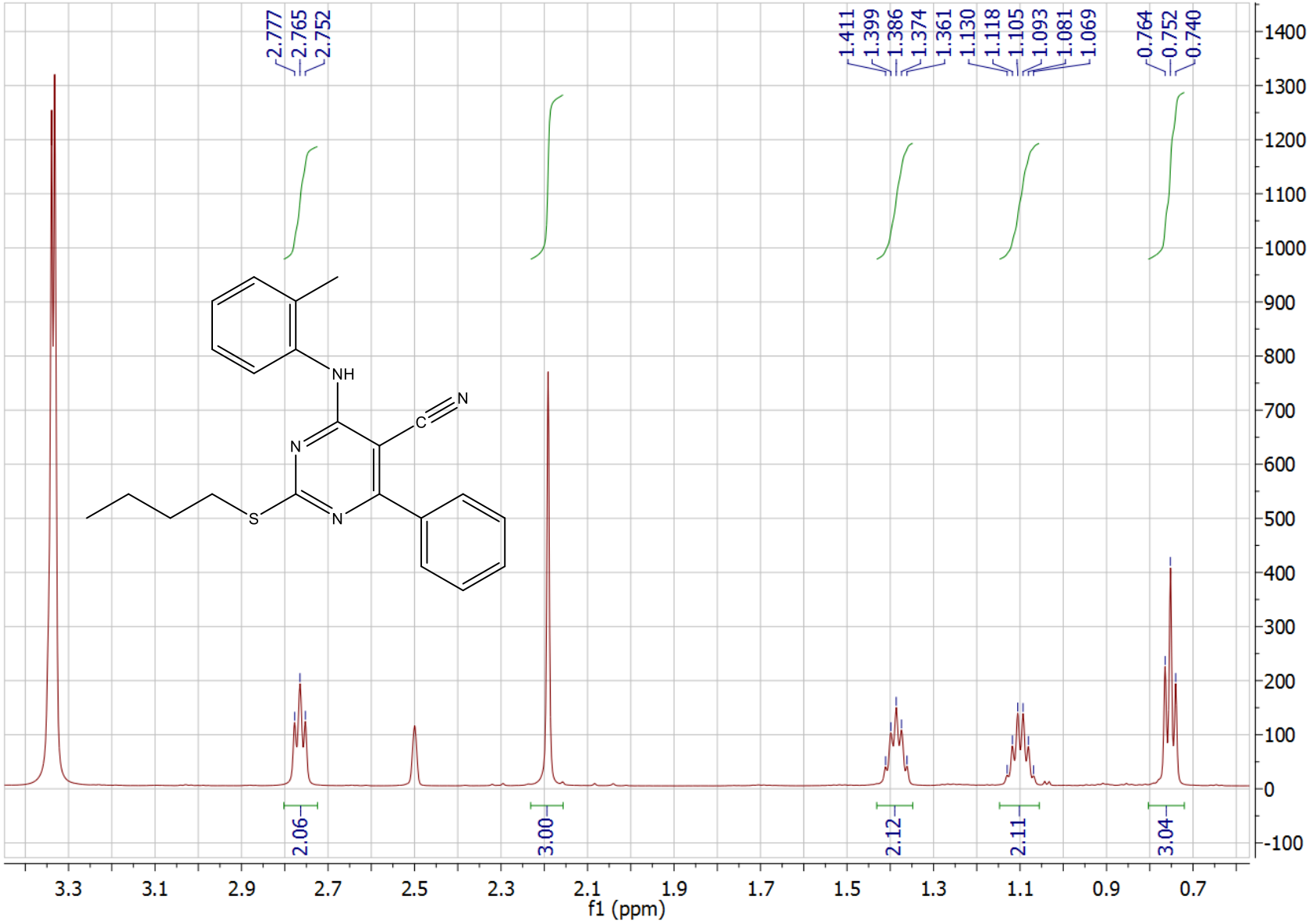
APT of compound 10_C



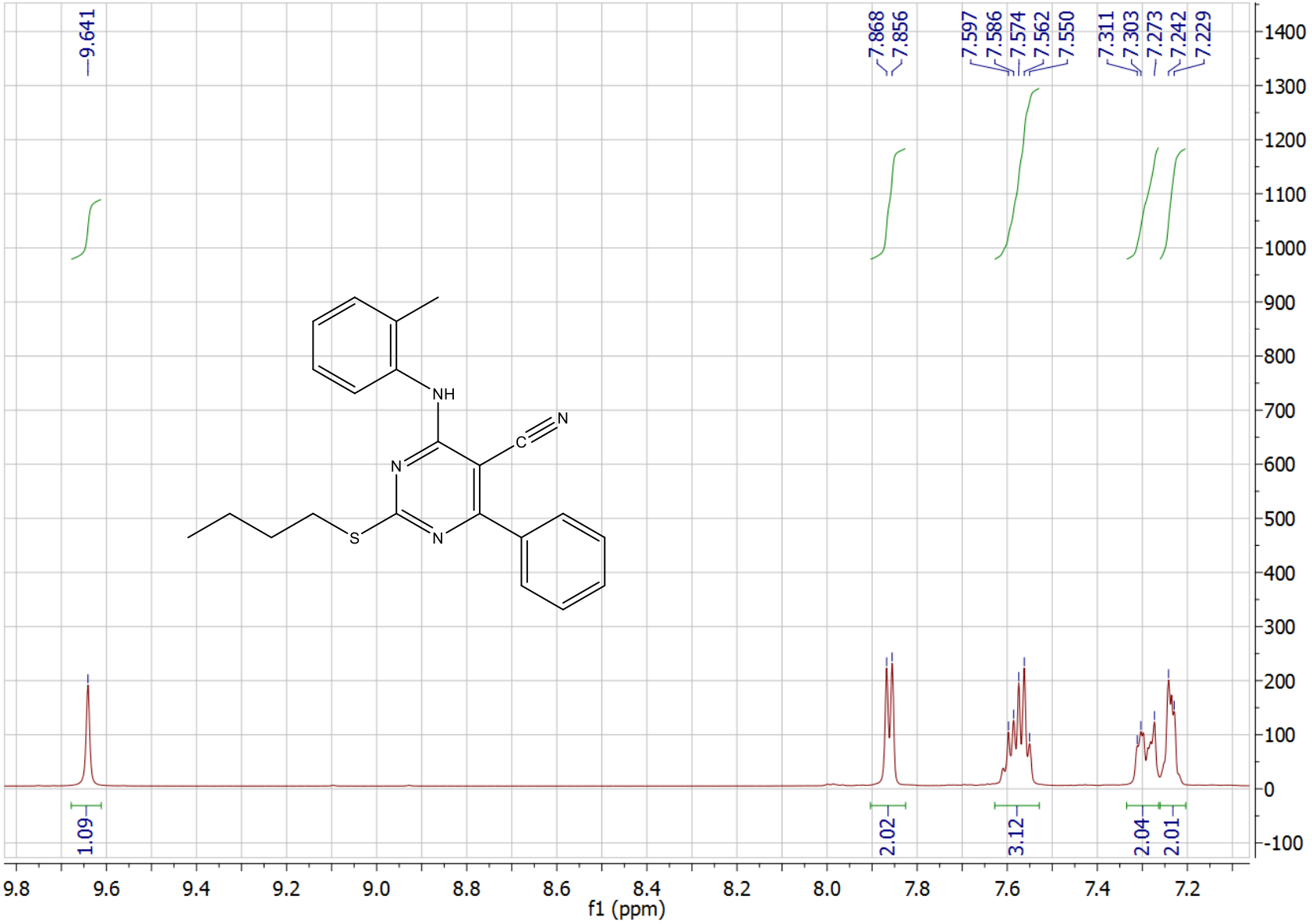
^{1}H NMR of compound 10_d



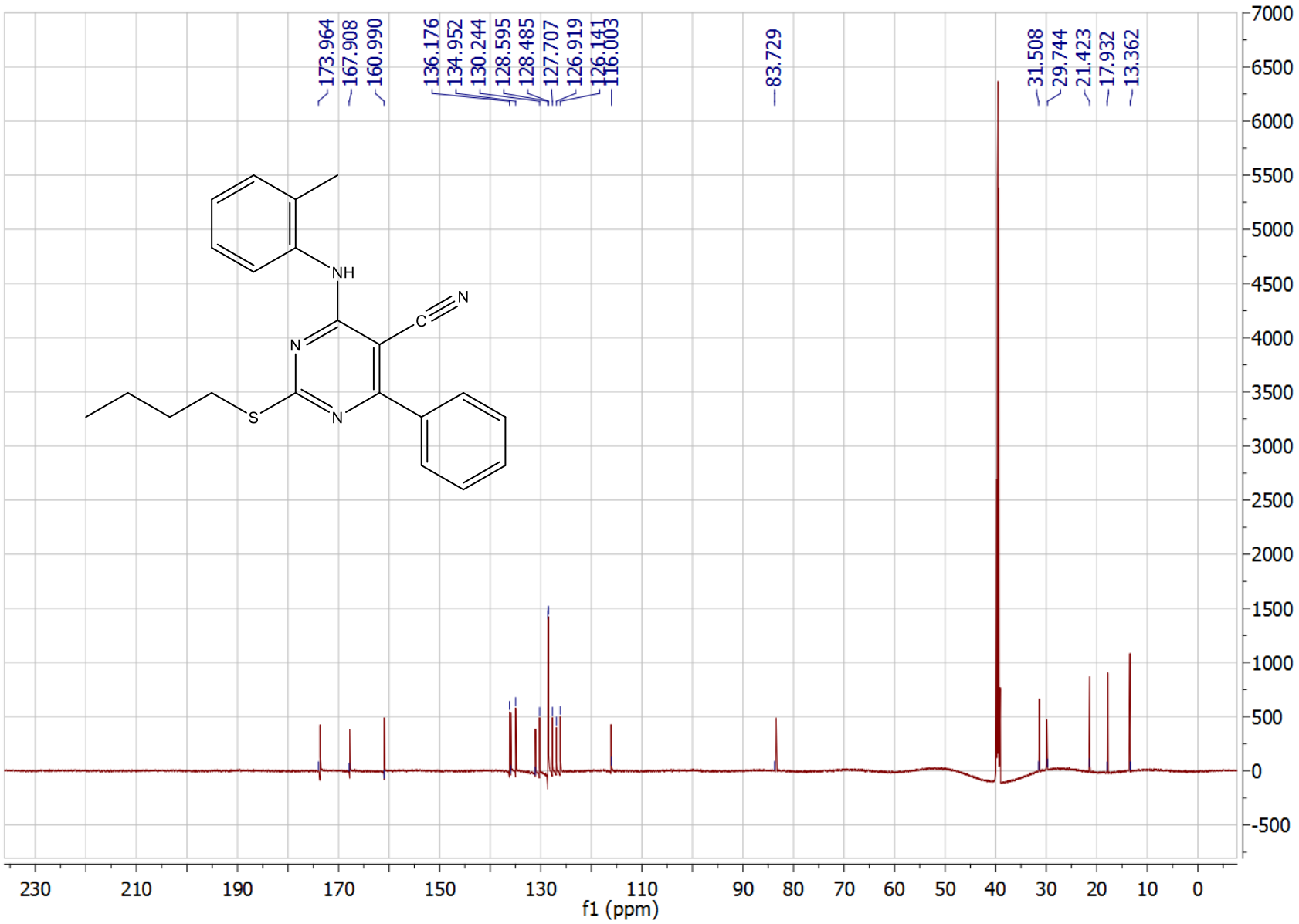
^{1}H NMR of compound 10_d



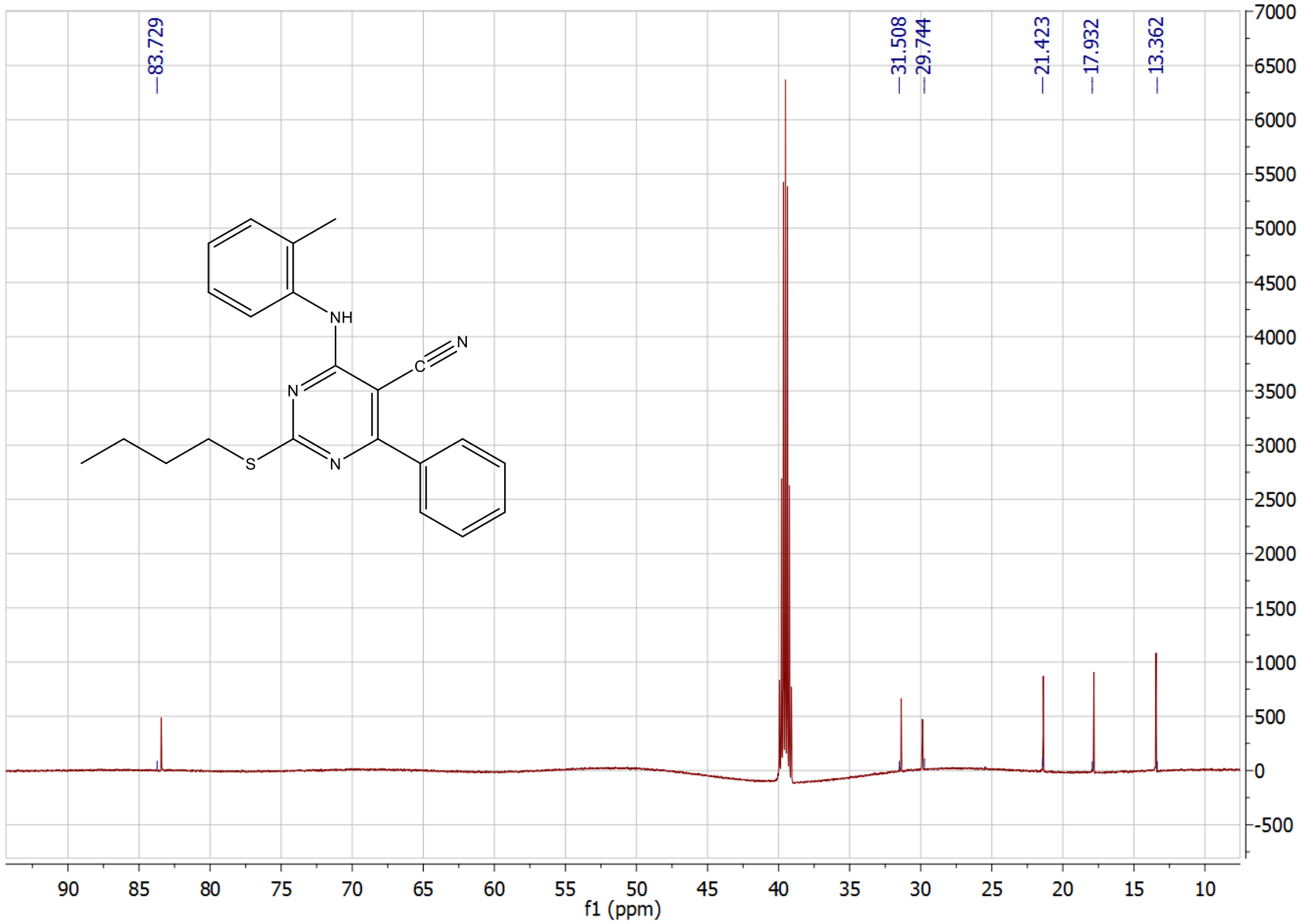
^{1}H NMR of compound 10_d



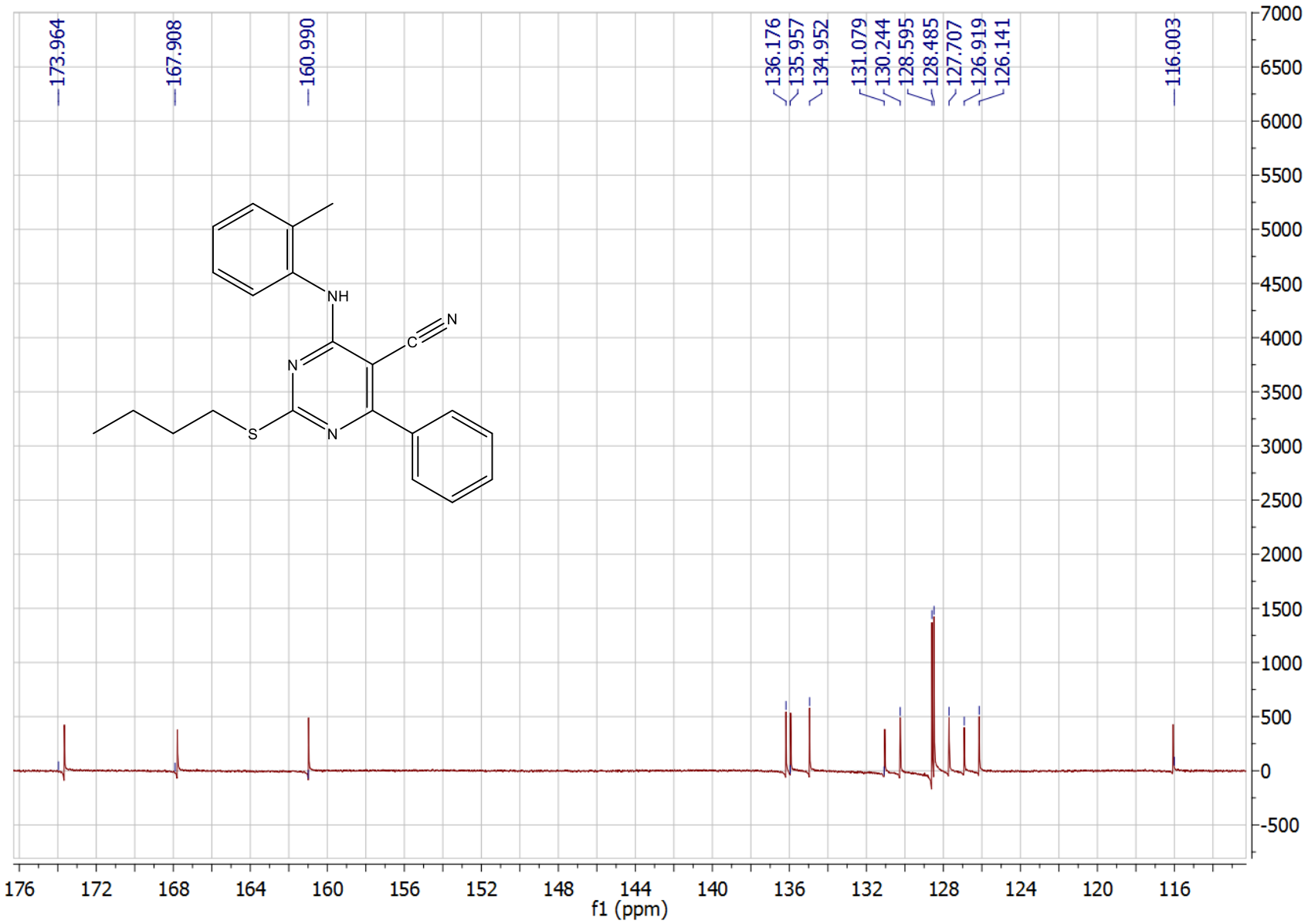
^{13}C NMR of compound 10_d



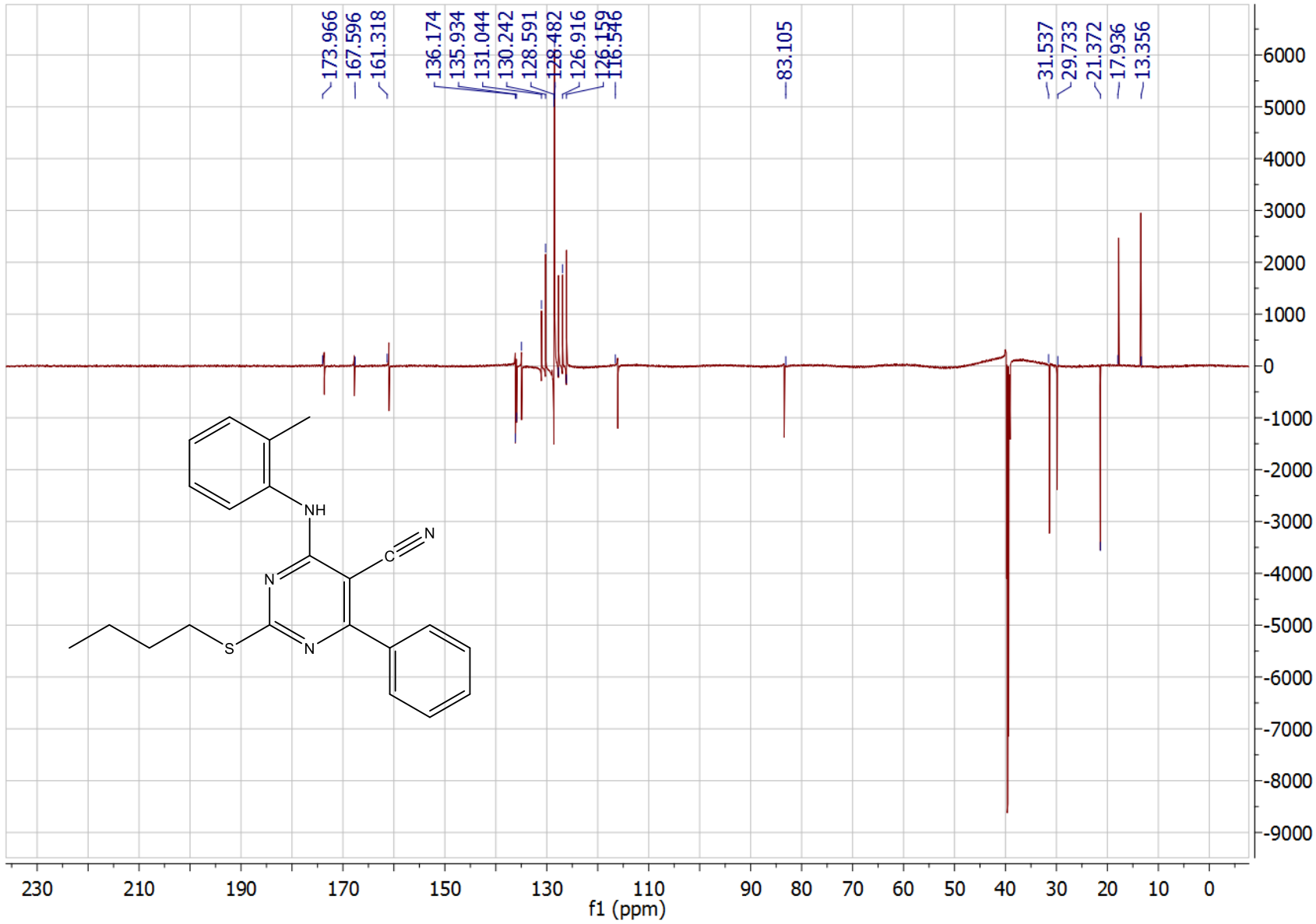
^{13}C NMR of compound 10_d



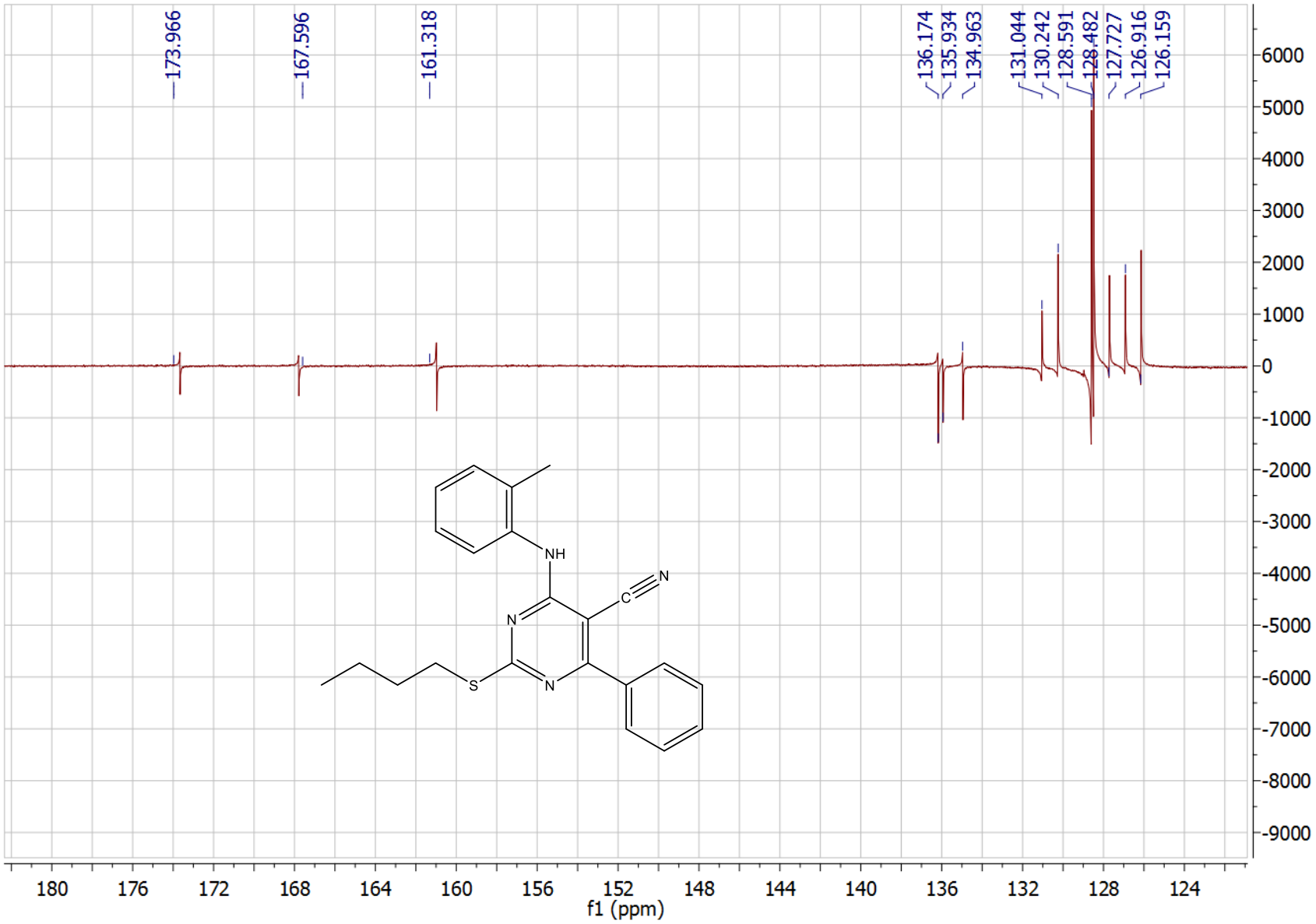
^{13}C NMR of compound 10_d



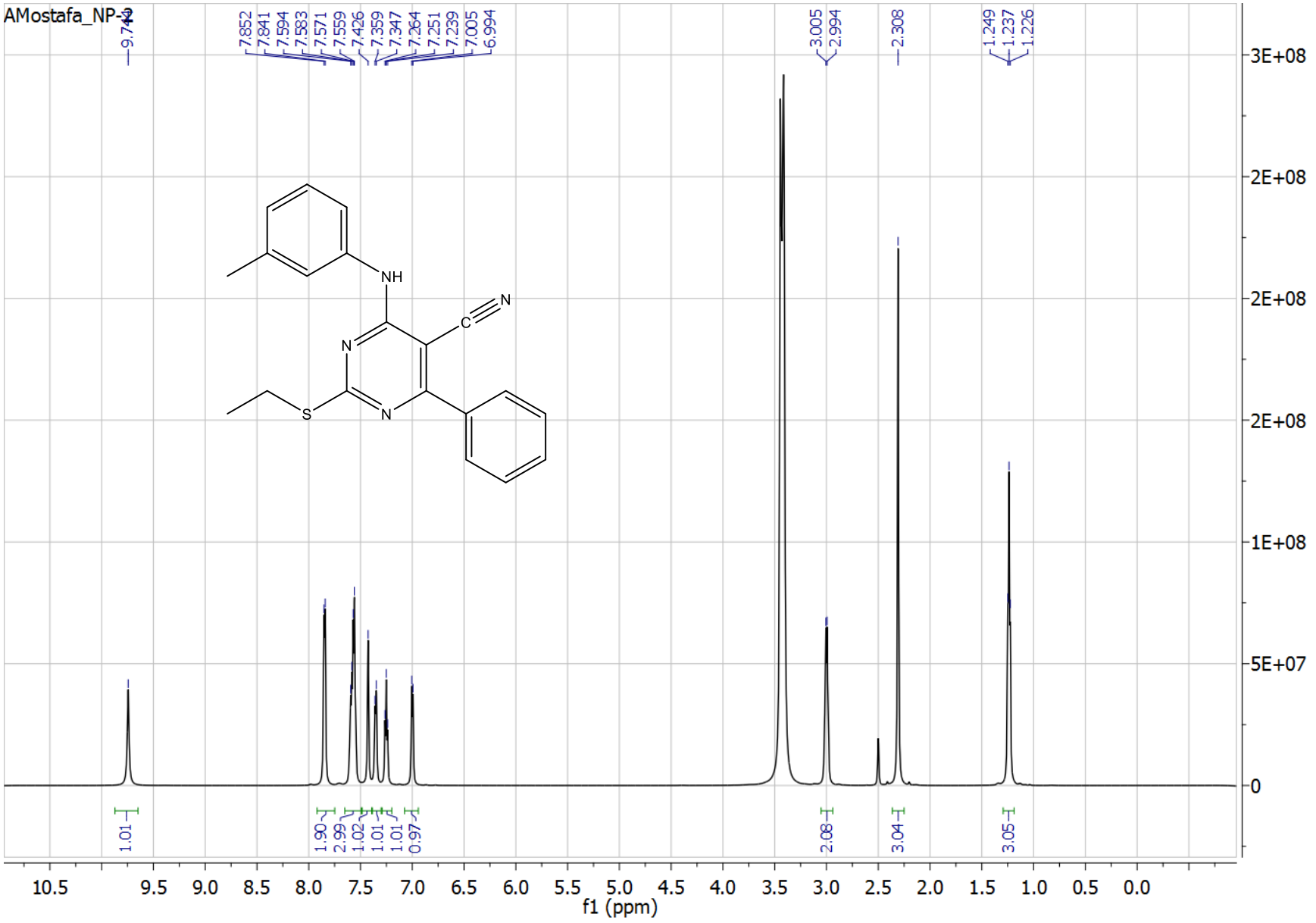
APT of compound 10_d



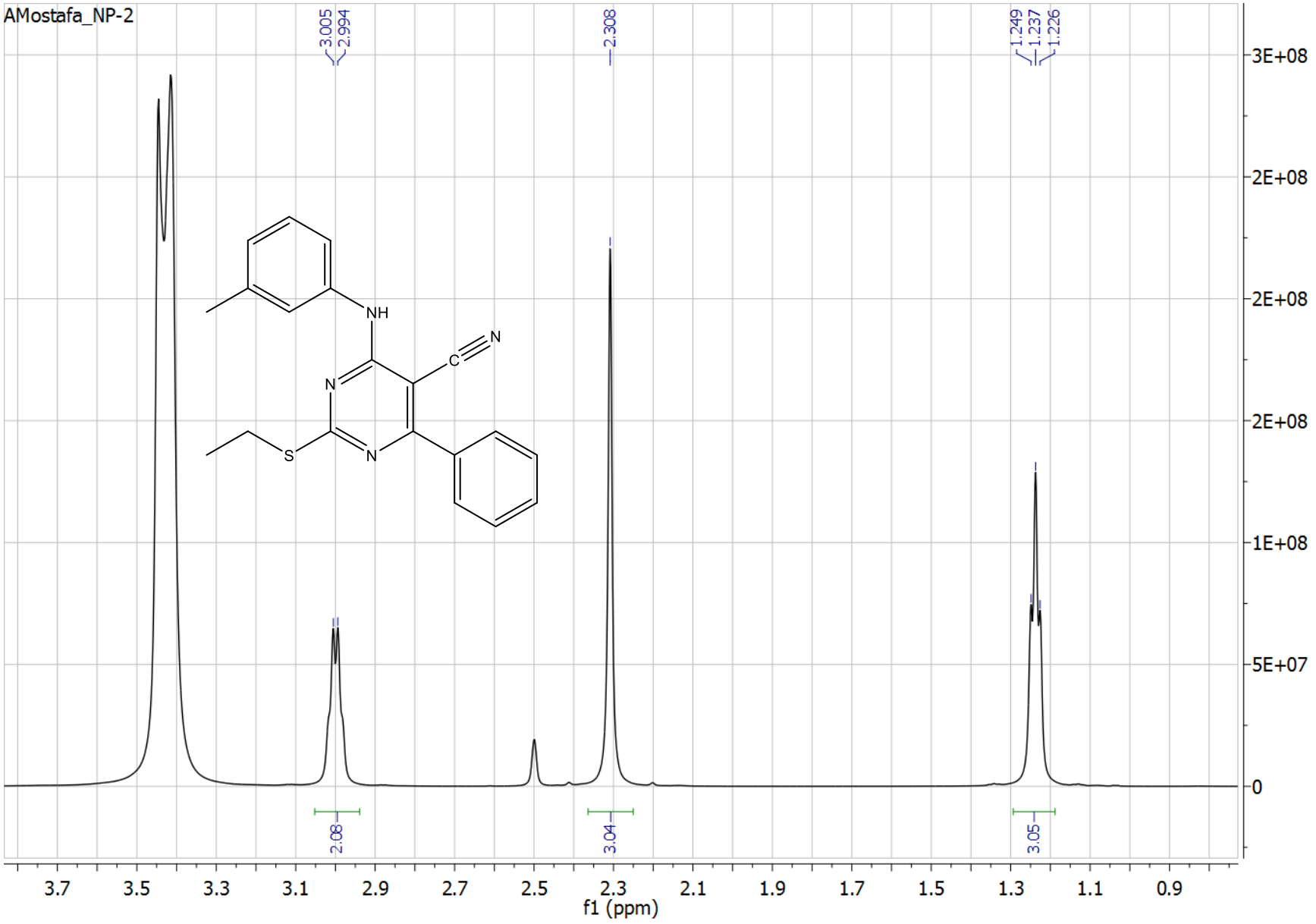
APT of compound 10_d



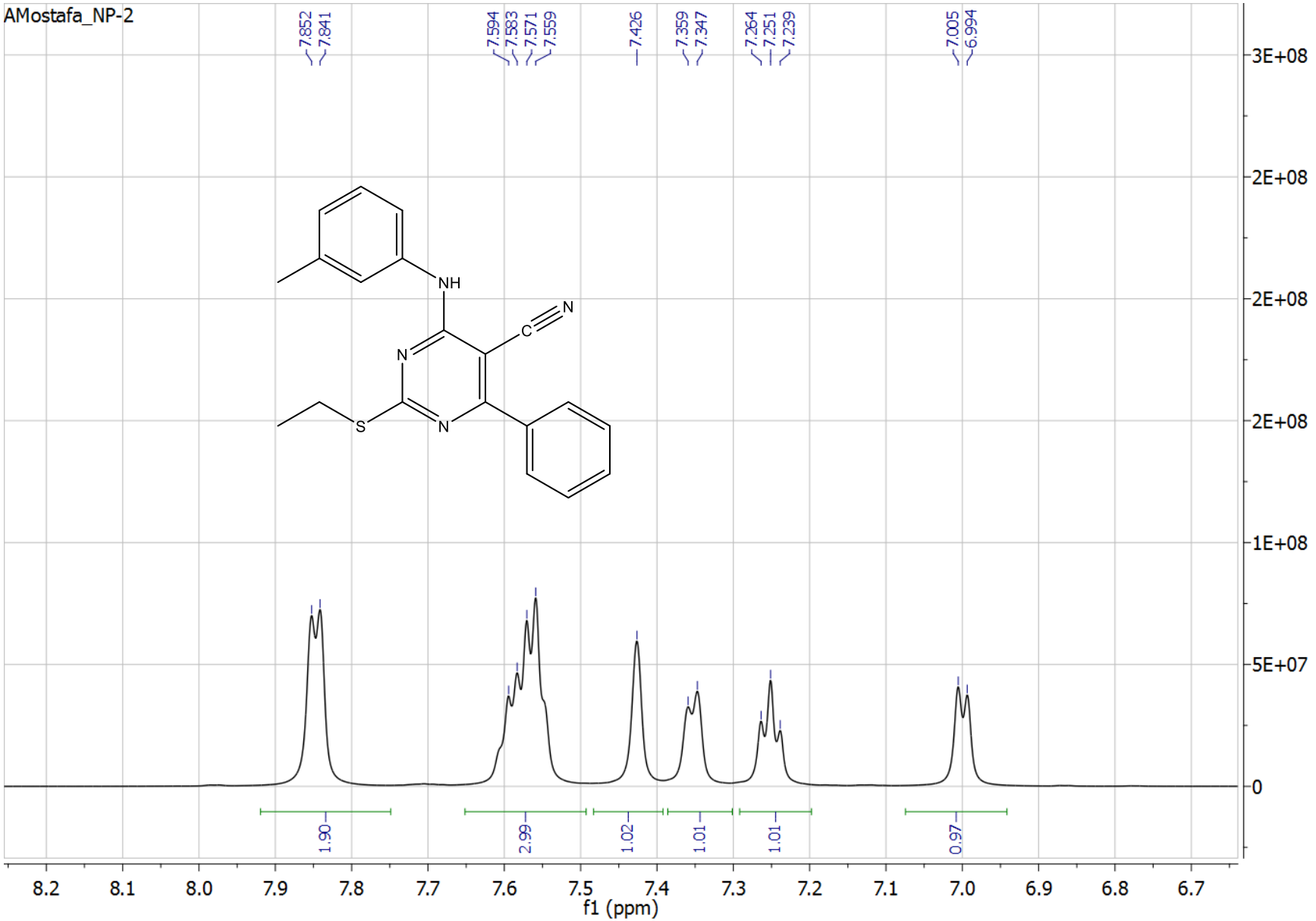
^{1}H NMR of compound 10_e



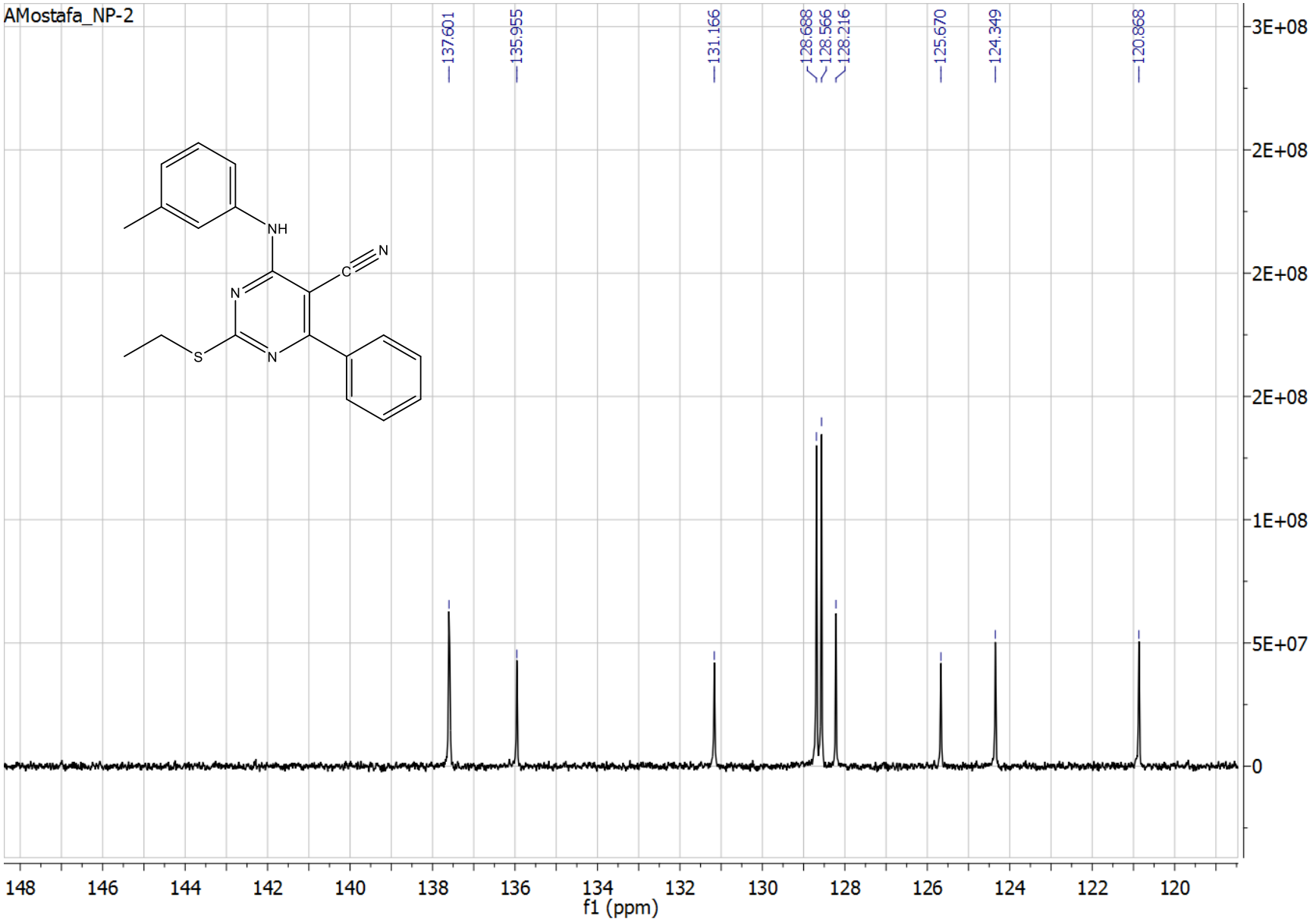
^{1}H NMR of compound 10_e



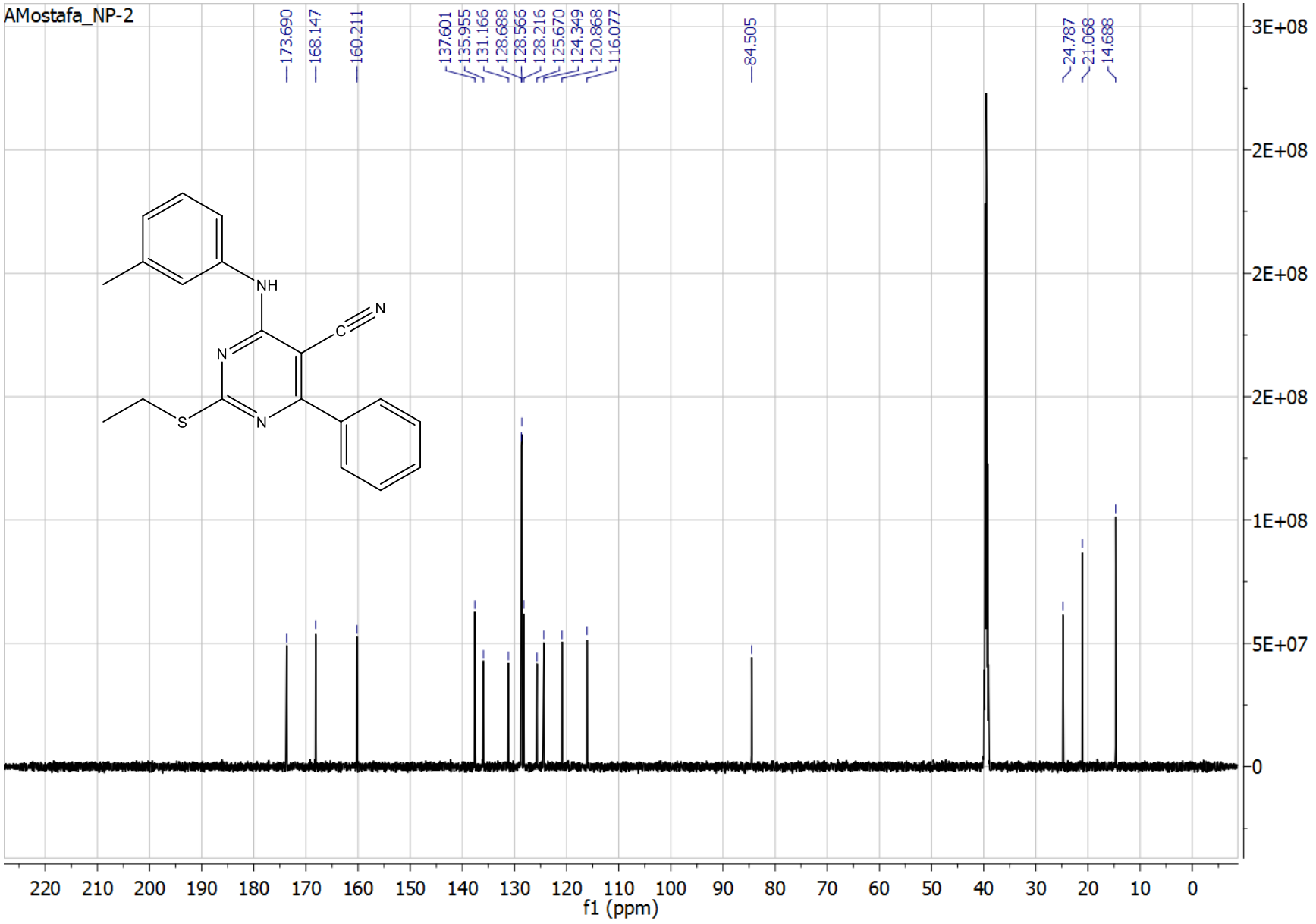
^{1}H NMR of compound 10_e



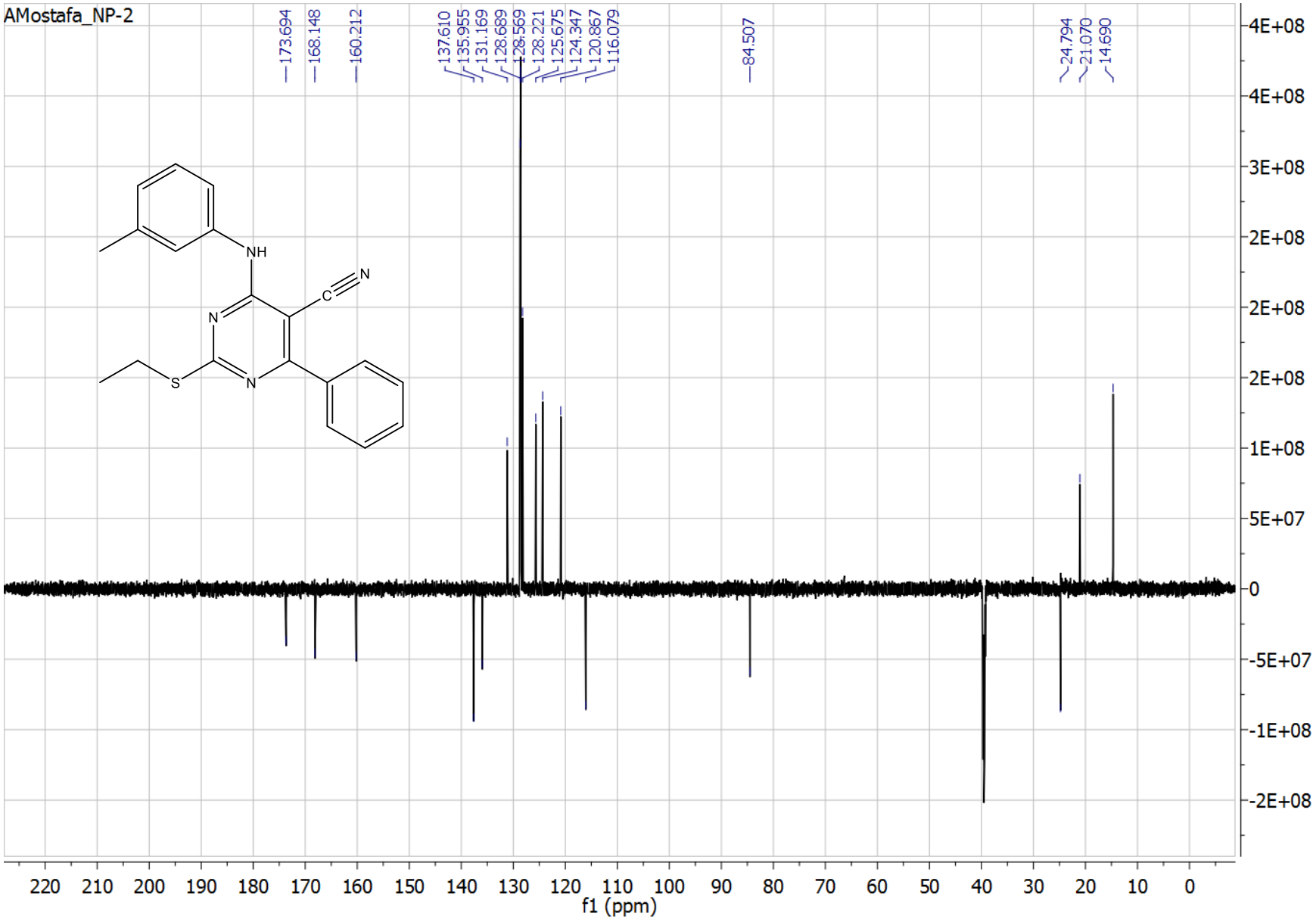
^{13}C NMR of compound 10_e



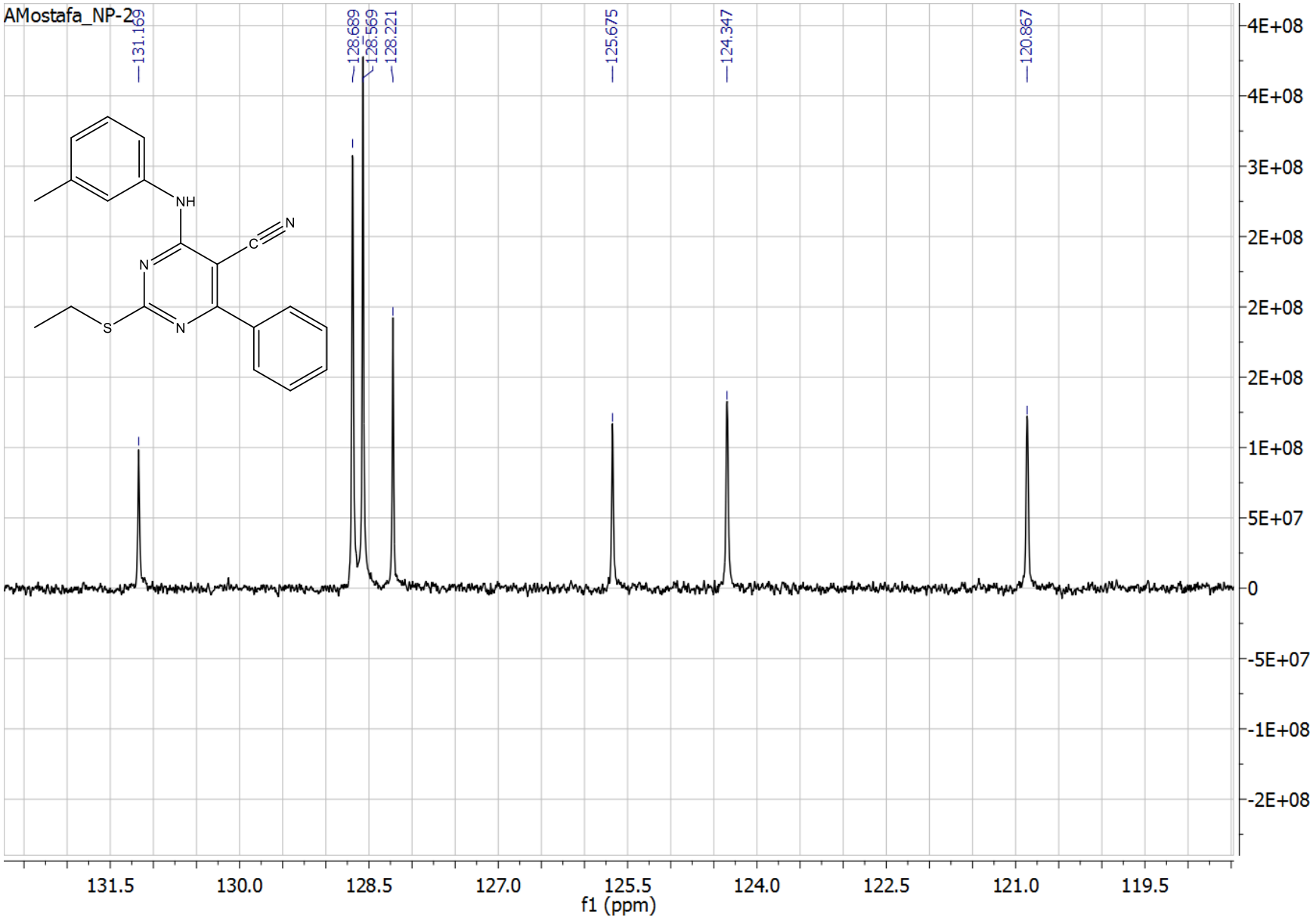
^{13}C NMR of compound 10_e



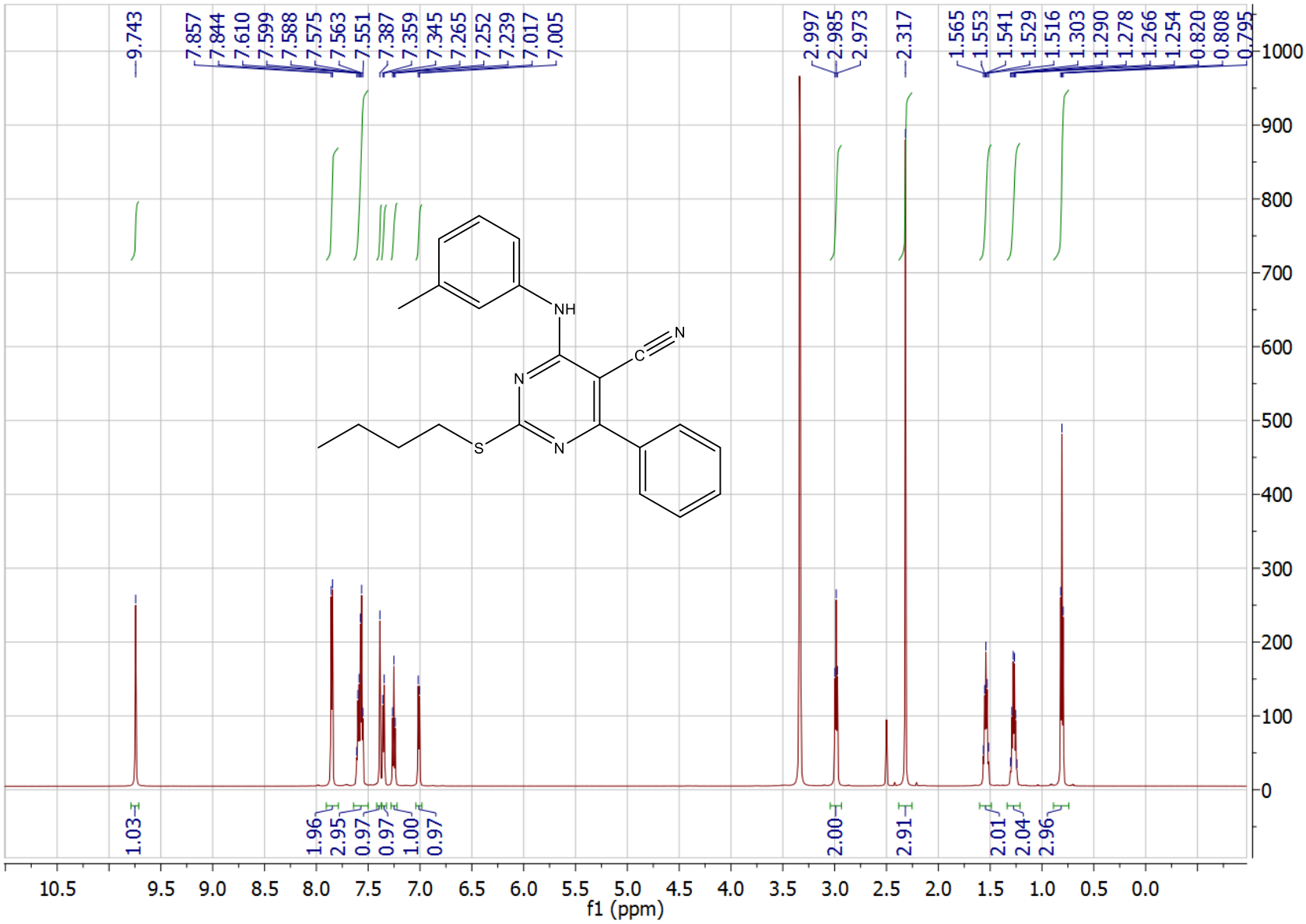
APT of compound 10_e



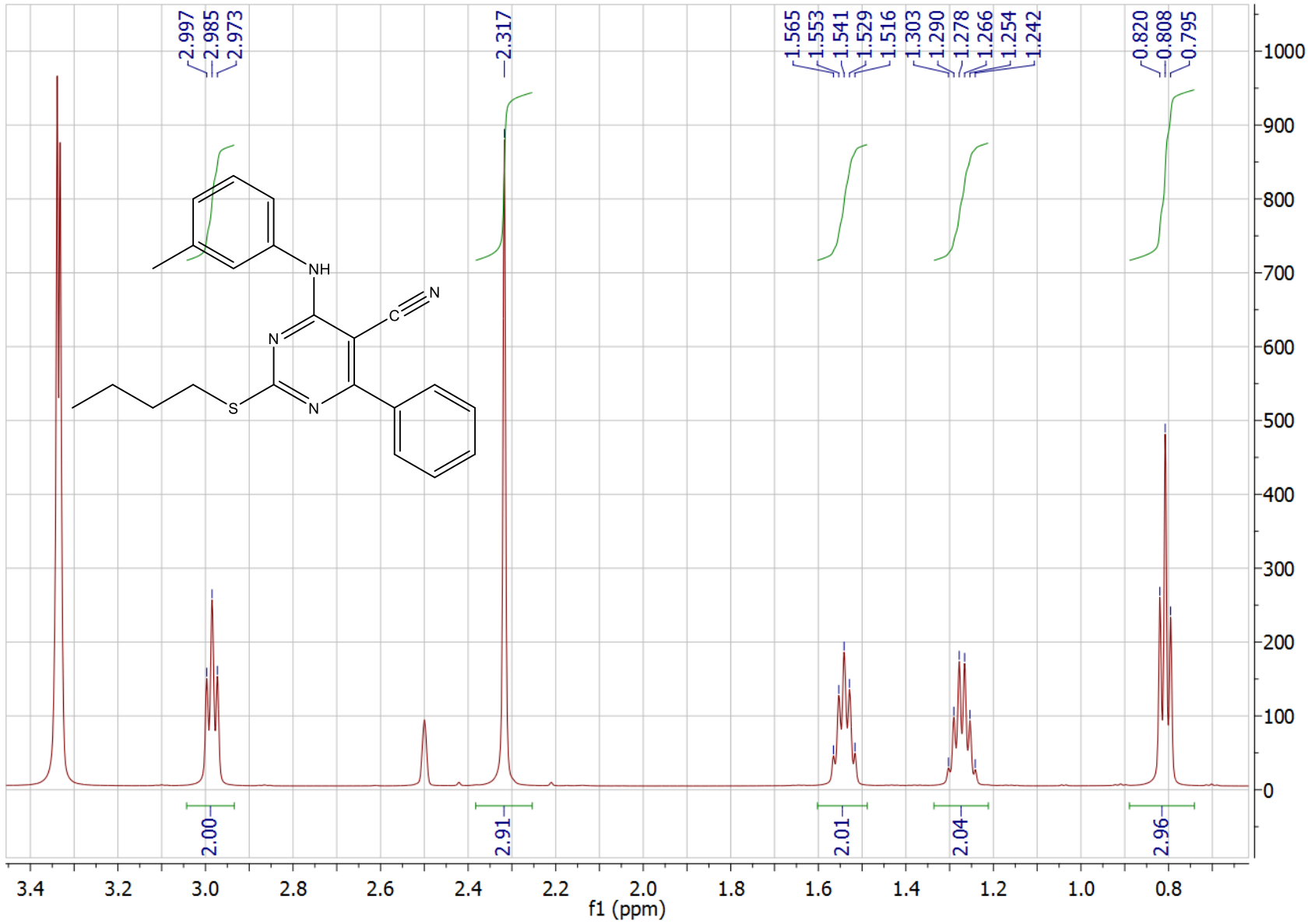
APT of compound 10_e



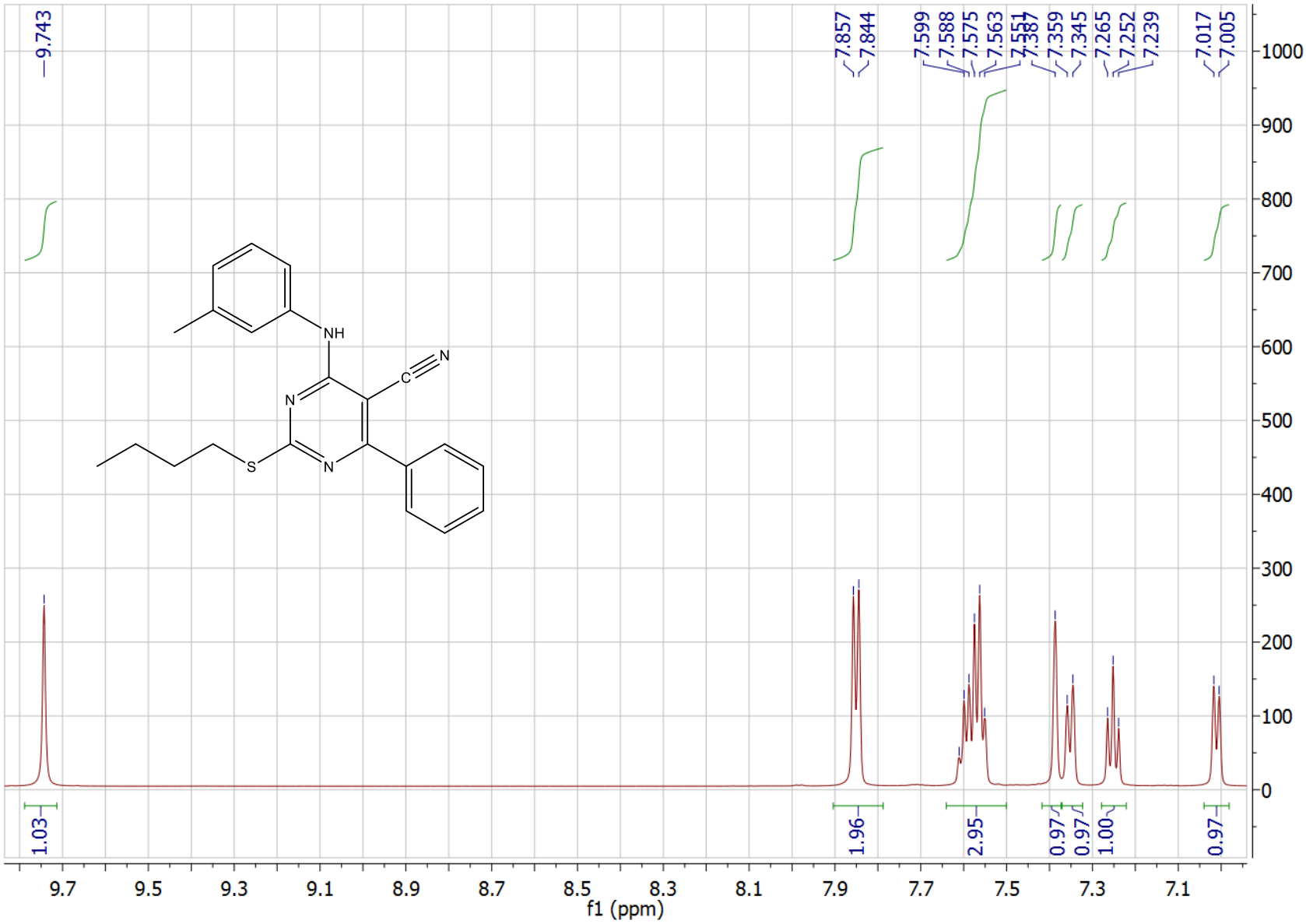
^{1}H NMR of compound 10_f



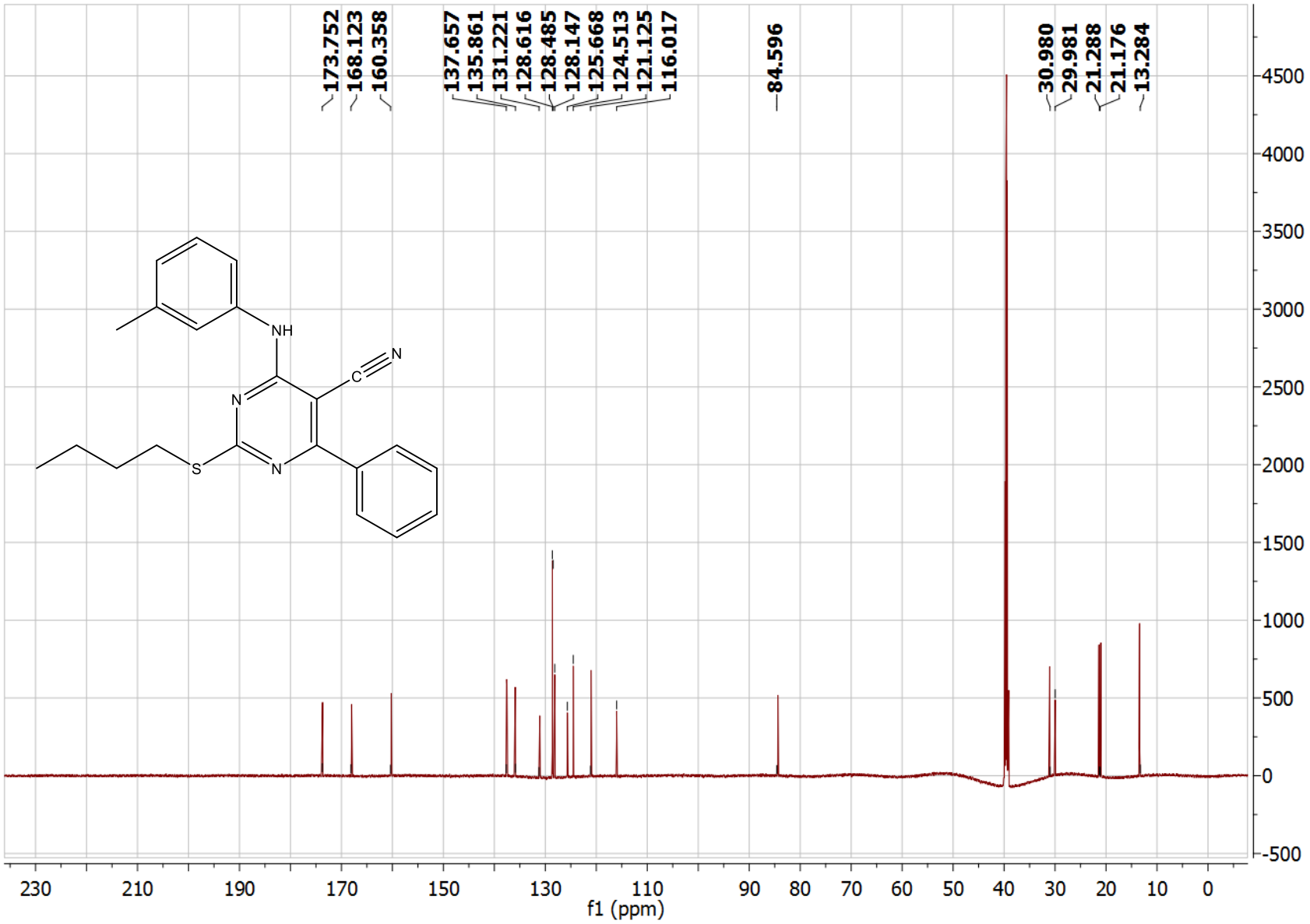
^{1}H NMR of compound 10_f



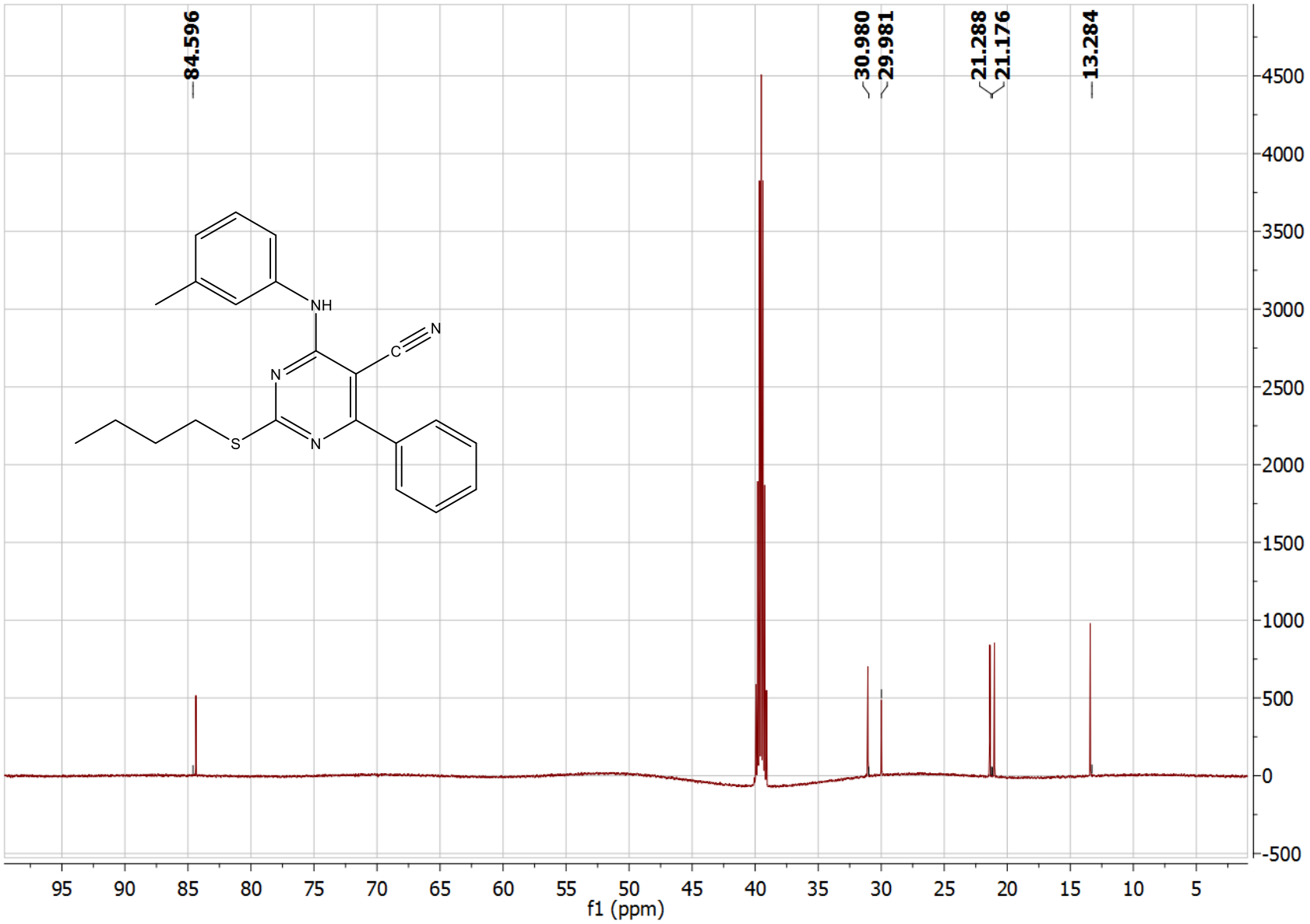
^{1}H NMR of compound 10_f



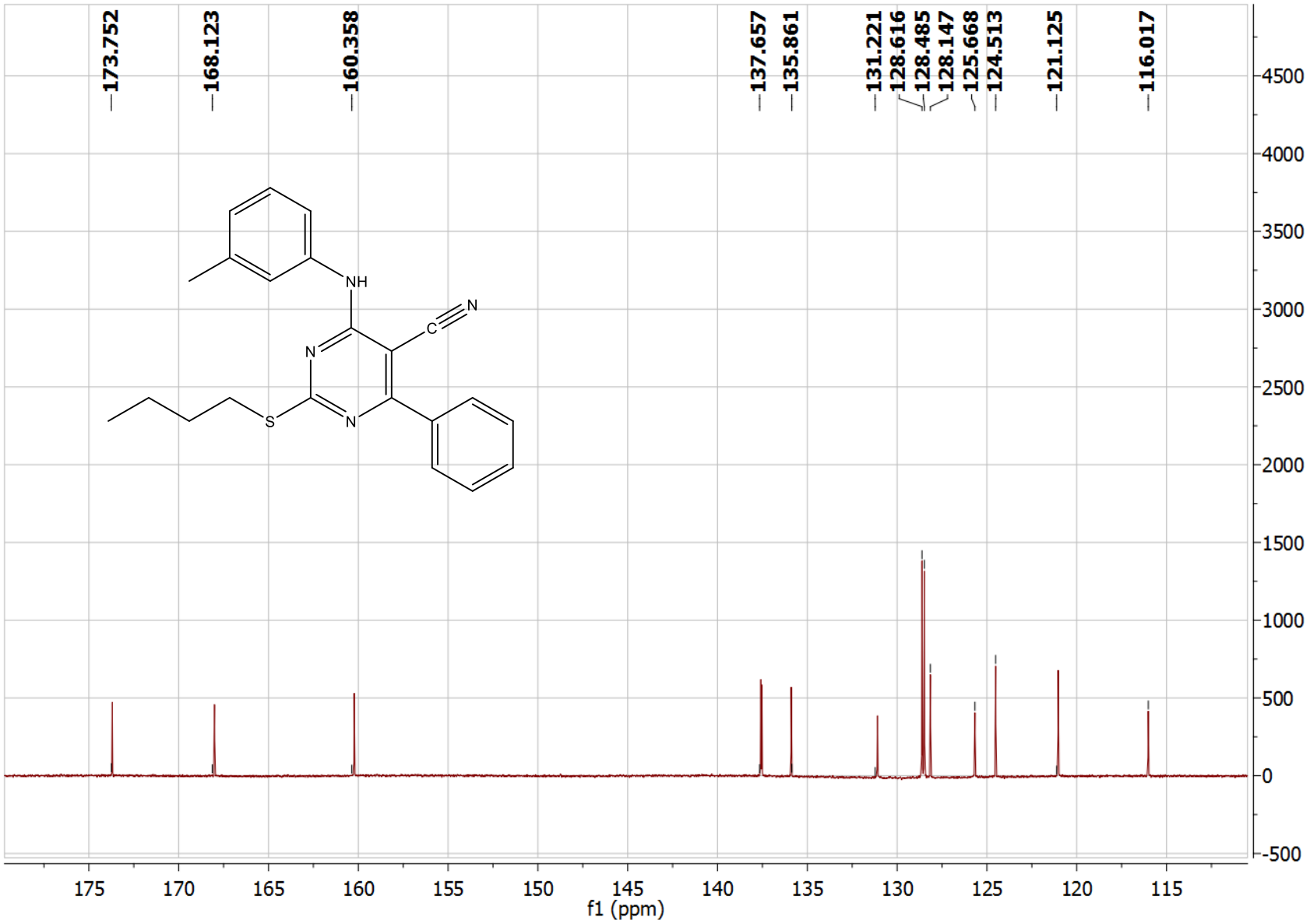
^{13}C NMR of compound 10_f



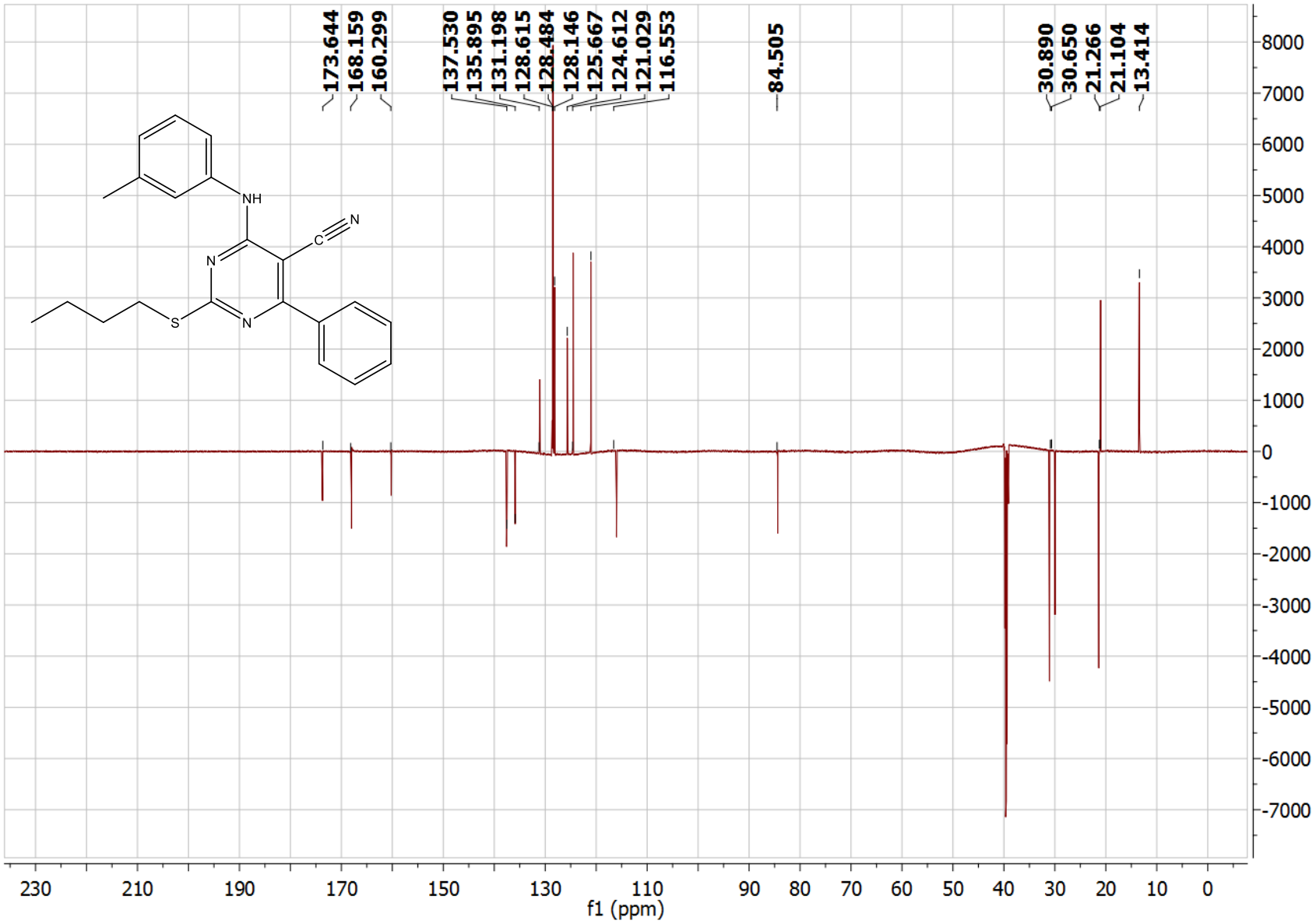
^{13}C NMR of compound 10_f



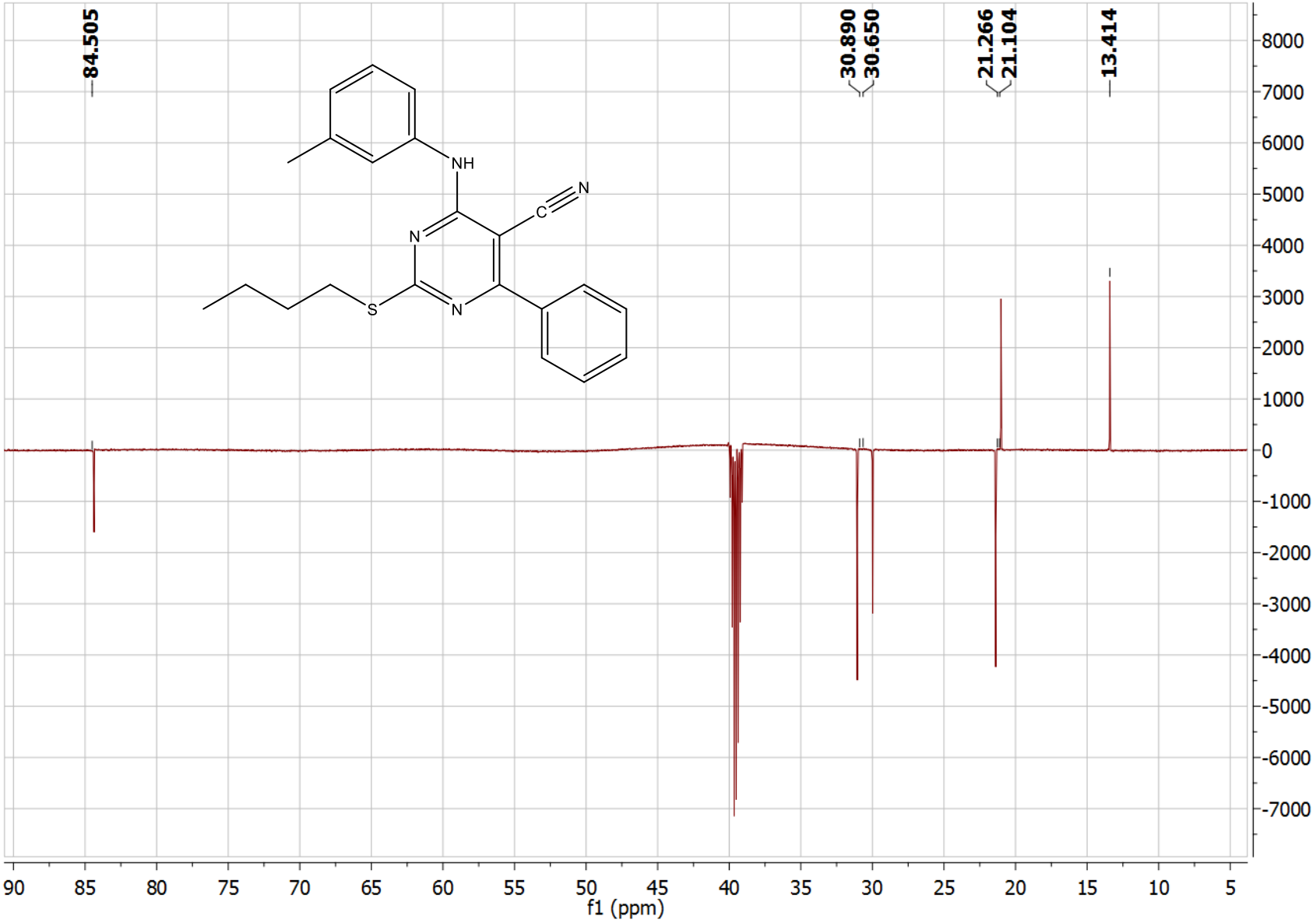
^{13}C NMR of compound 10_f



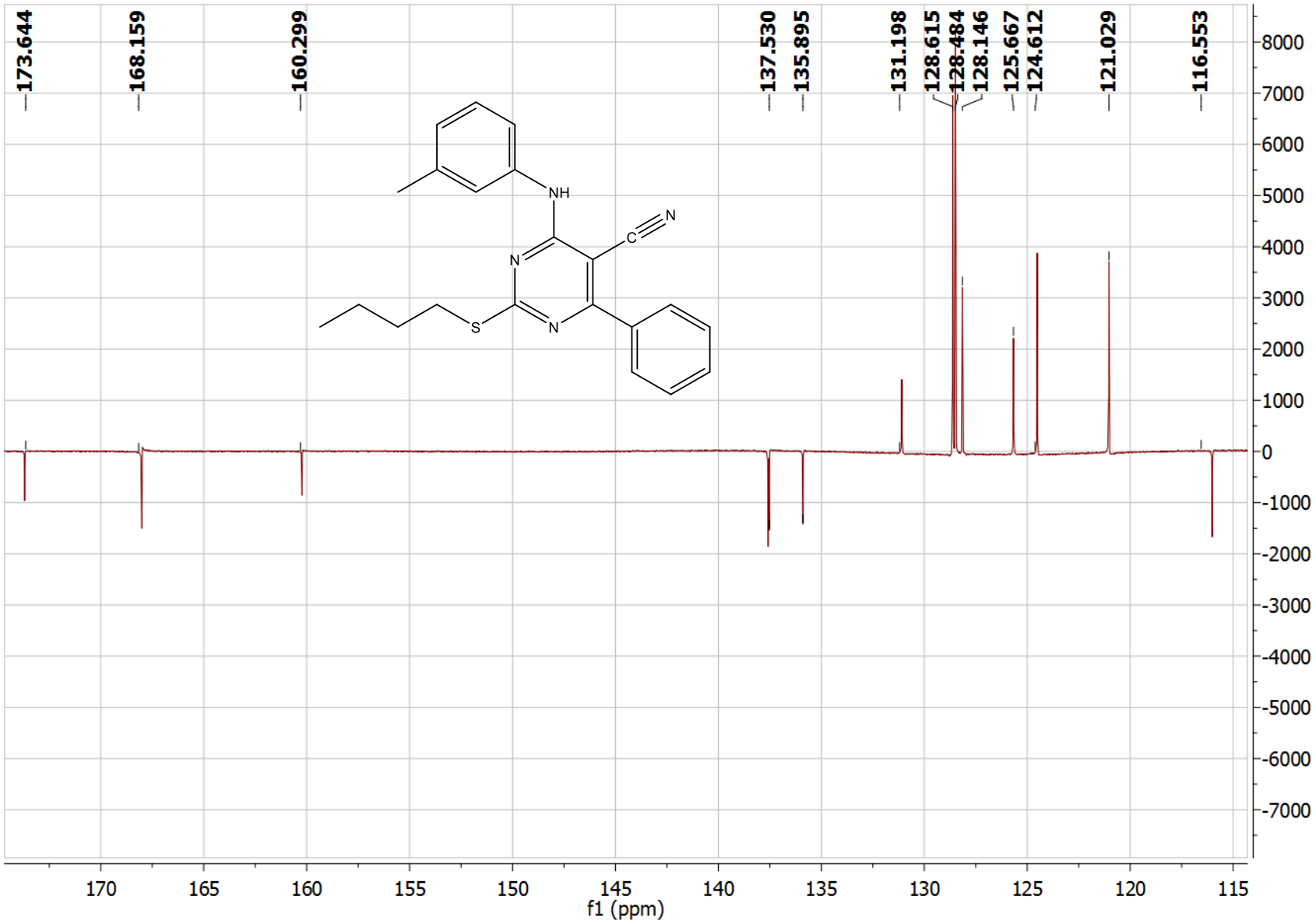
APT of compound 10_f



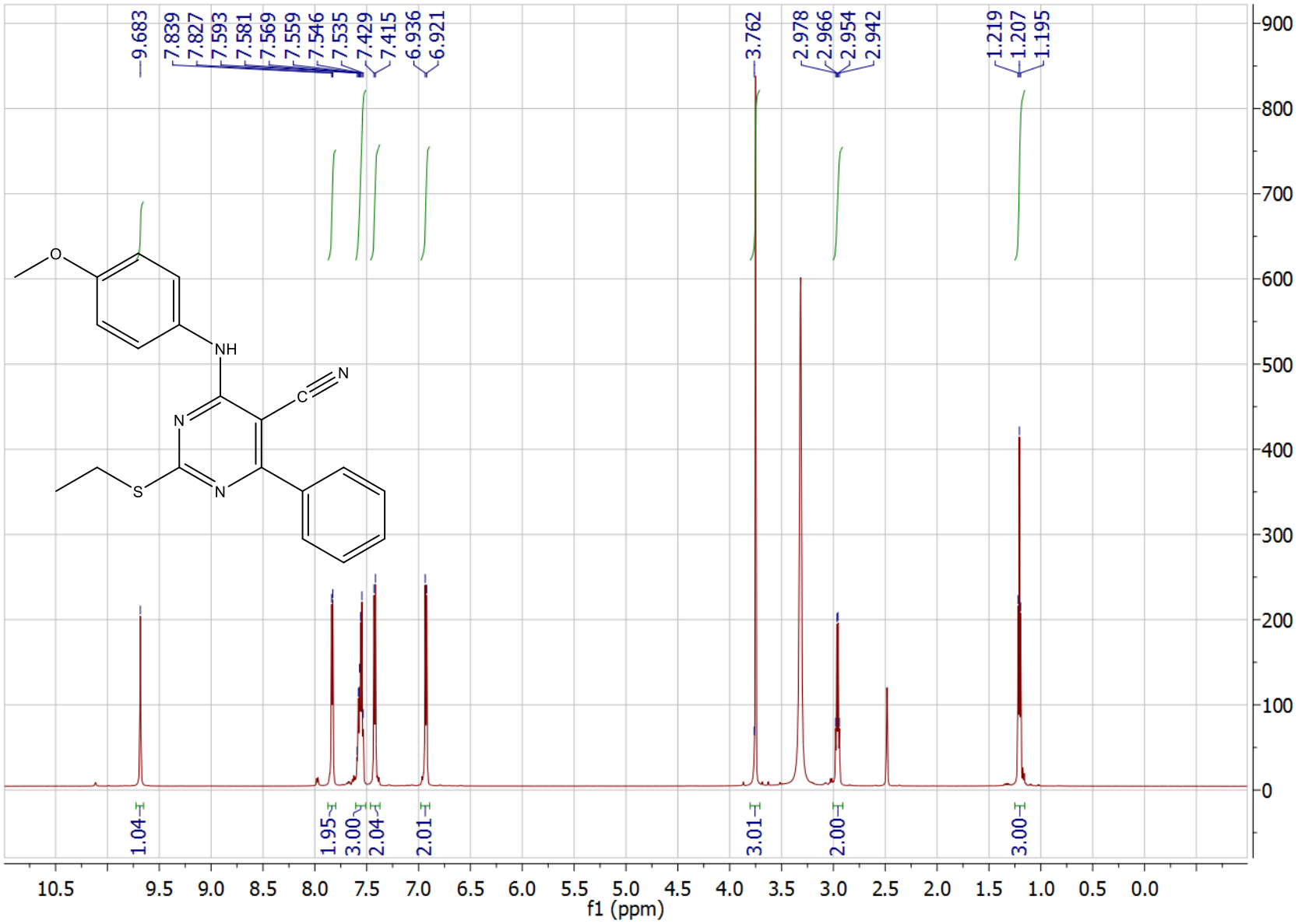
APT of compound 10_f



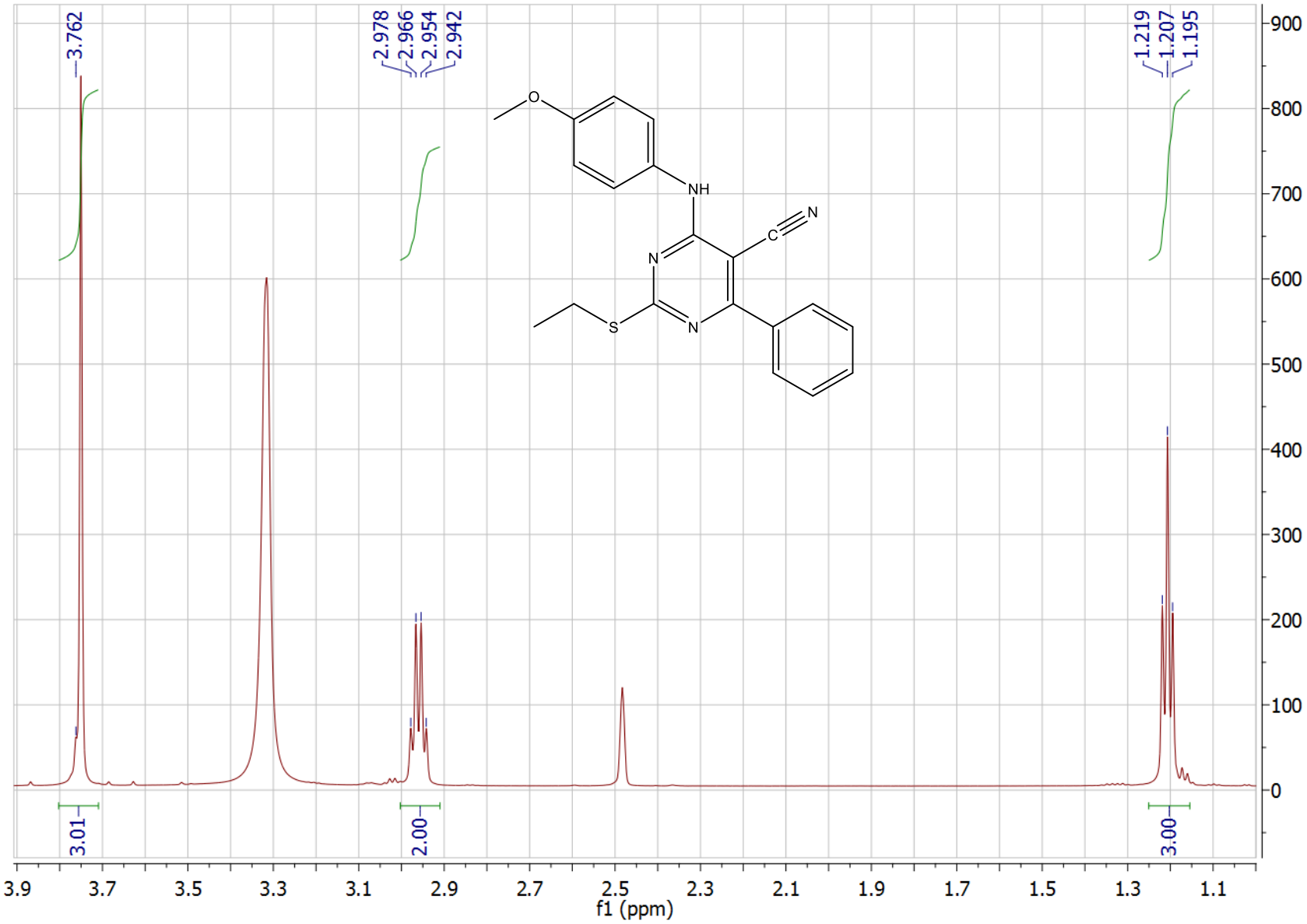
APT of compound 10_f



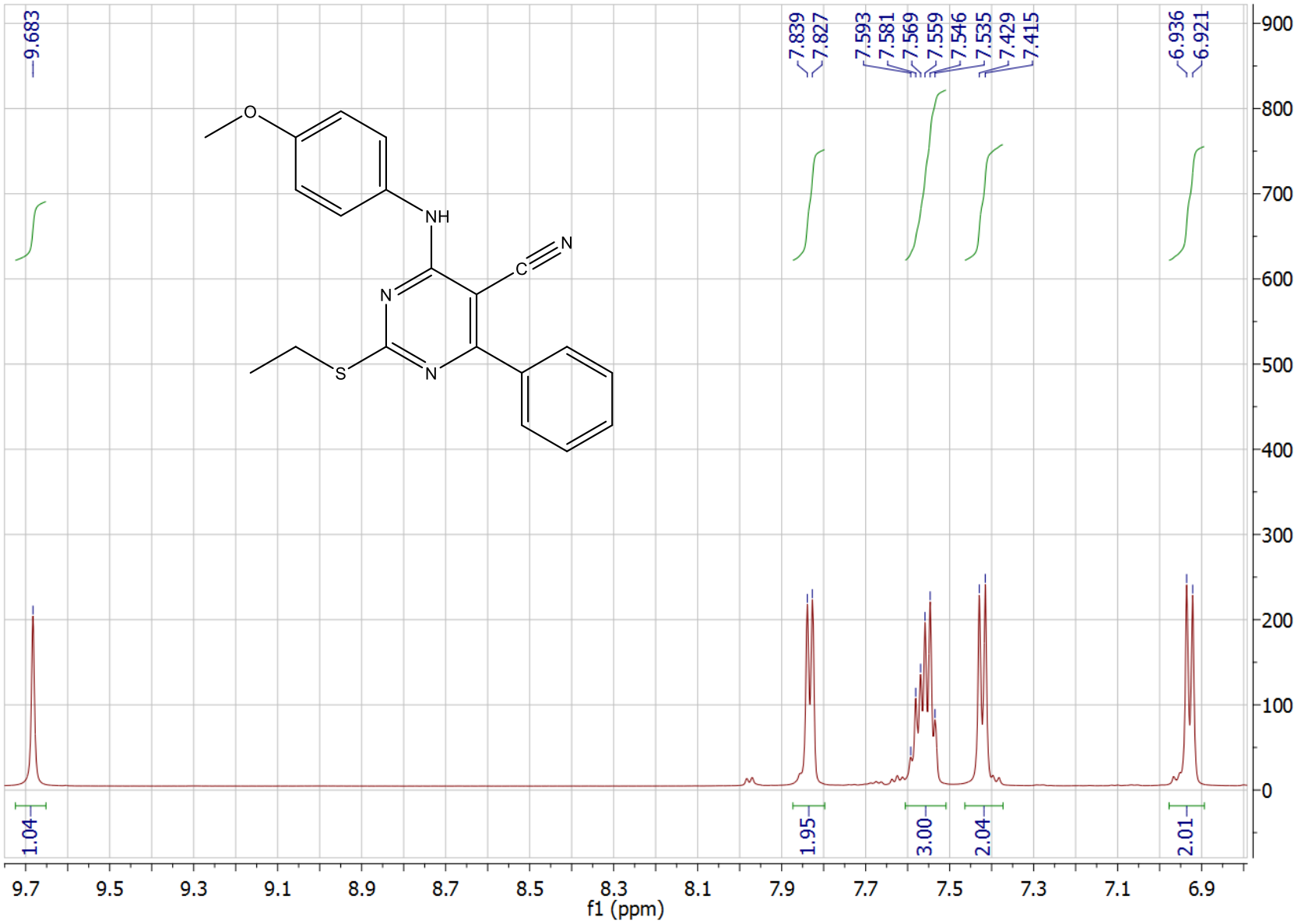
^{1}H NMR of compound 10_g



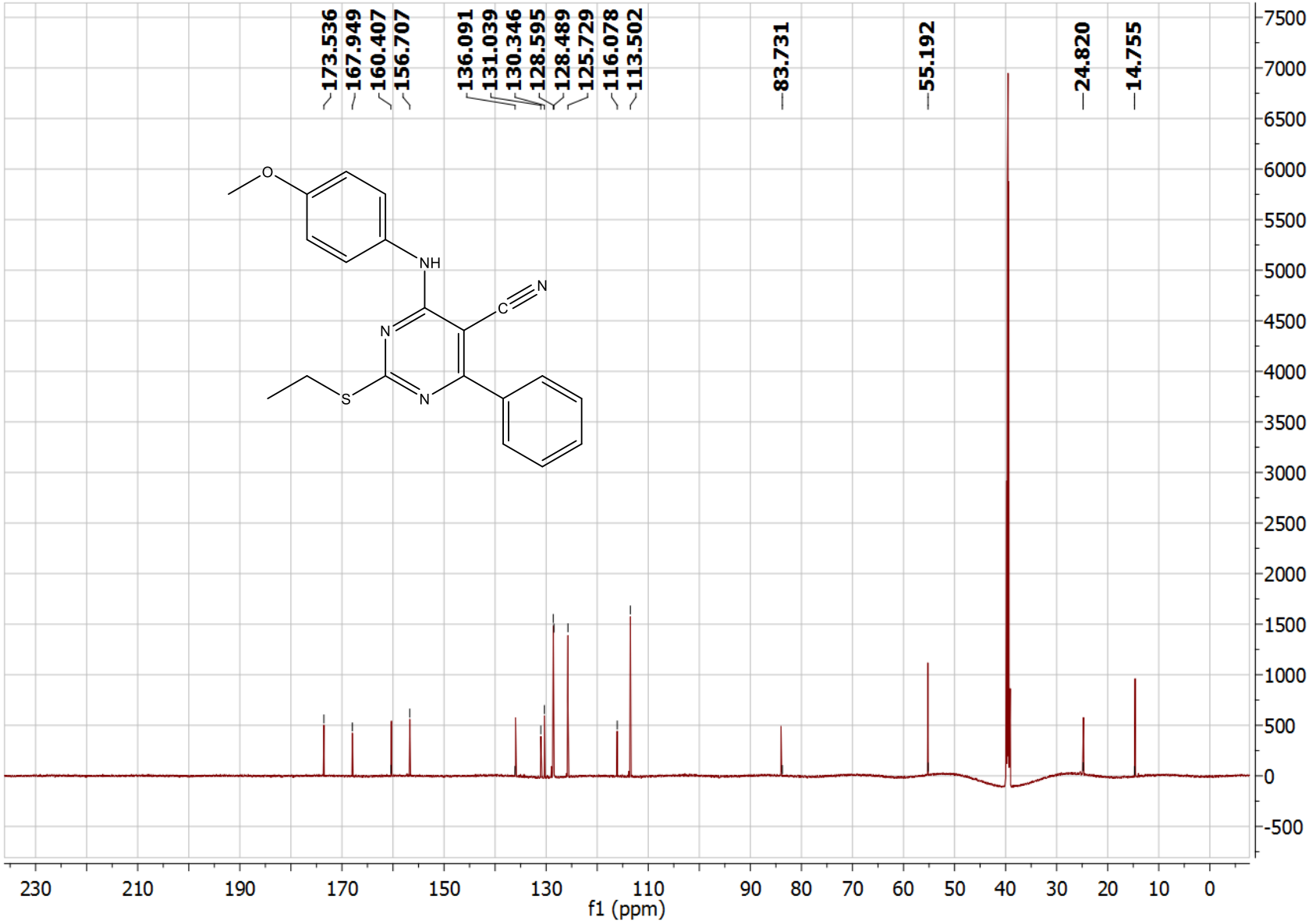
^{1}H NMR of compound 10_g



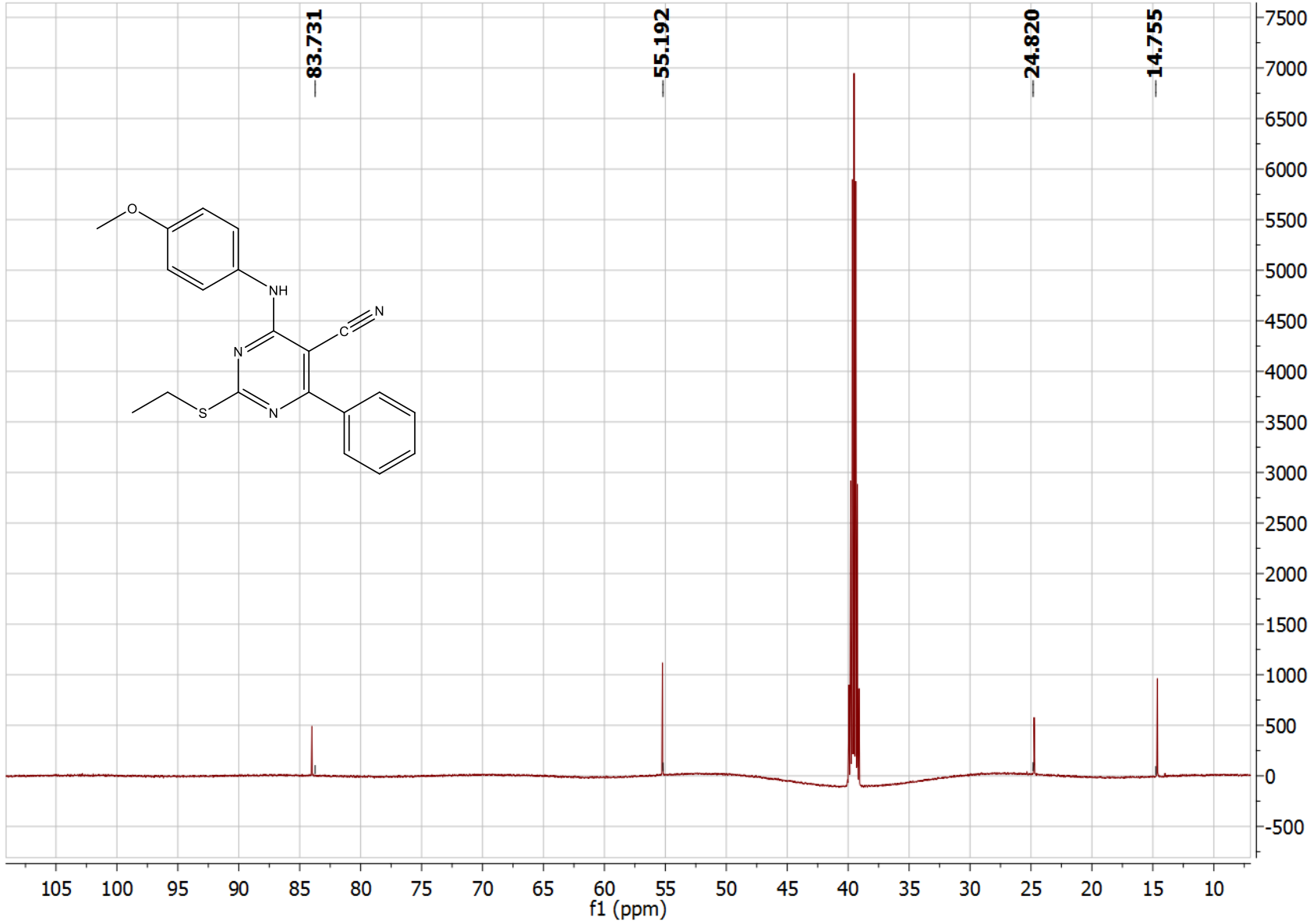
^{1}H NMR of compound 10_g



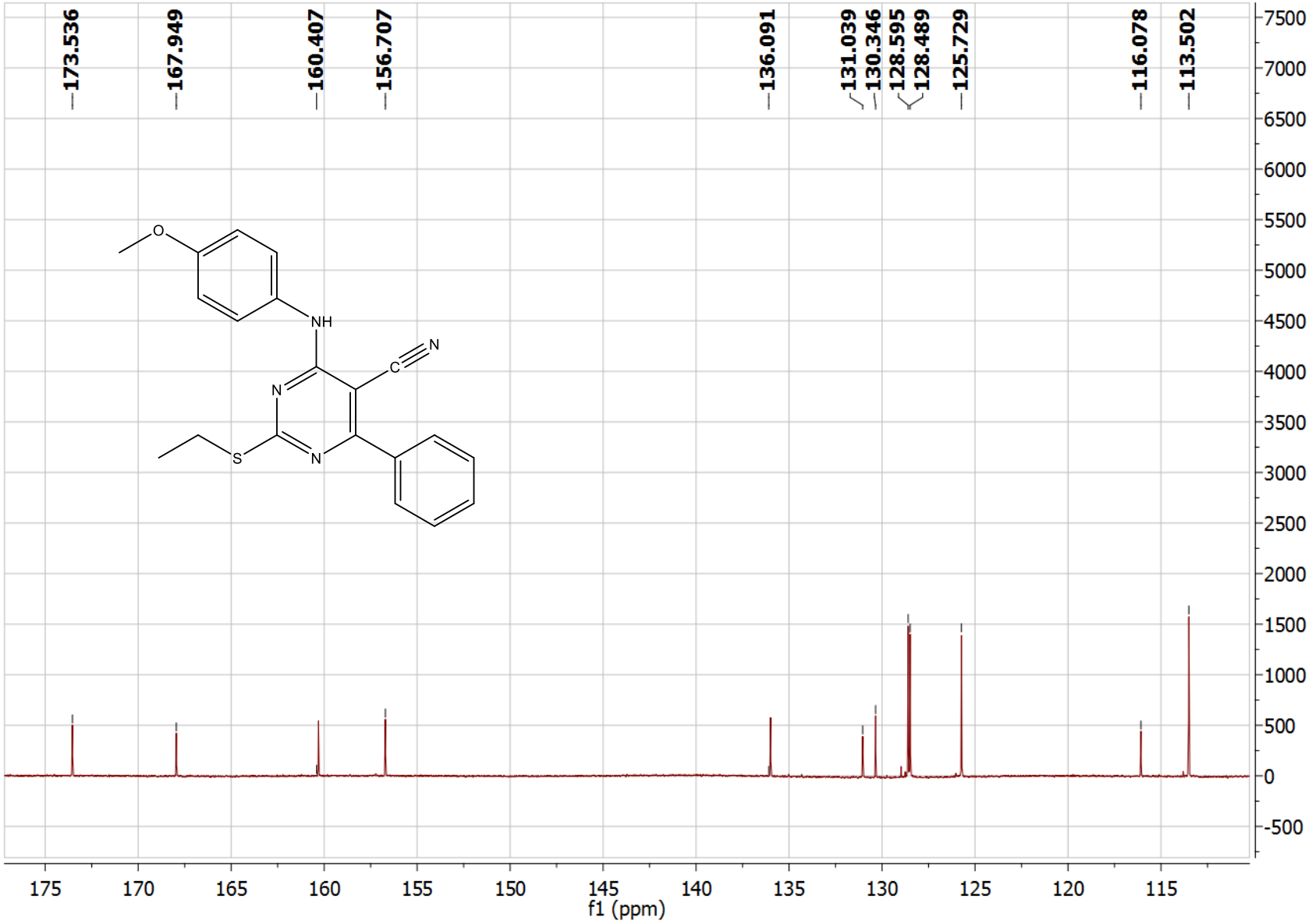
^{13}C NMR of compound 10_g



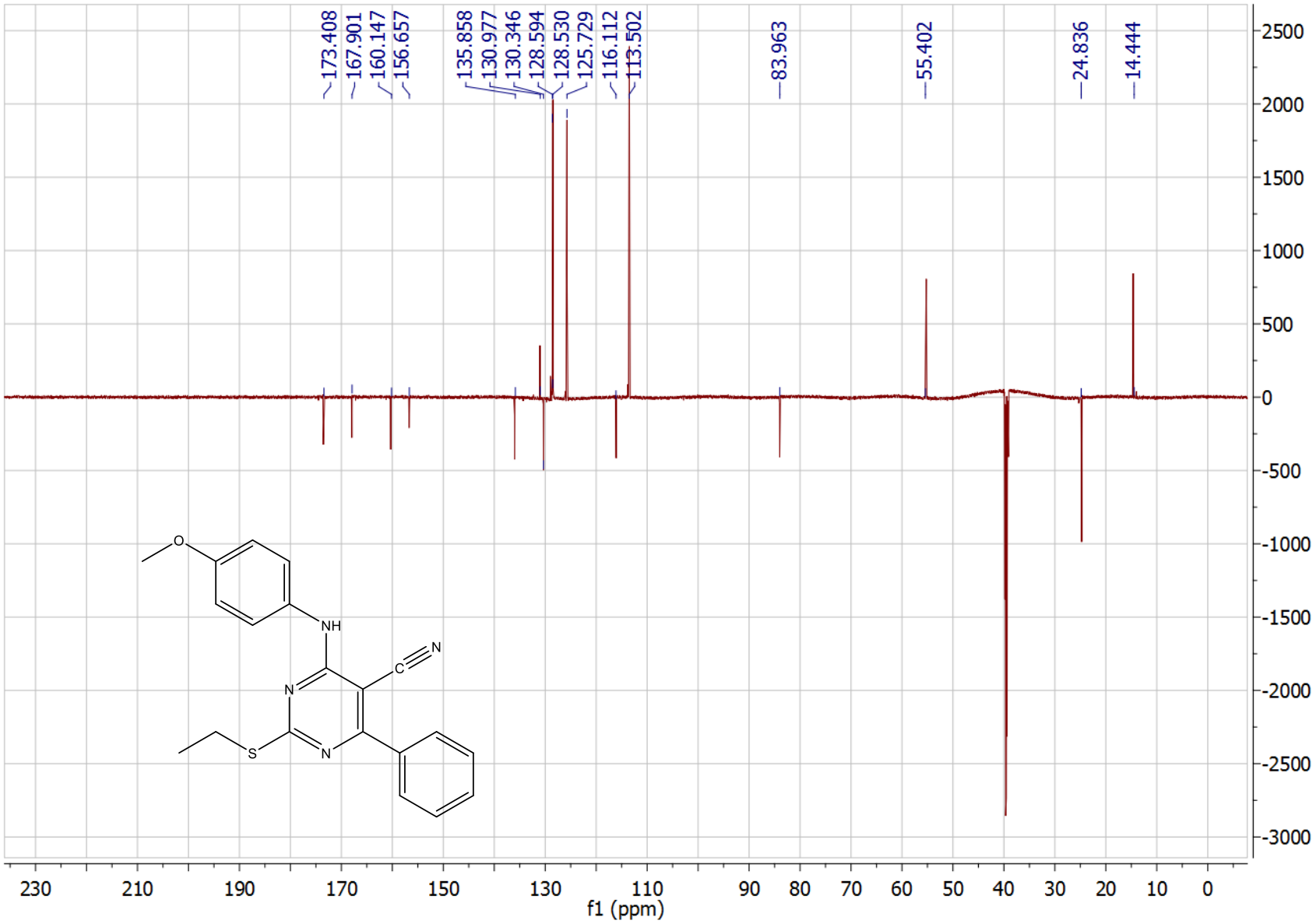
^{13}C NMR of compound 10_g



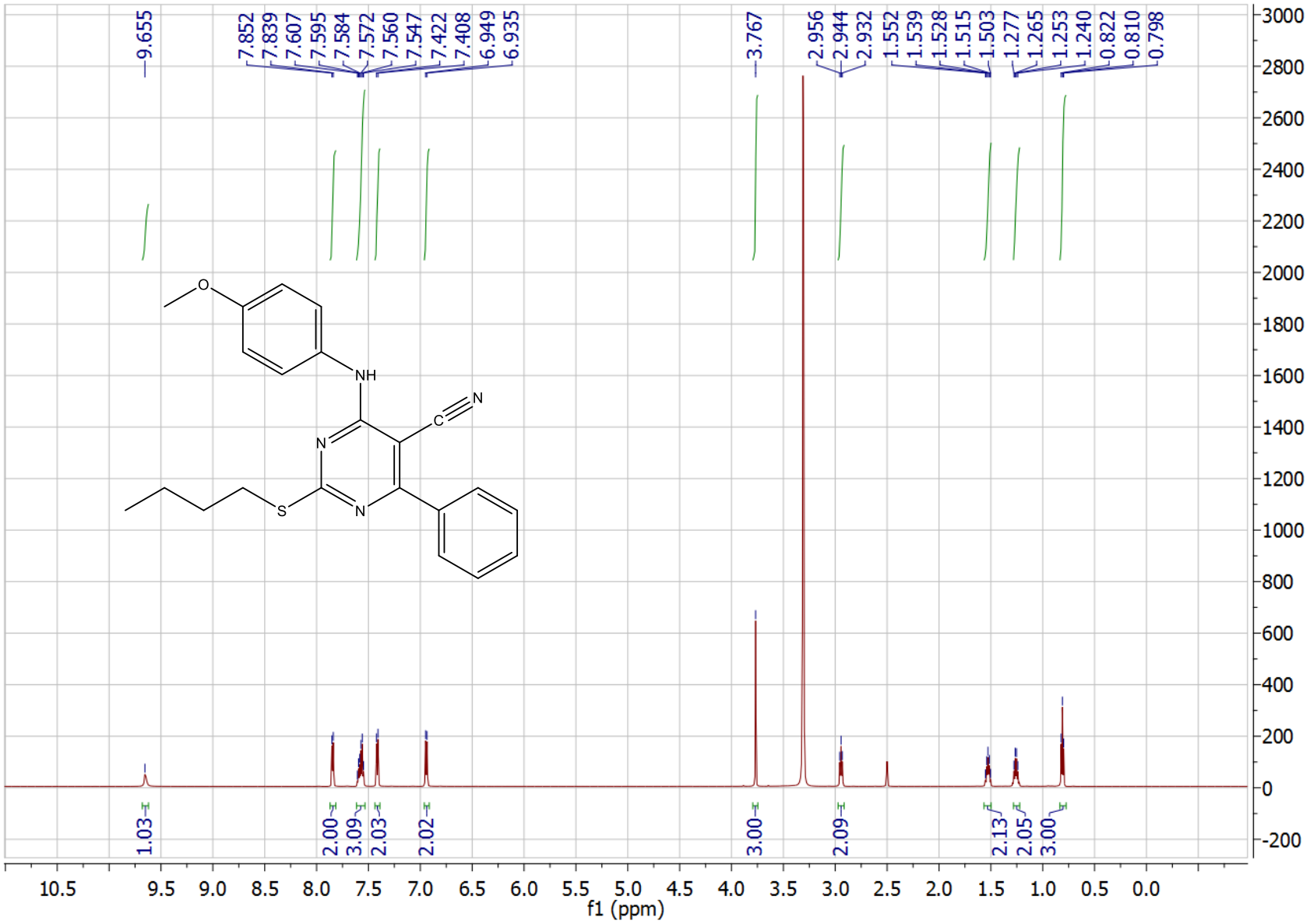
^{13}C NMR of compound 10_g



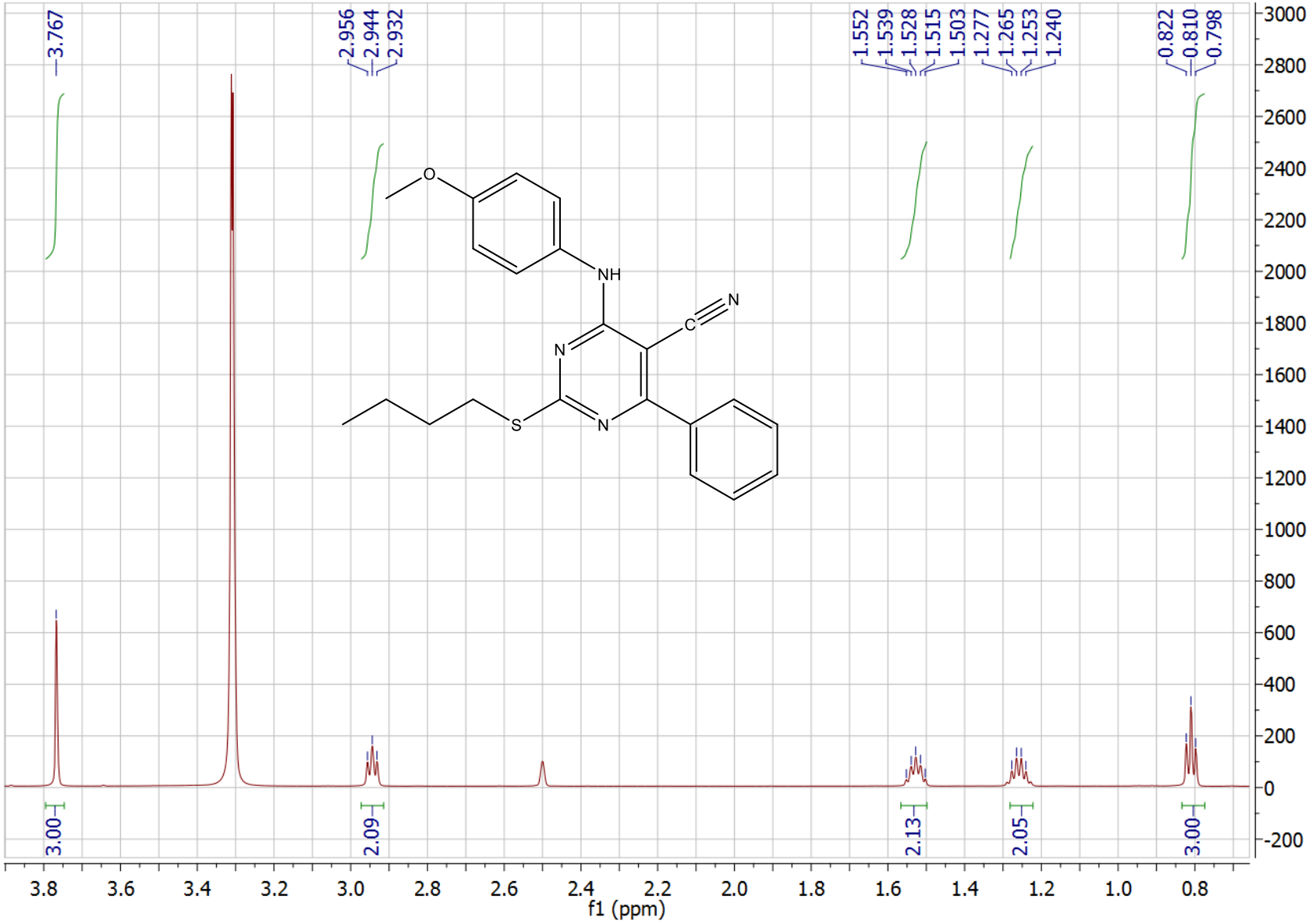
APT of compound 10_g



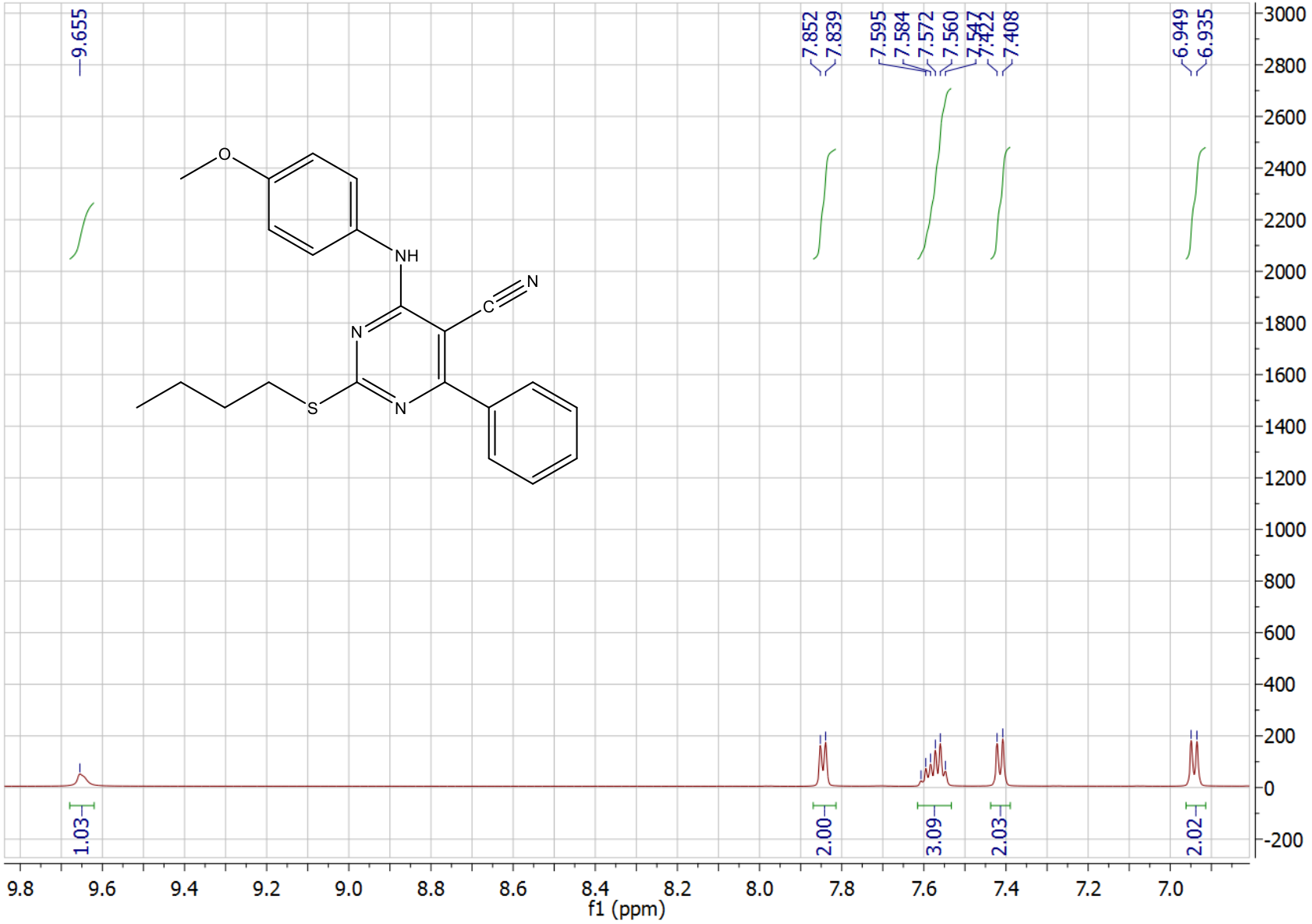
^{1}H NMR of compound 10_h



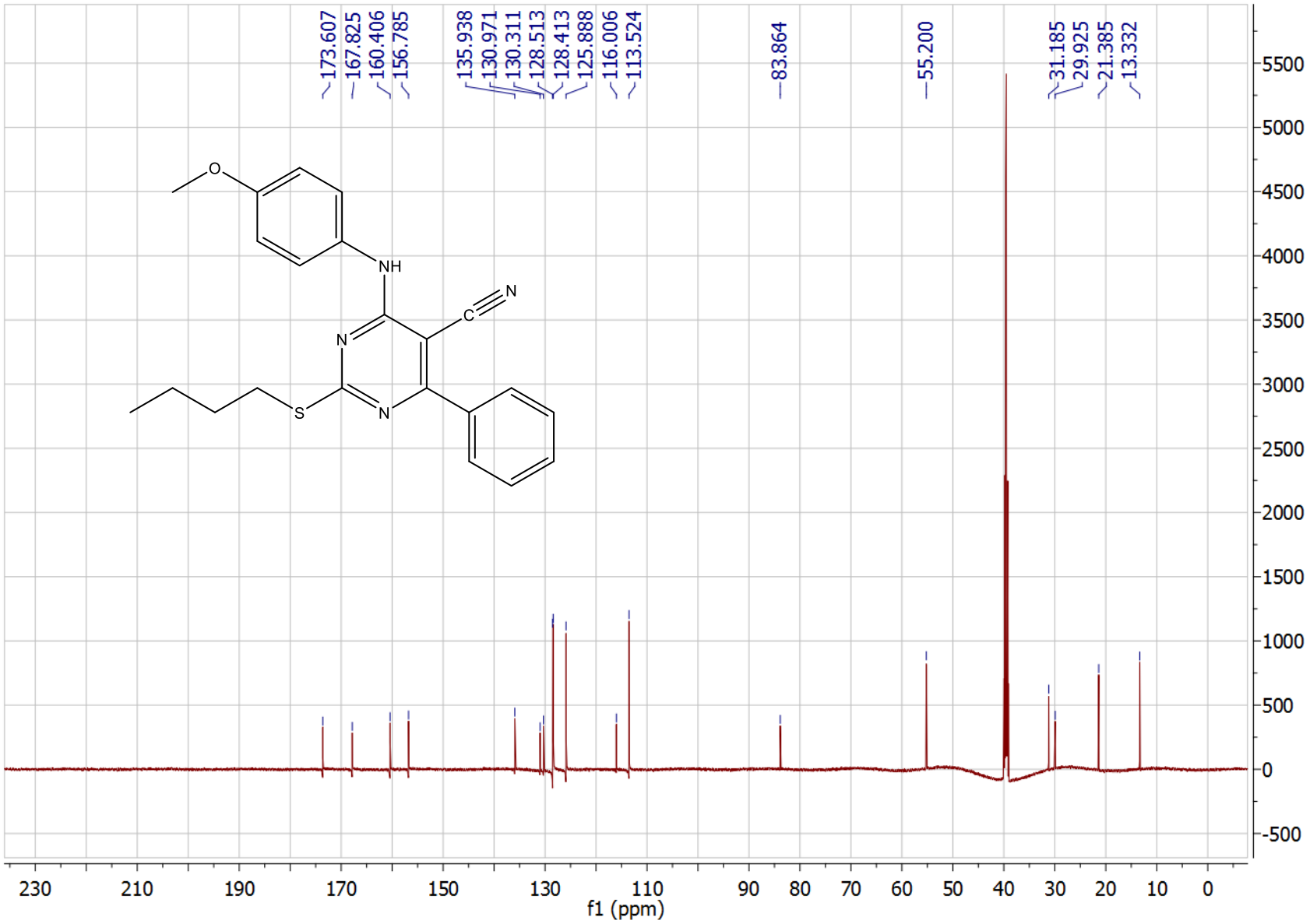
^{1}H NMR of compound 10_h



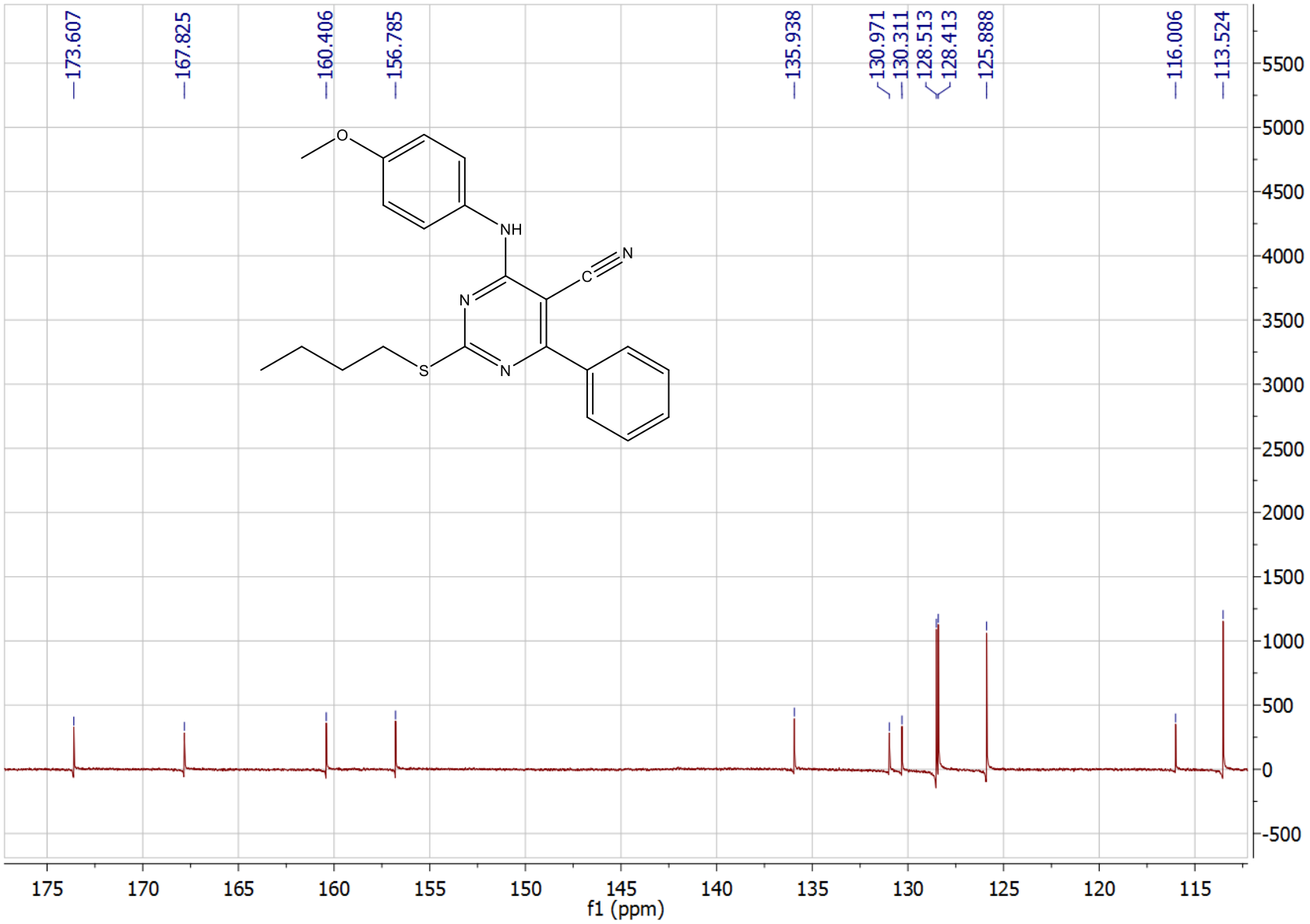
^{1}H NMR of compound 10_h



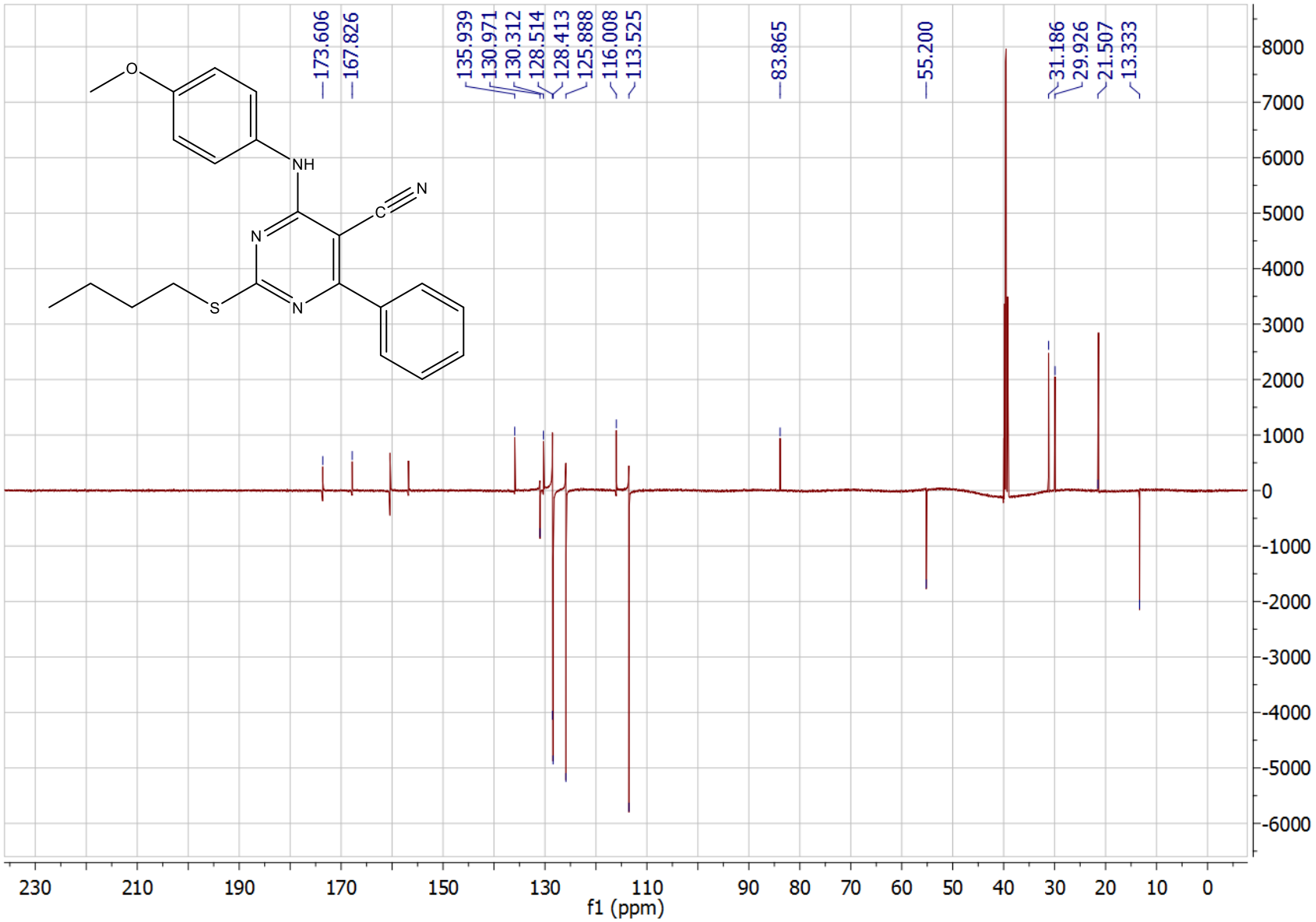
^{13}C NMR of compound 10_h



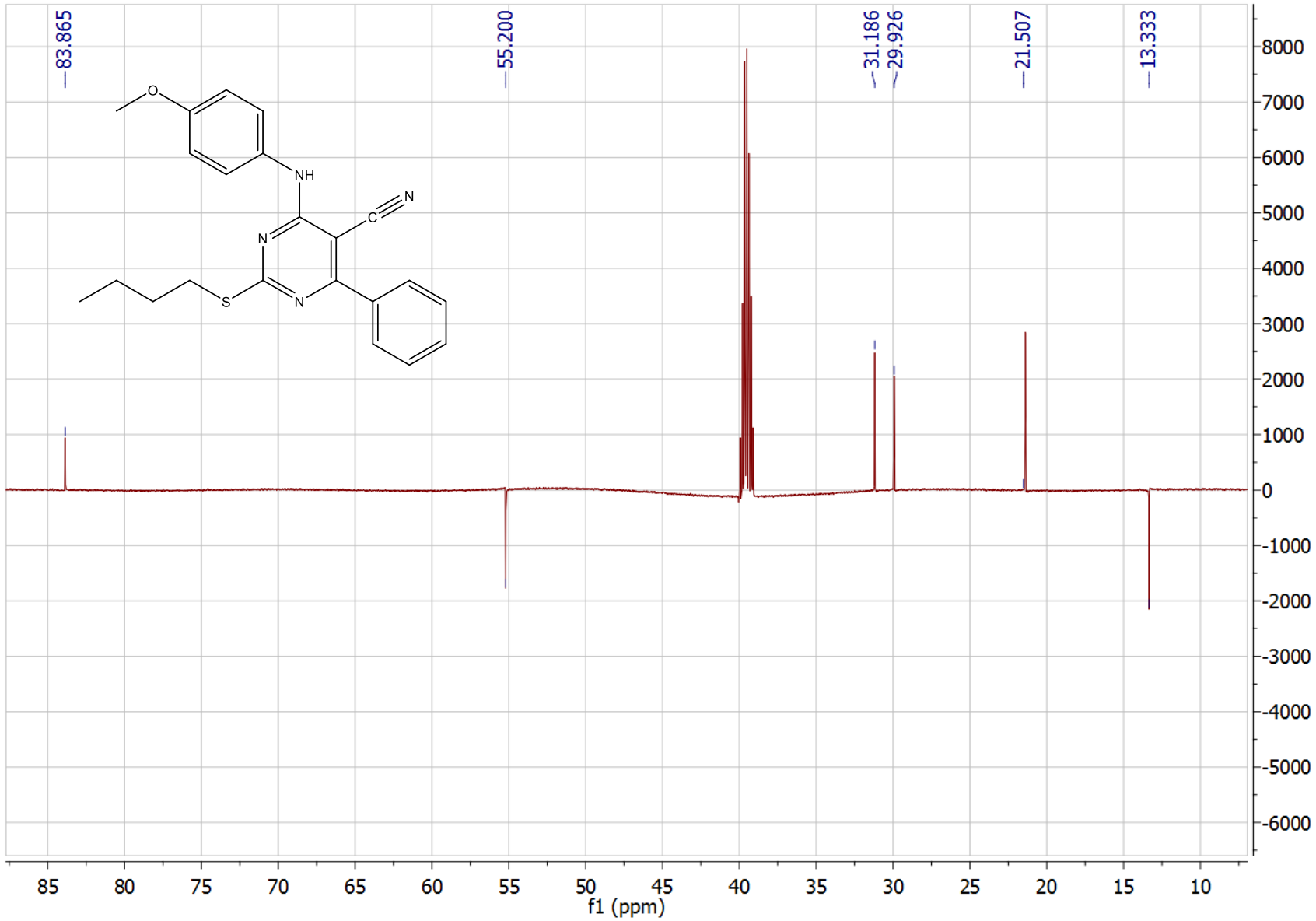
^{13}C NMR of compound 10_h



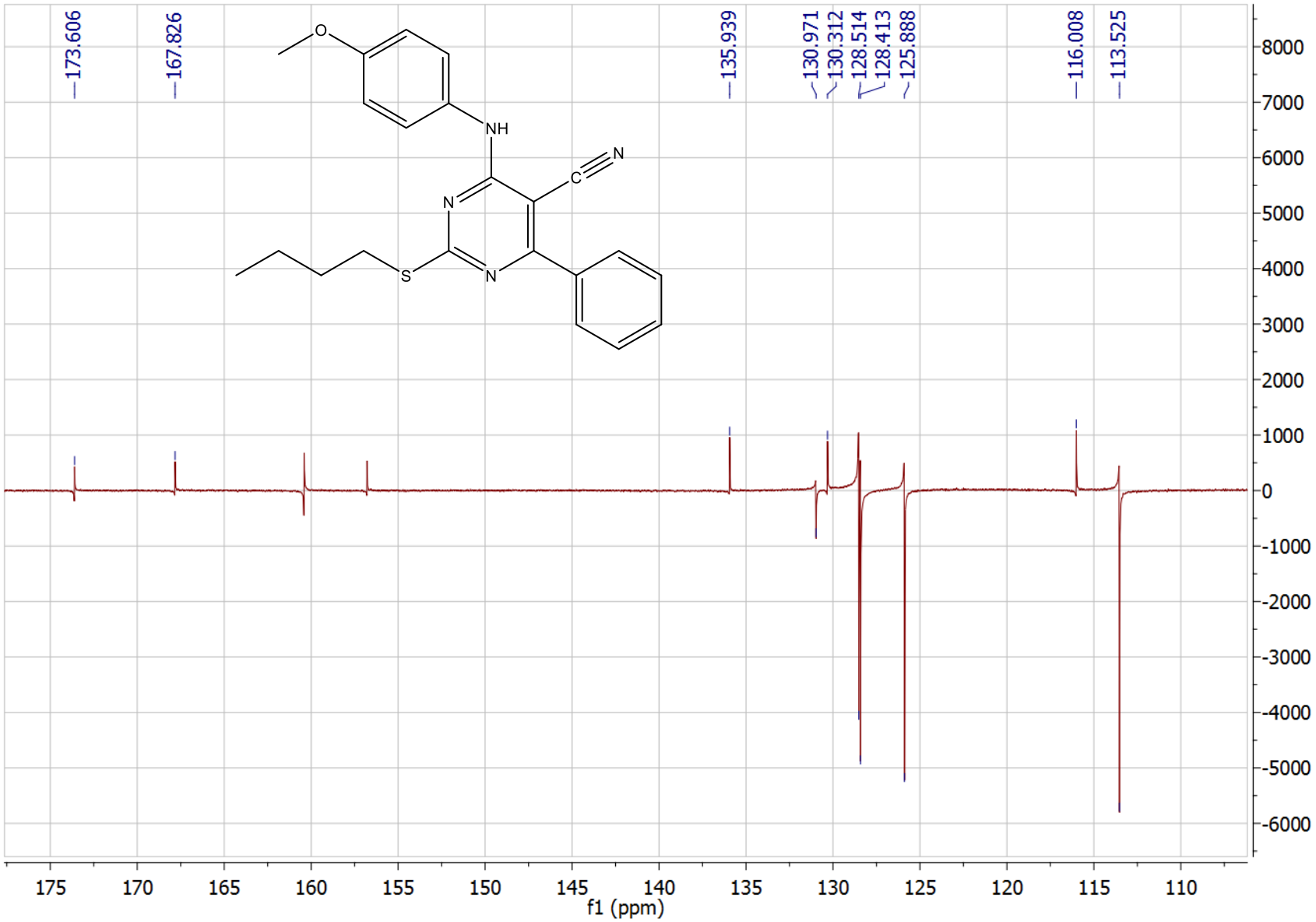
APT of compound 10_h



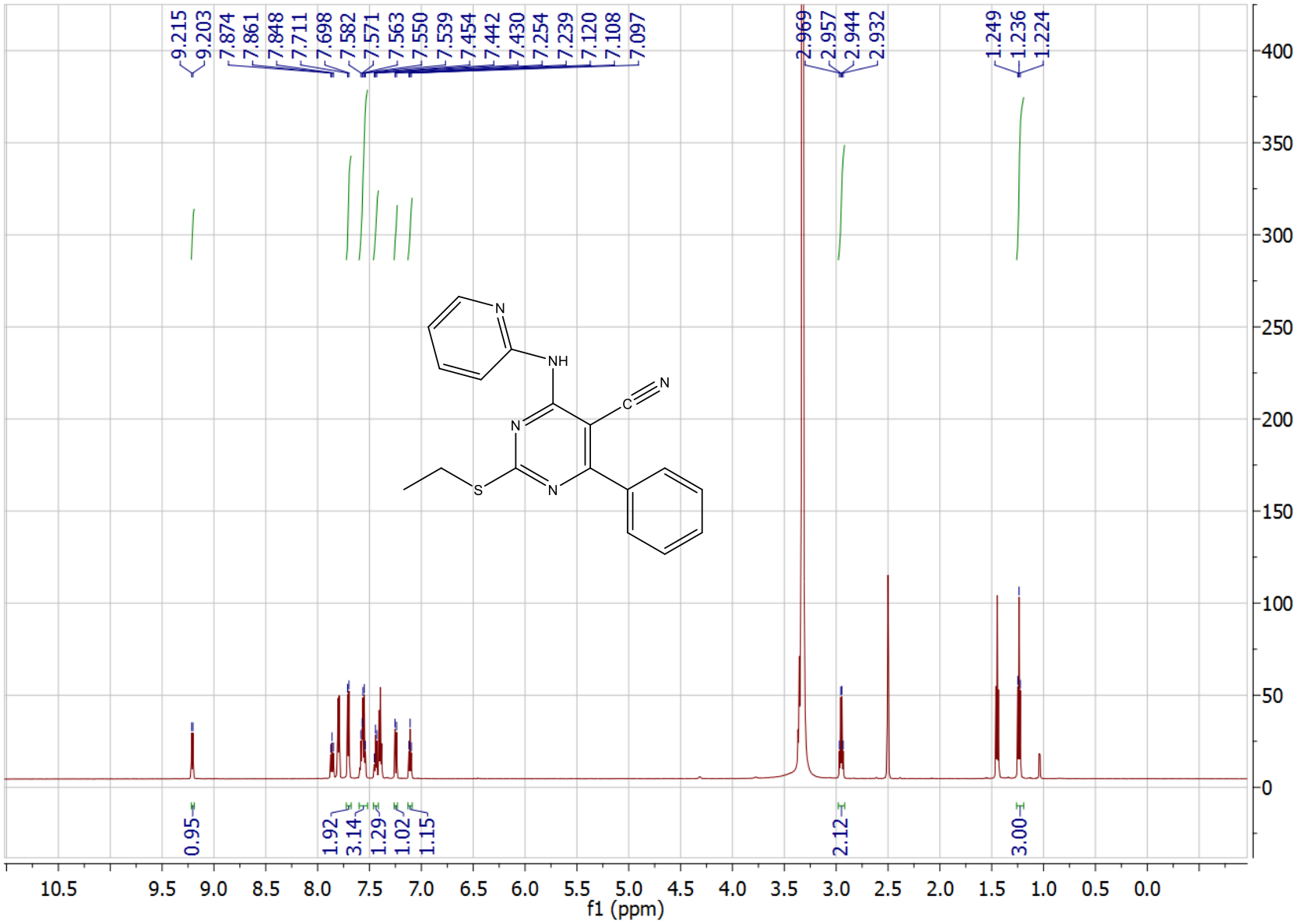
APT of compound 10_h



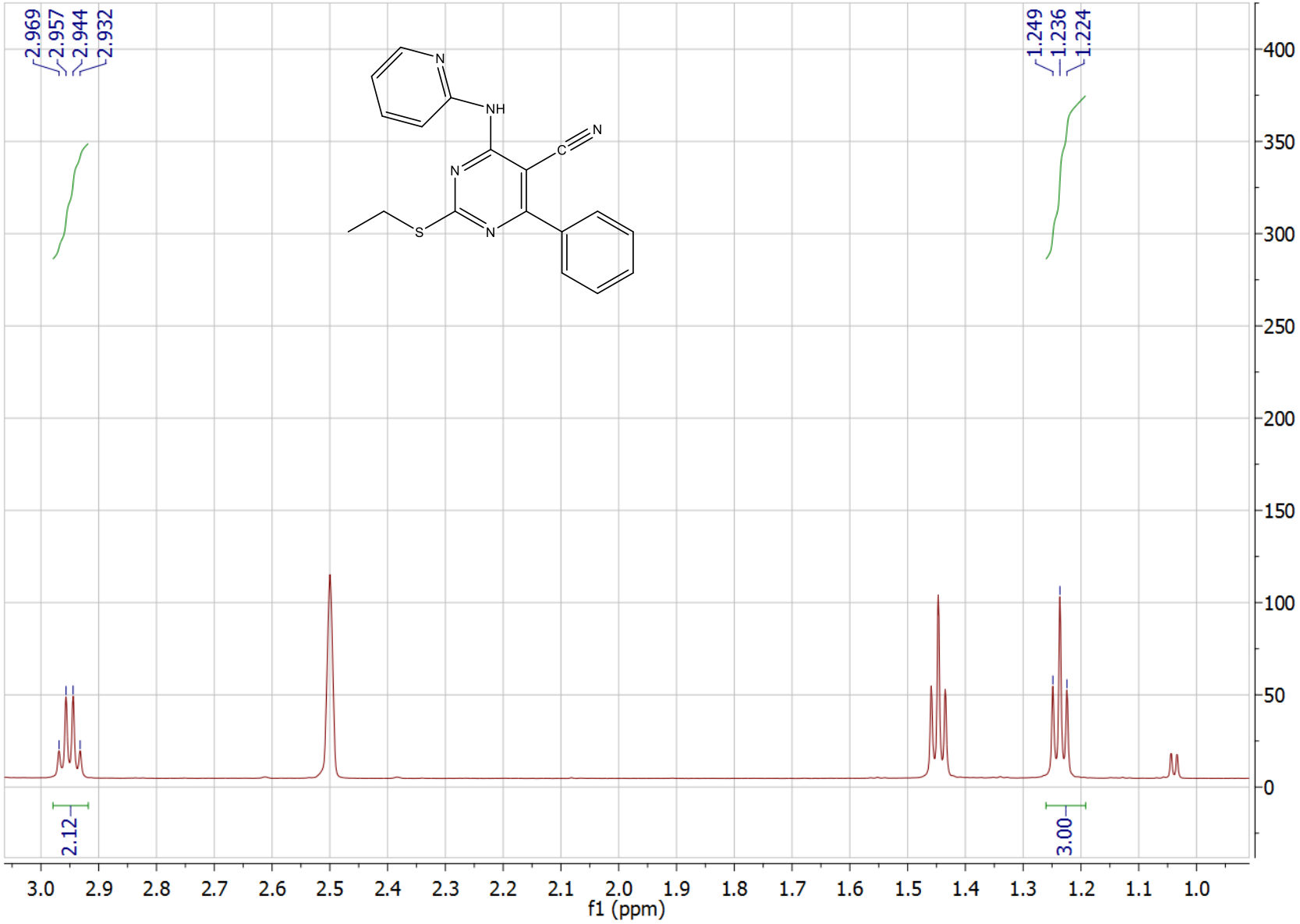
APT of compound 10_h



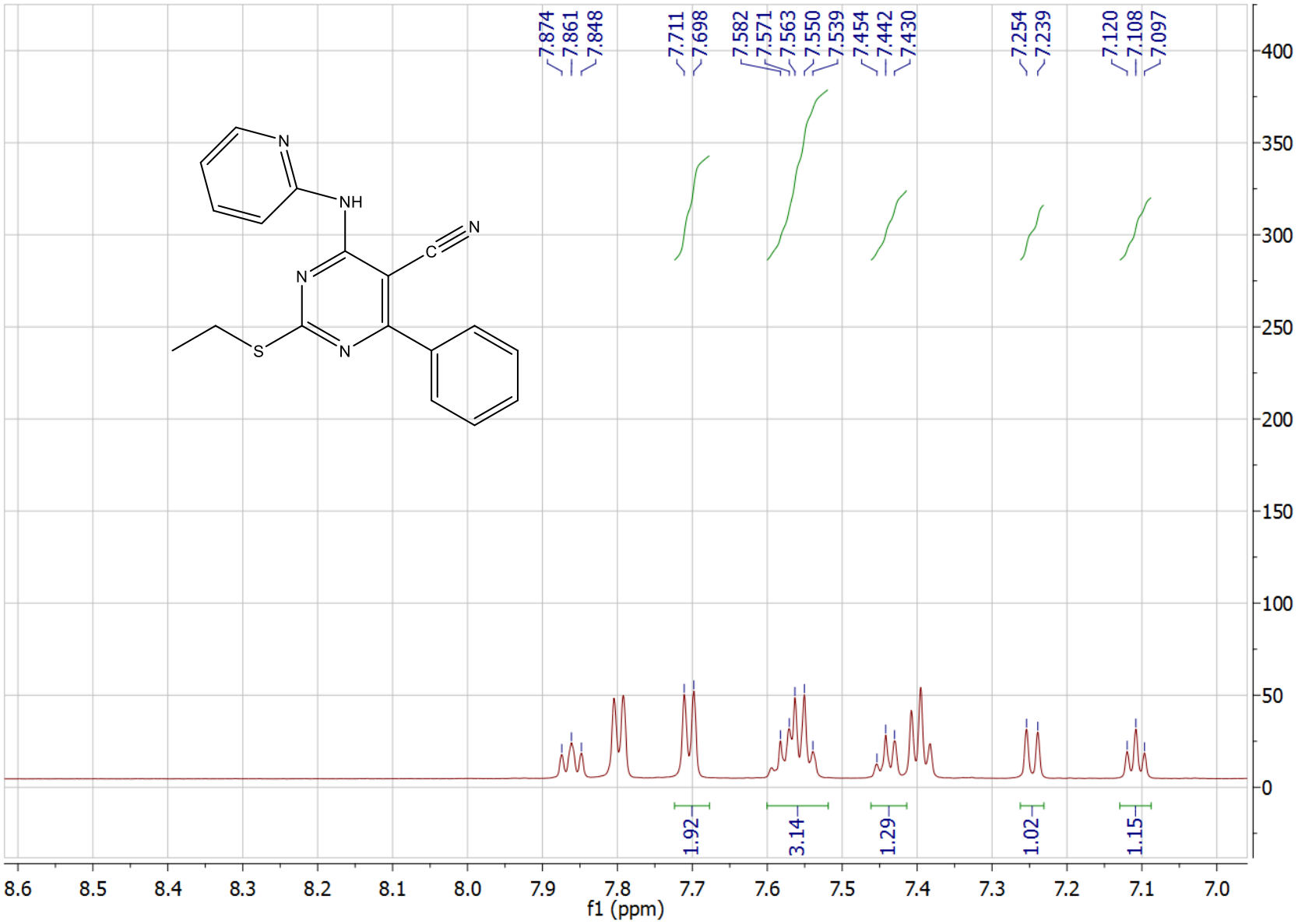
^{1}H NMR of compound 11_a



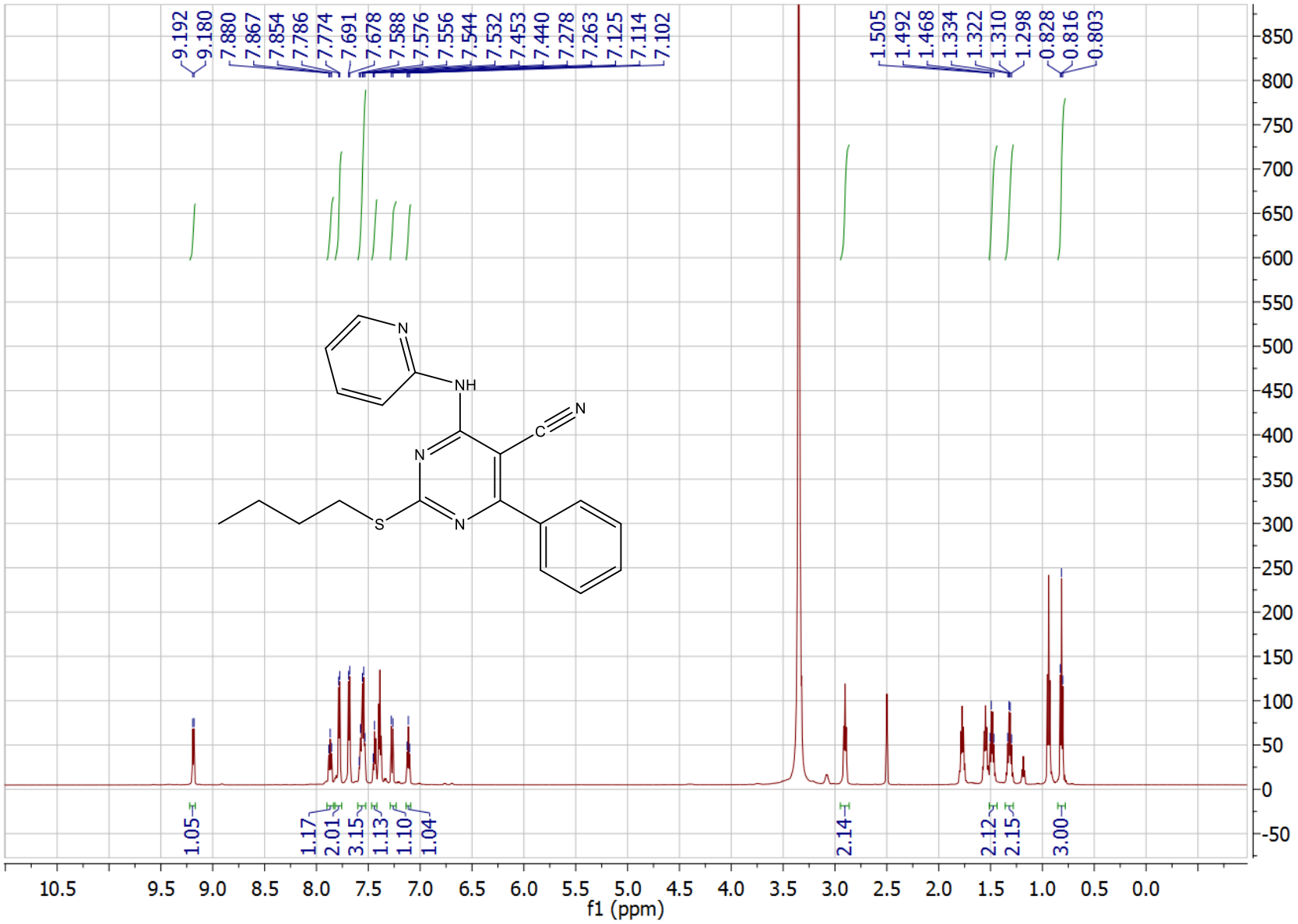
^{1}H NMR of compound 11_a



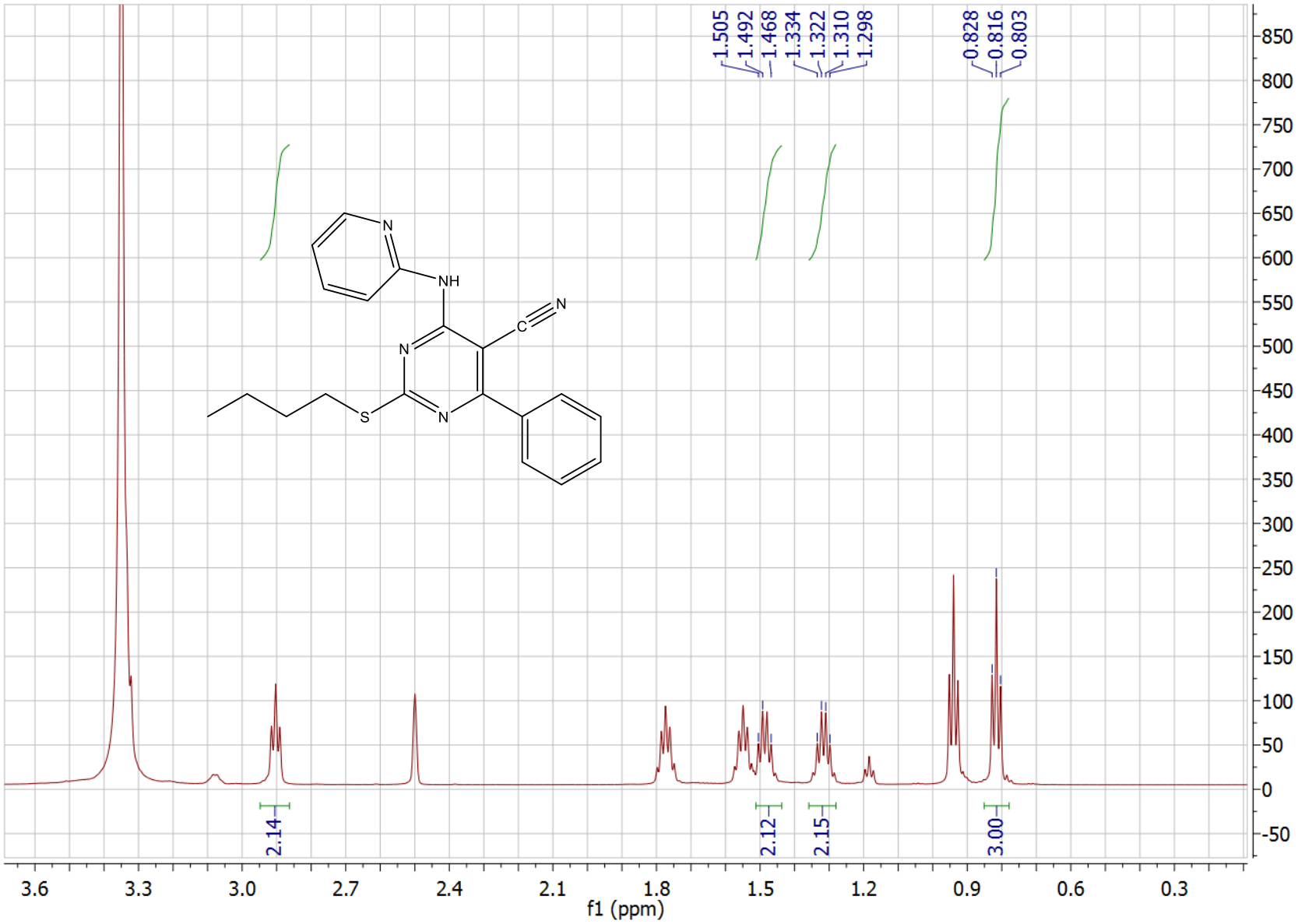
^{1}H NMR of compound 11_a



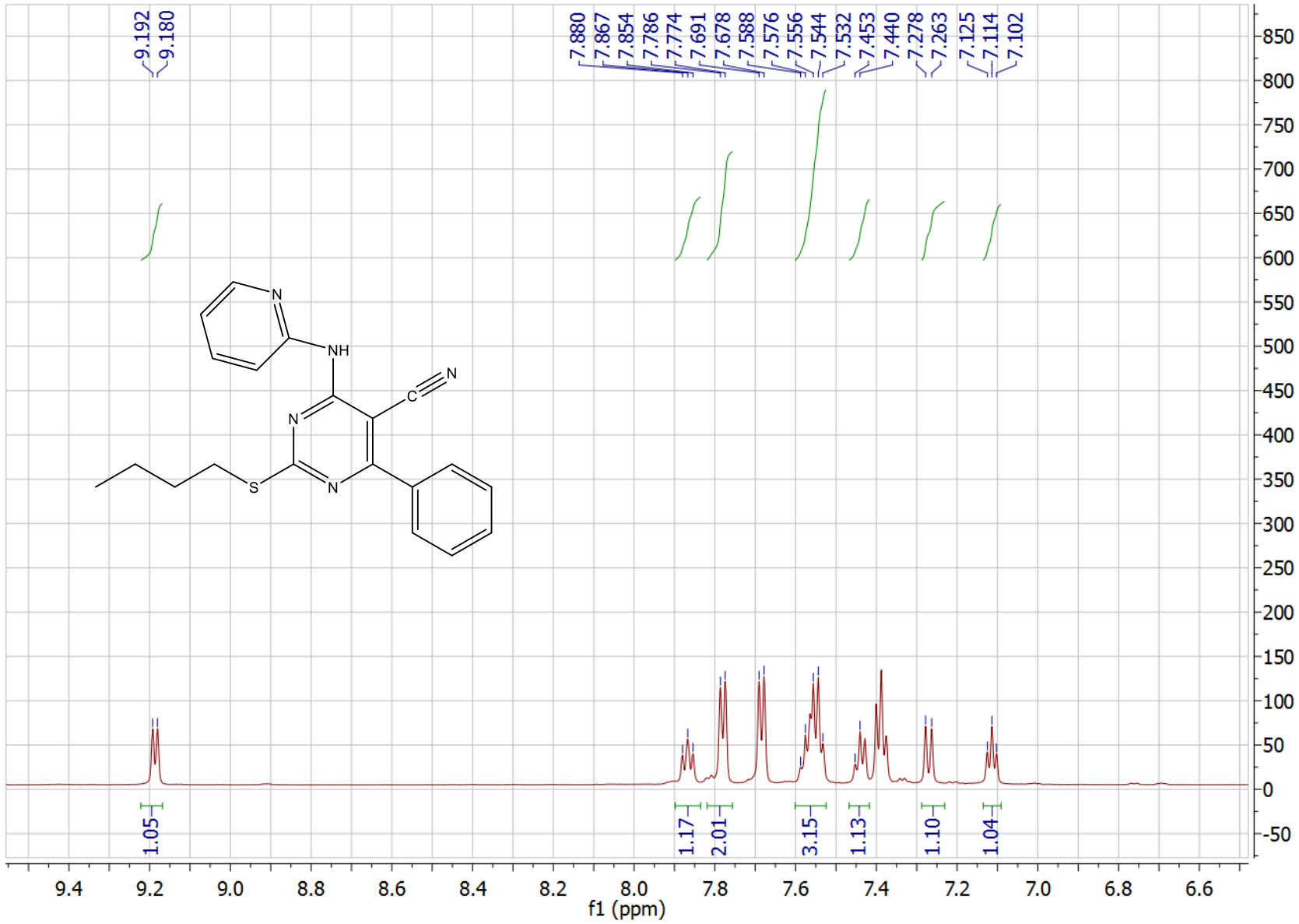
^{1}H NMR of compound 11_b



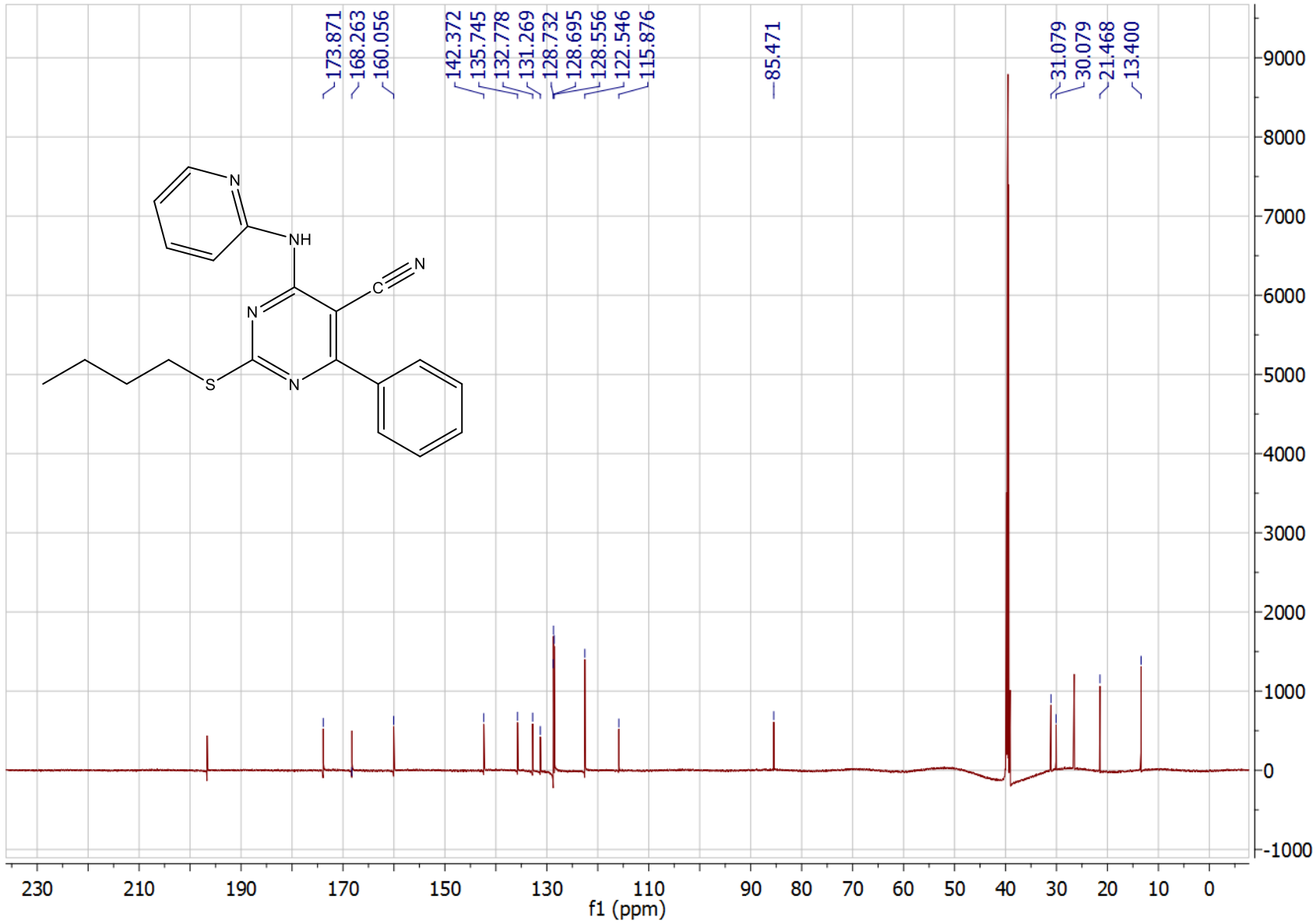
^{1}H NMR of compound 11_b



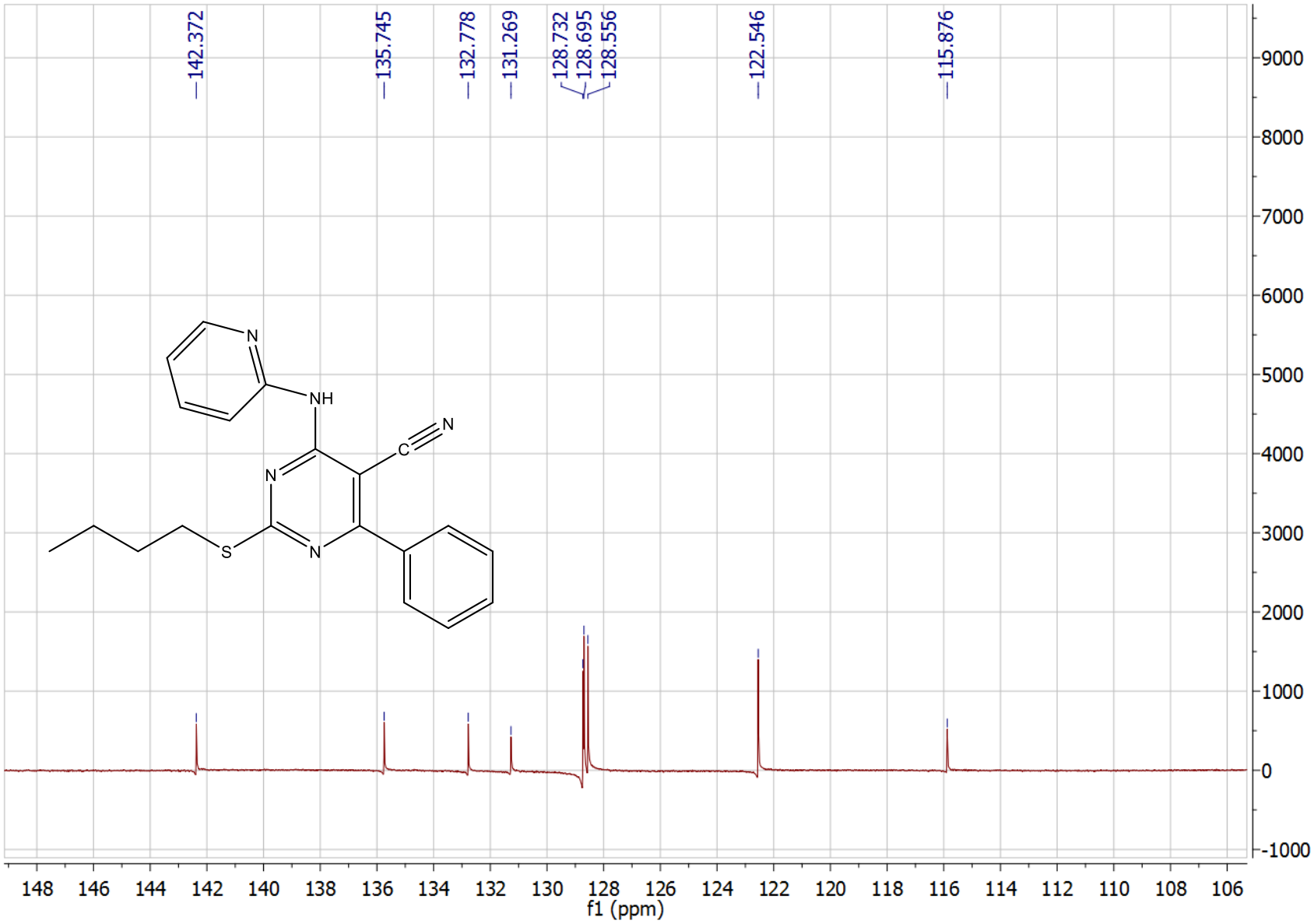
^{1}H NMR of compound 11_b



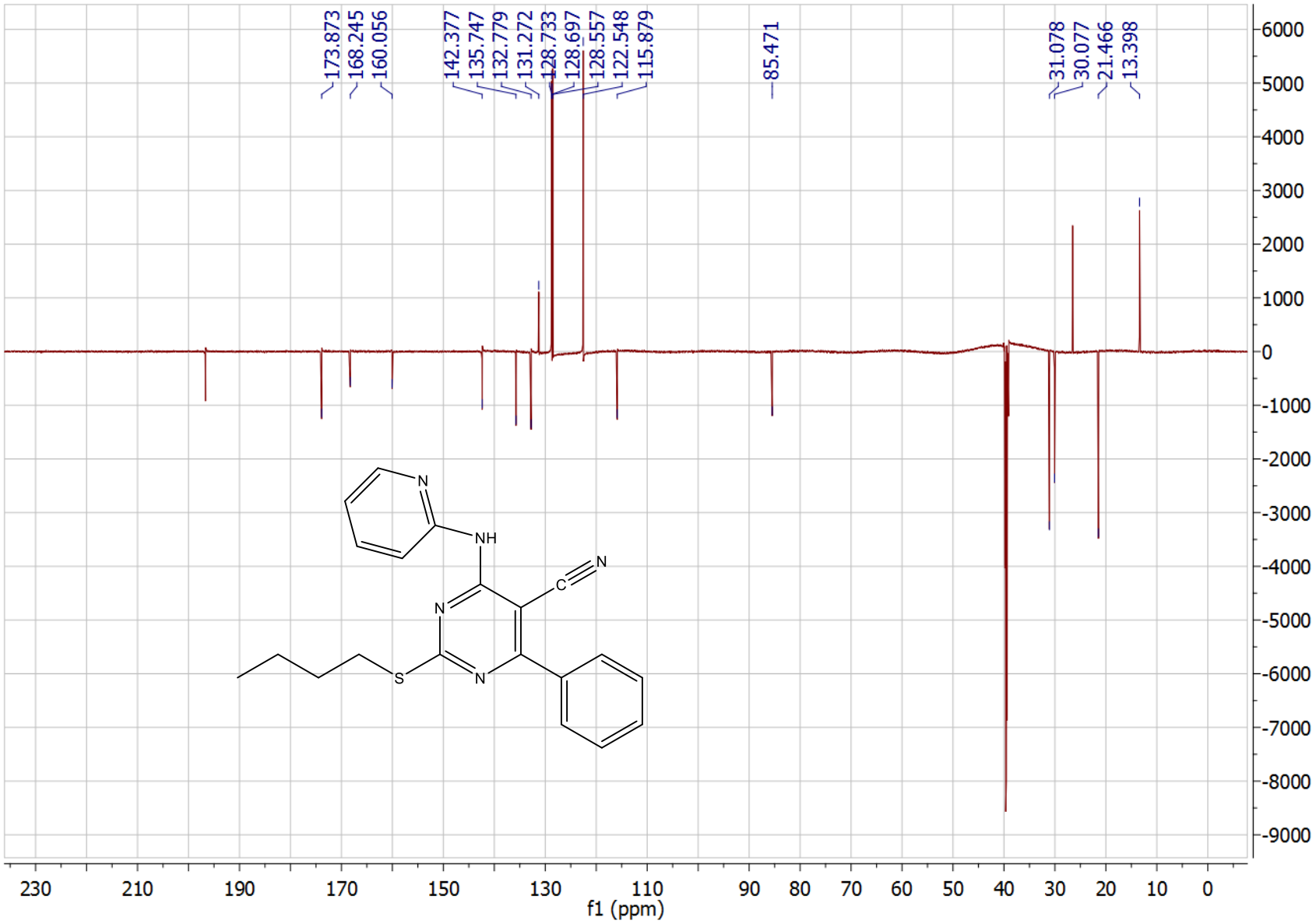
^{13}C NMR of compound 11_b



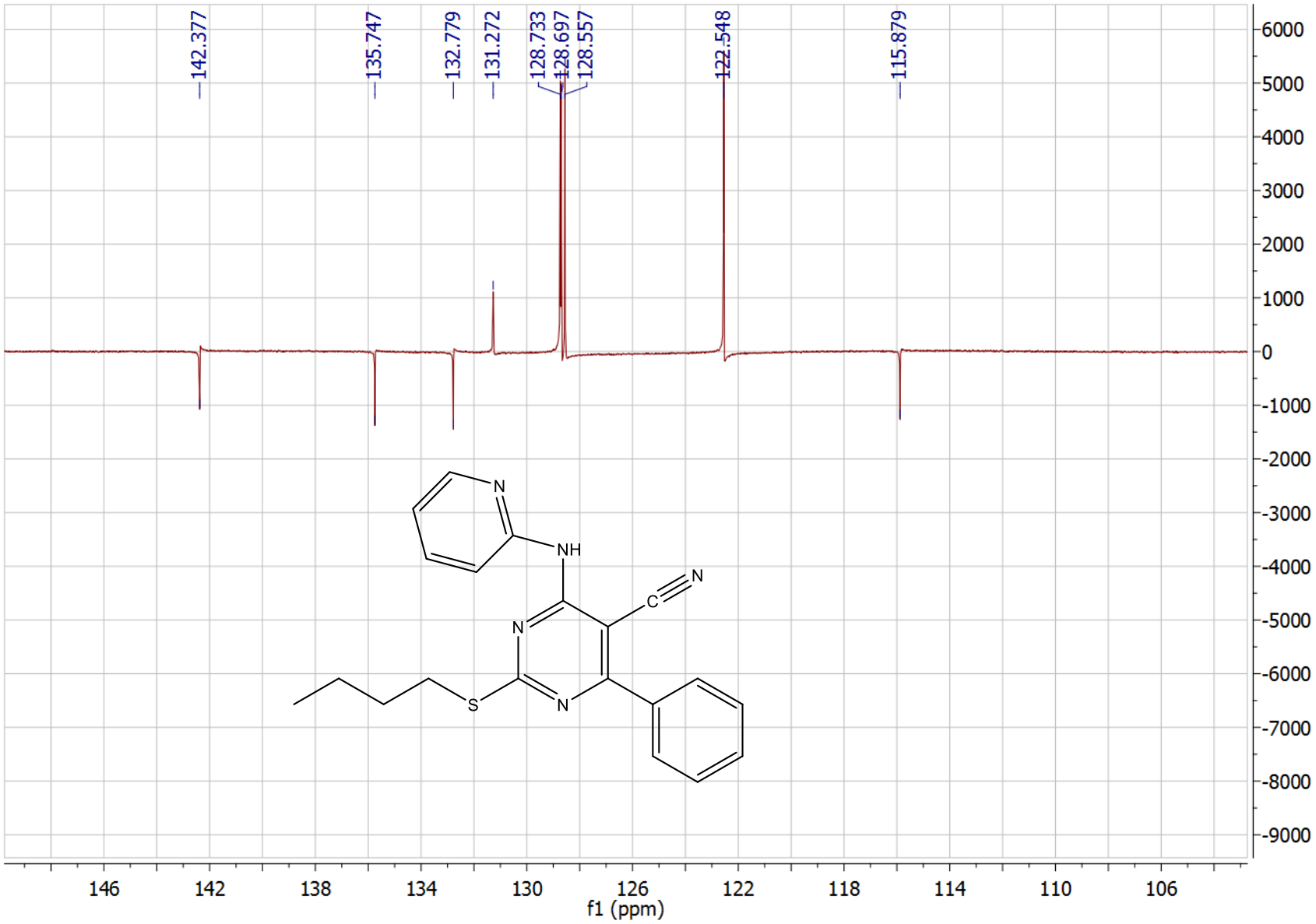
^{13}C NMR of compound 11_b



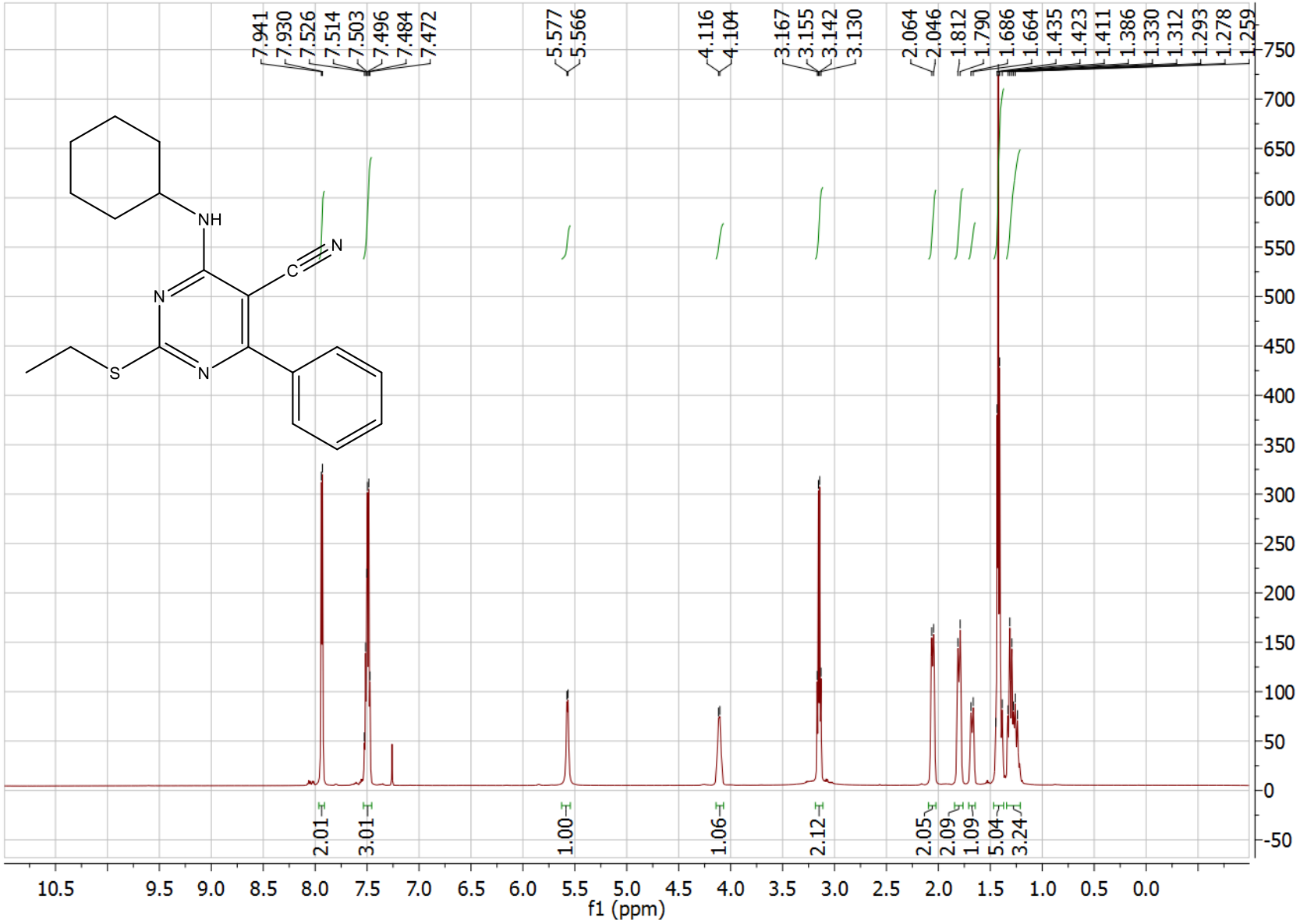
APT of compound 11_b



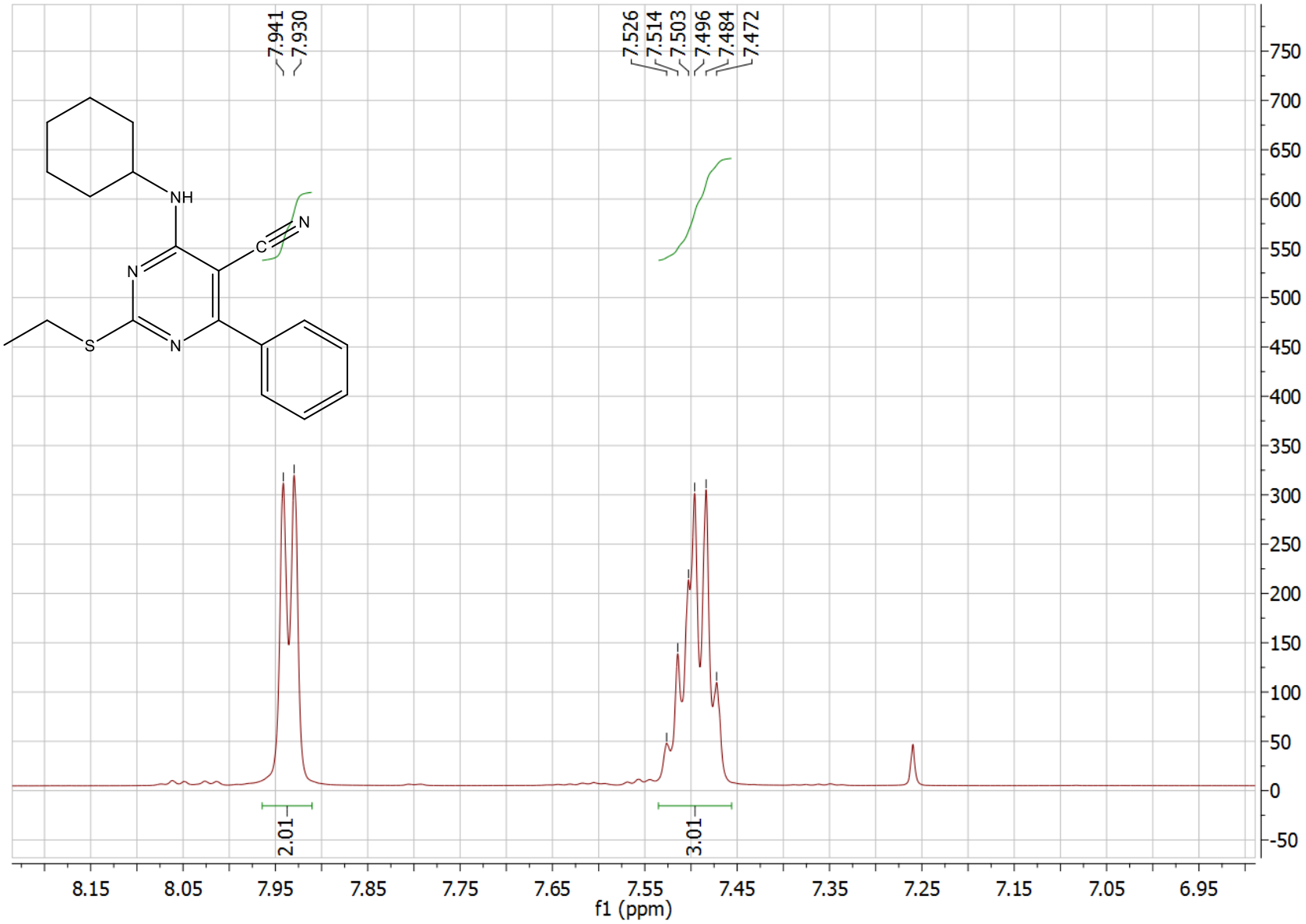
APT of compound 11_b



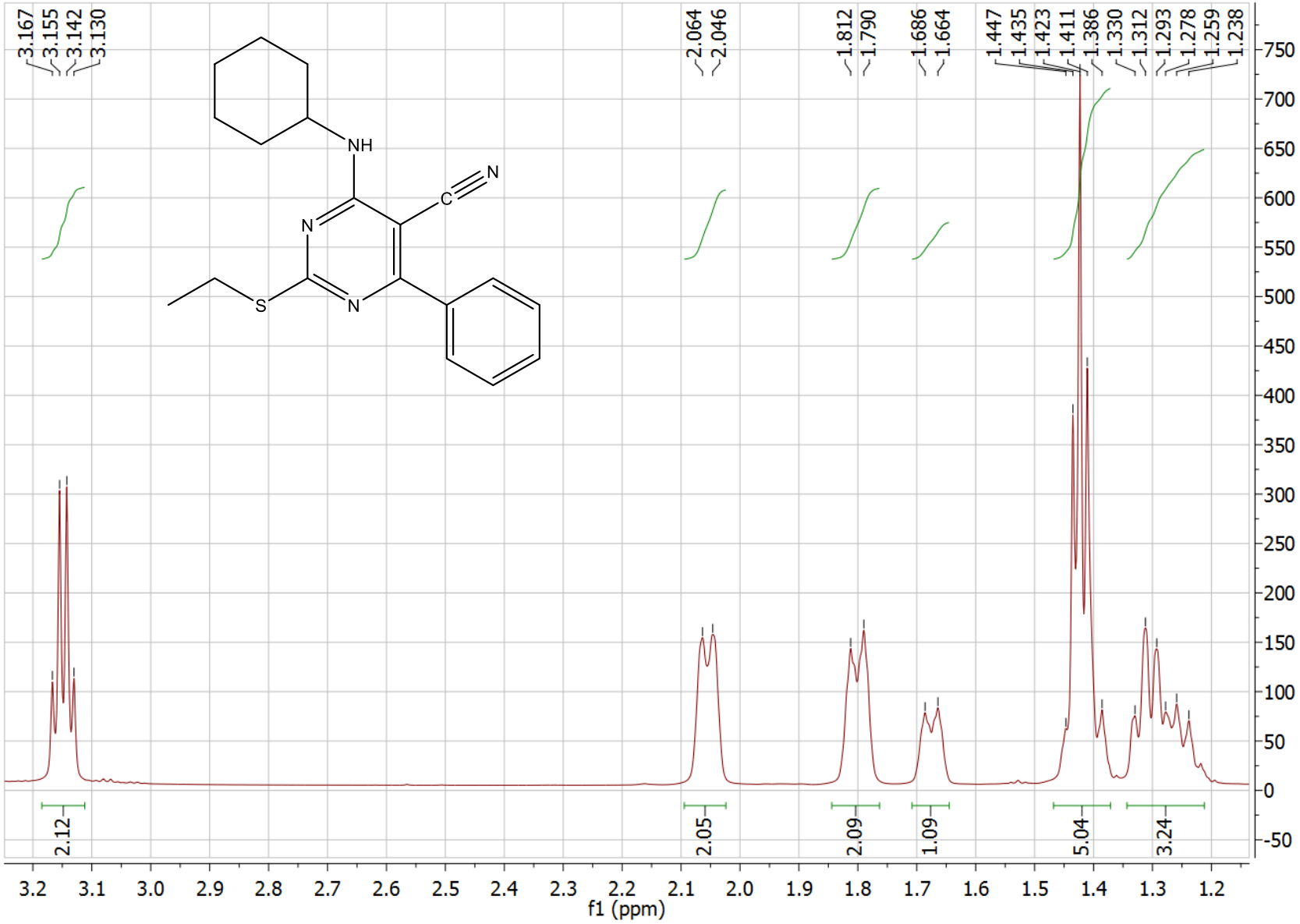
^{1}H NMR of compound 12_a



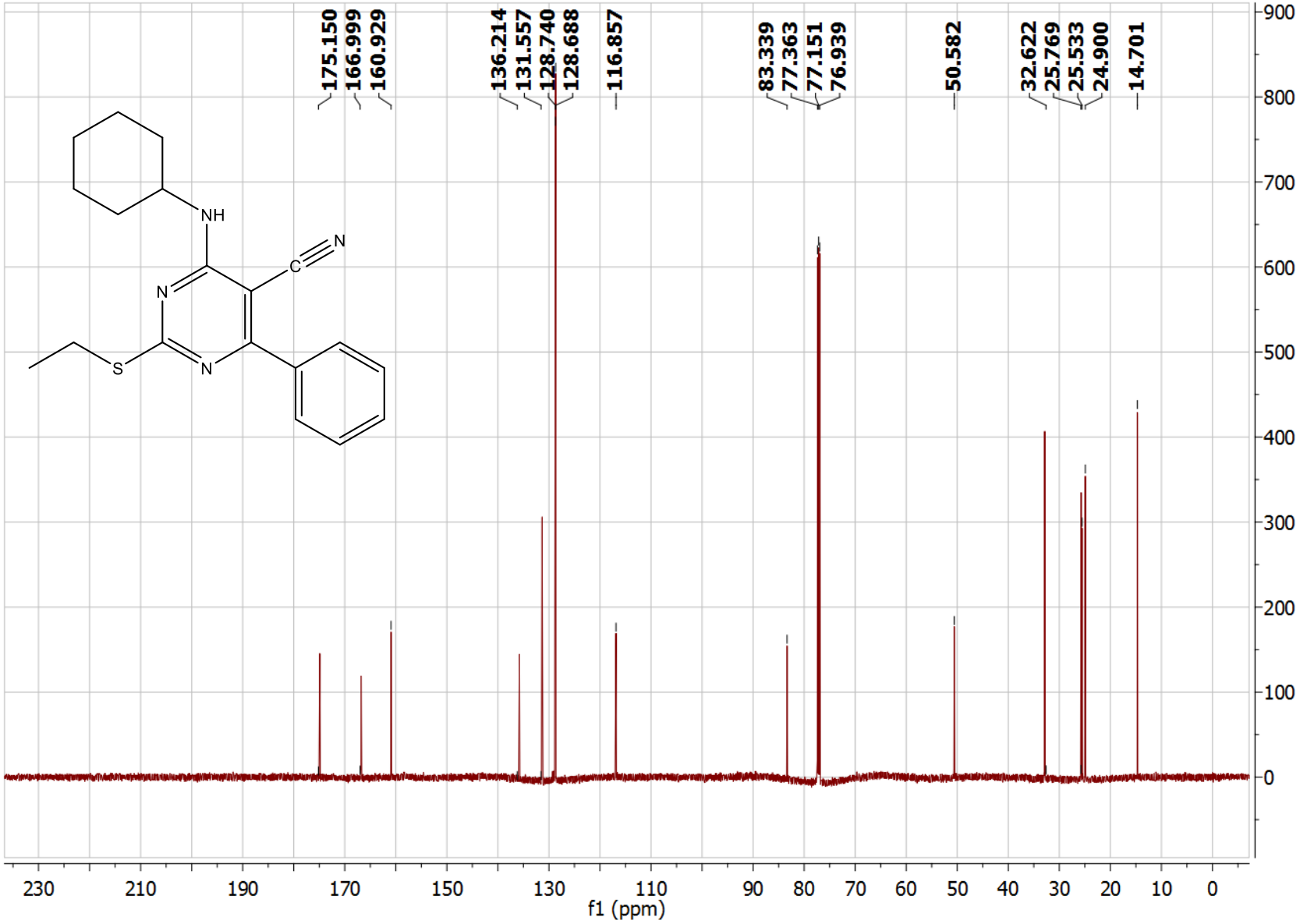
^{1}H NMR of compound 12$_{a}$



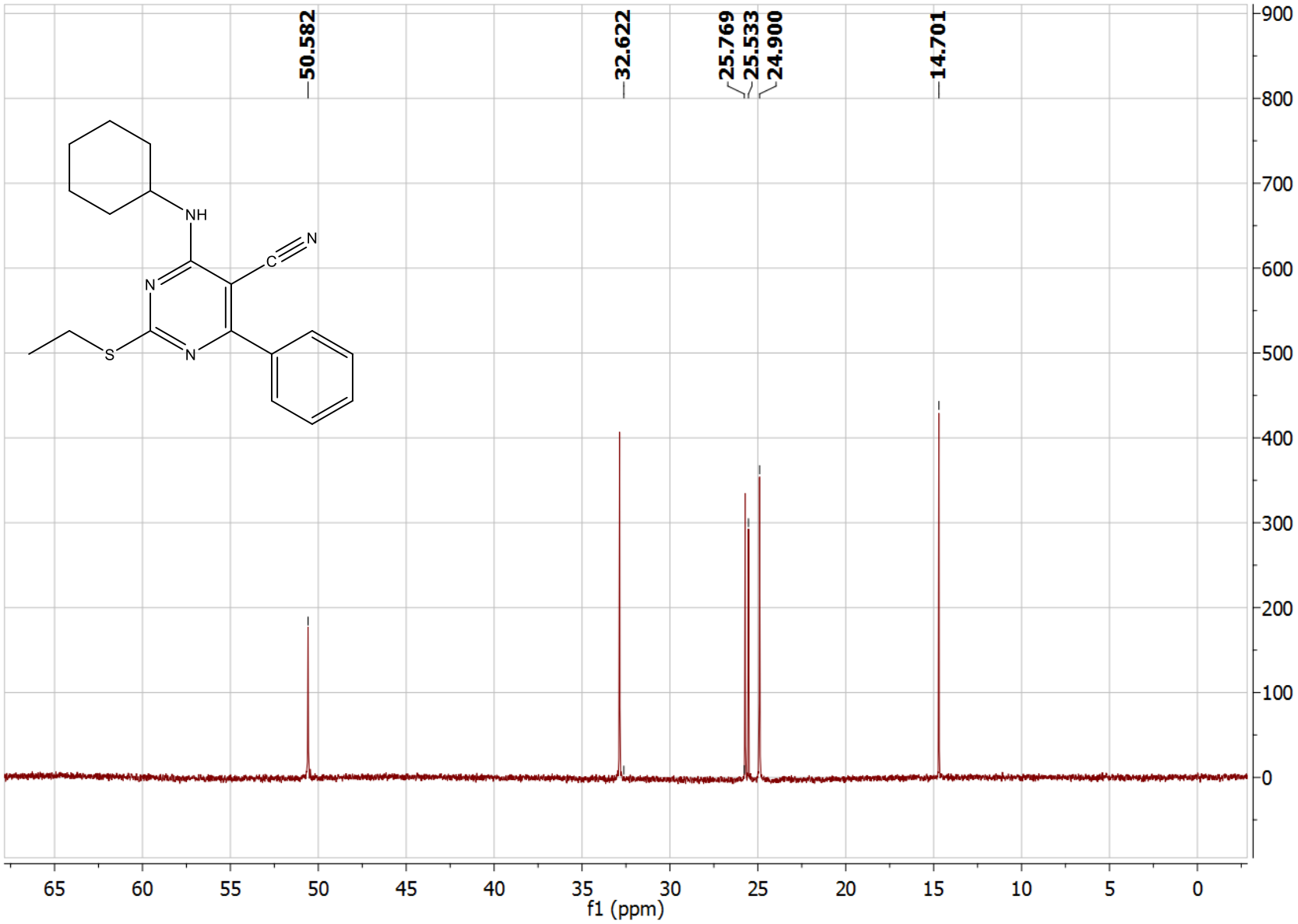
^{1}H NMR of compound 12$_{a}$



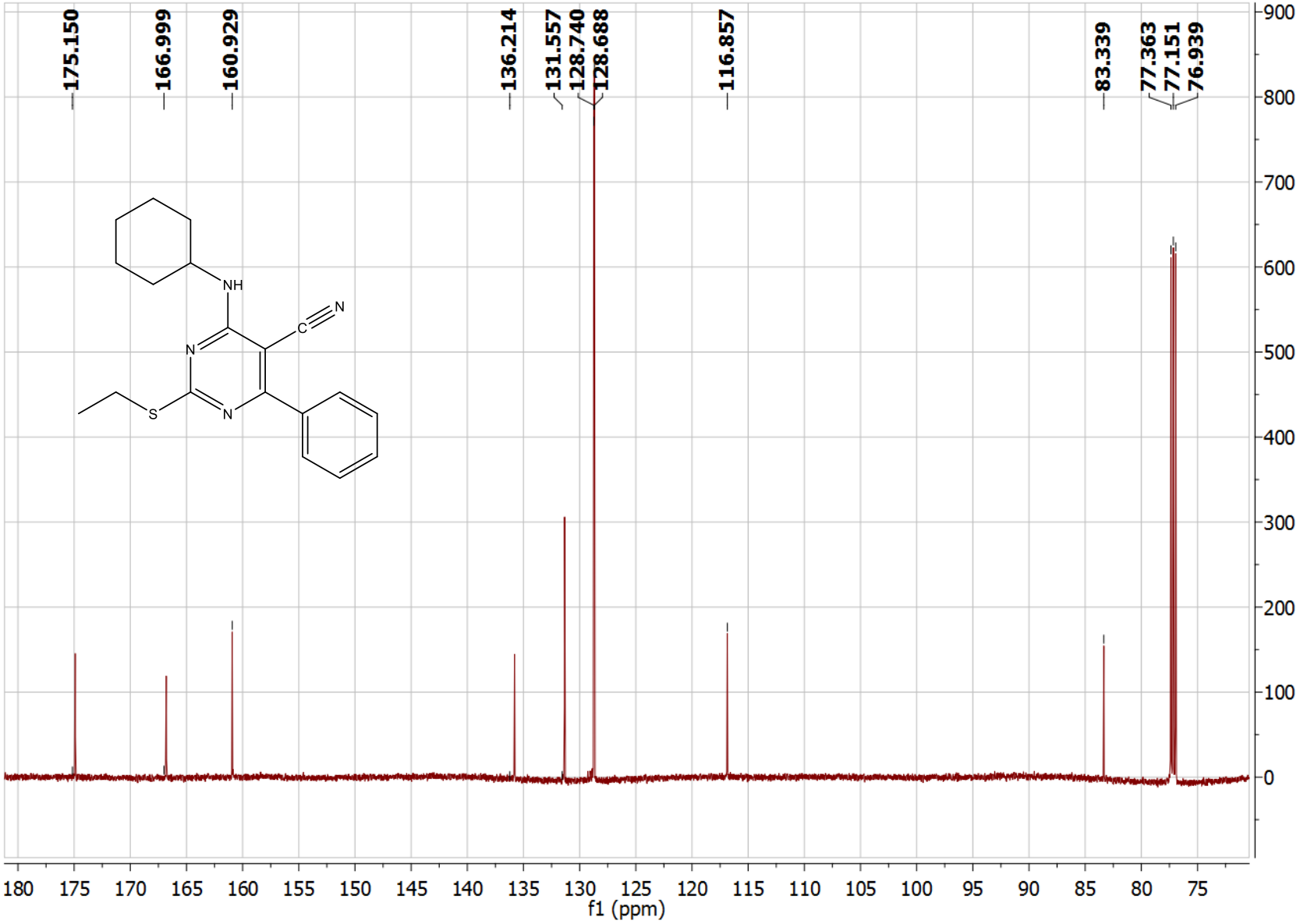
^{13}C NMR of compound 12$_a$



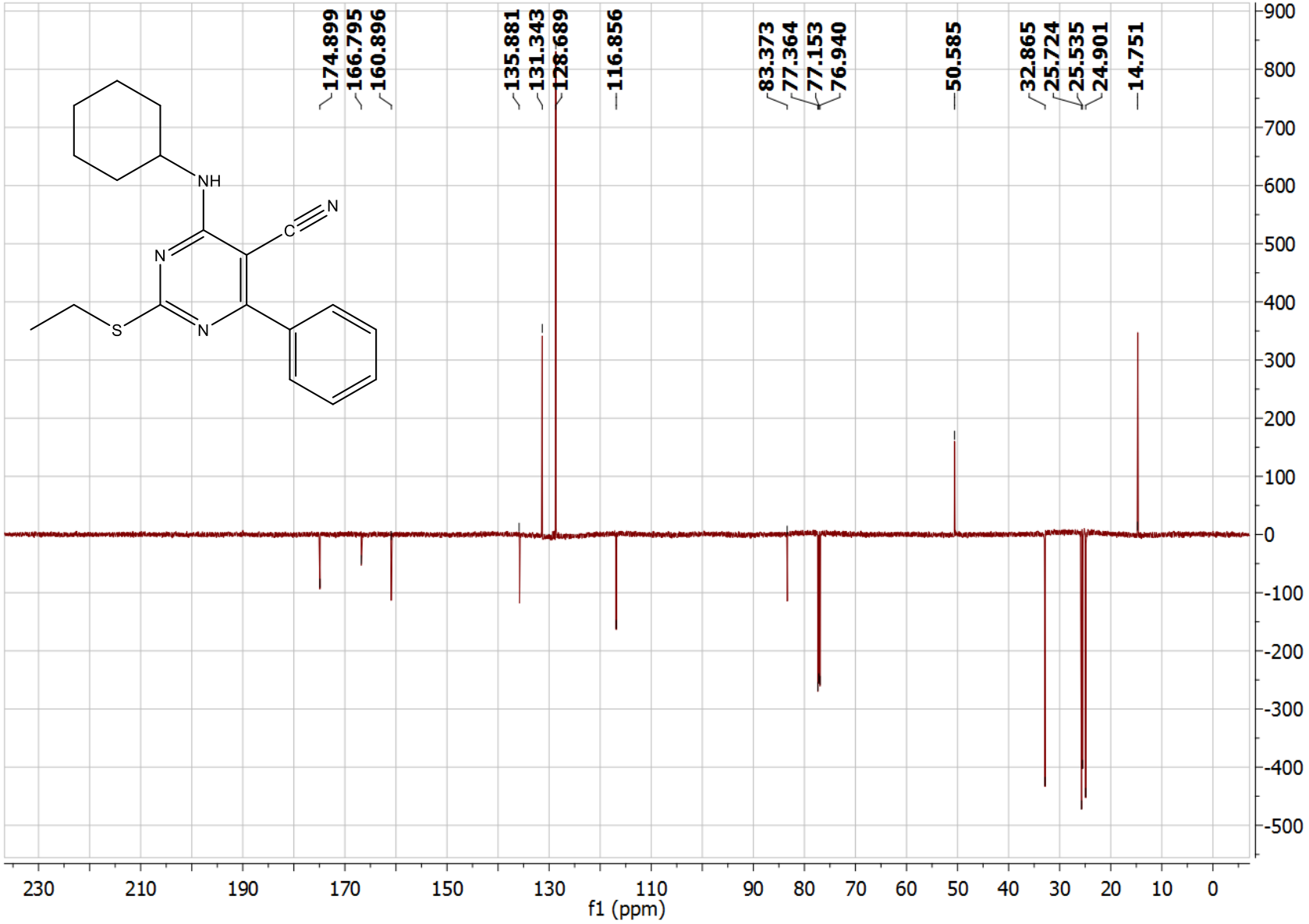
^{13}C NMR of compound 12_a



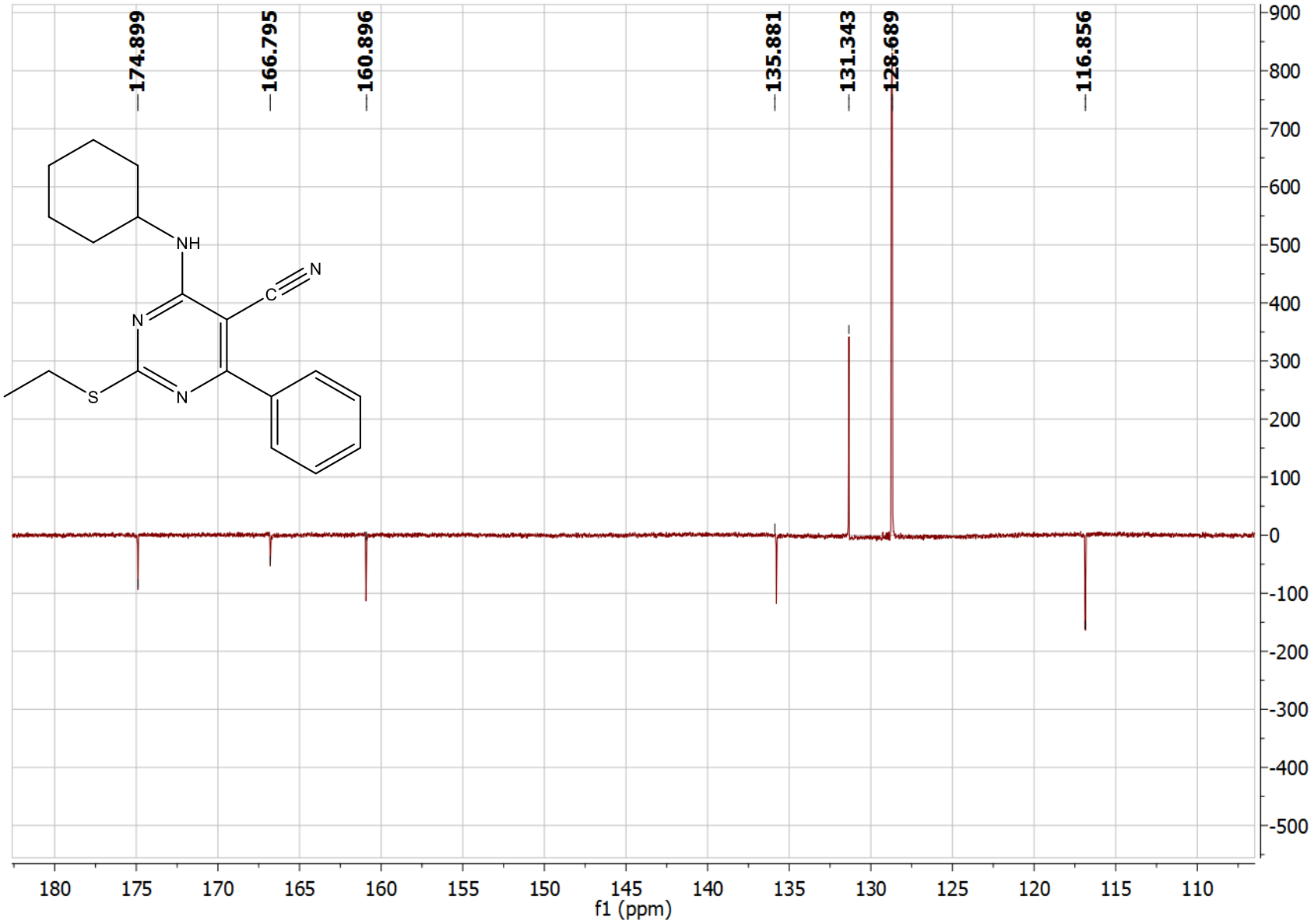
^{13}C NMR of compound 12_a



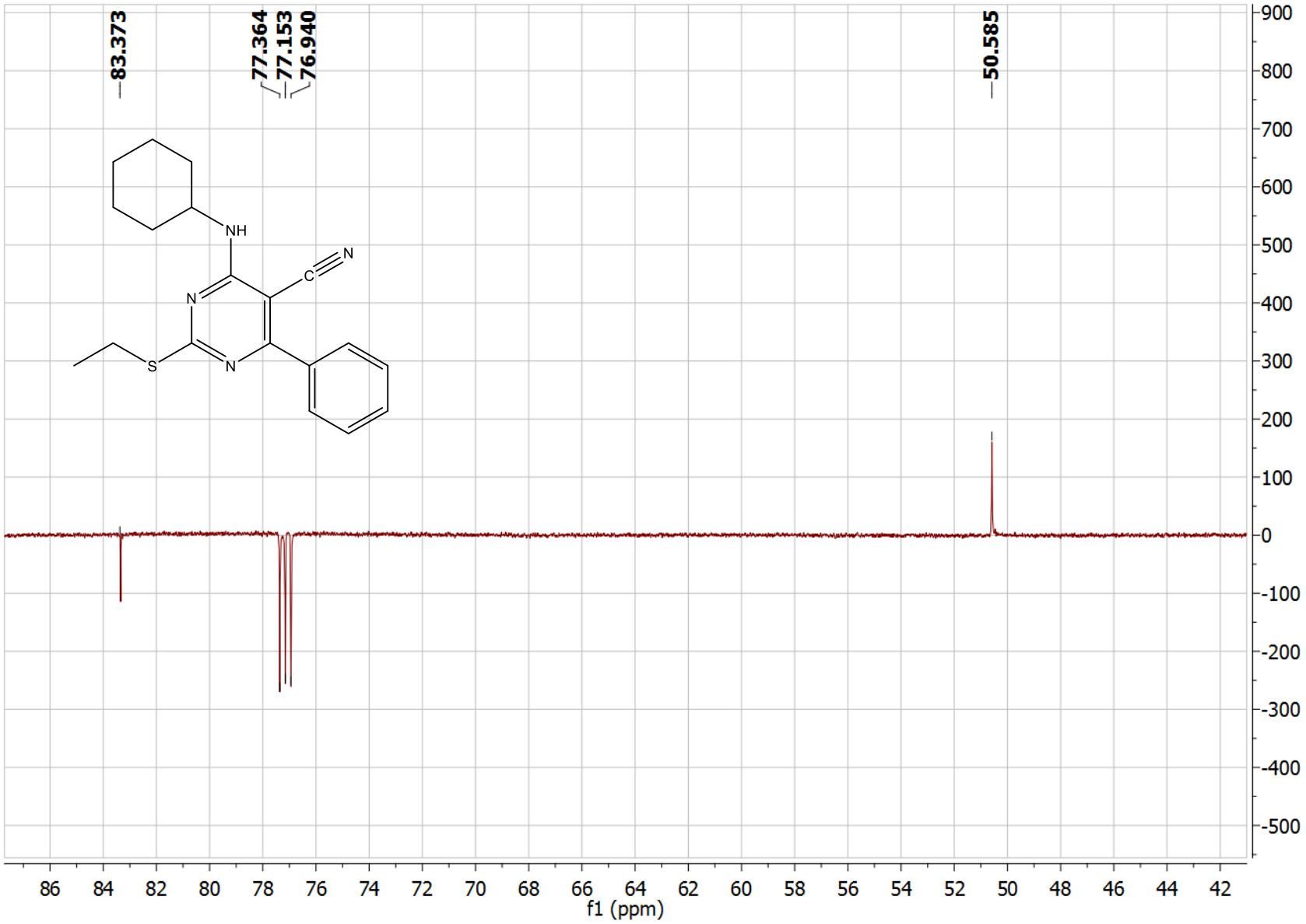
APT of compound 12_a



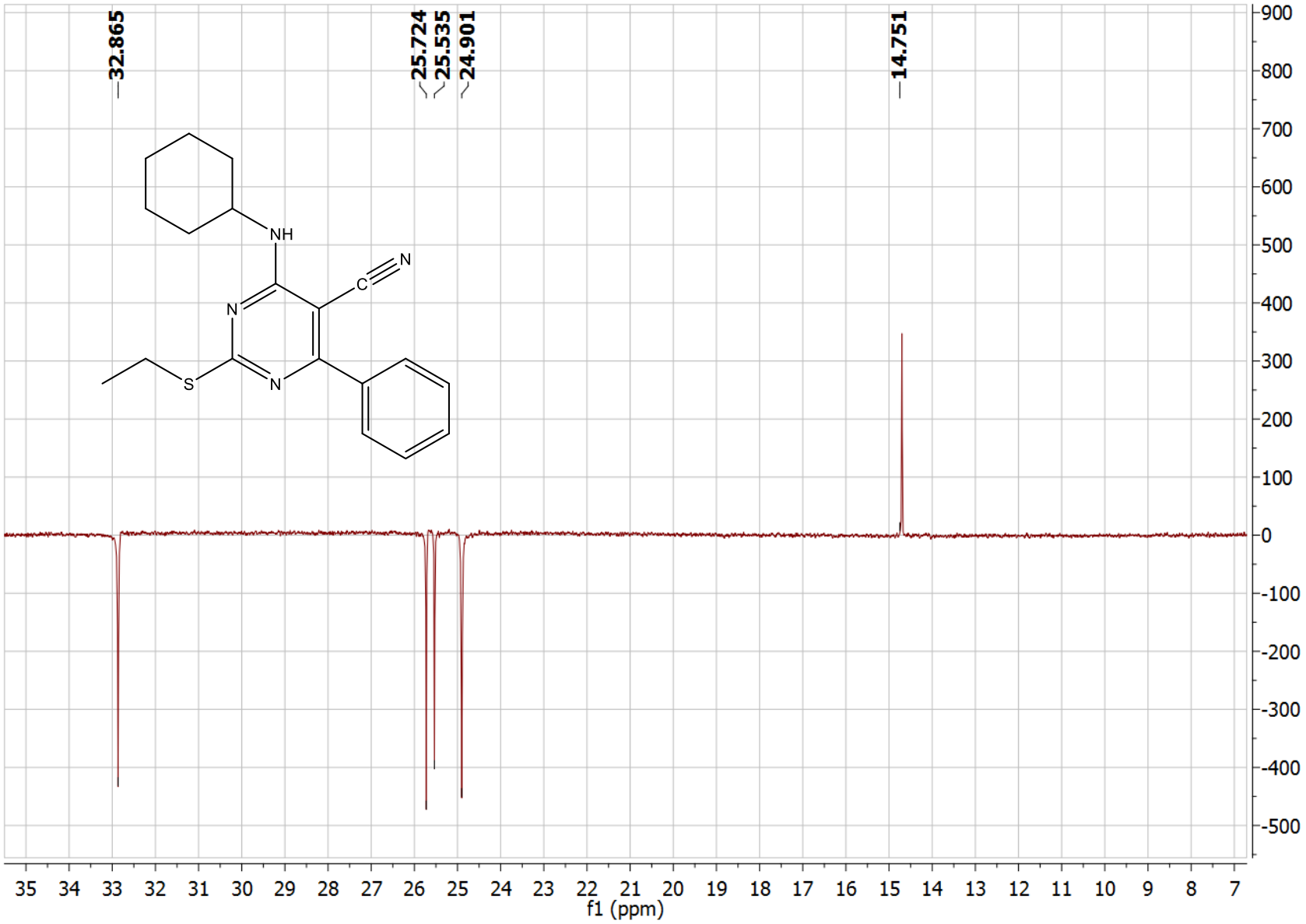
APT of compound 12_a



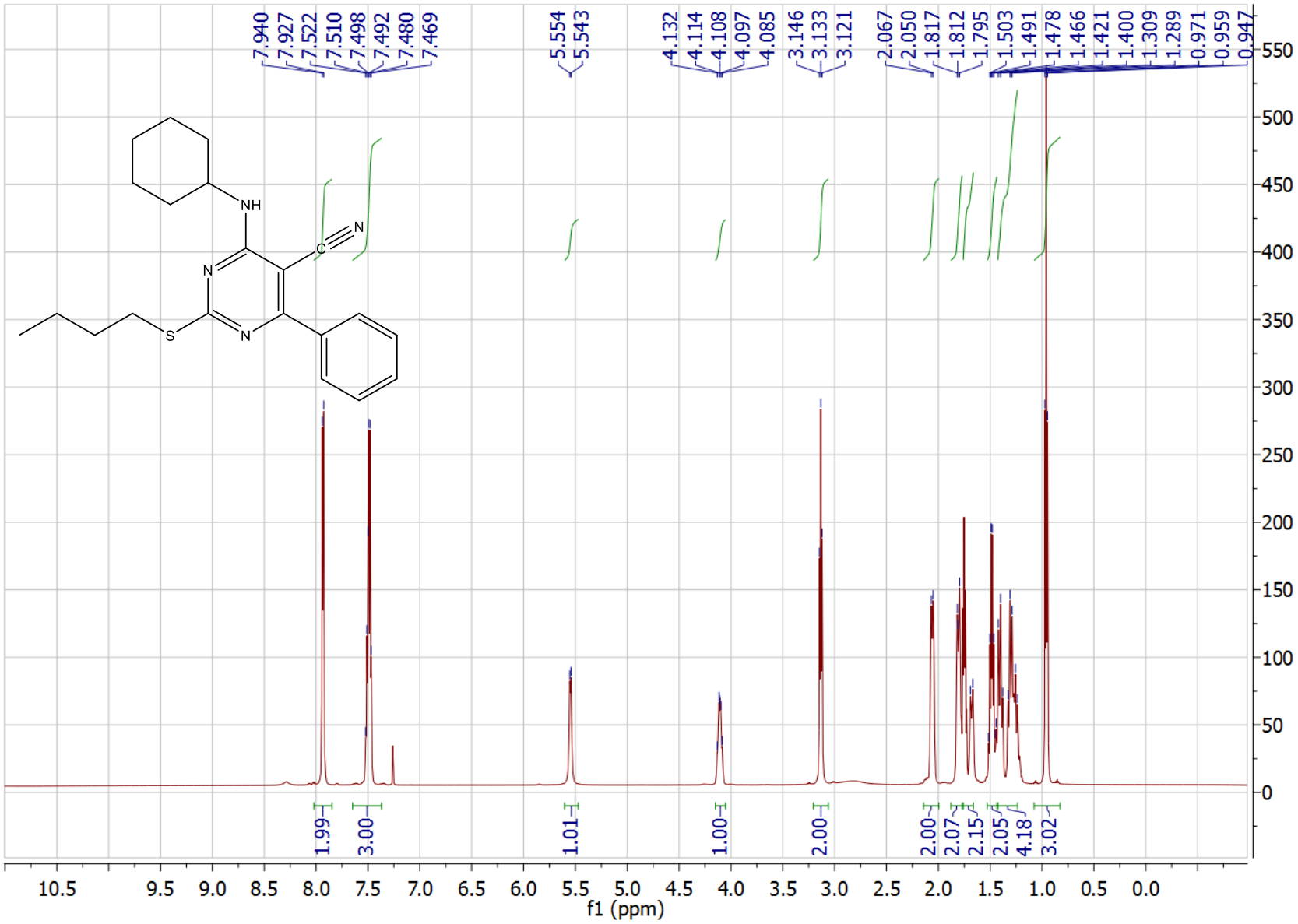
APT of compound 12_a



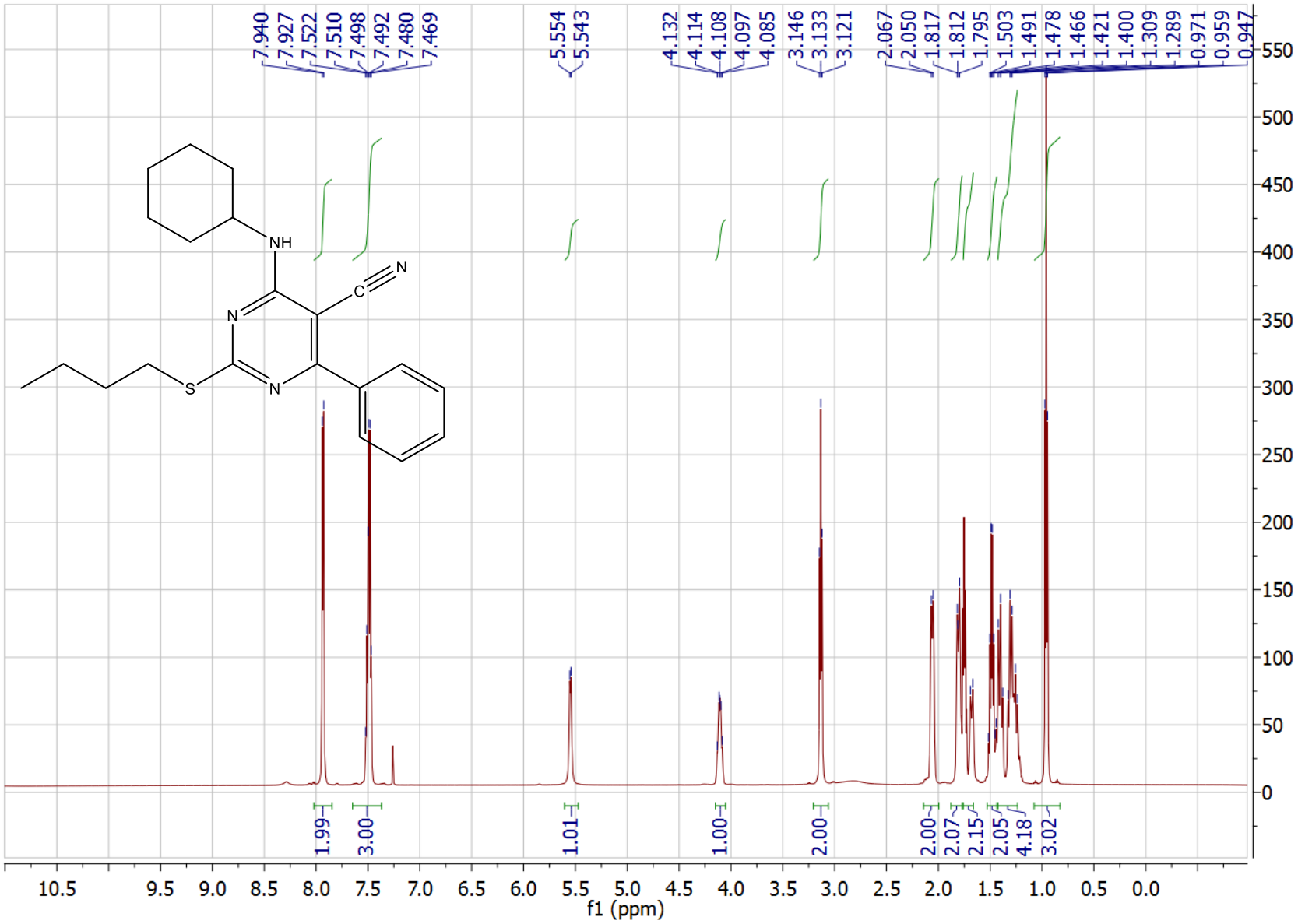
APT of compound 12_a



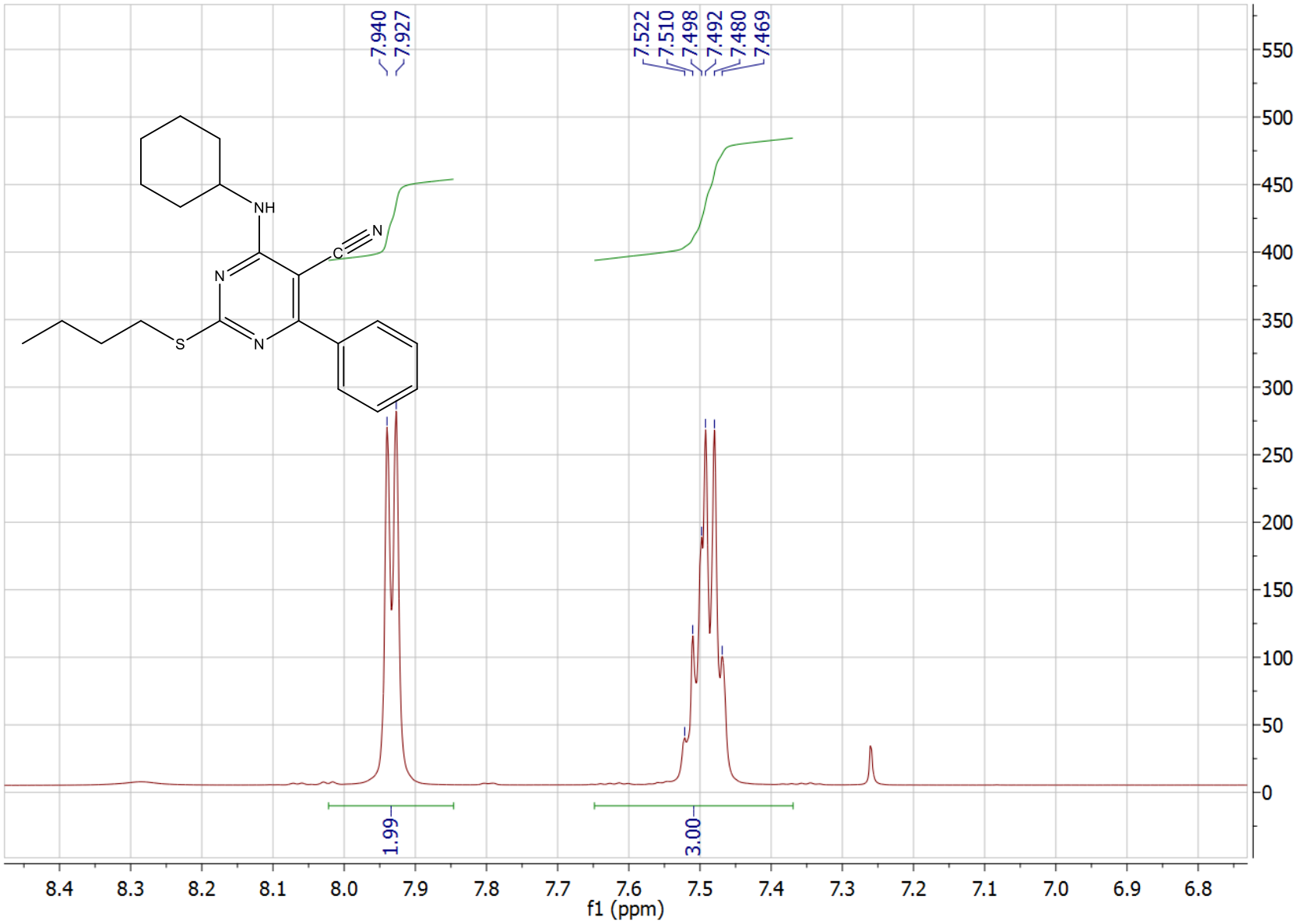
^{1}H NMR of compound 12_b



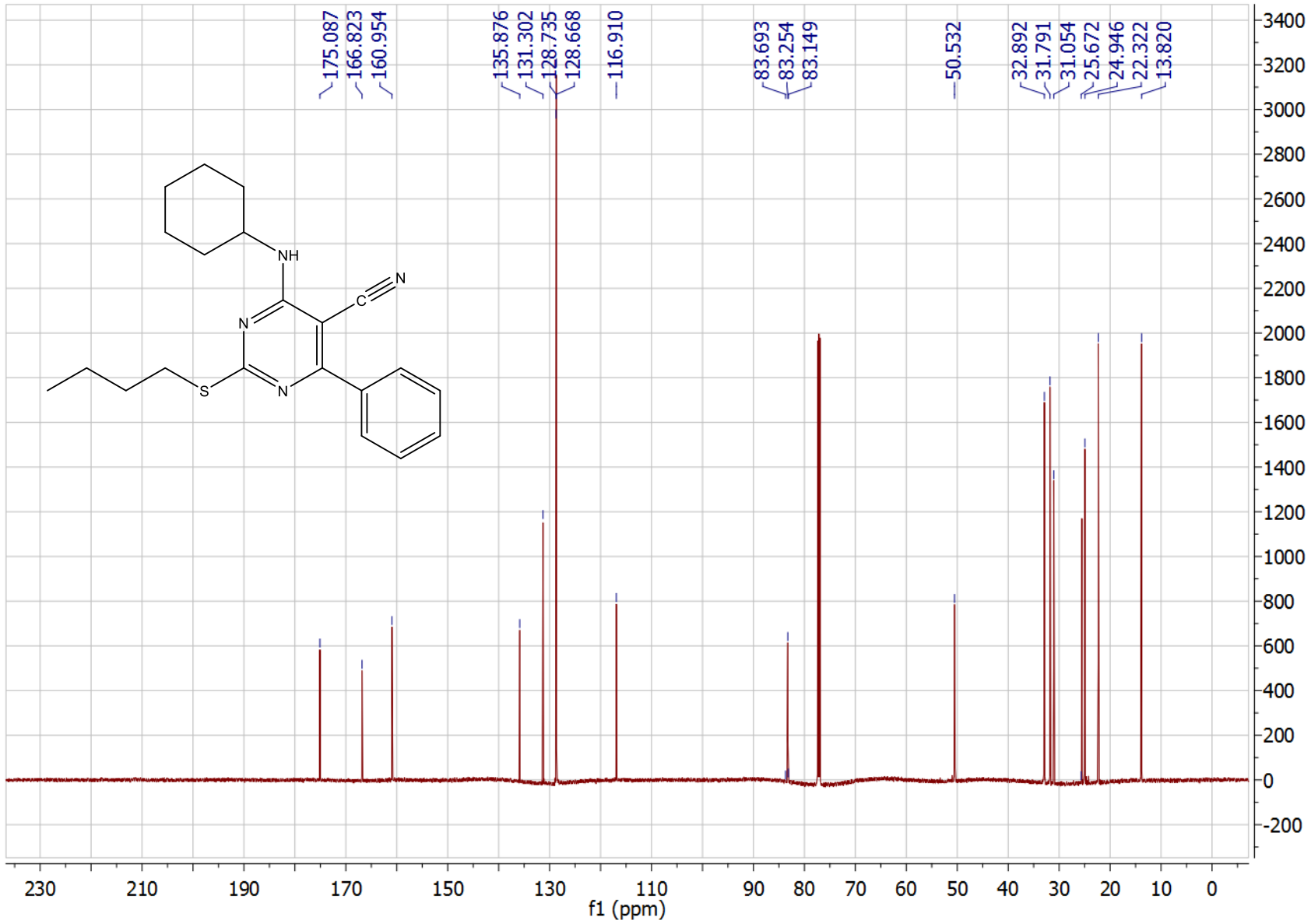
^{1}H NMR of compound 12_b



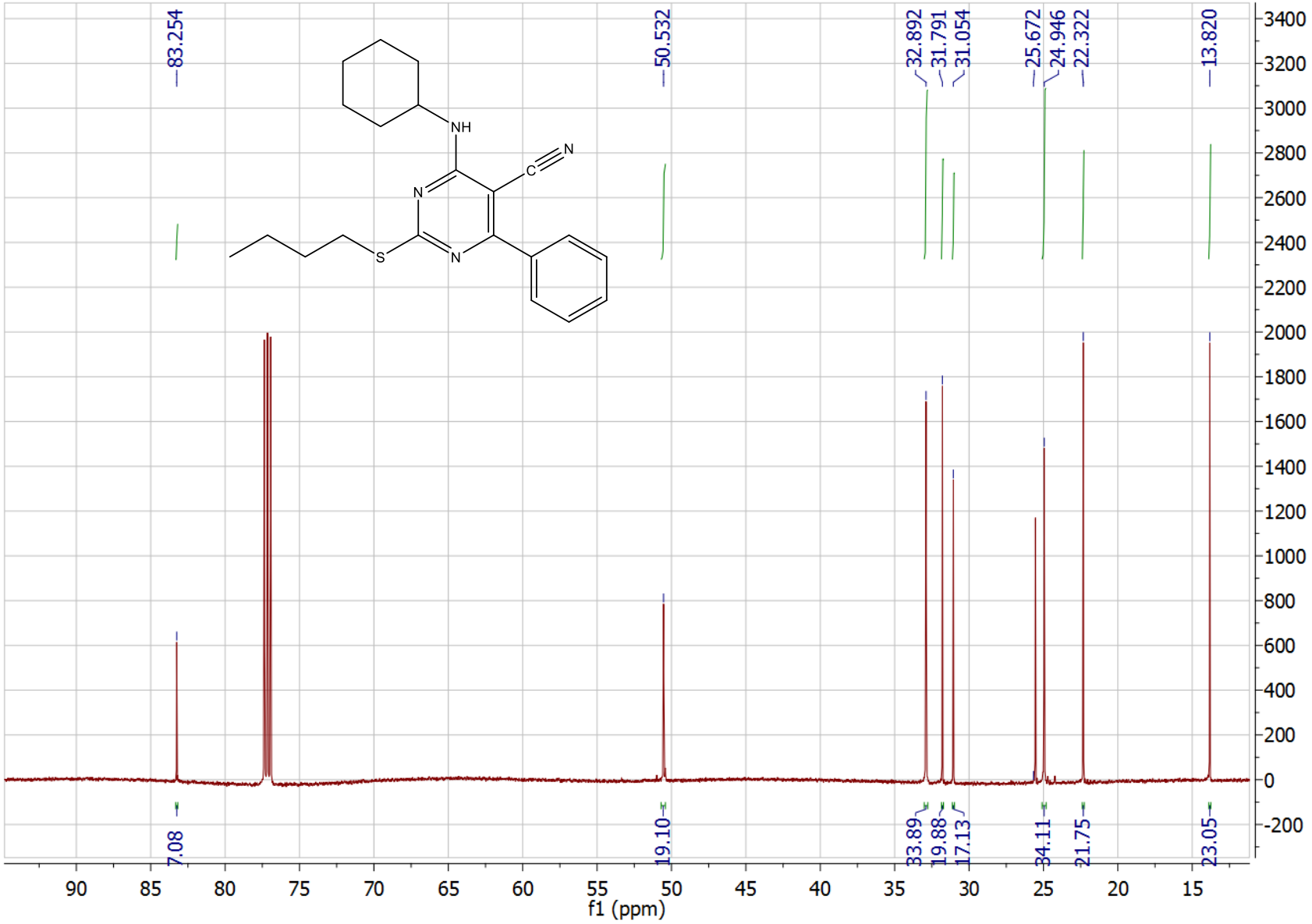
^{1}H NMR of compound 12_b



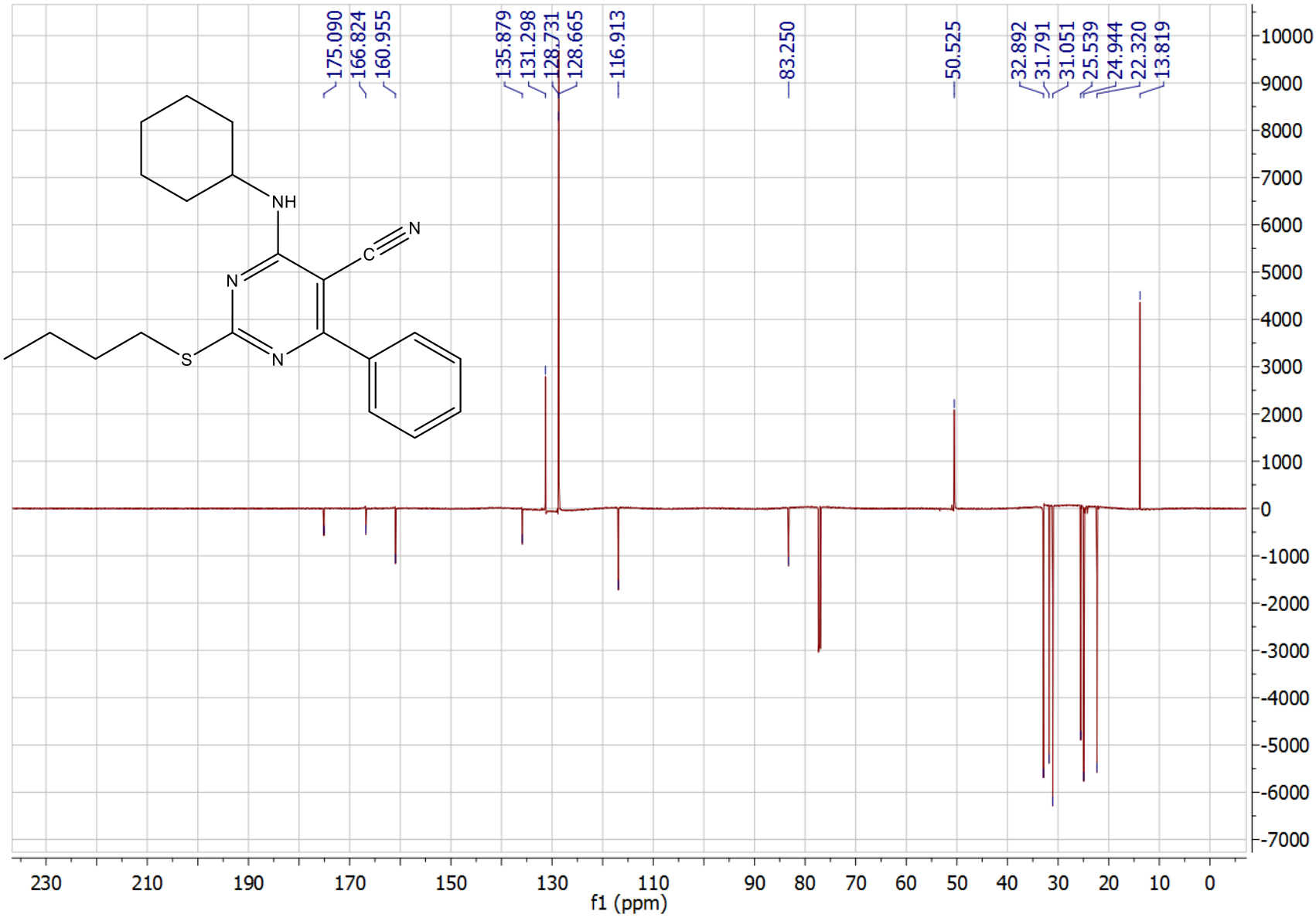
^{13}C NMR of compound 12_b



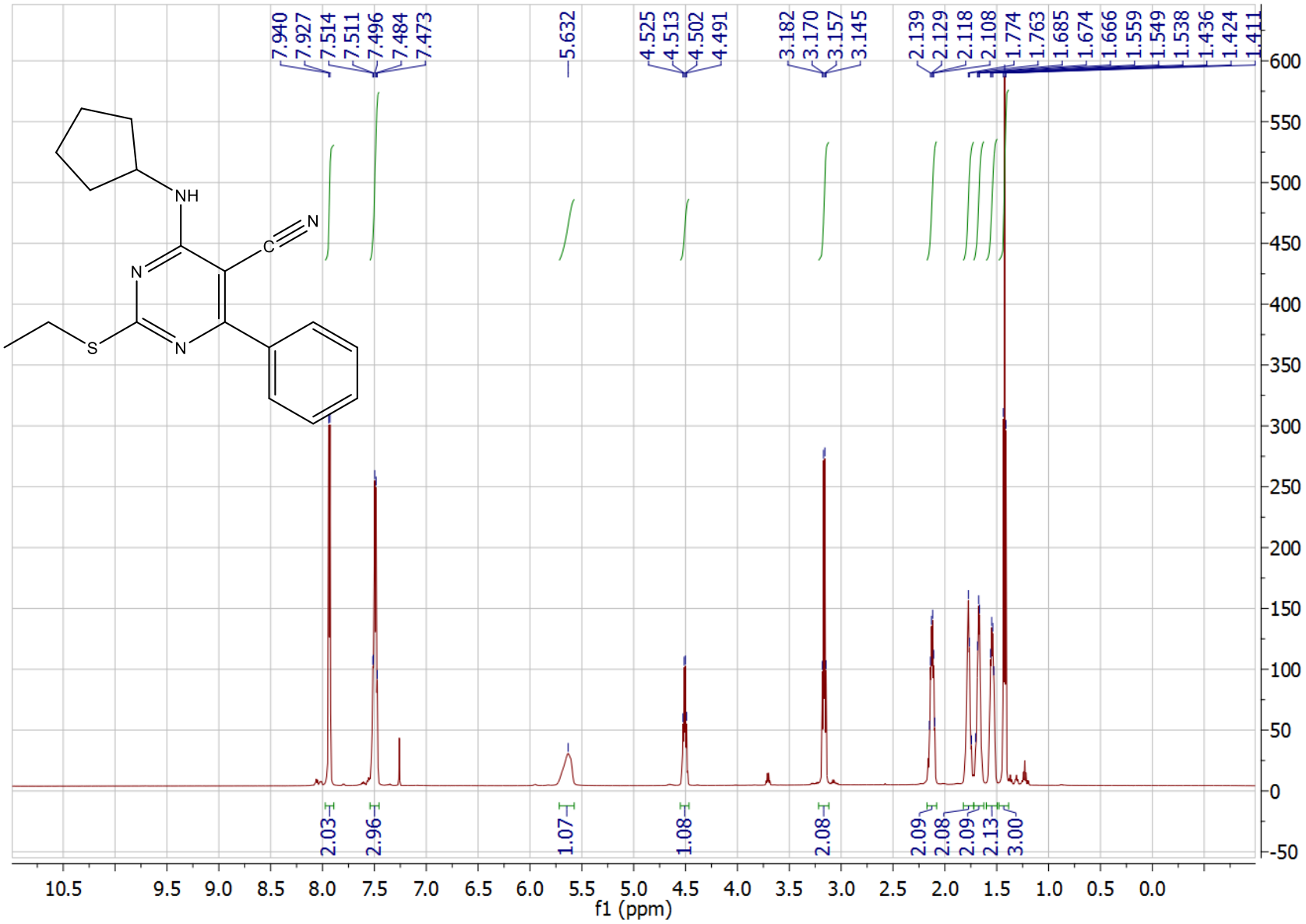
^{13}C NMR of compound 12_b



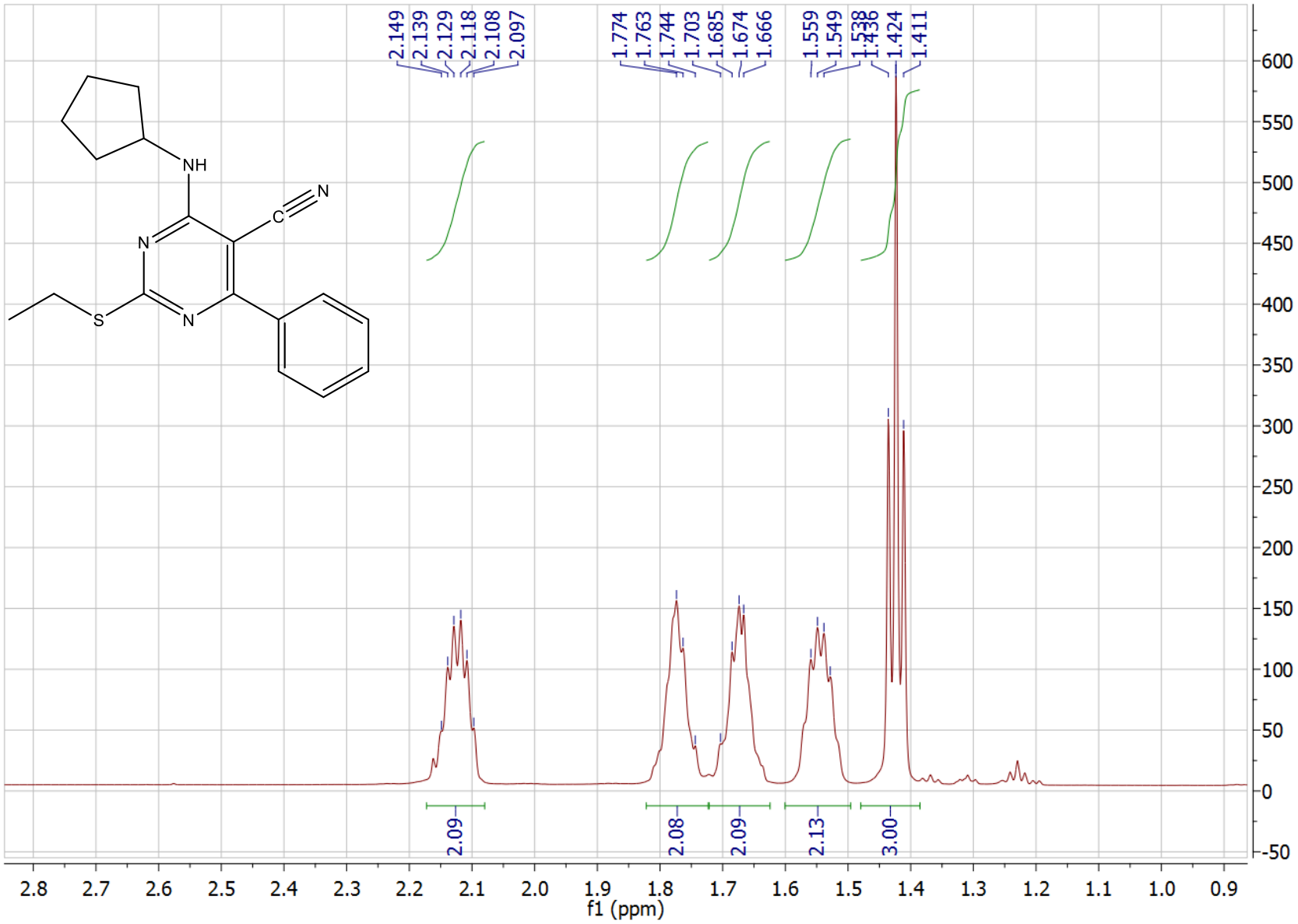
APT of compound 12_b



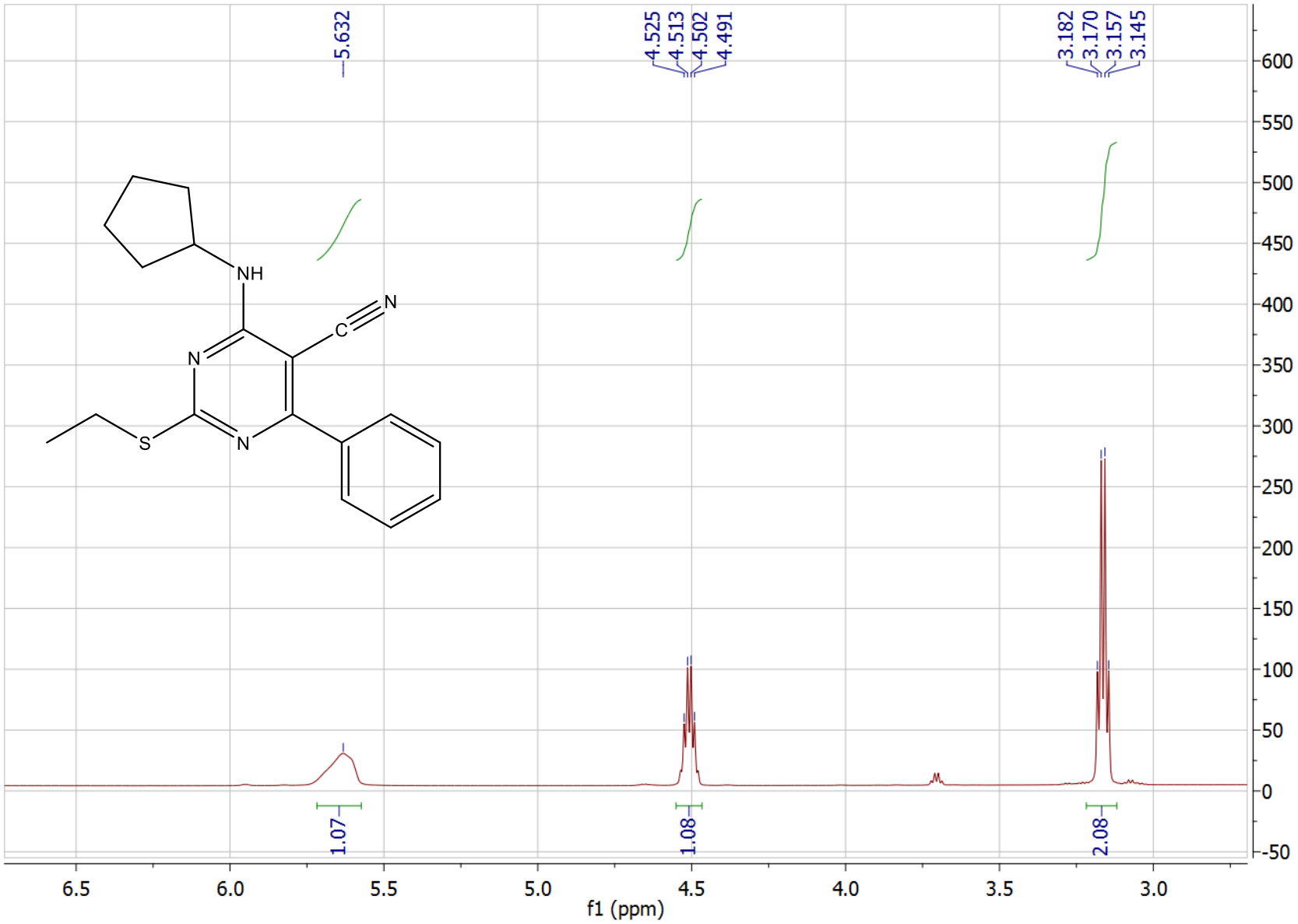
^{1}H NMR of compound 13$_a$



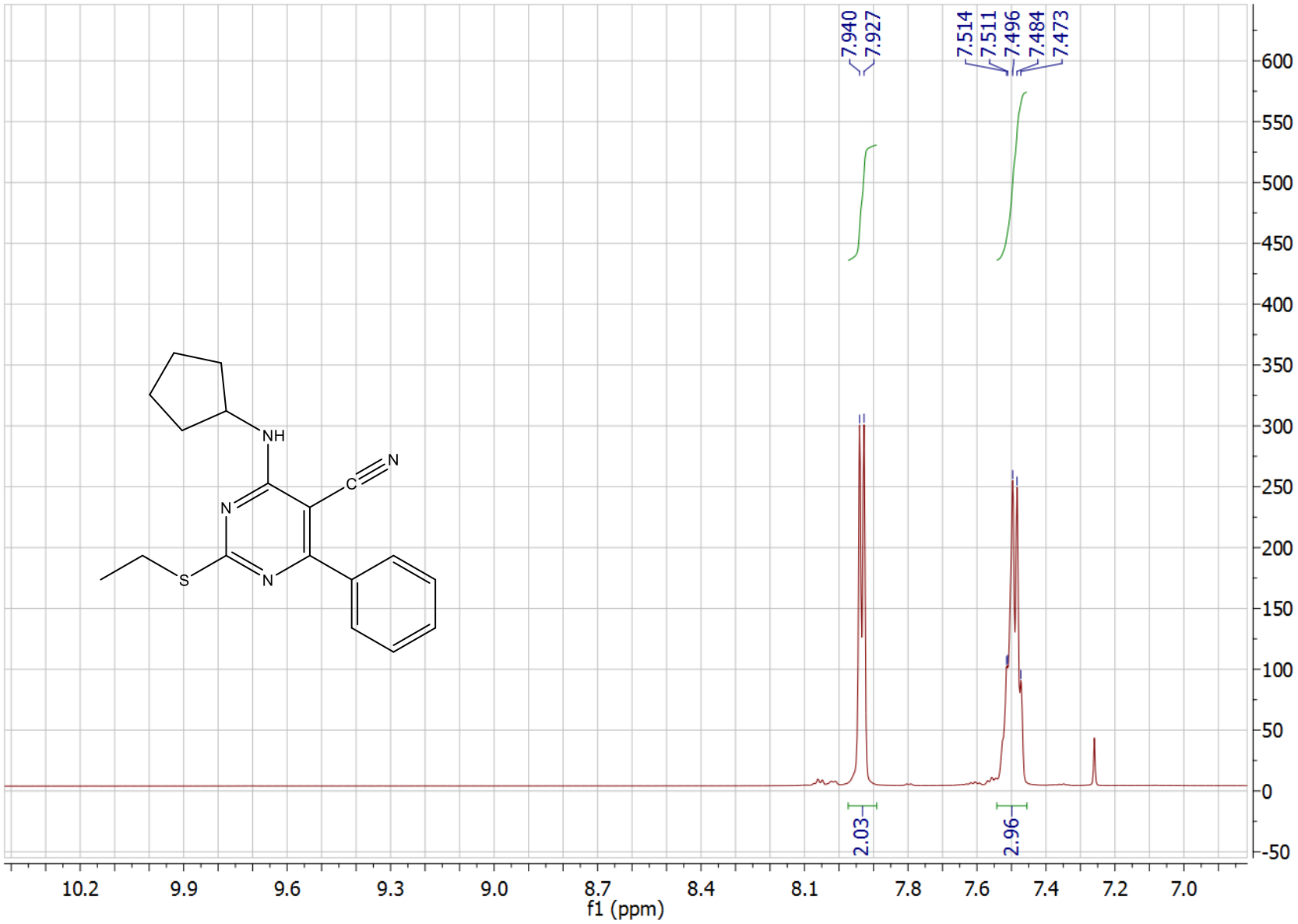
^{1}H NMR of compound 13$_a$



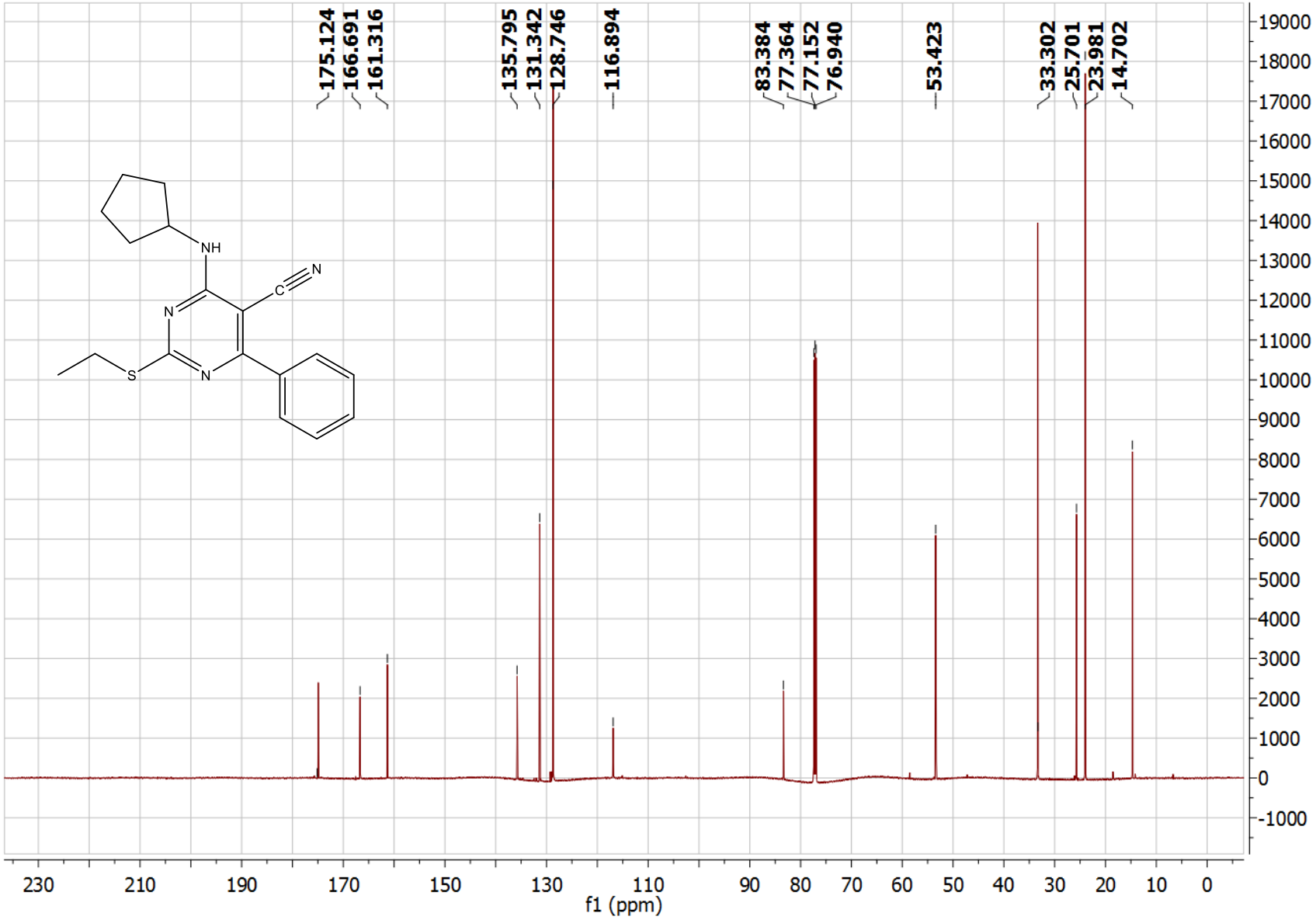
^{1}H NMR of compound 13$_a$



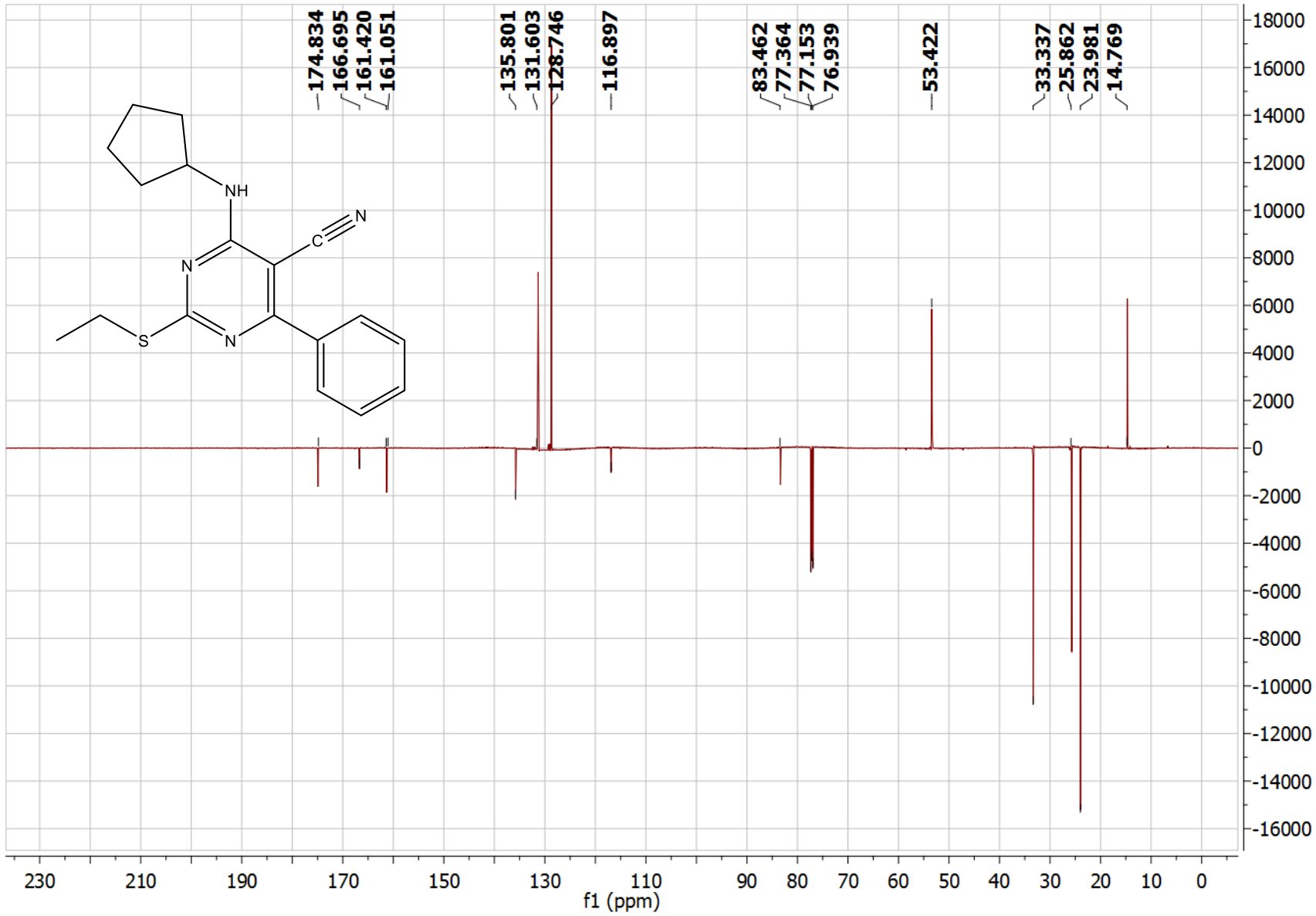
^{1}H NMR of compound 13_a



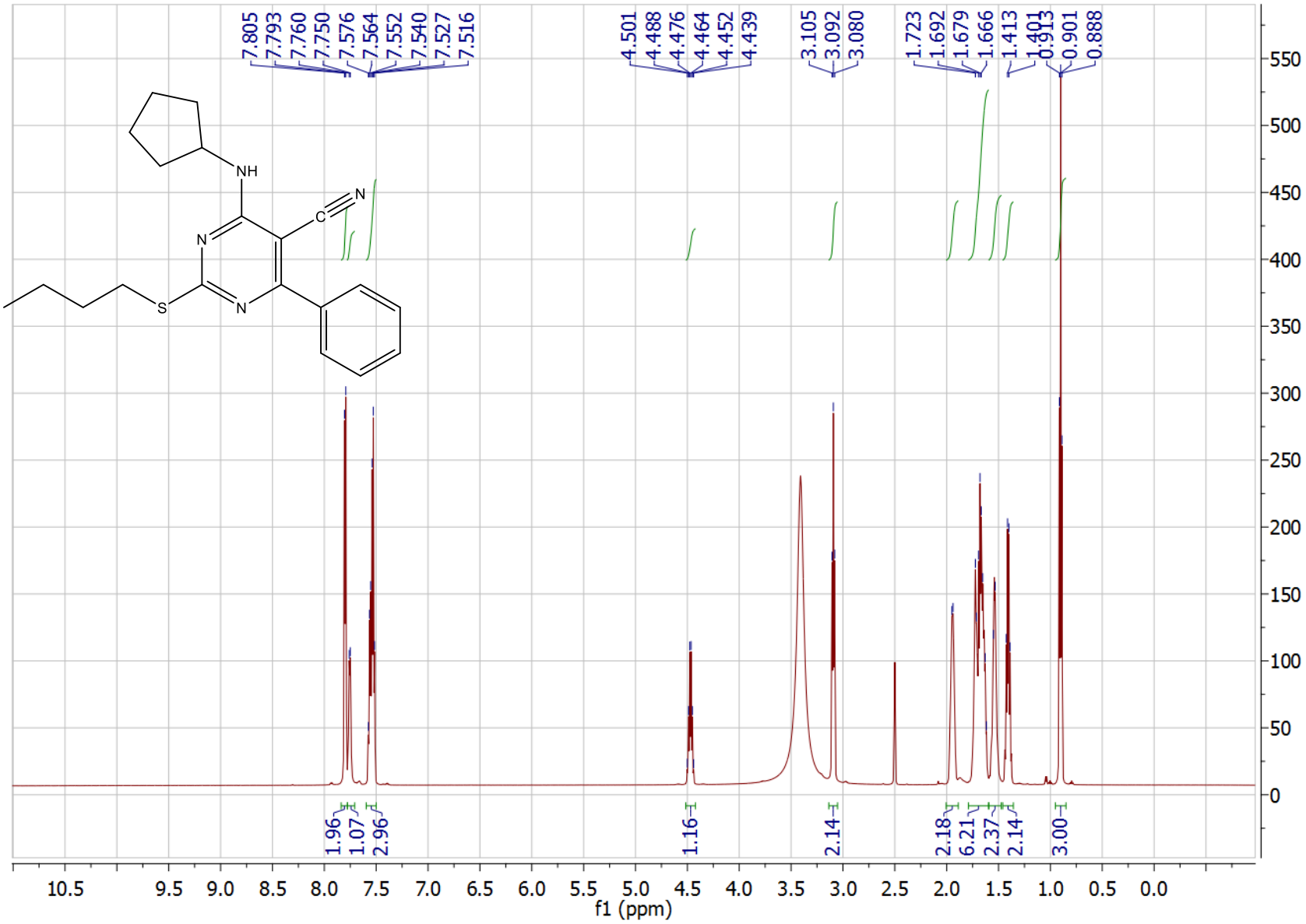
^{13}C NMR of compound 13_a



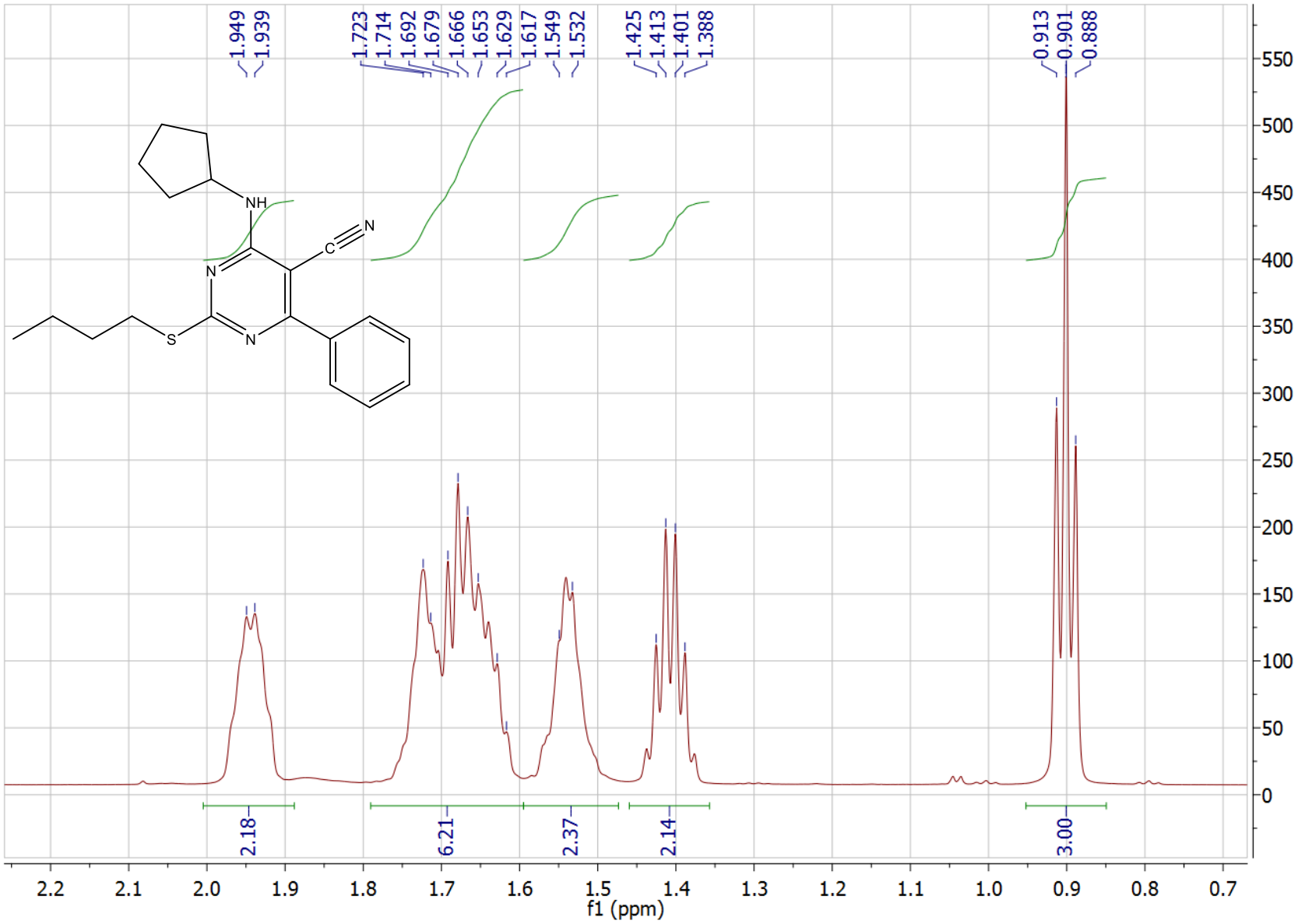
APT of compound 13_a



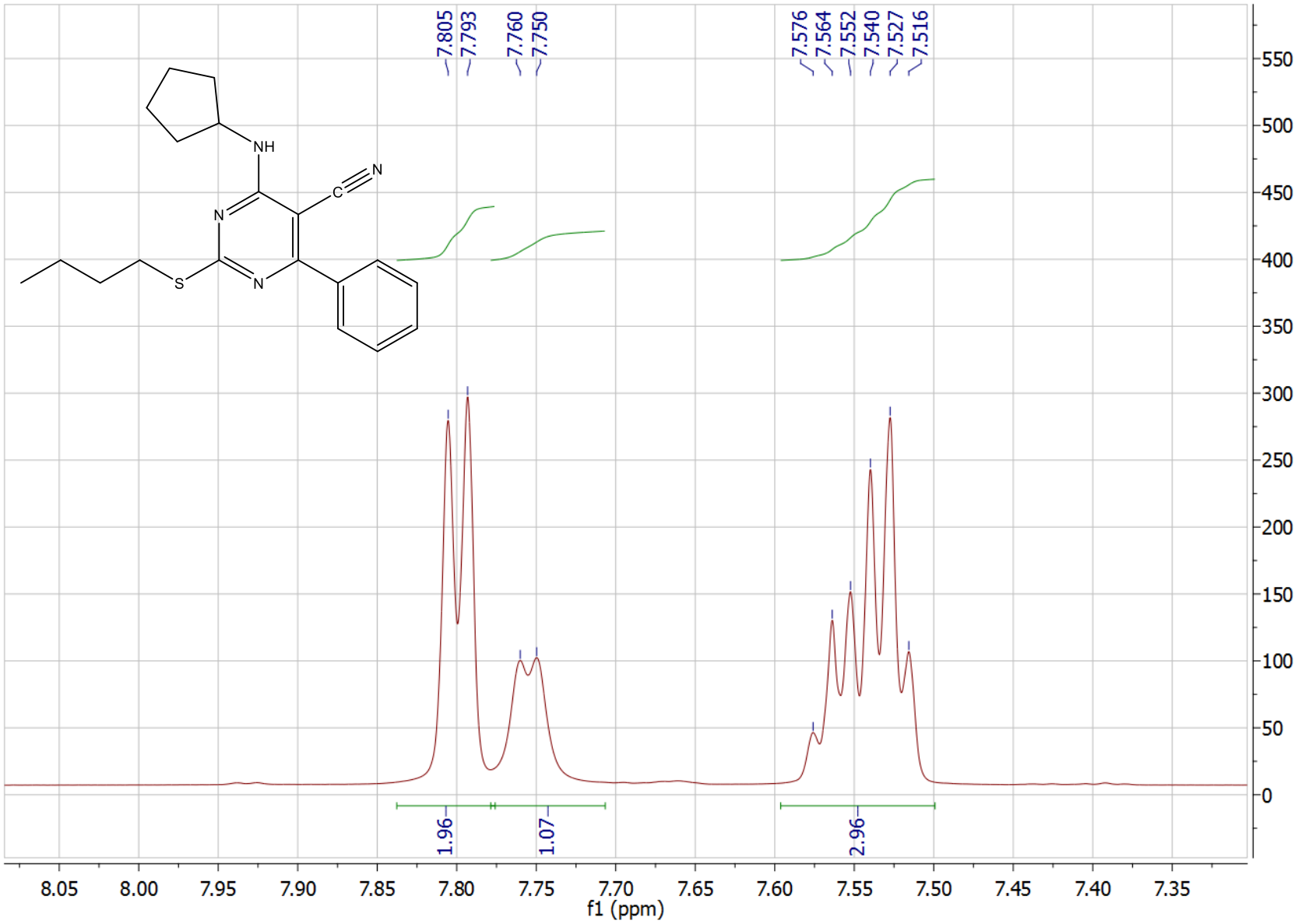
^{1}H NMR of compound 13_b



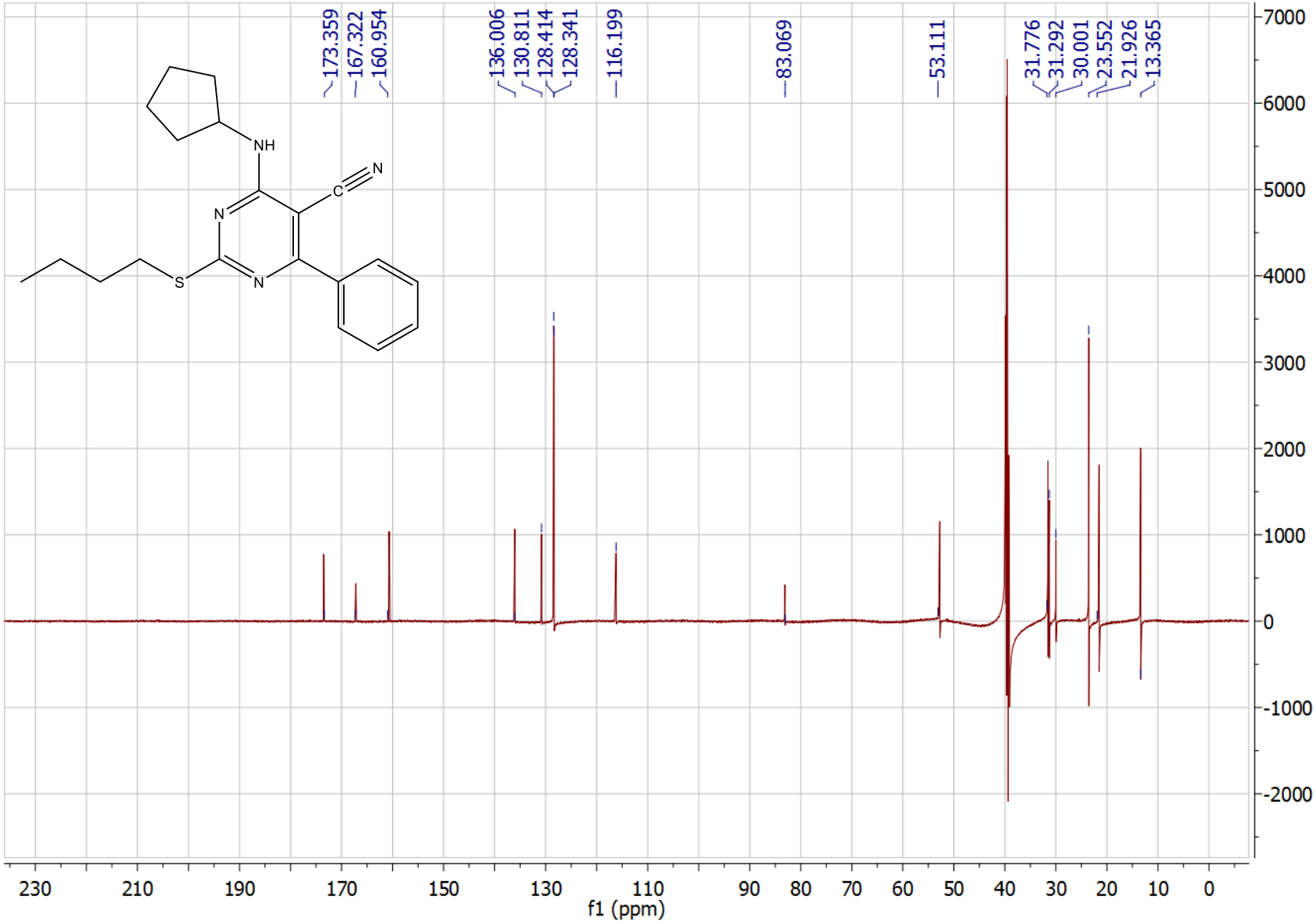
^{1}H NMR of compound 13_b



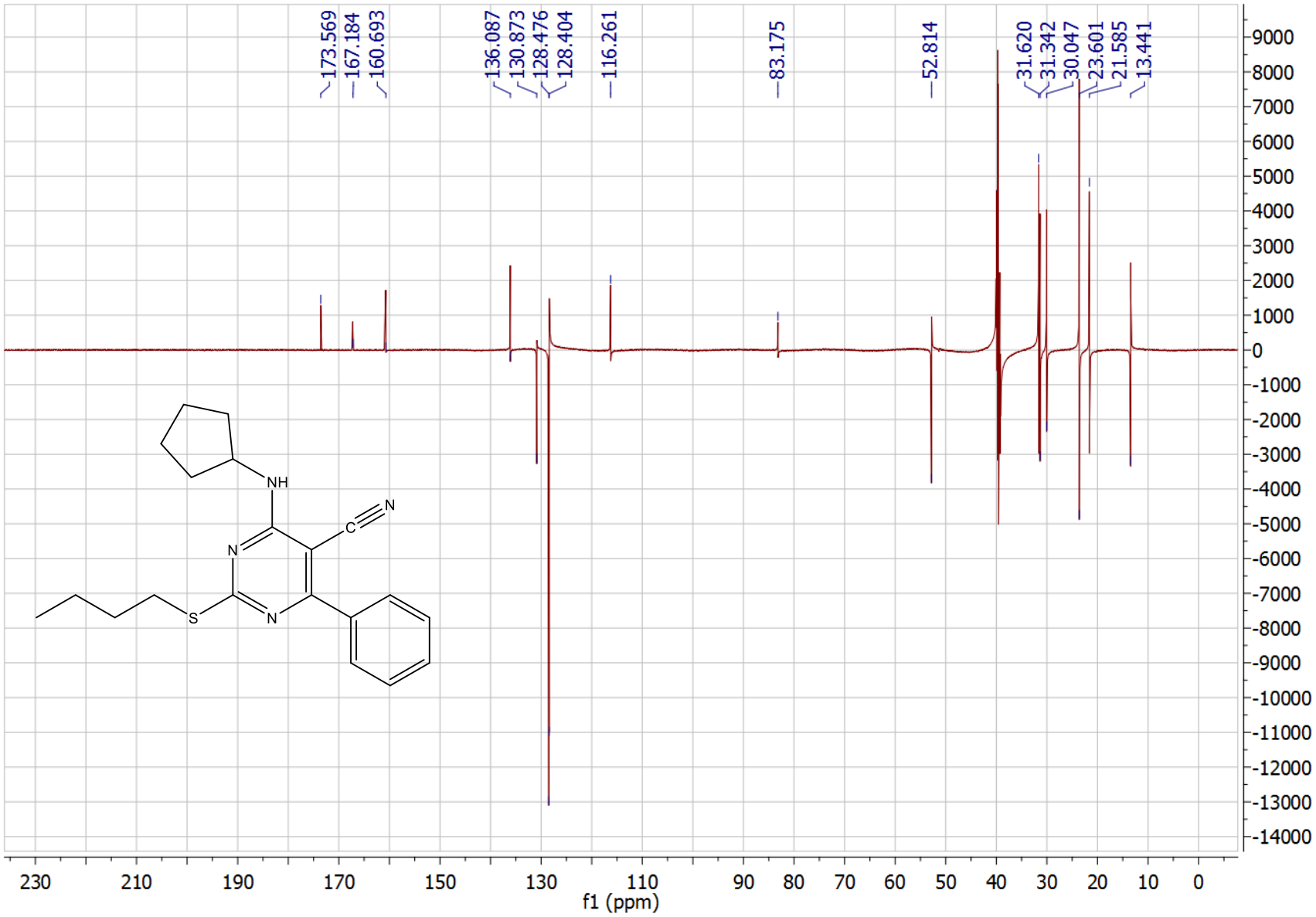
^{1}H NMR of compound 13_b



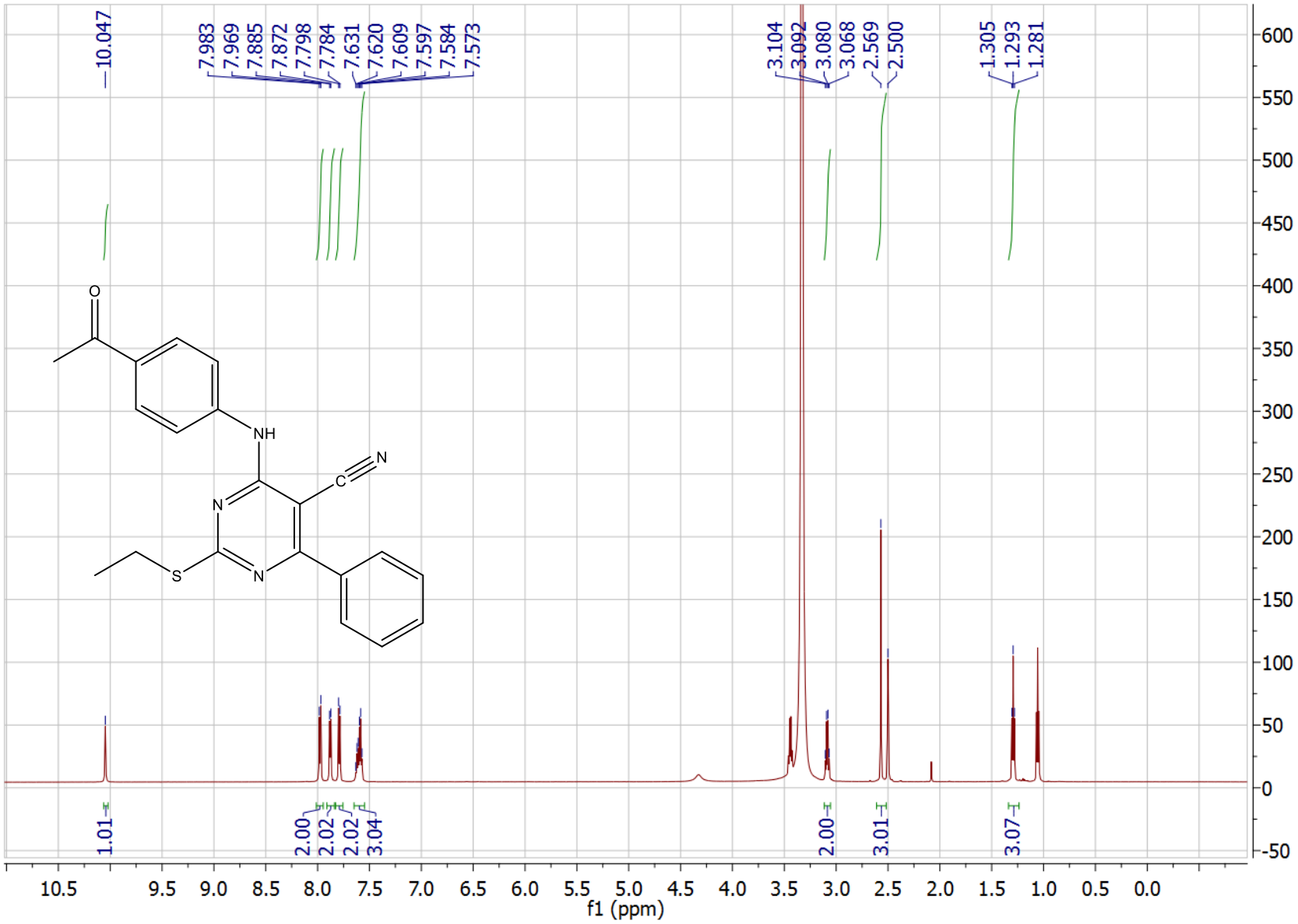
^{13}C NMR of compound 13_b



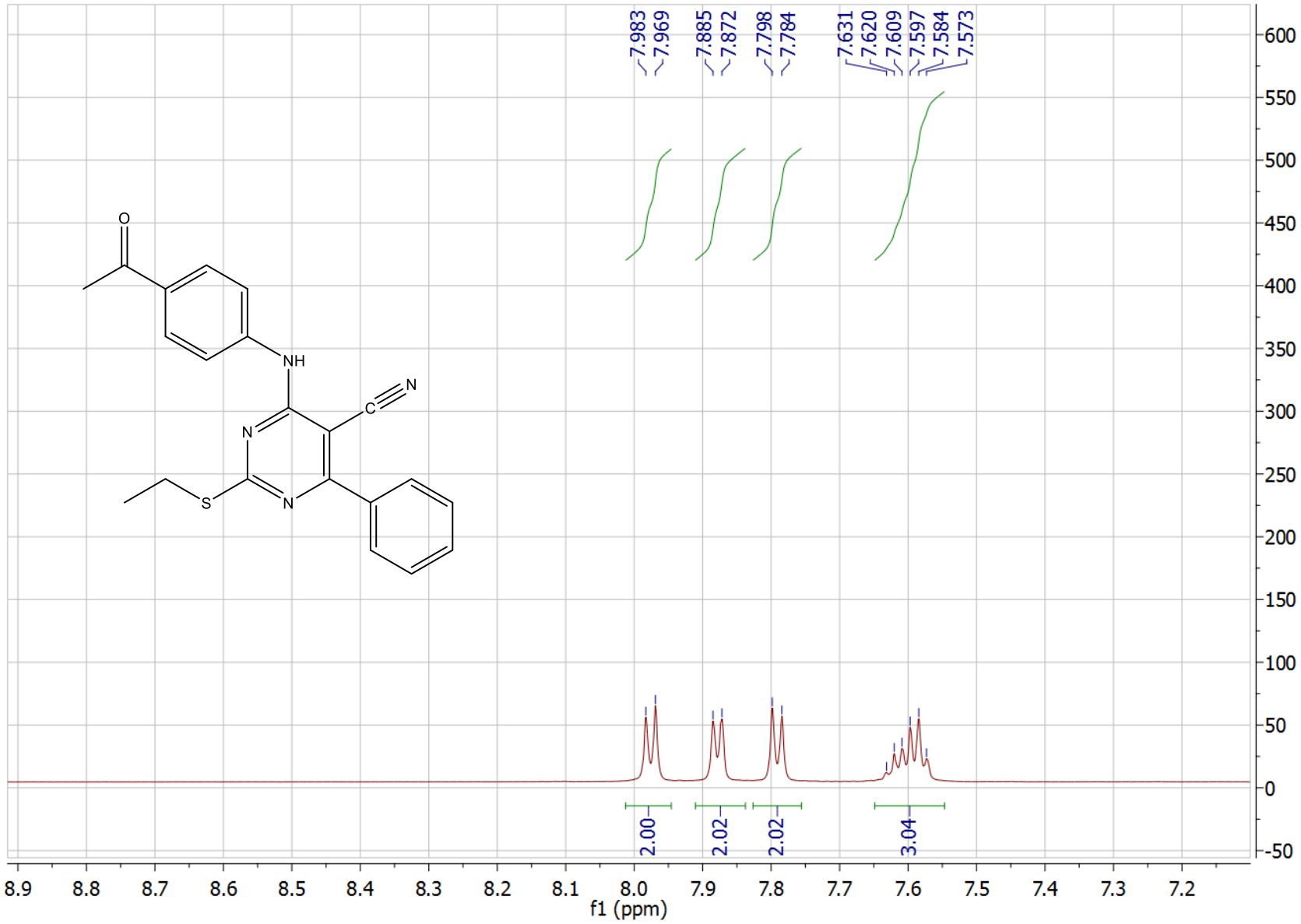
APT of compound 13_b



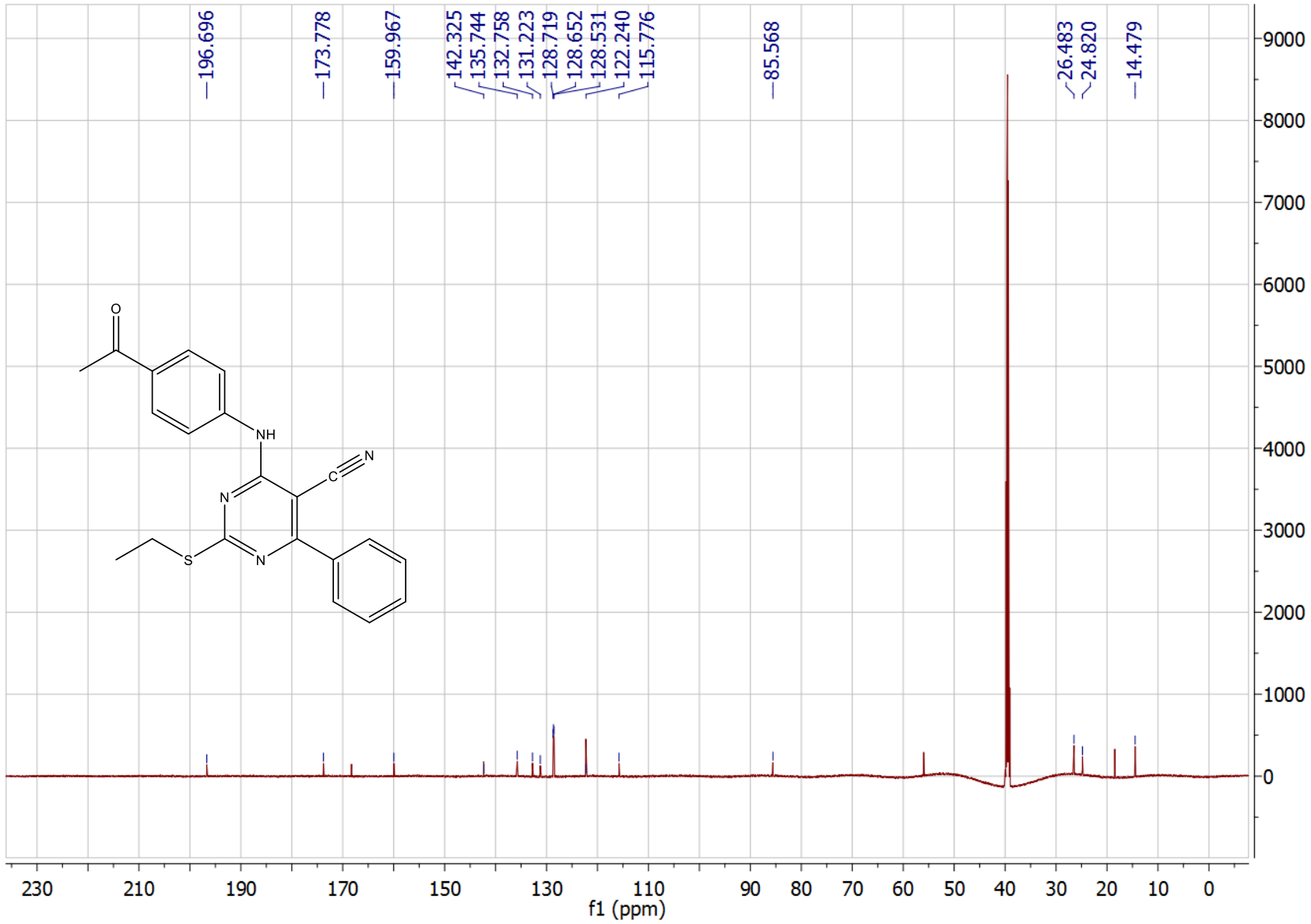
^{1}H NMR of compound 14_a



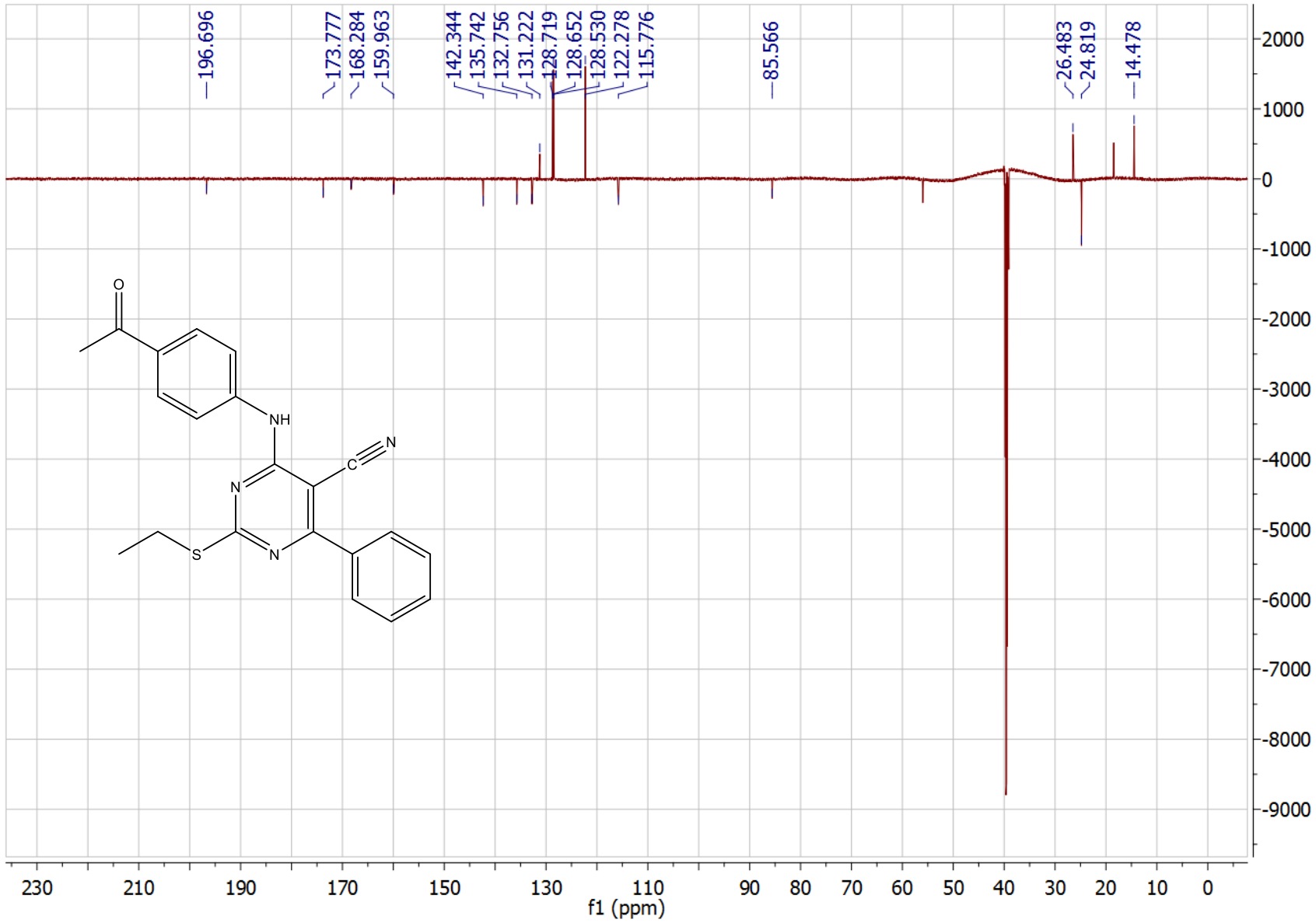
^{1}H NMR of compound 14_a



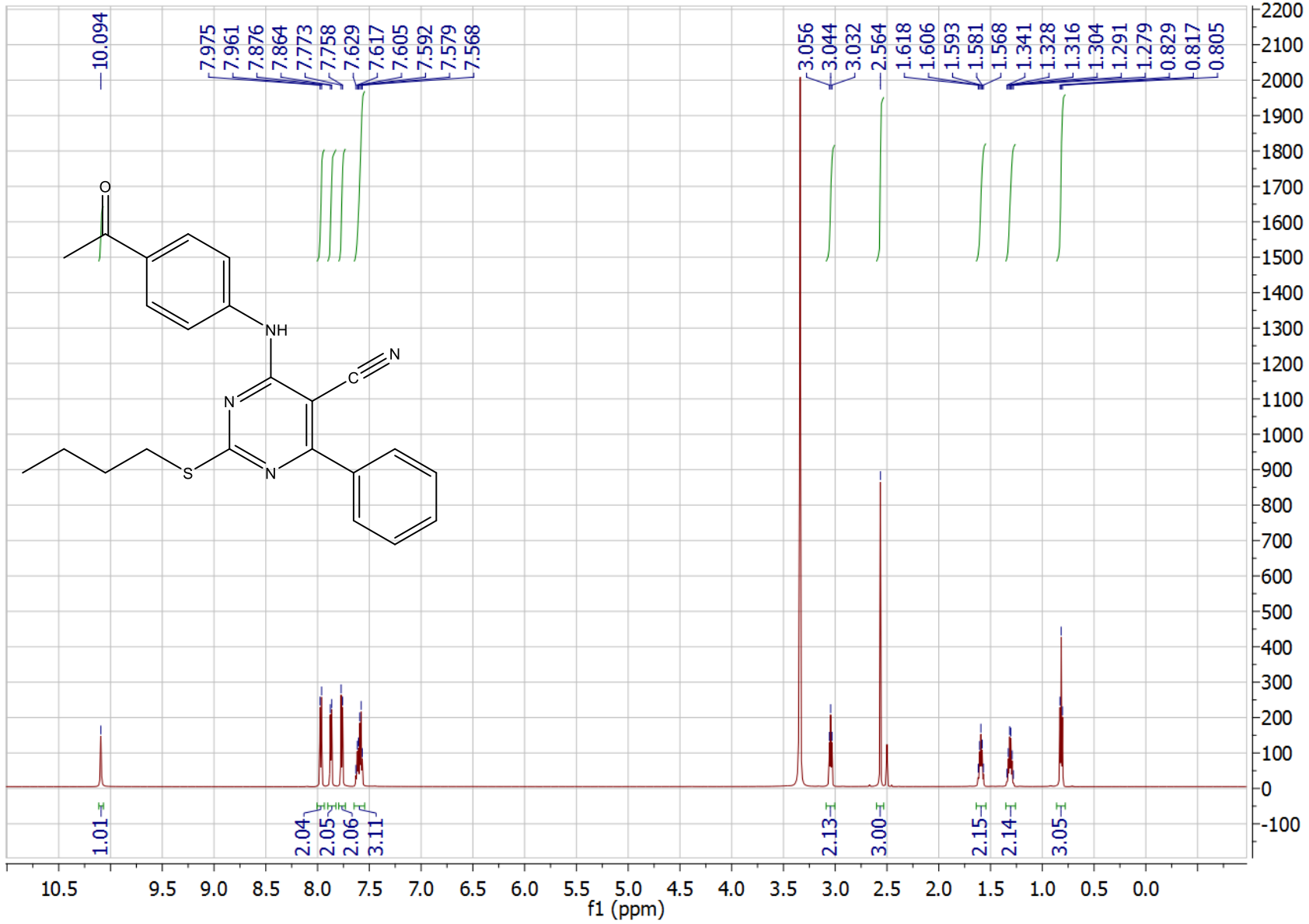
^{13}C NMR of compound 14_a



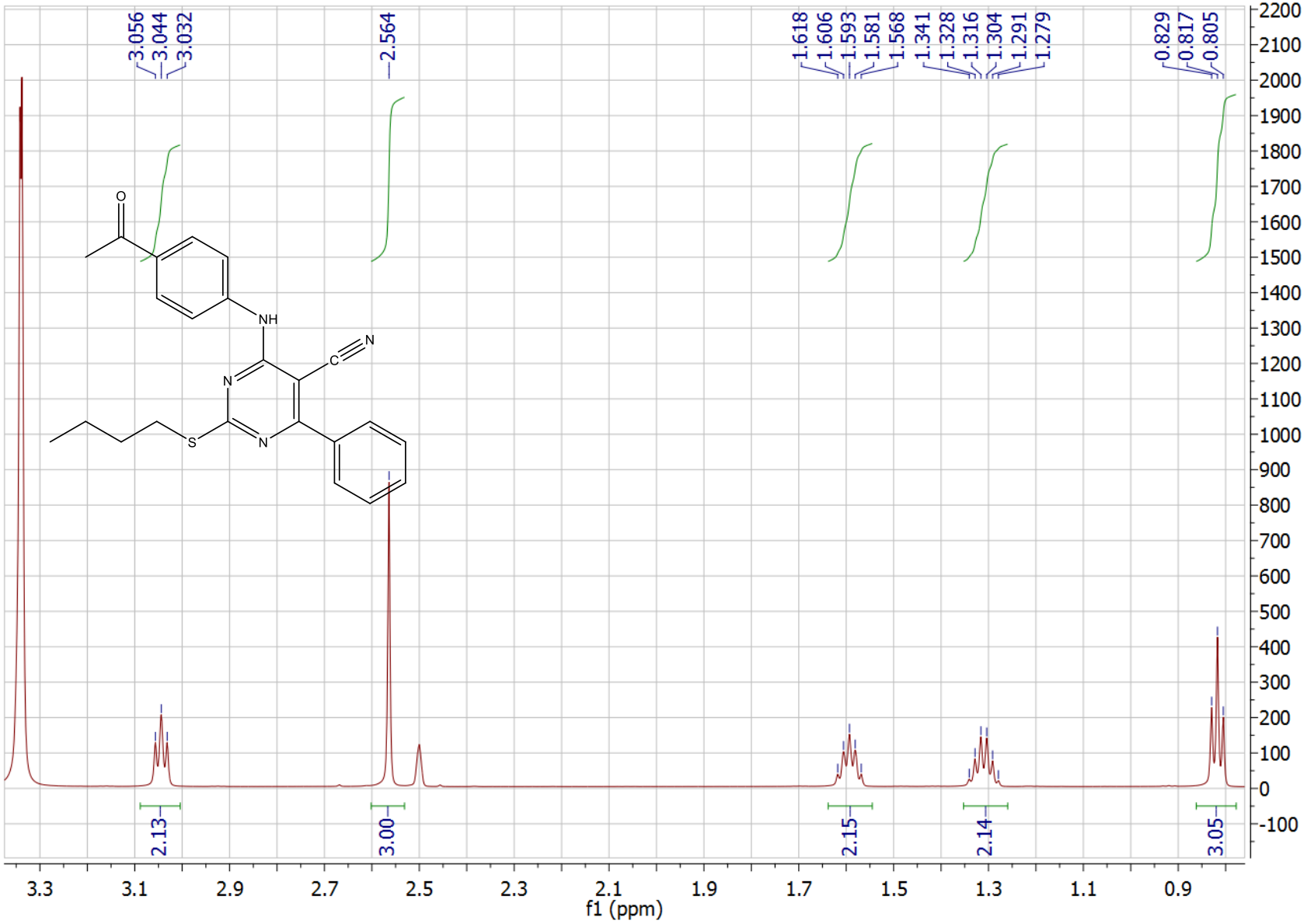
APT of compound 14_a



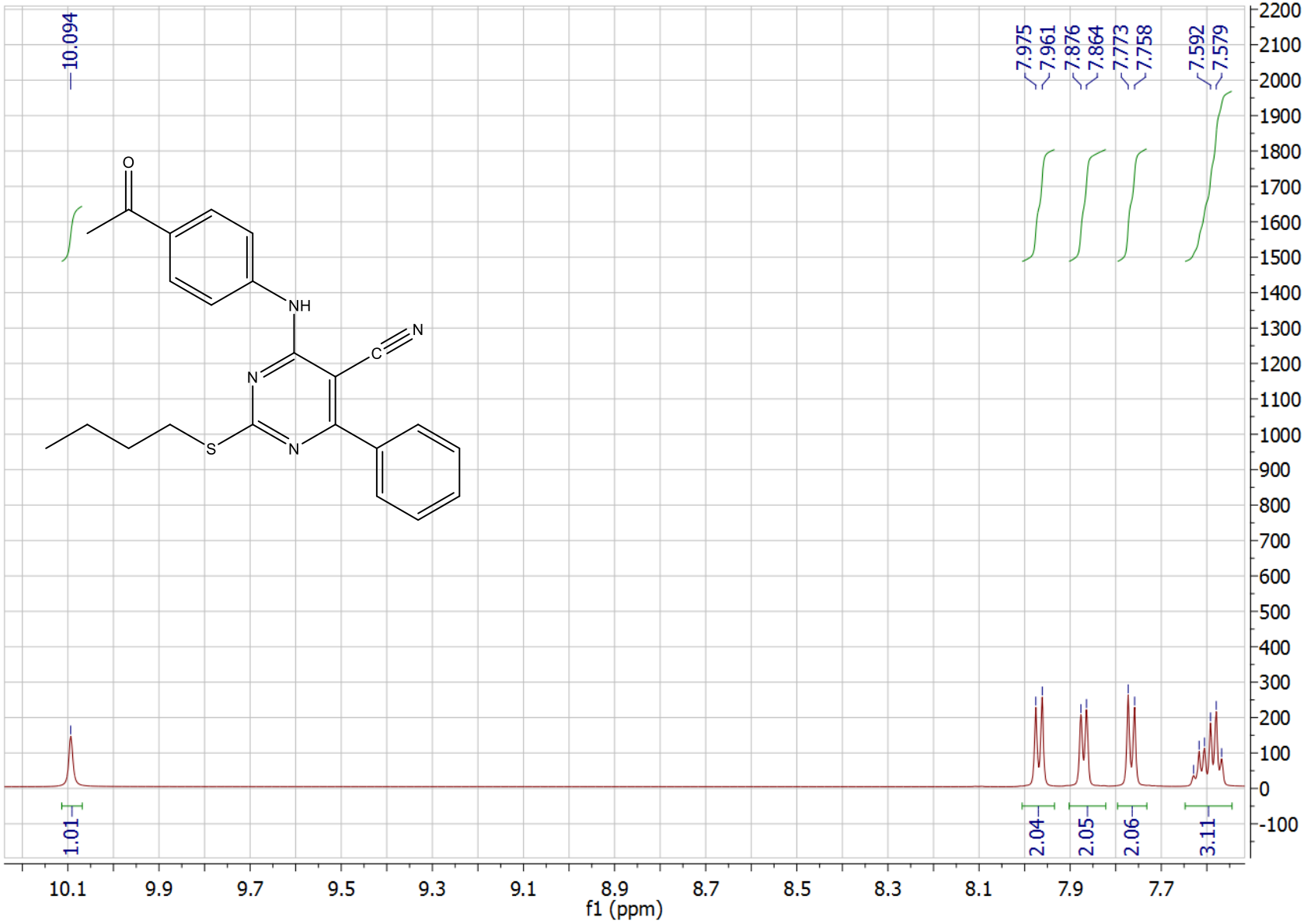
^{1}H NMR of compound 14_b



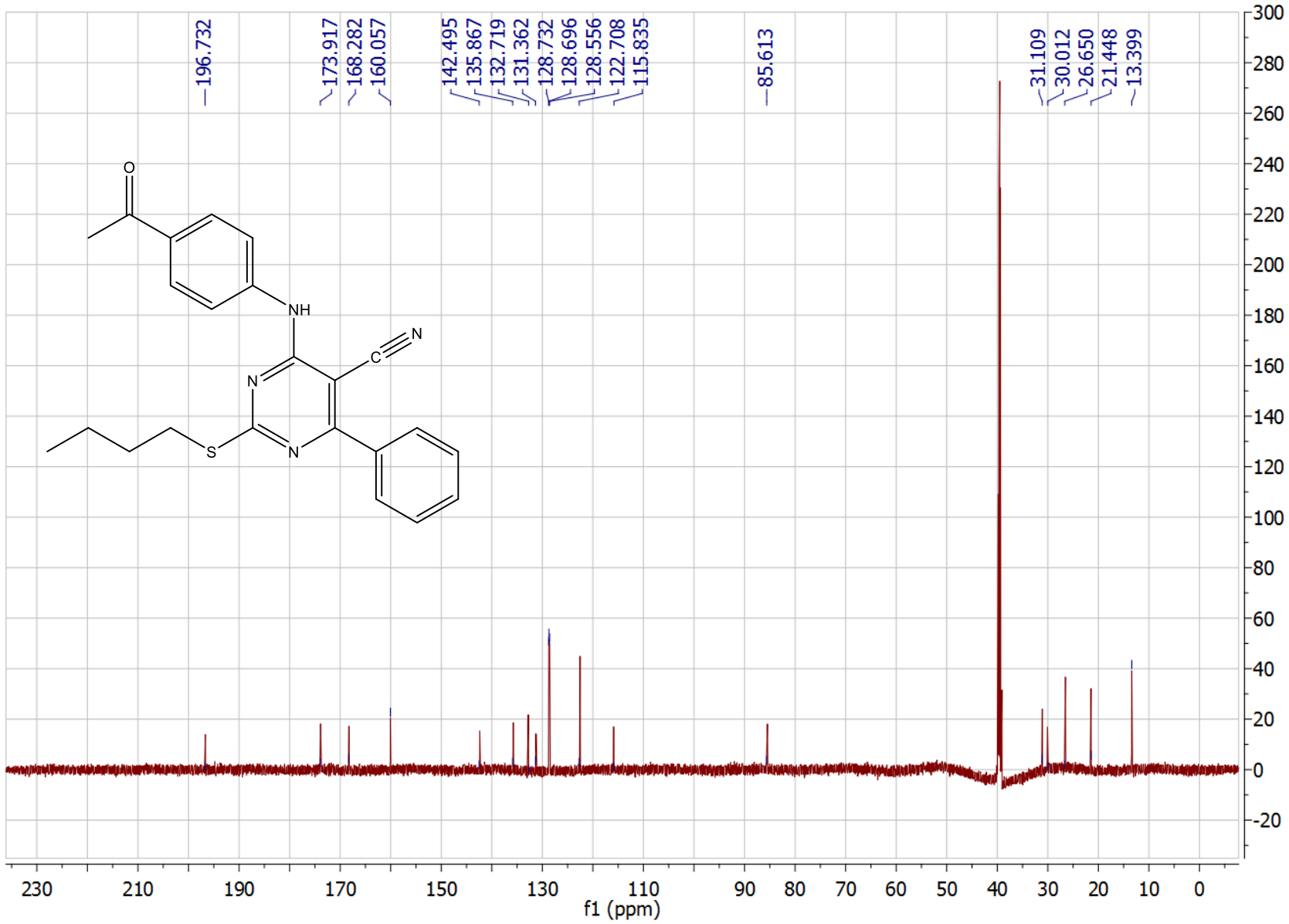
^{1}H NMR of compound 14_b



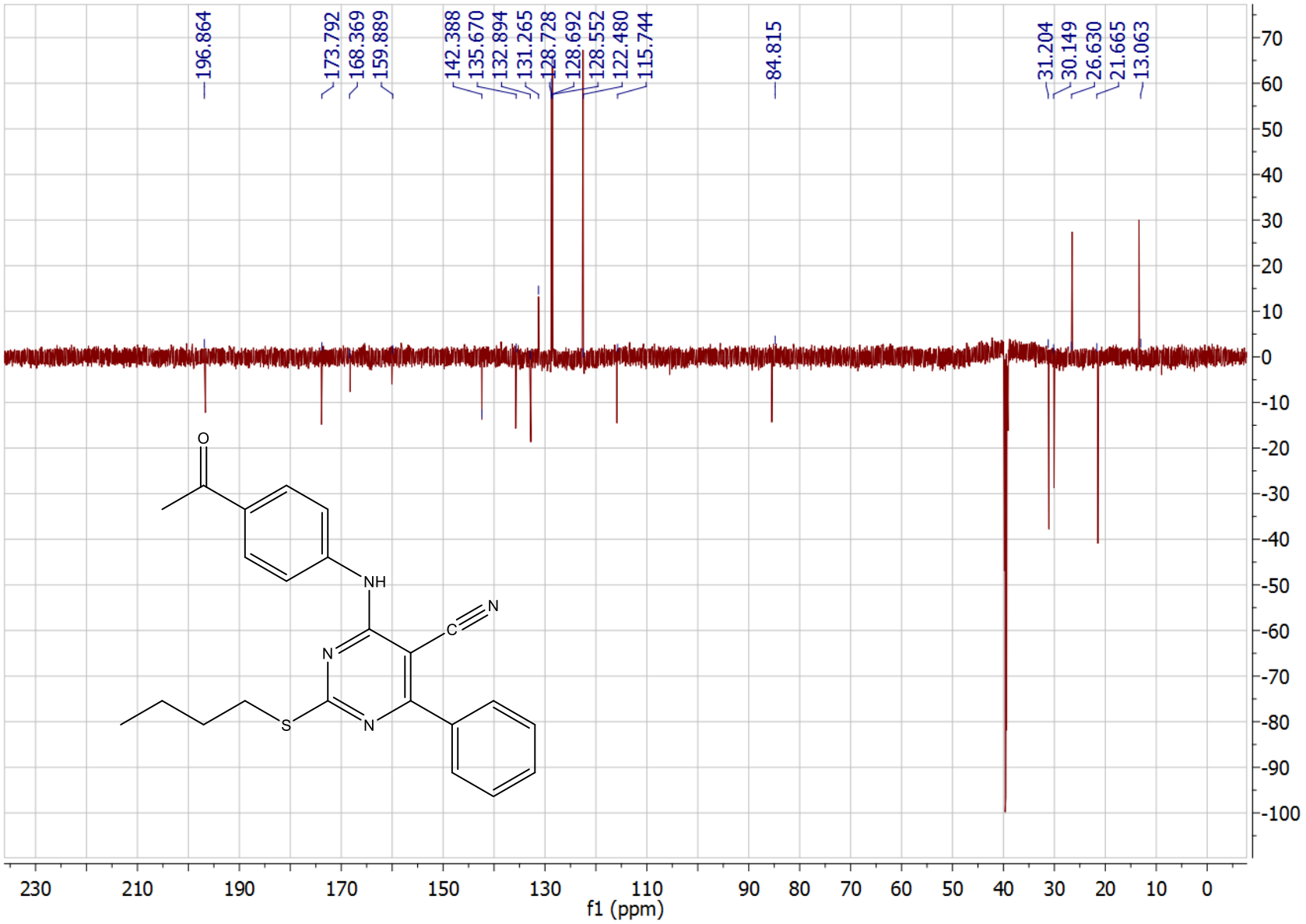
^{1}H NMR of compound 14_b



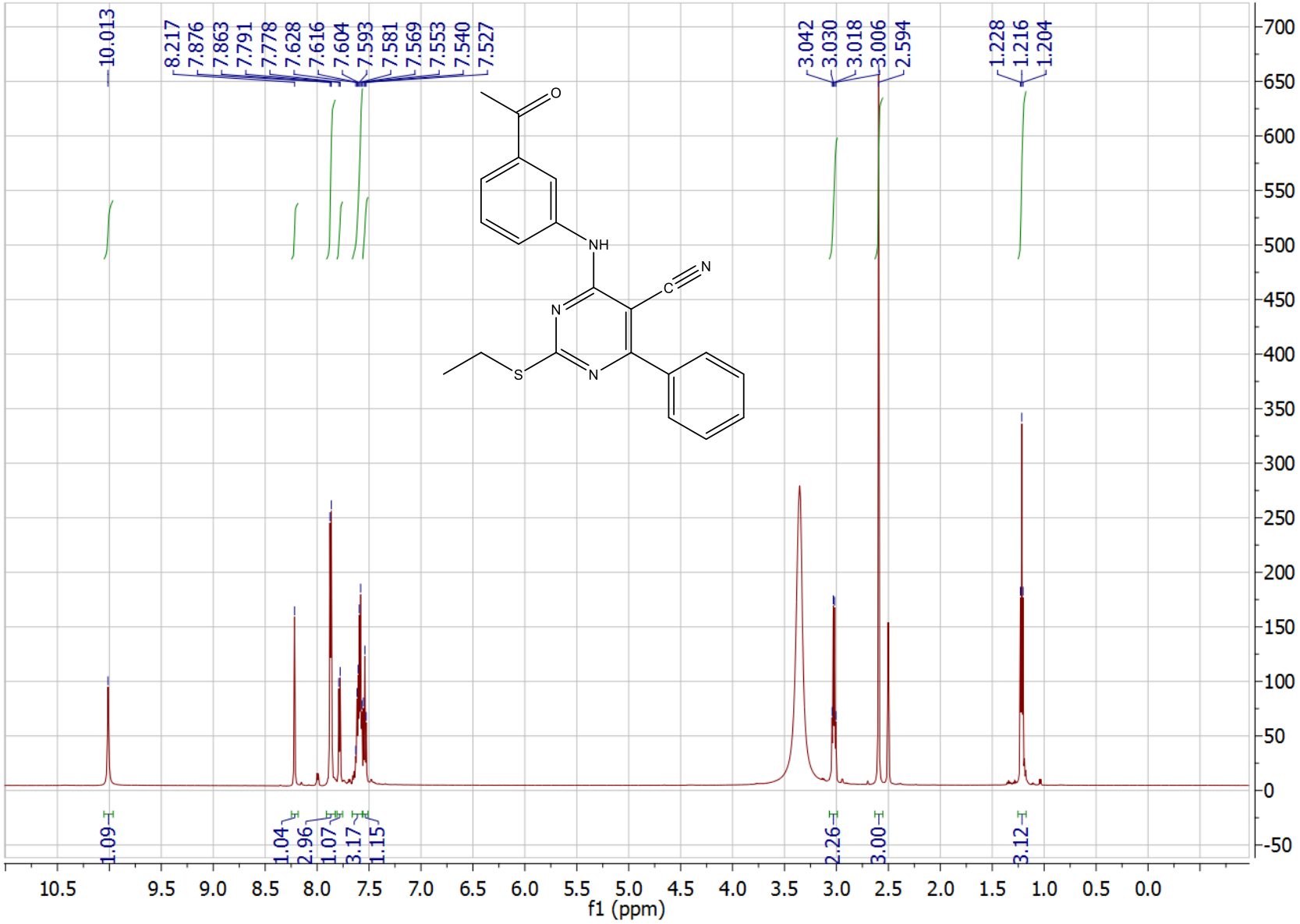
^{13}C NMR of compound 14_b



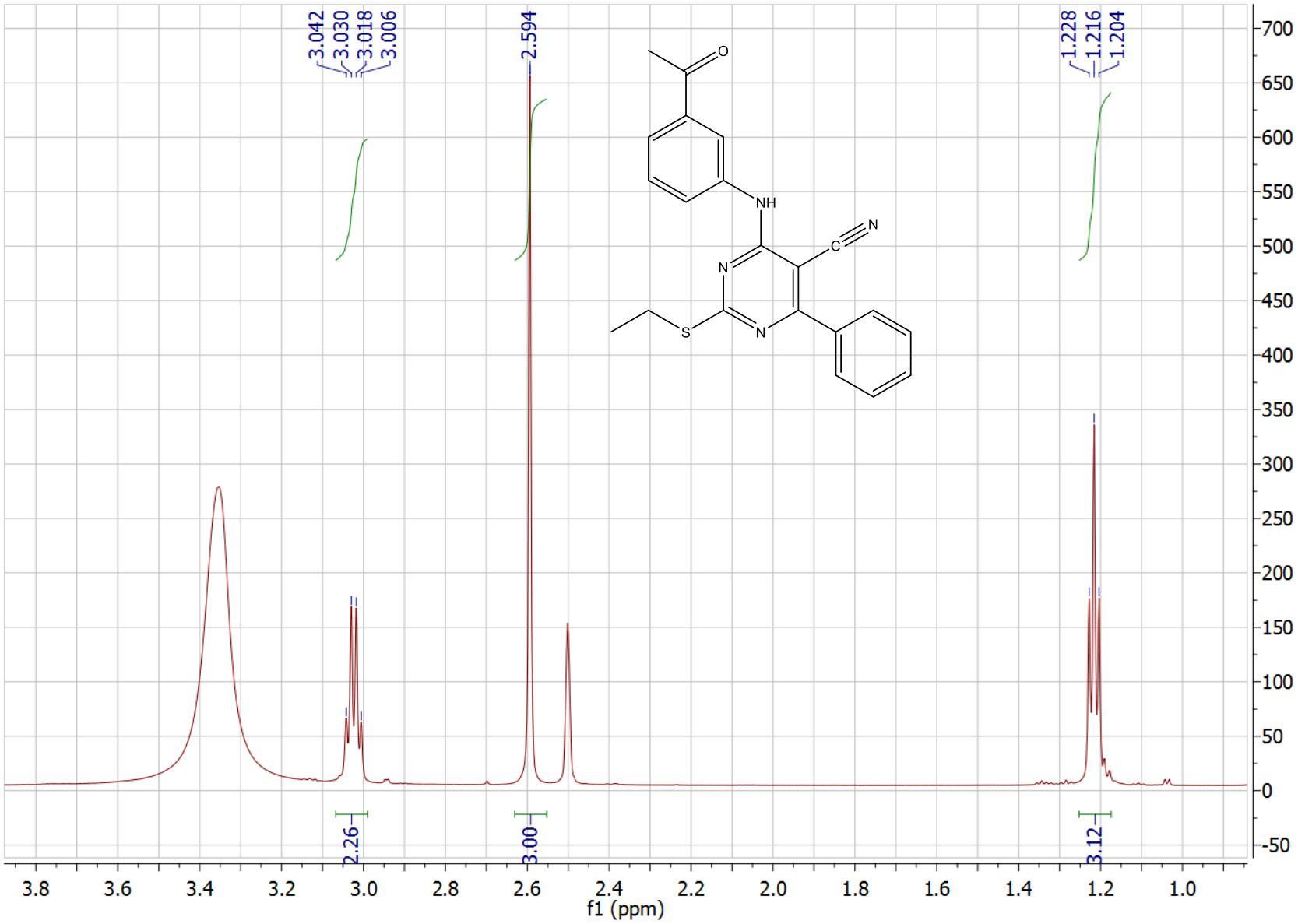
APT of compound 14_b



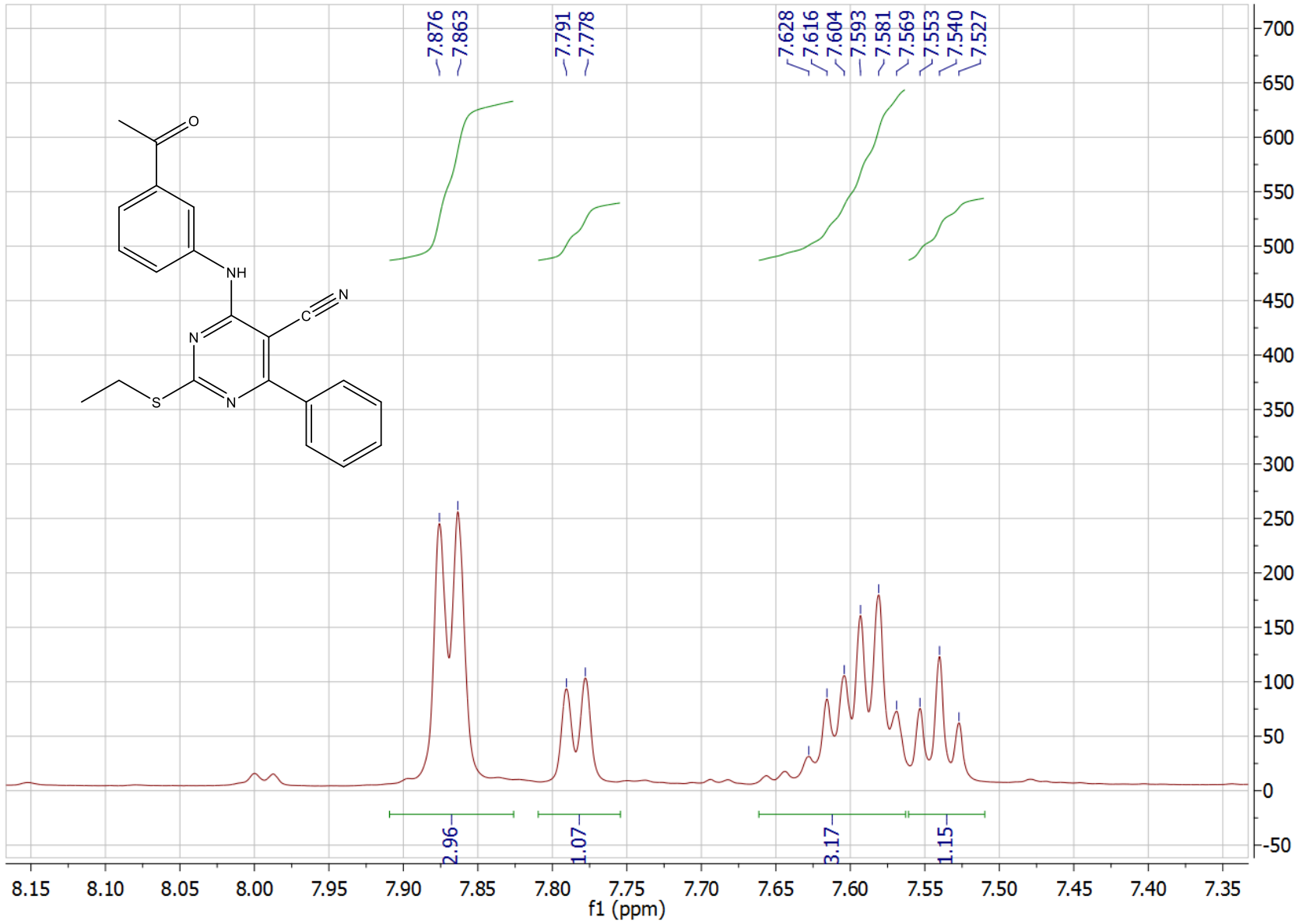
^{1}H NMR of compound 15$_a$



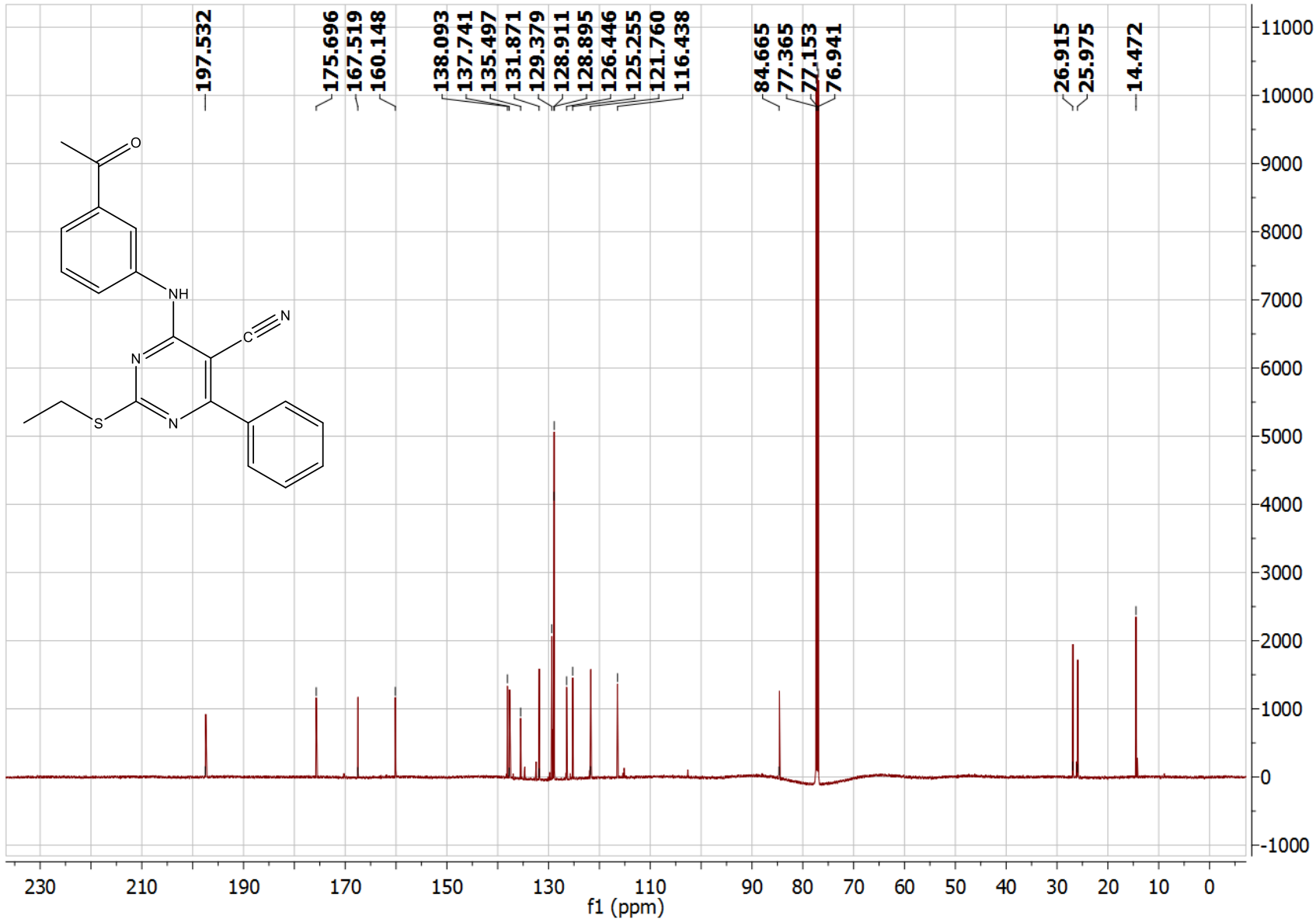
^{1}H NMR of compound 15_a



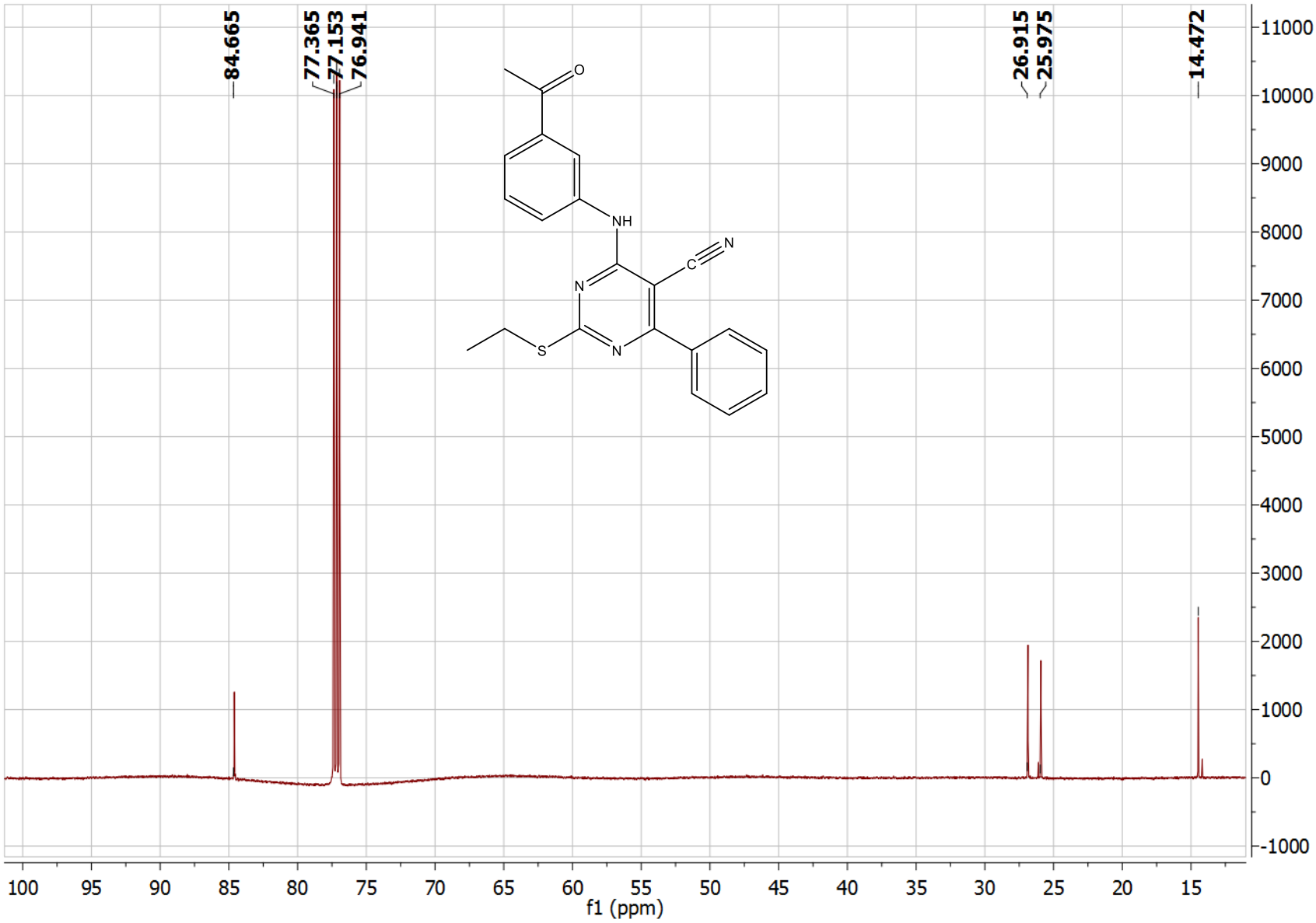
^{1}H NMR of compound 15$_a$



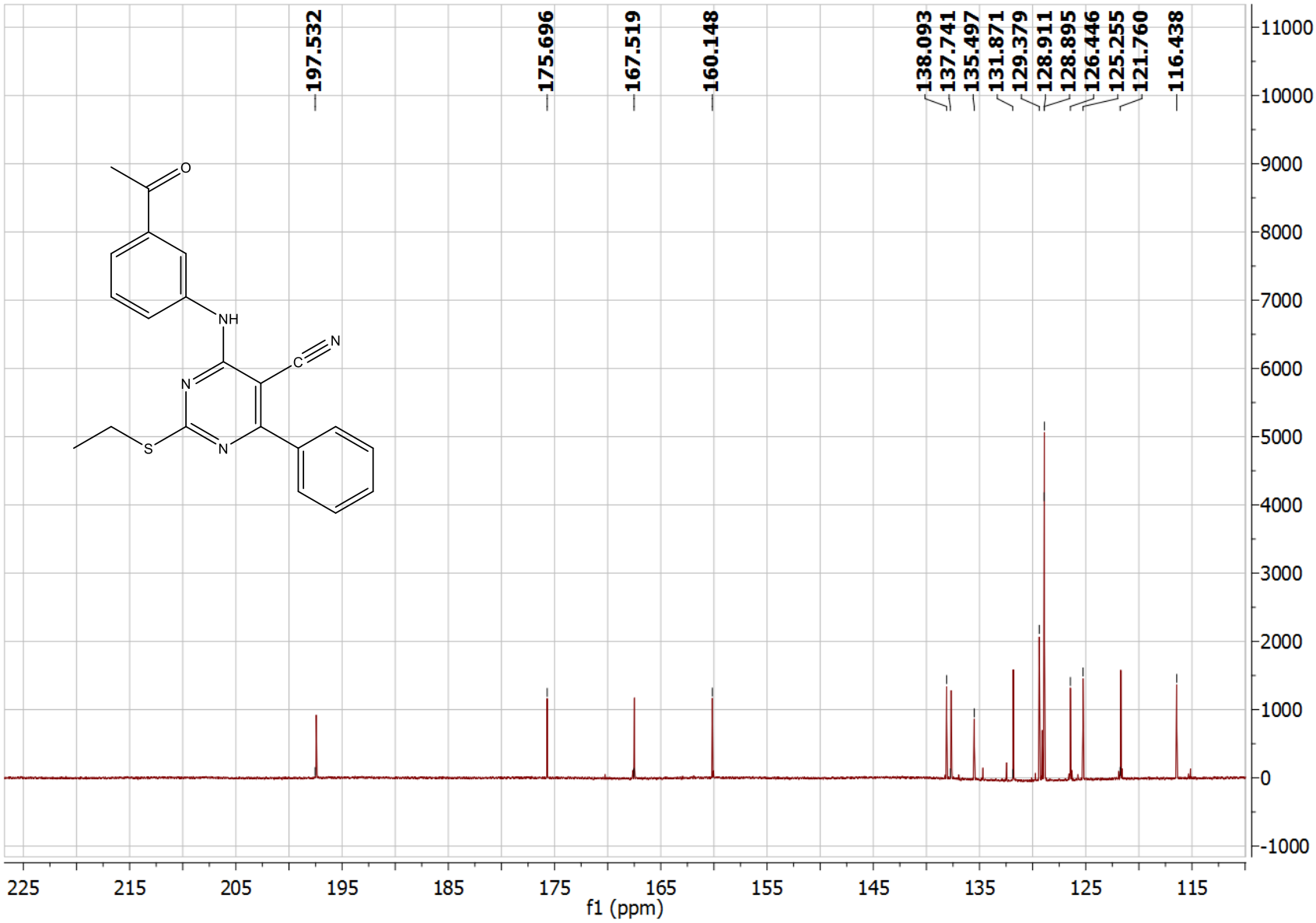
^{13}C NMR of compound 15$_a$



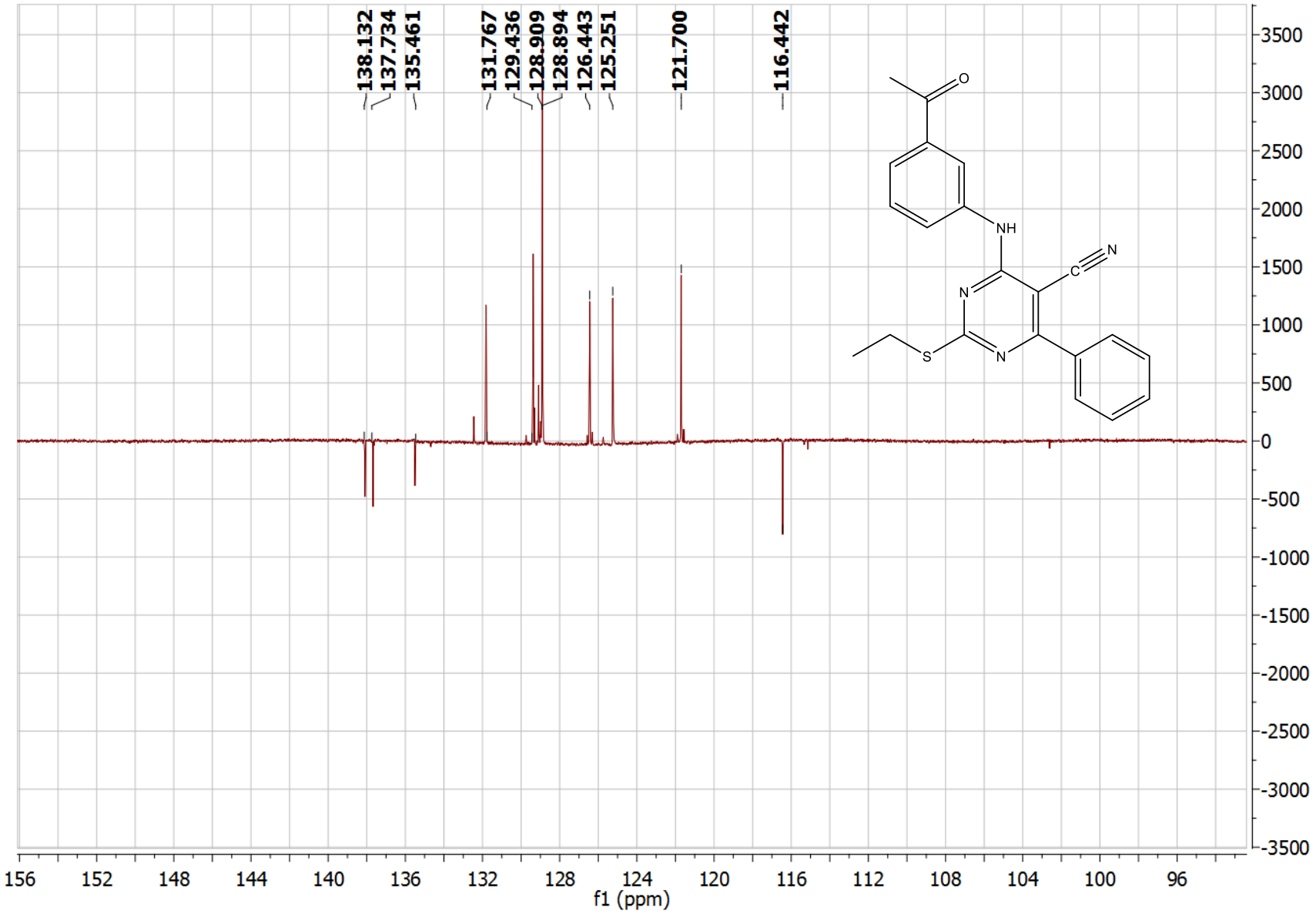
^{13}C NMR of compound 15_a



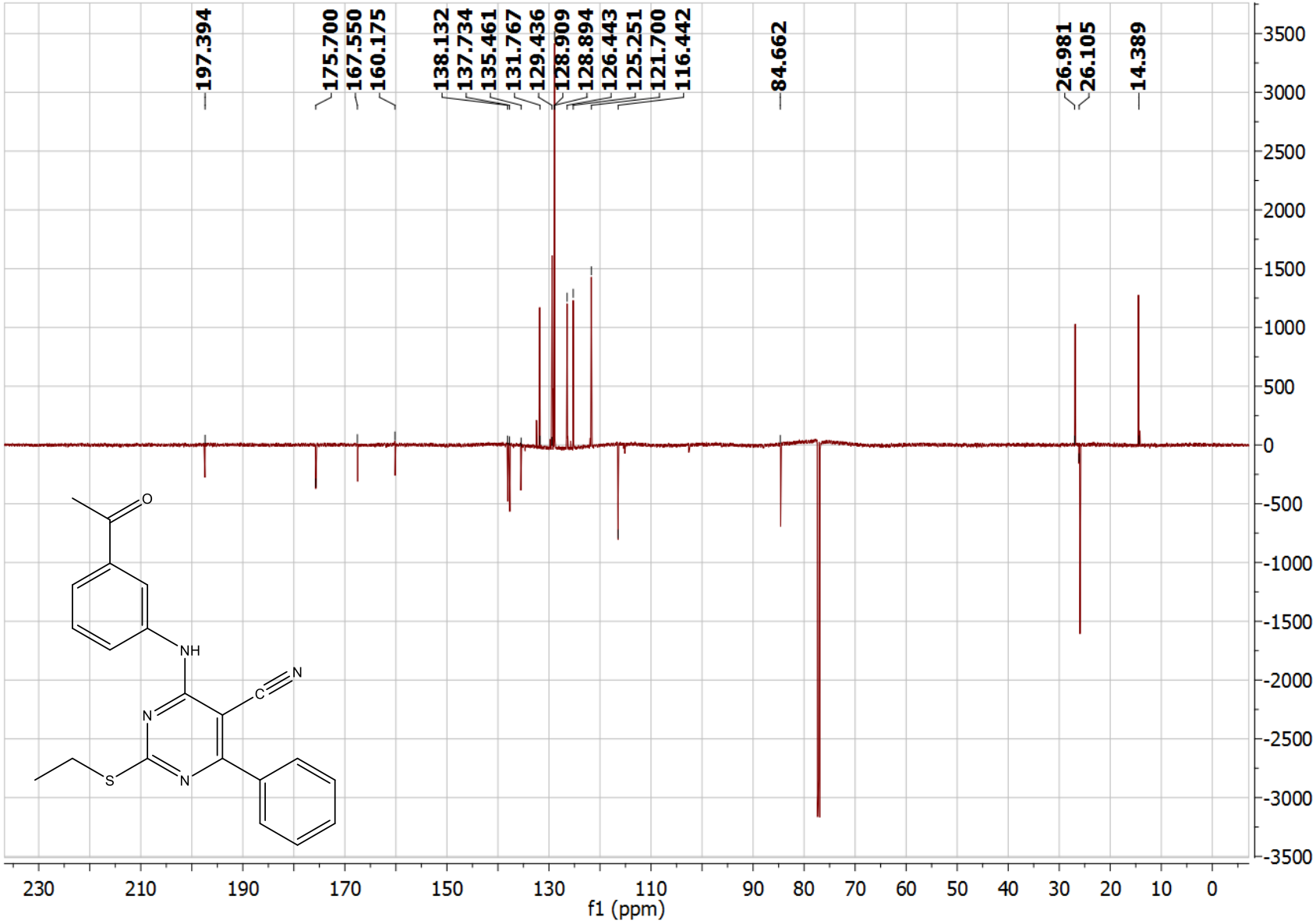
^{13}C NMR of compound 15_a



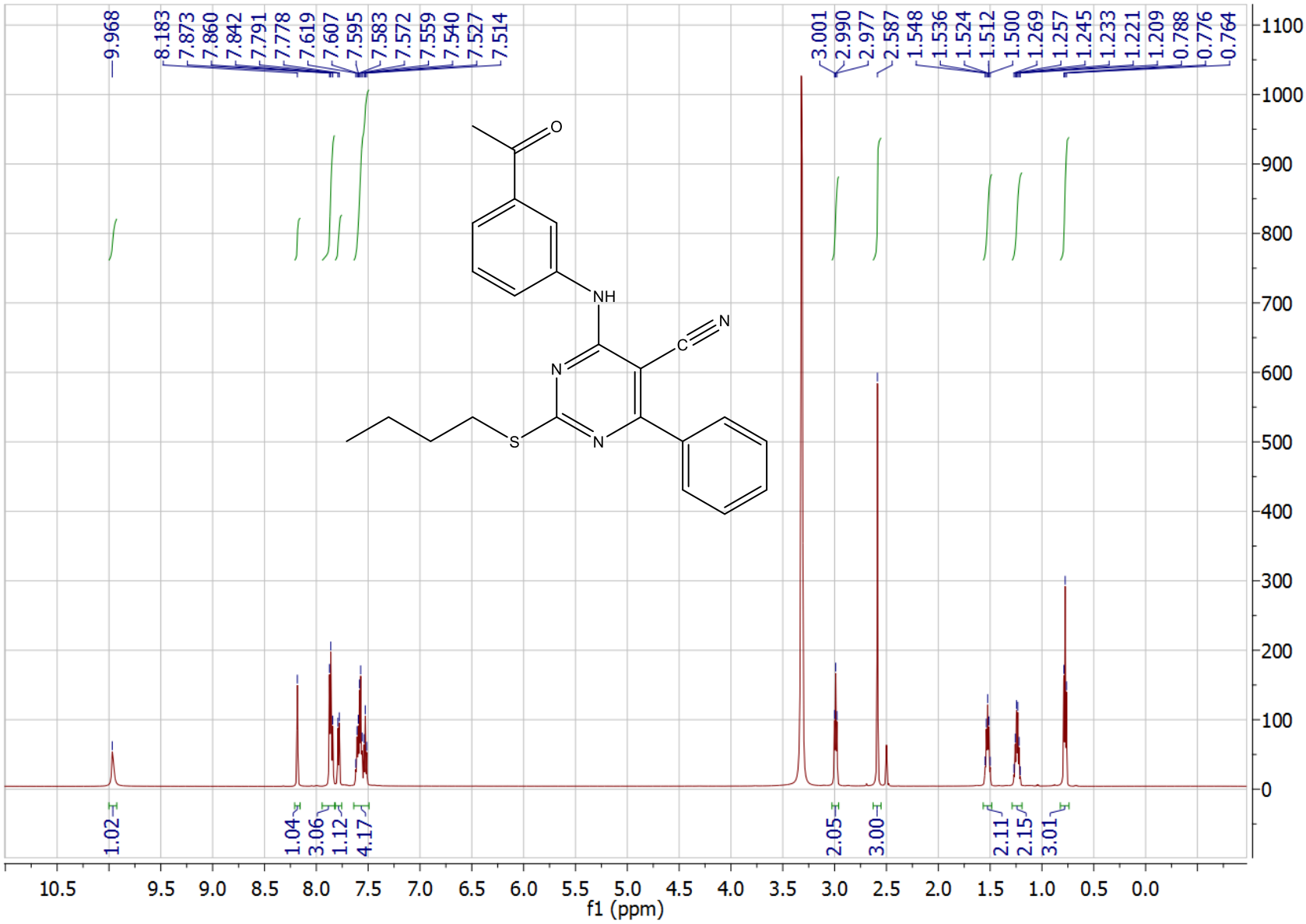
APT of compound 15_a



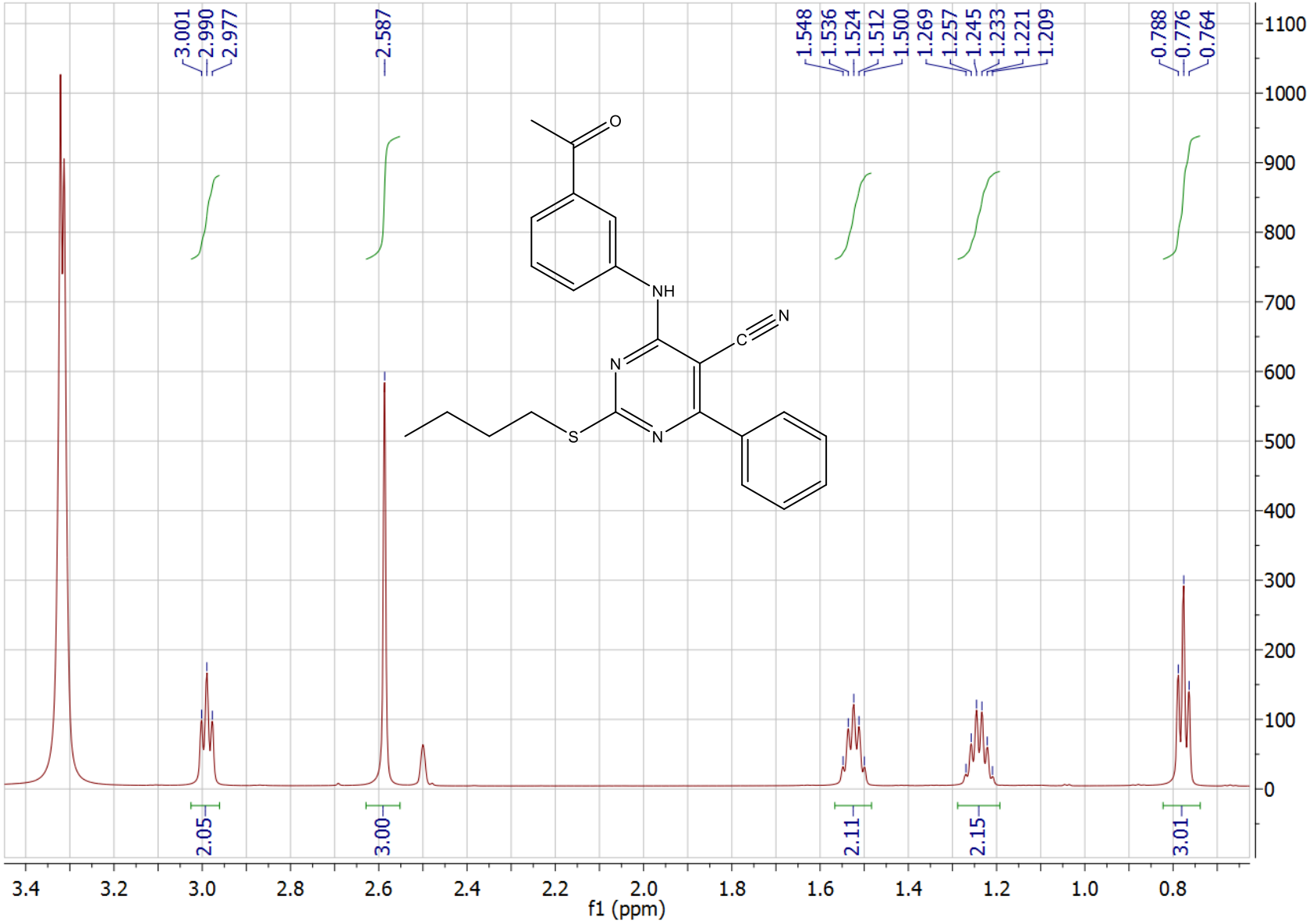
APT of compound 15_a



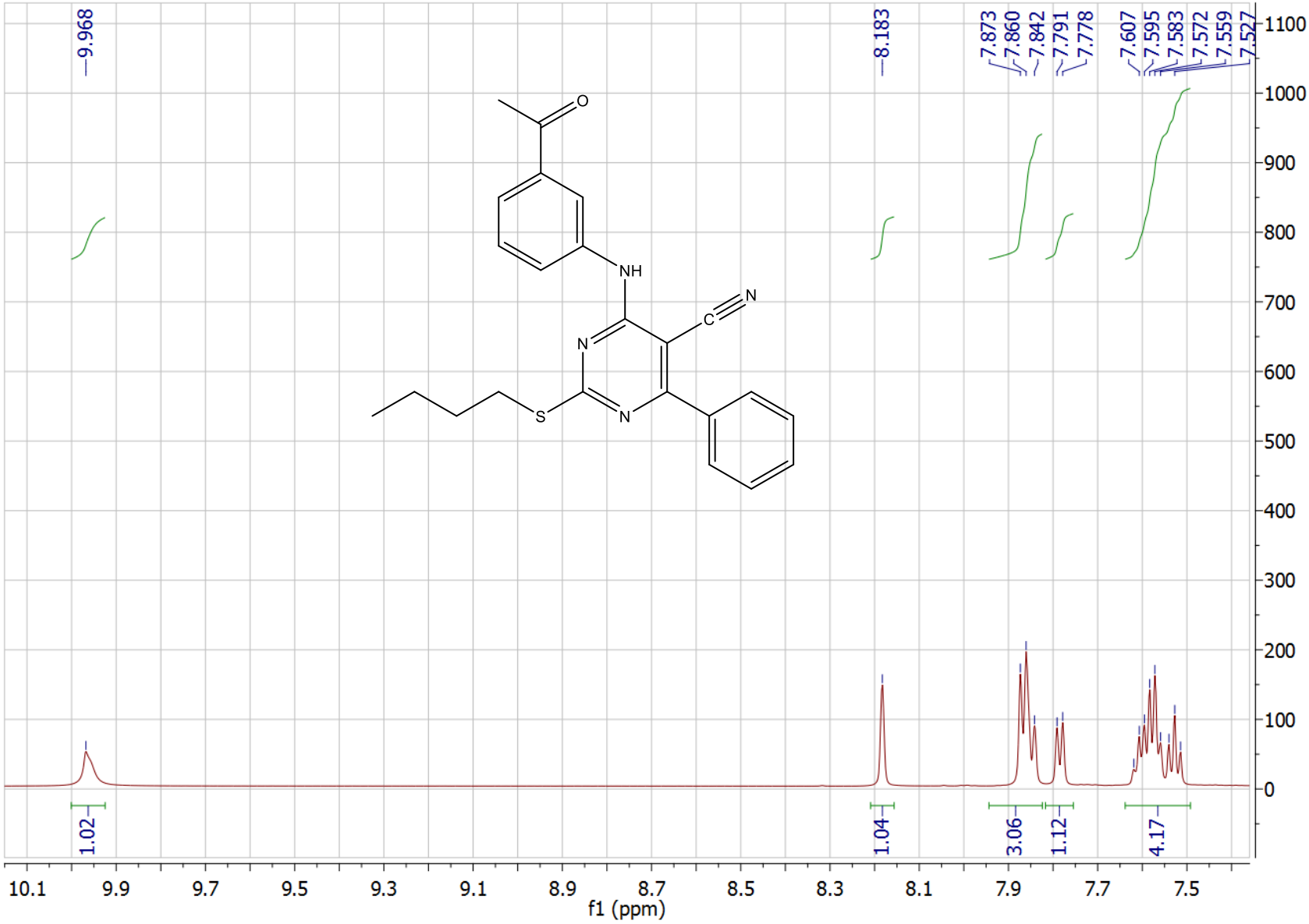
^{1}H NMR of compound 15_b



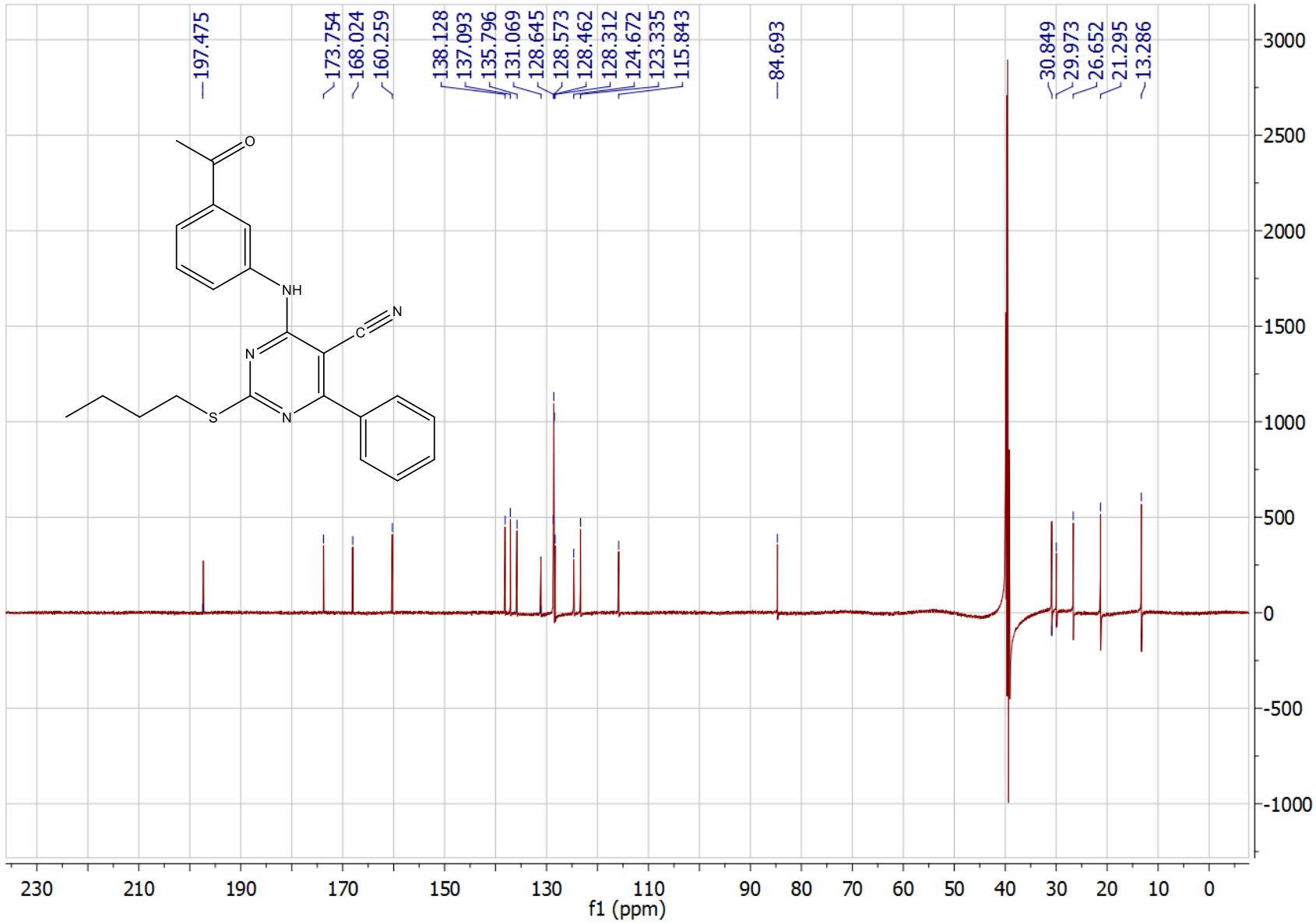
^{1}H NMR of compound 15_b



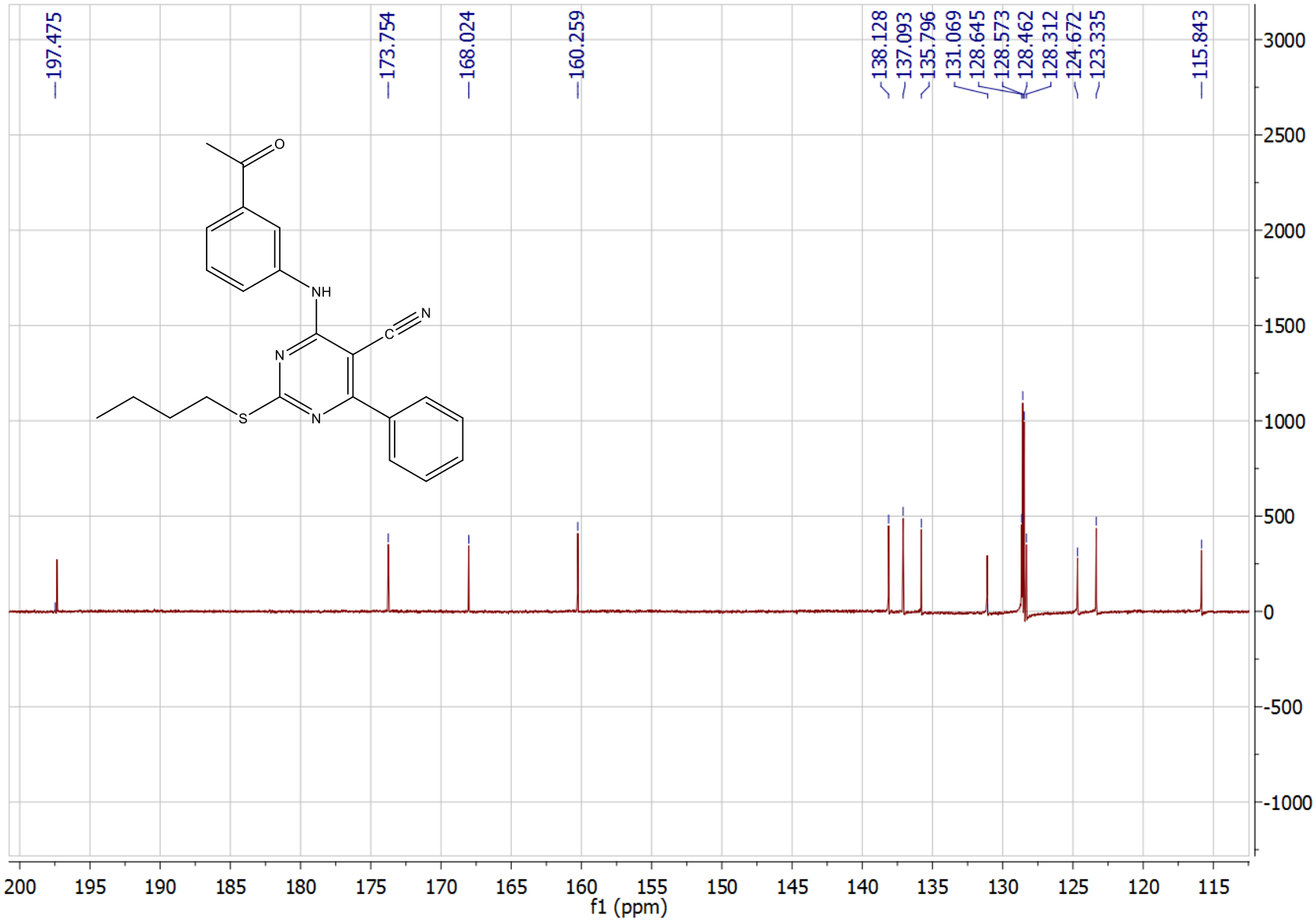
^{1}H NMR of compound 15_b



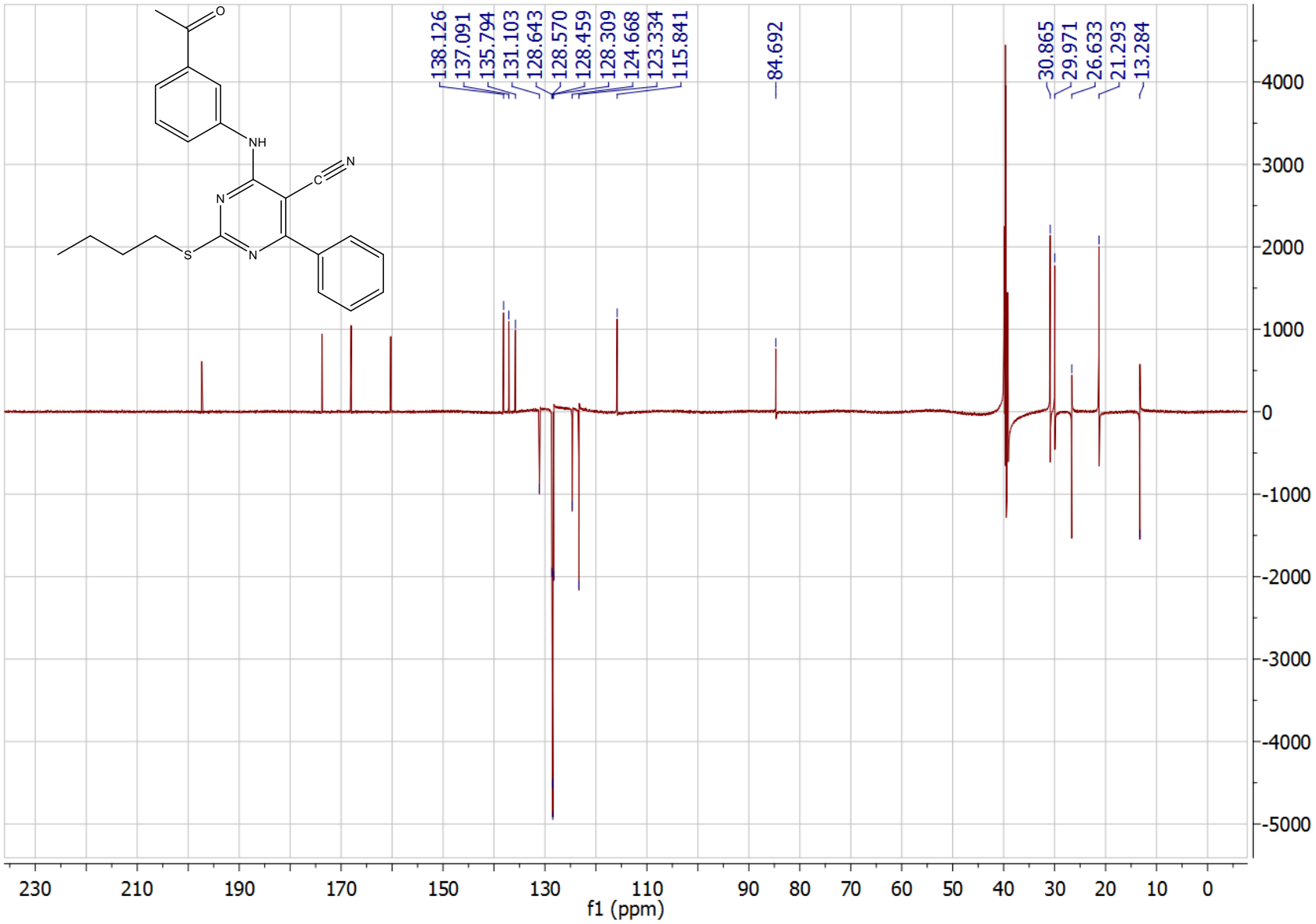
^{1}H NMR of compound 15_b



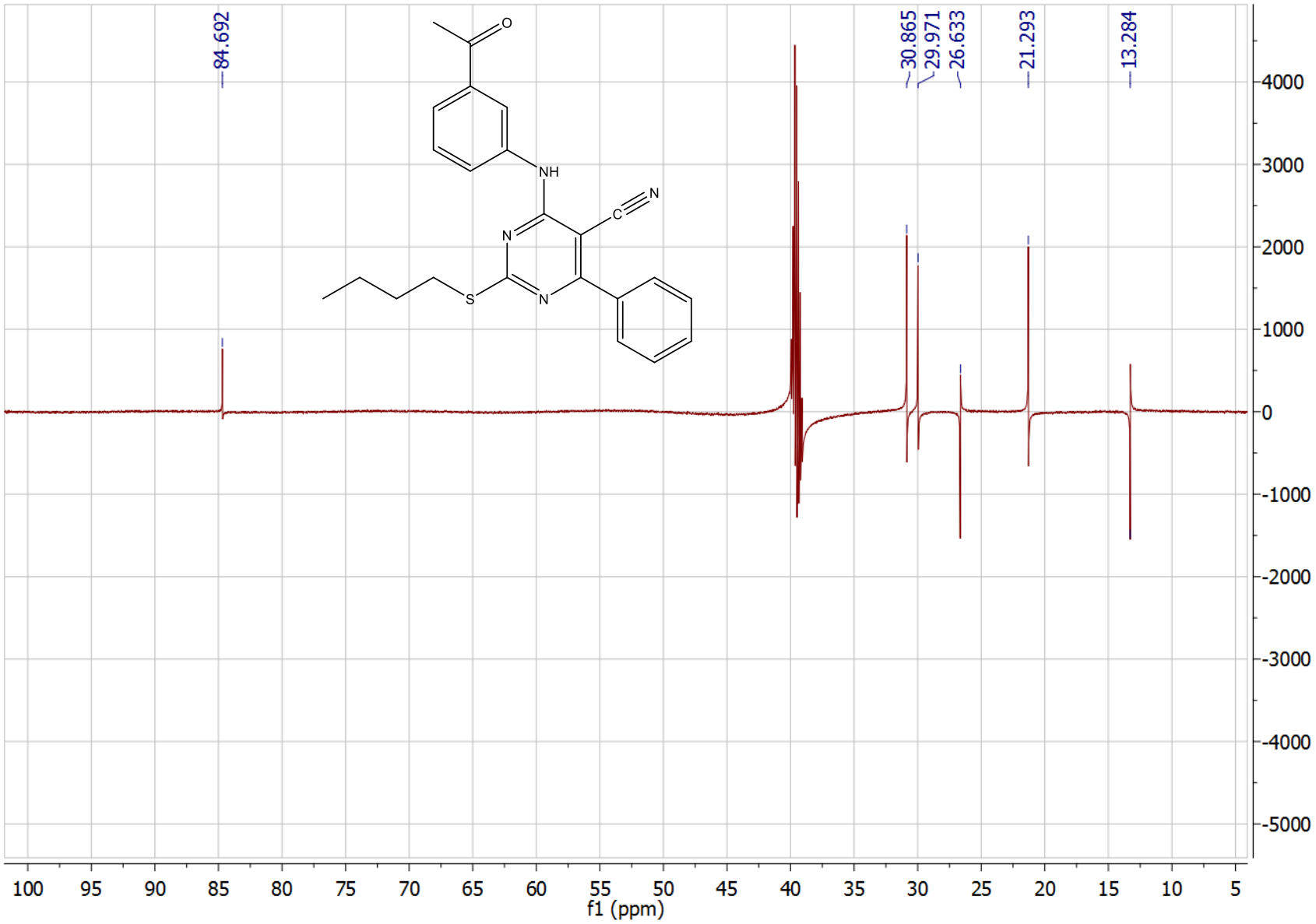
^{1}H NMR of compound 15_b



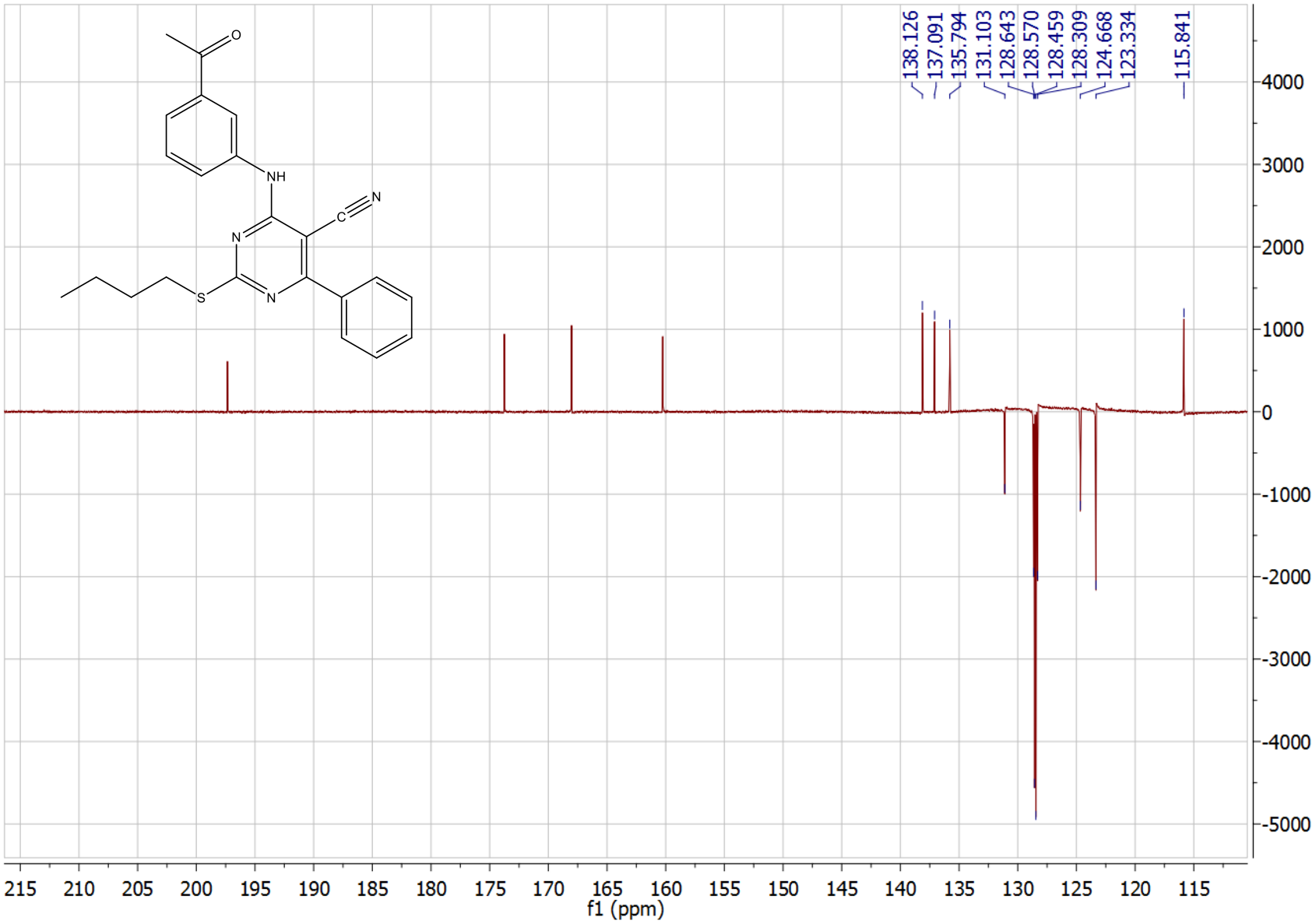
^{1}H NMR of compound 15_b



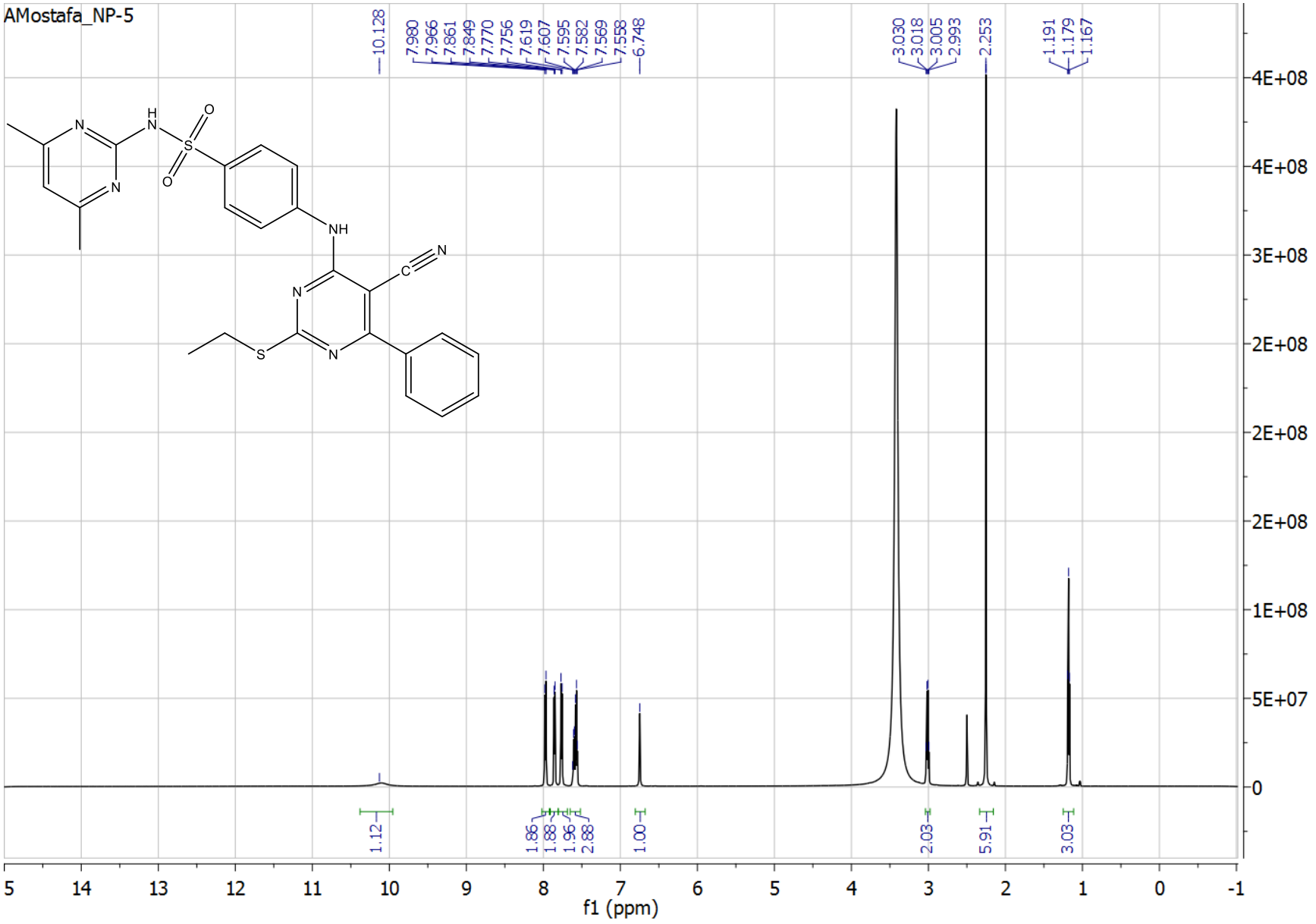
^{1}H NMR of compound 15_b



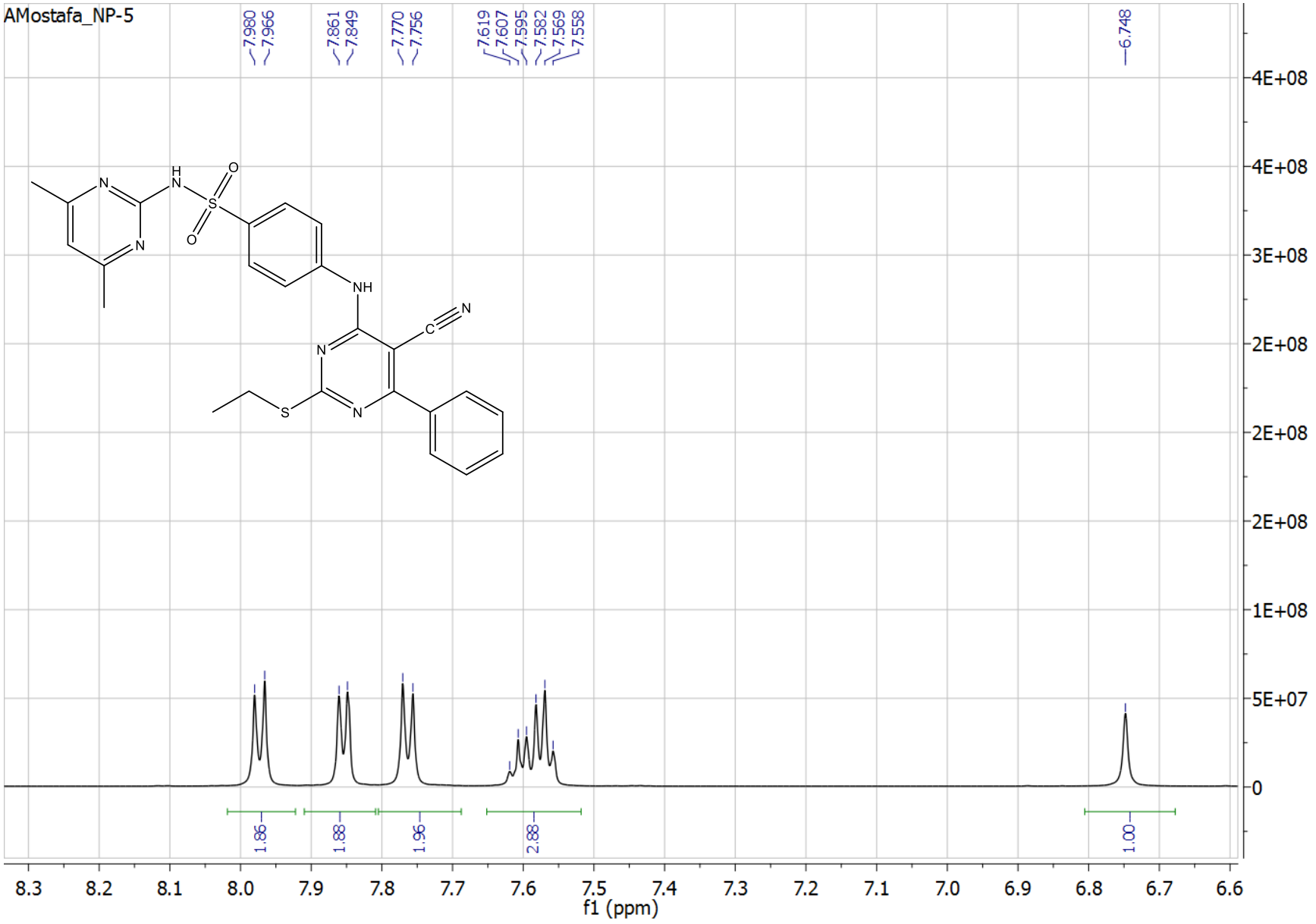
^{1}H NMR of compound 15_b



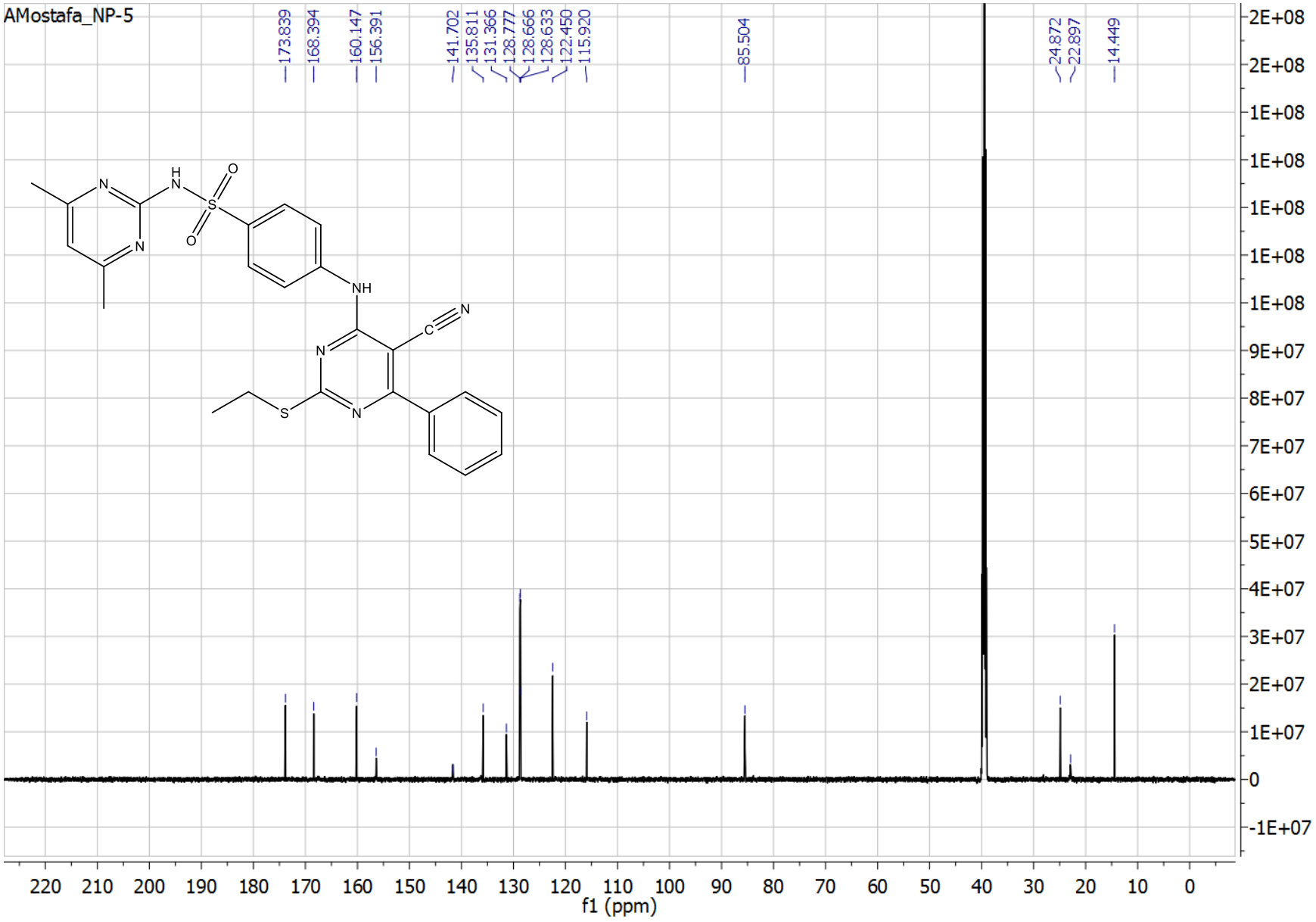
^{1}H NMR of compound 16$_{a}$



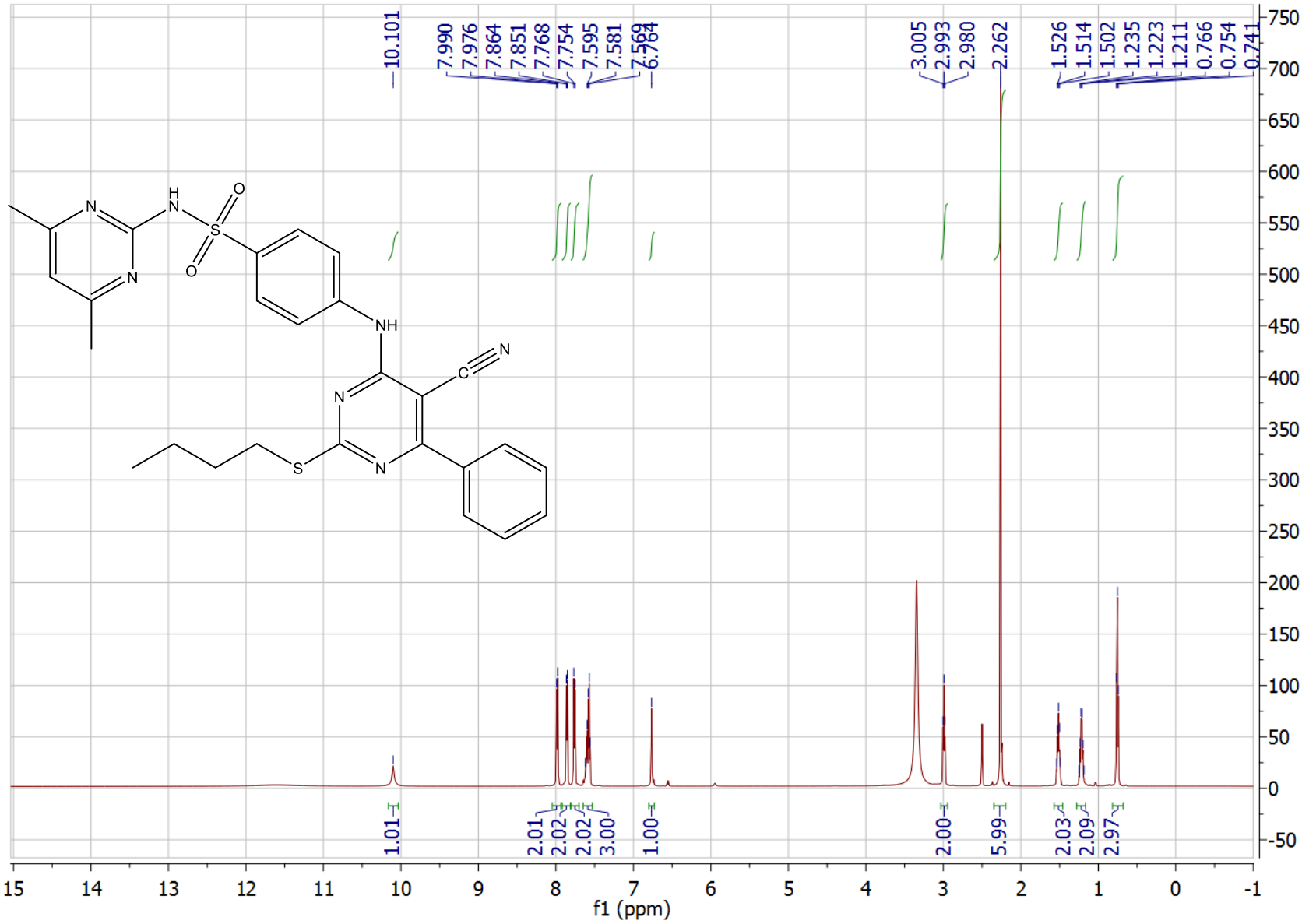
^{1}H NMR of compound 16$_{a}$



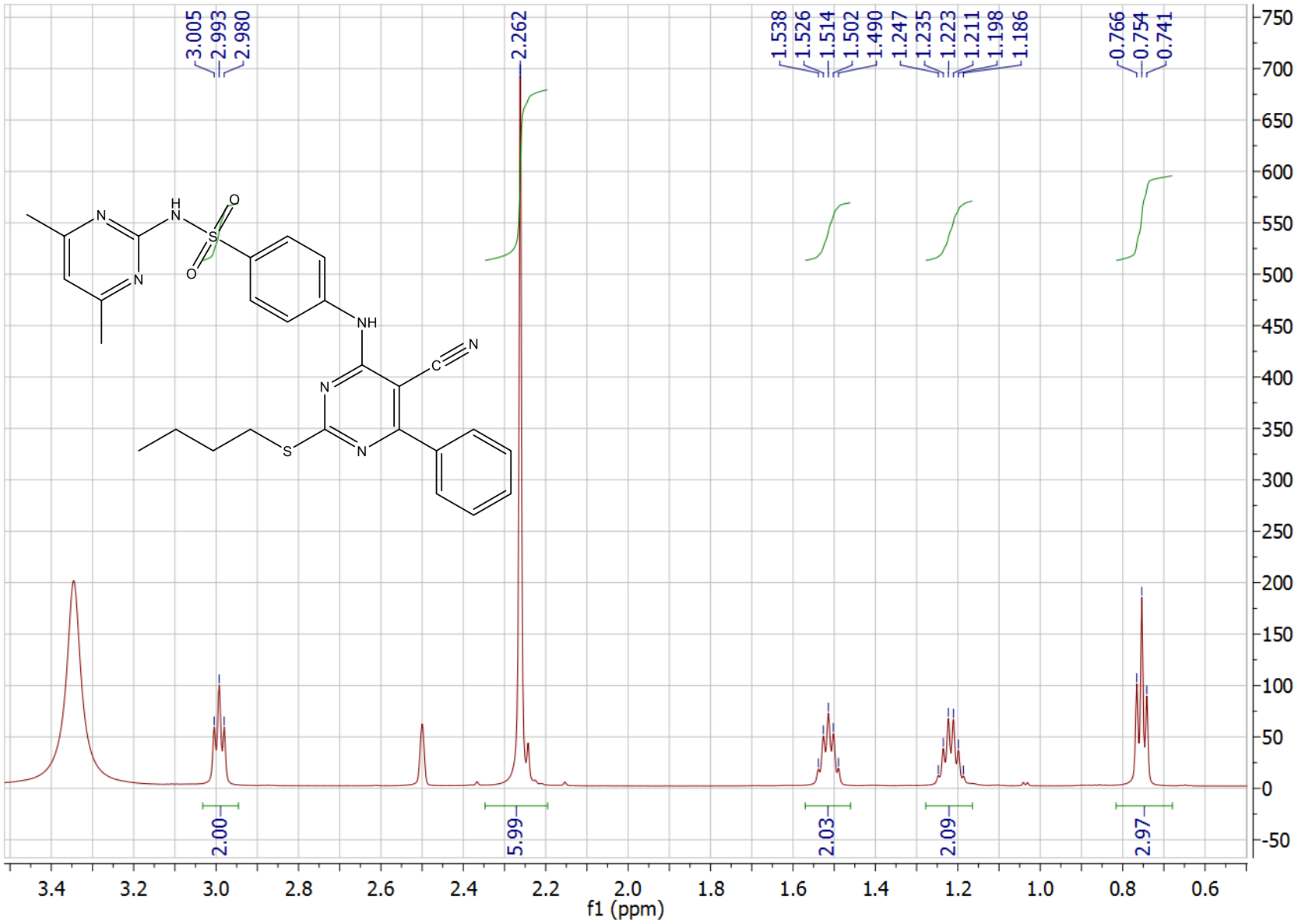
^{13}C NMR of compound 16$_a$



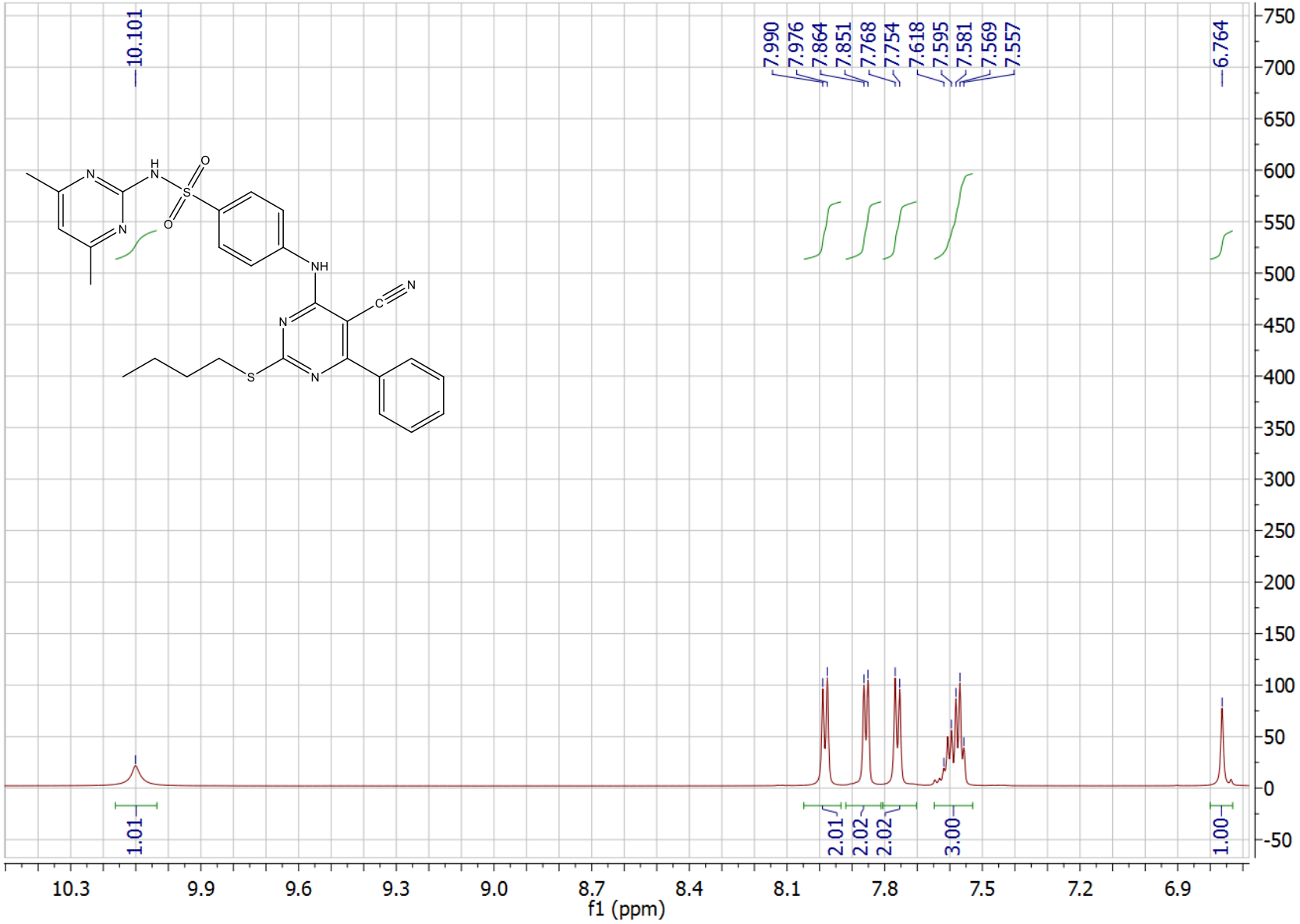
^{1}H NMR of compound 16_b



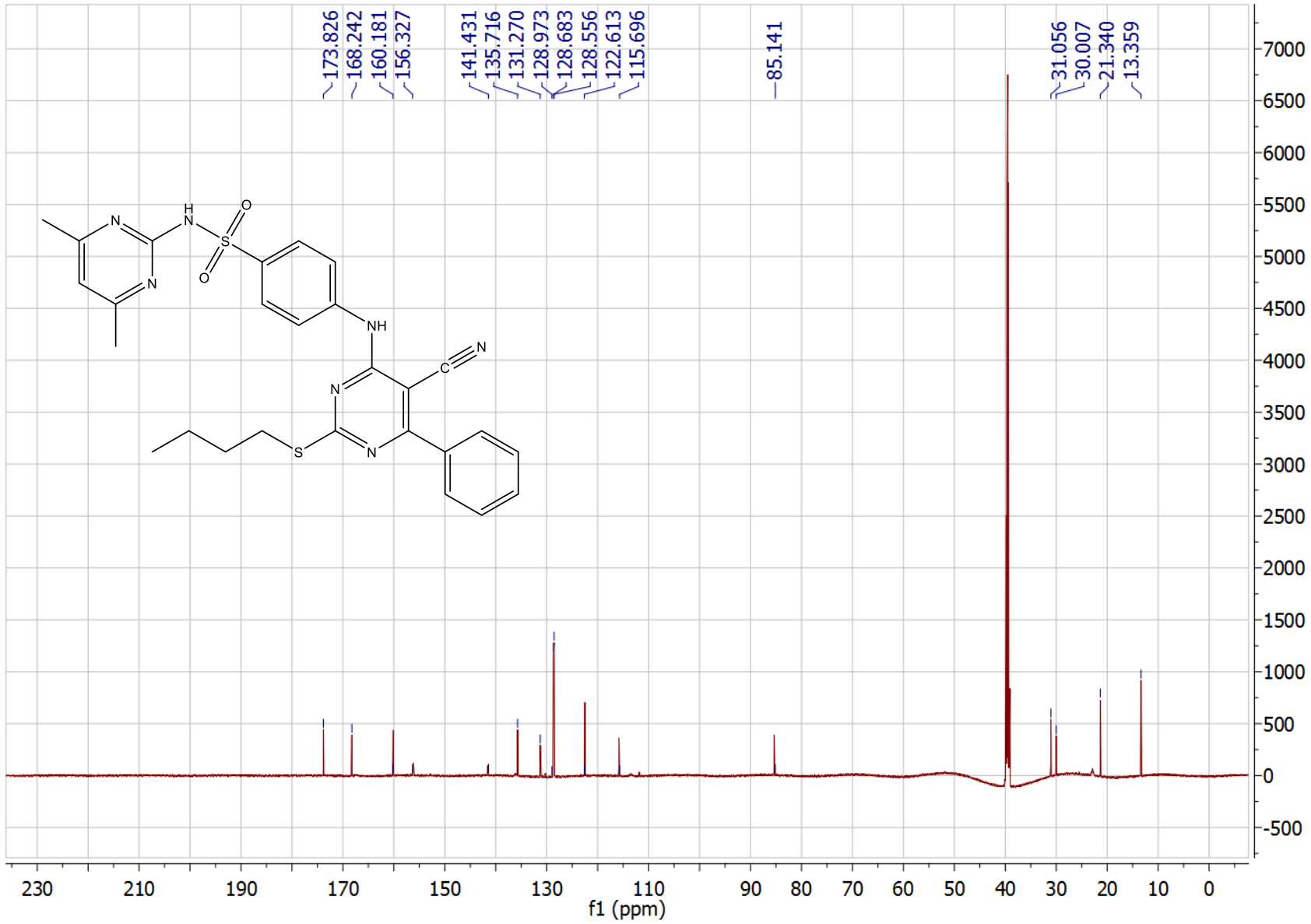
^{1}H NMR of compound 16_b



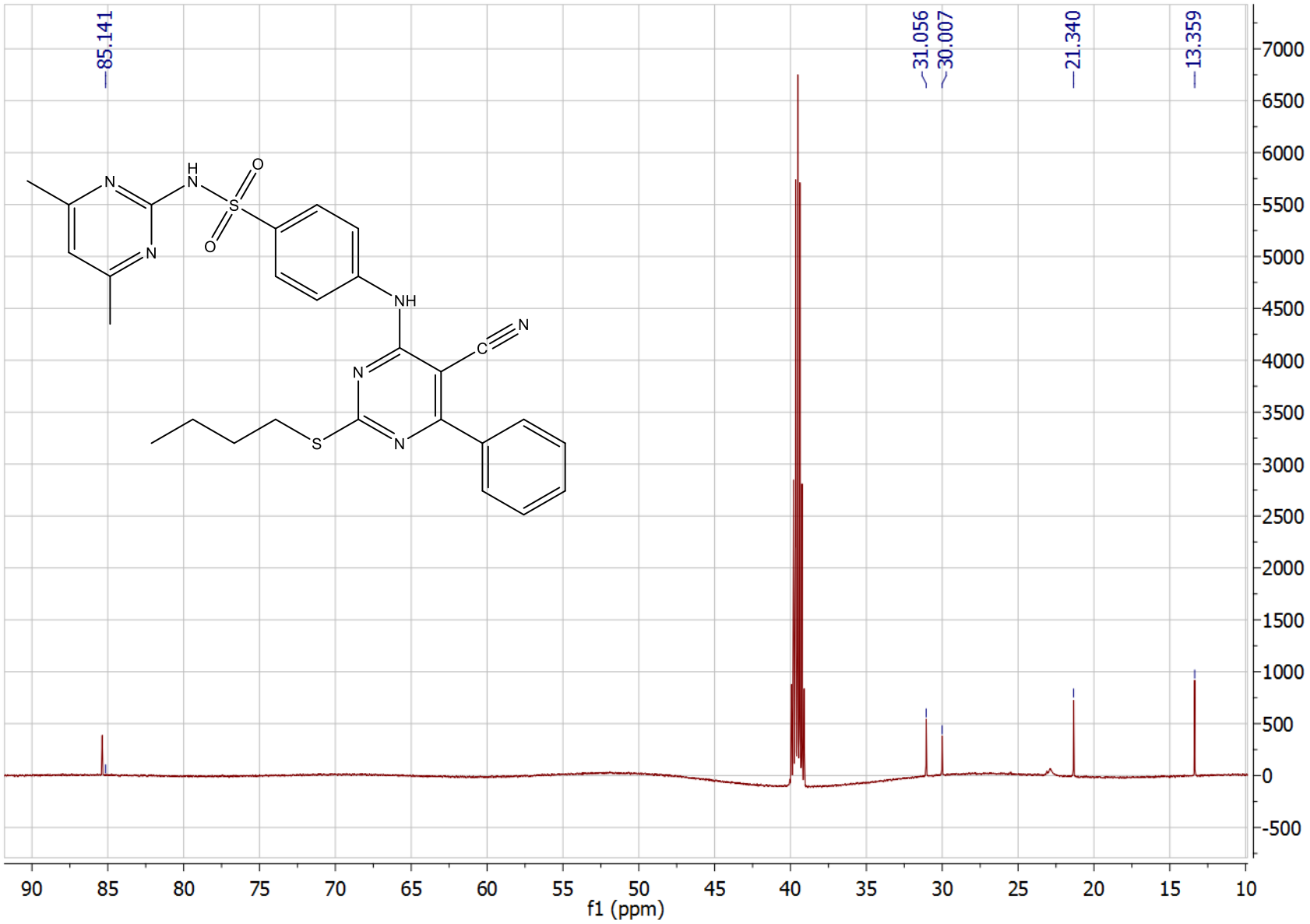
^{1}H NMR of compound 16_b



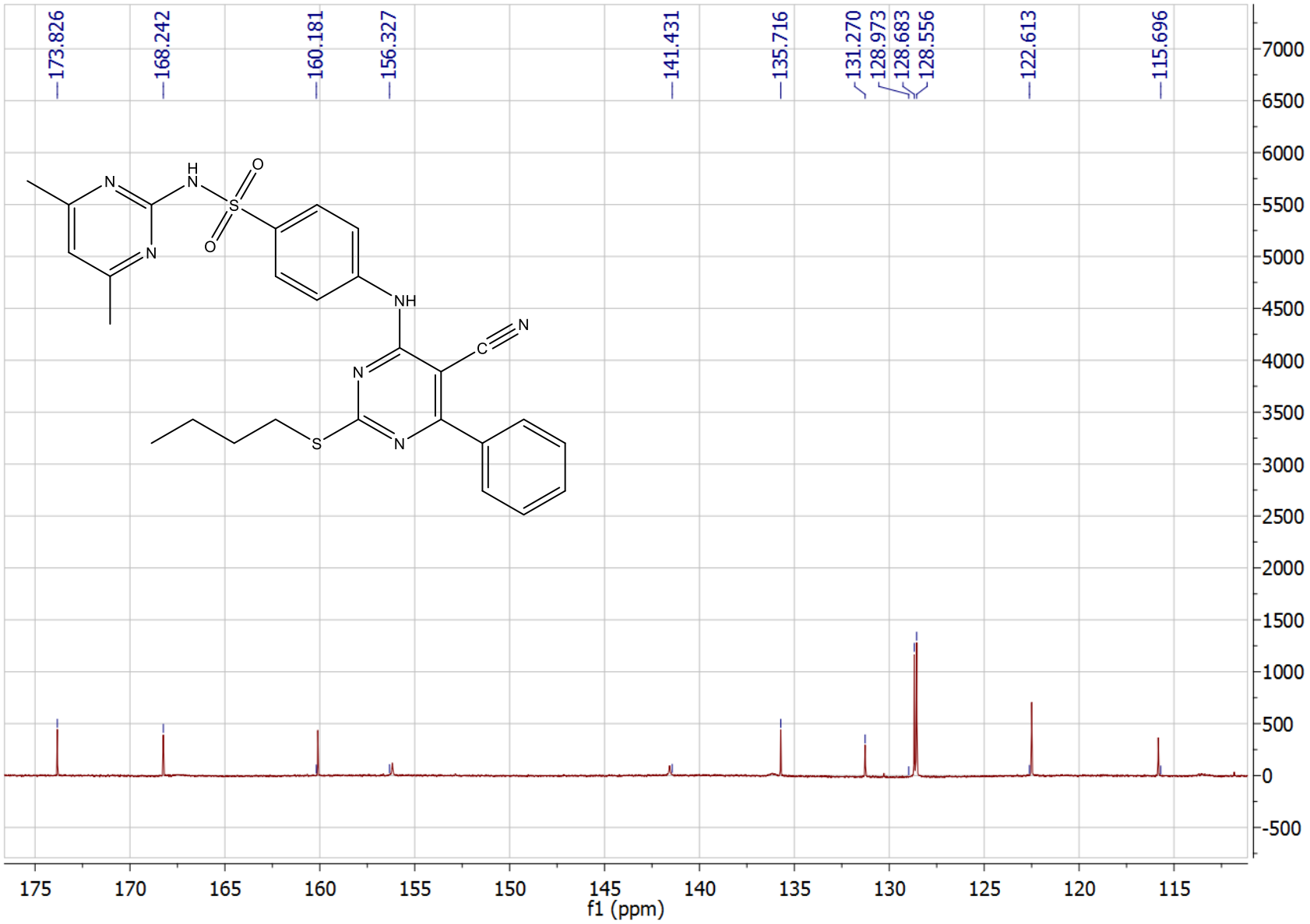
^{13}C NMR of compound 16_b



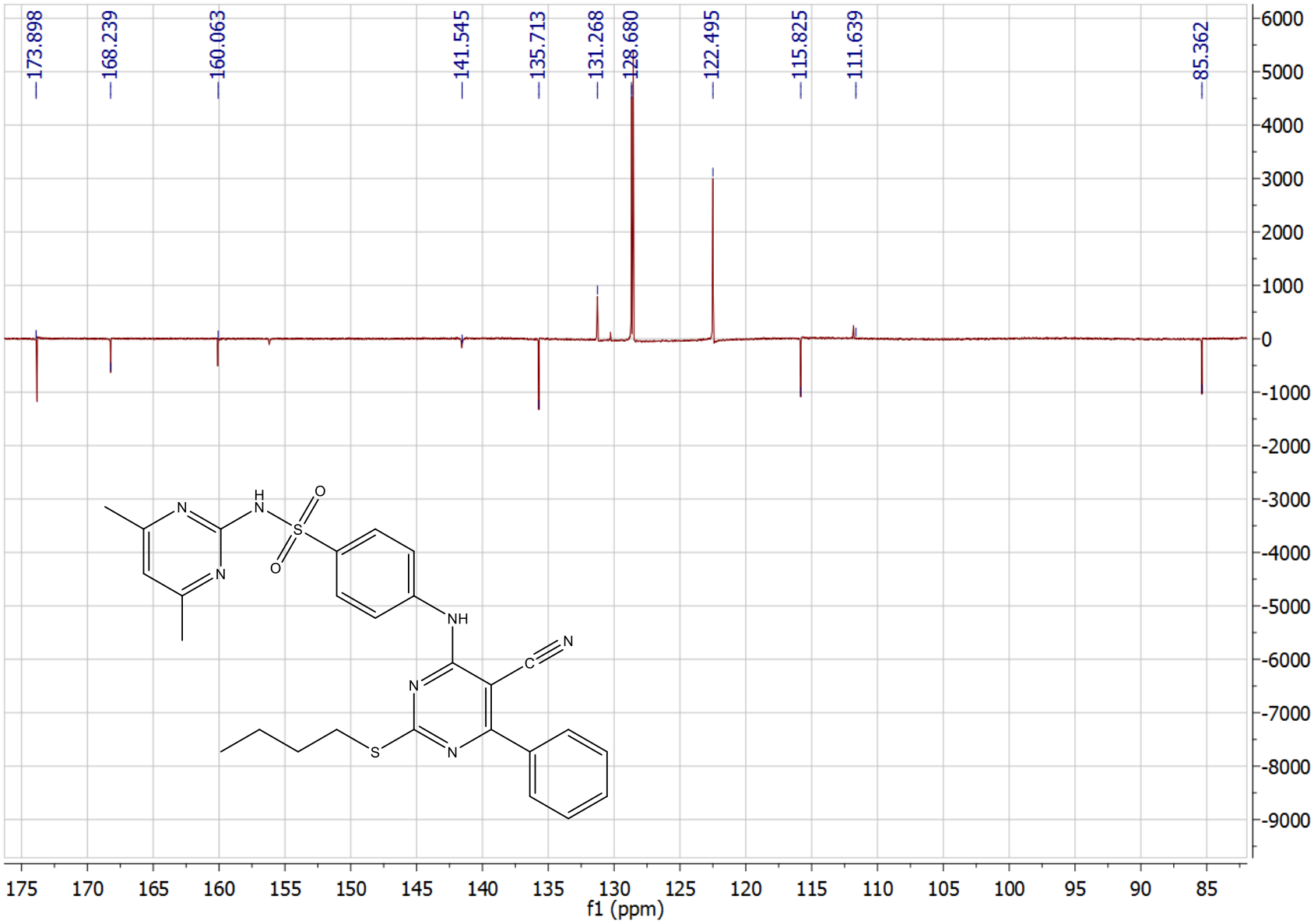
^{13}C NMR of compound 16$_b$



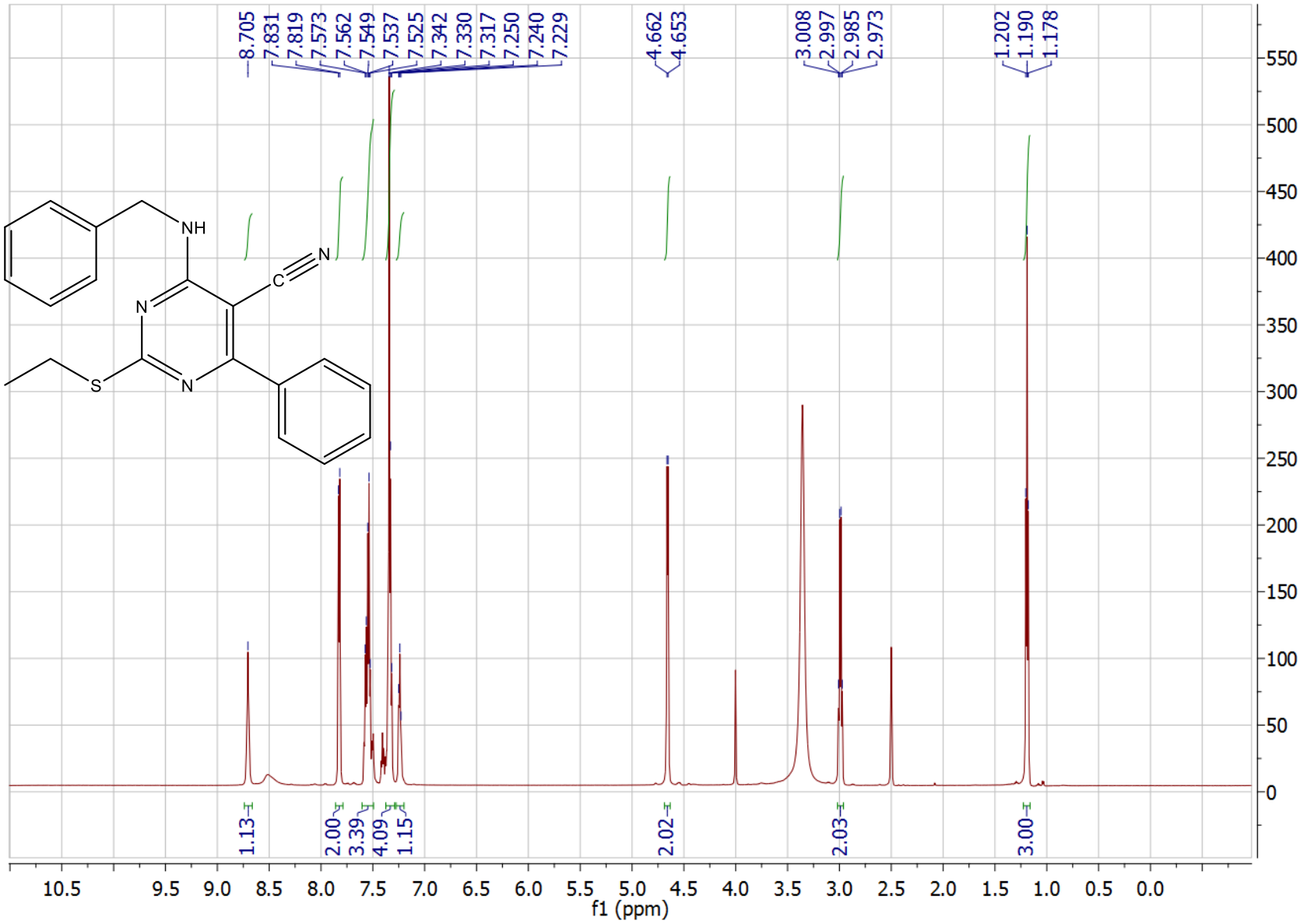
^{13}C NMR of compound 16_b



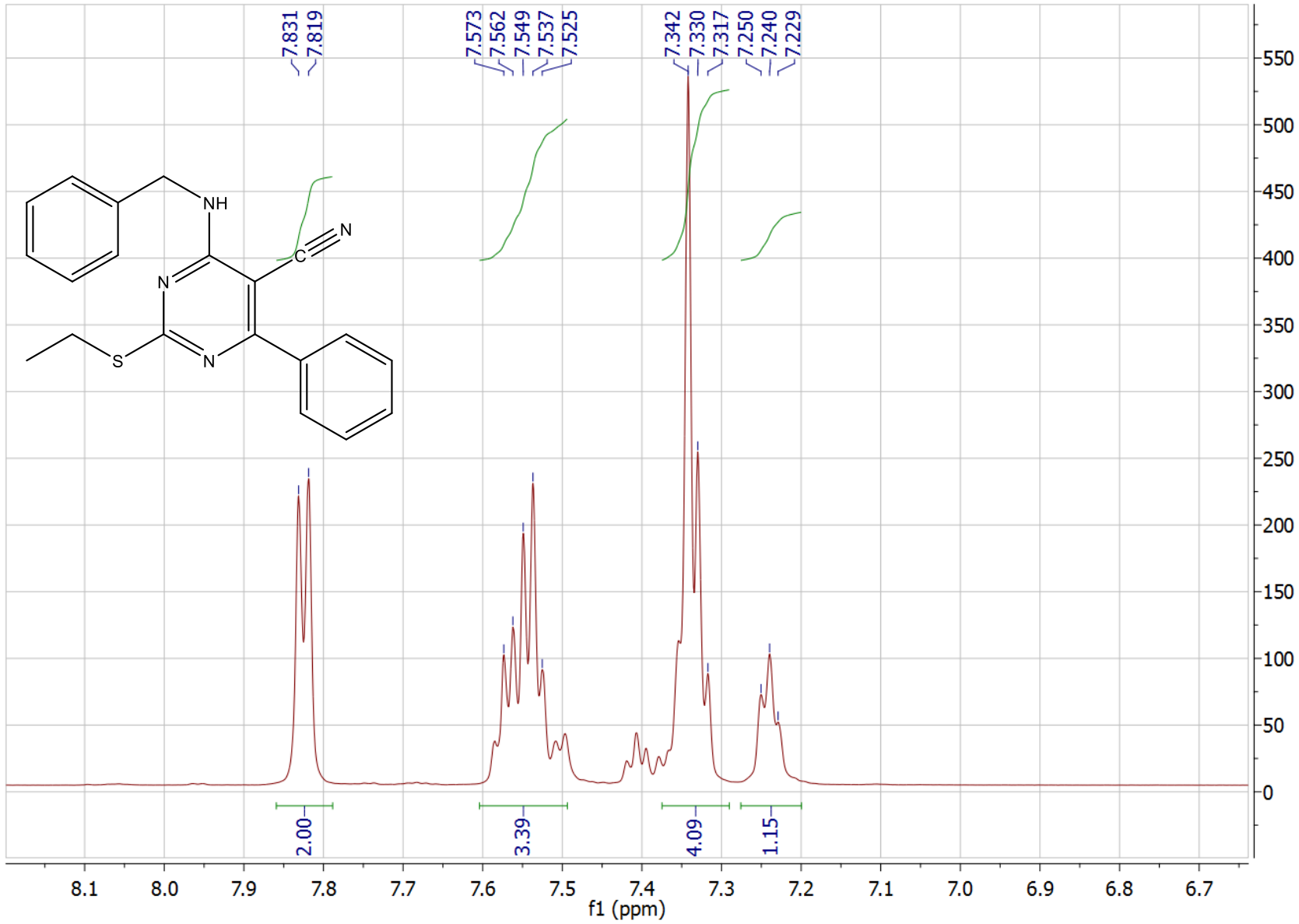
APT of compound 16_b



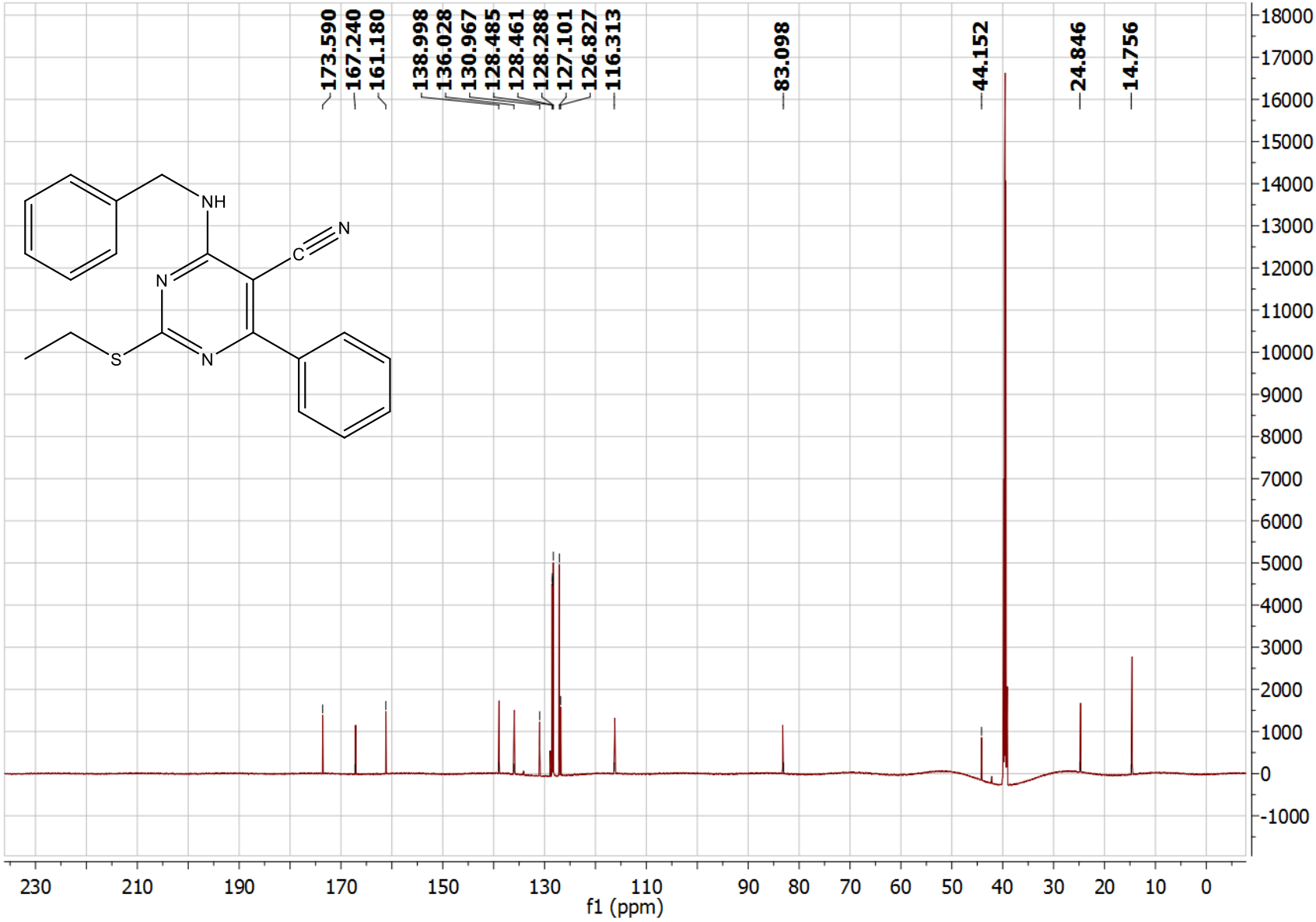
^{1}H NMR of compound 17$_{a}$



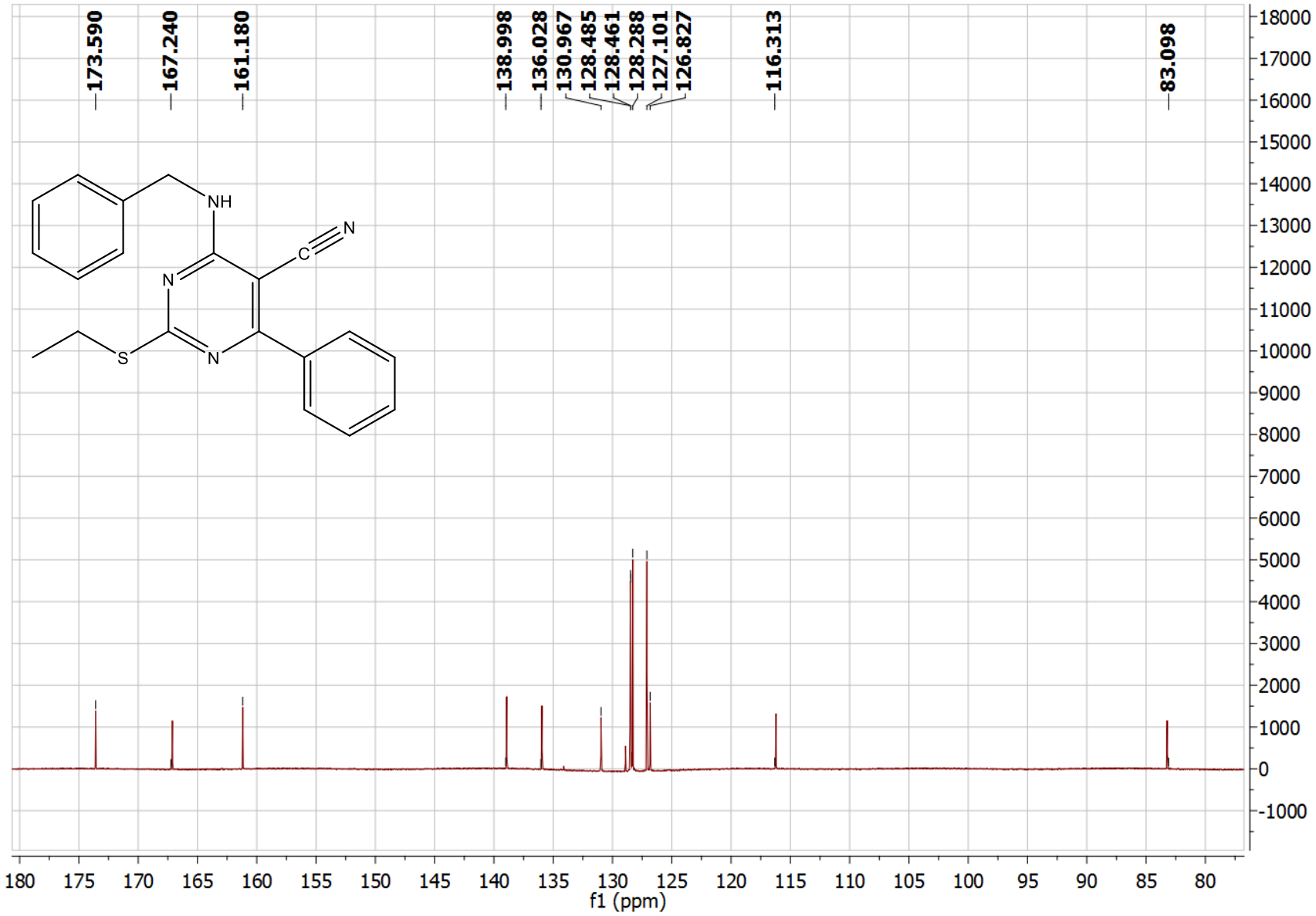
^{1}H NMR of compound 17$_{a}$



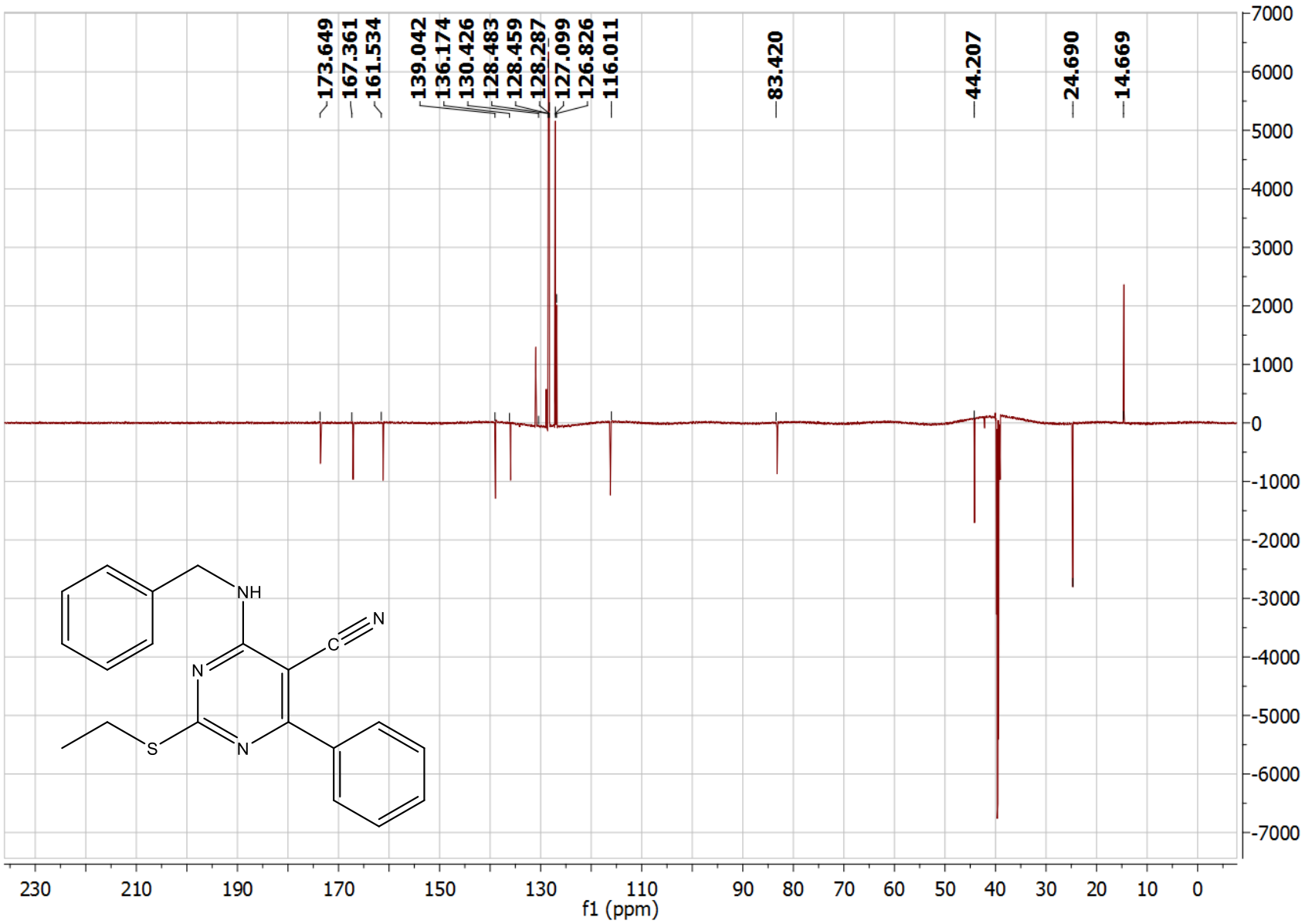
^{13}C NMR of compound 17_a



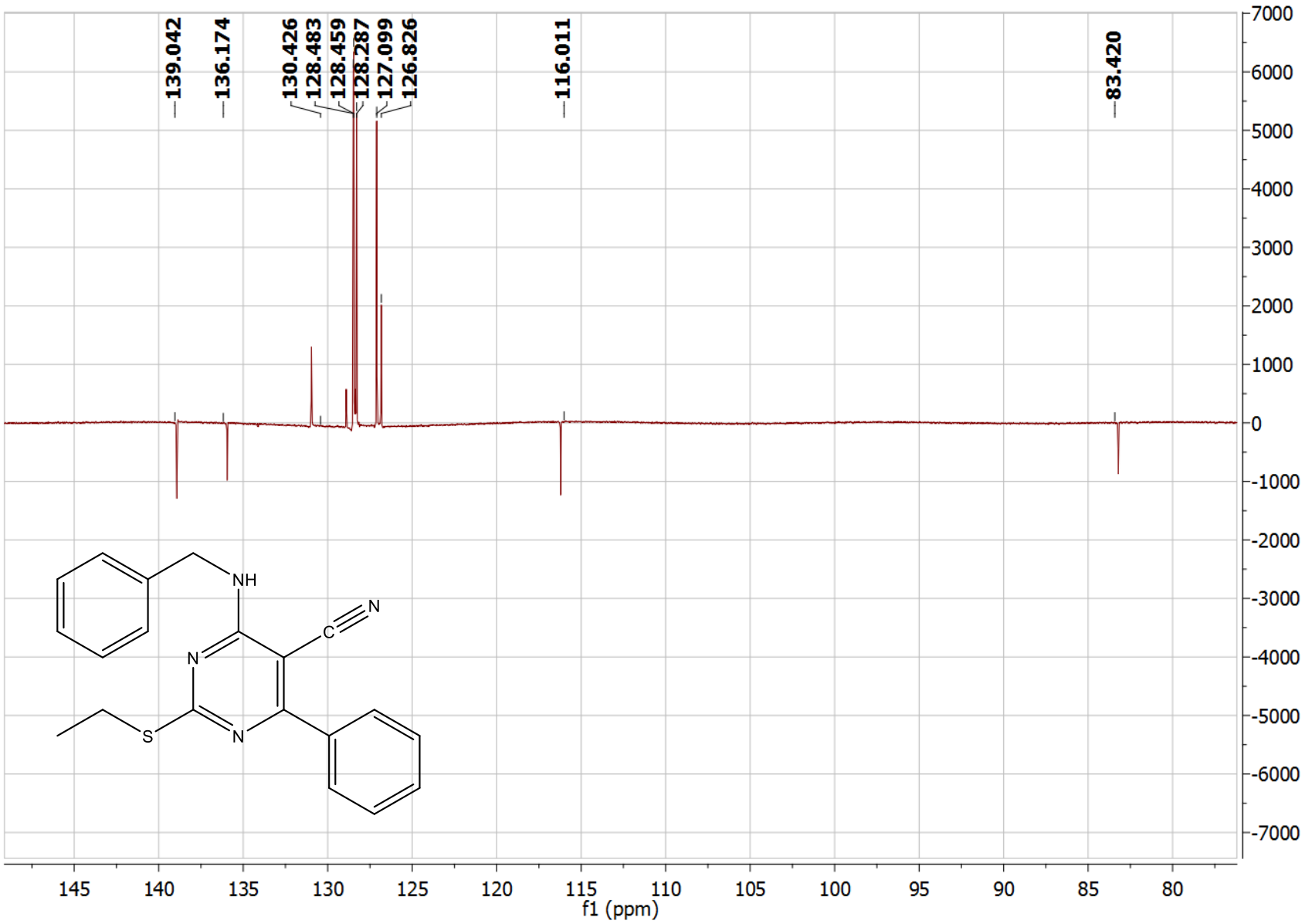
^{13}C NMR of compound 17$_a$



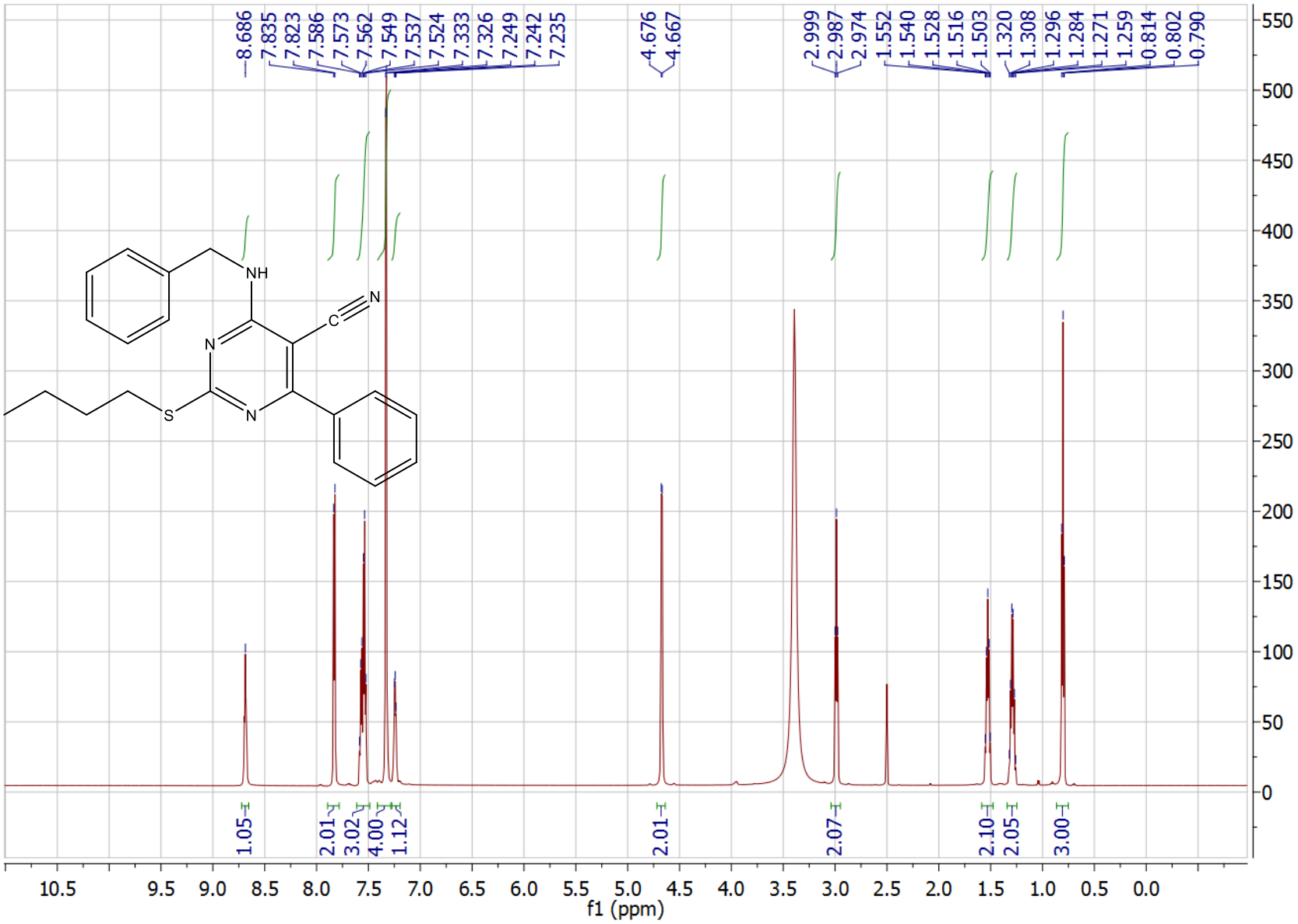
APT of compound 17_a



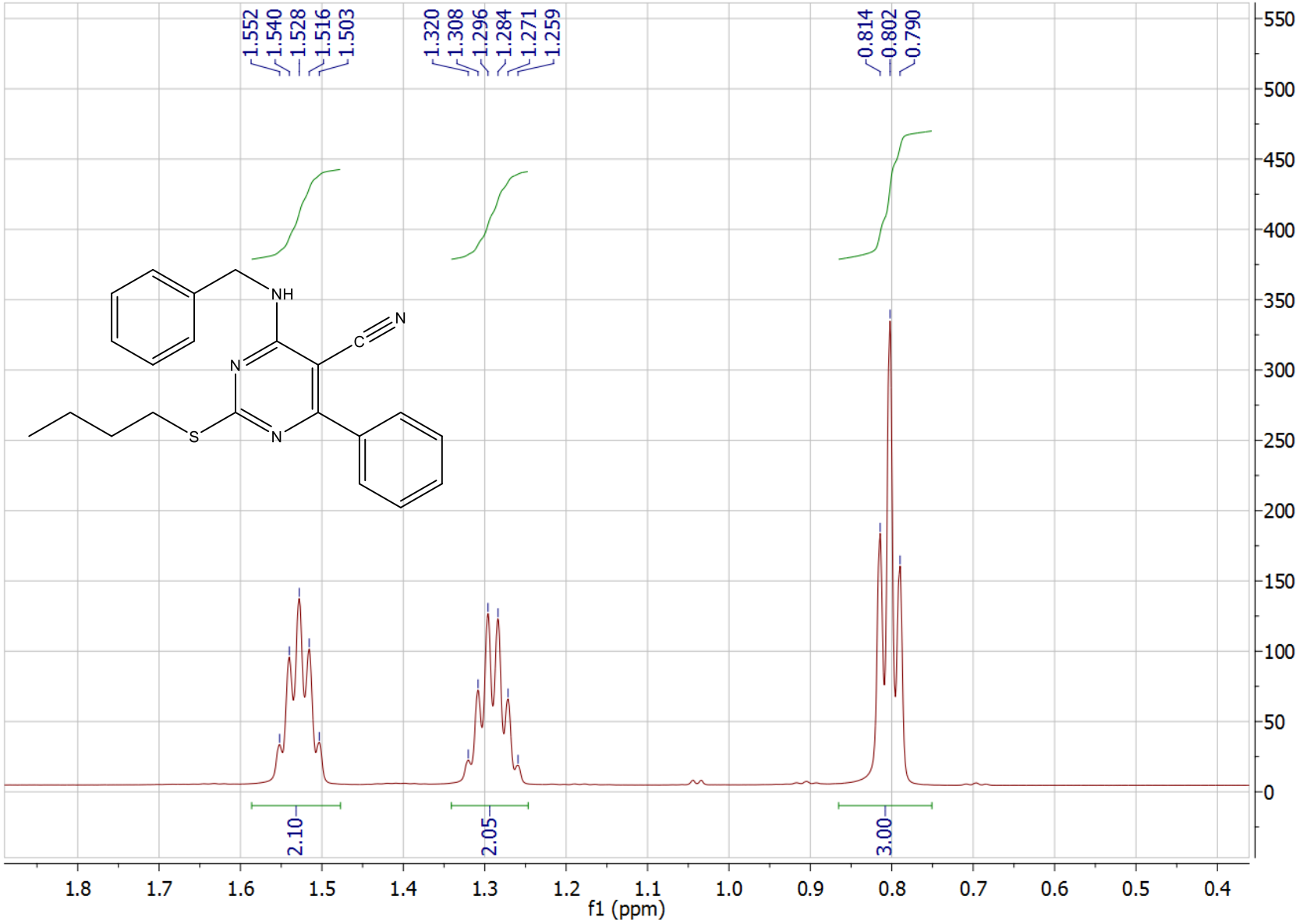
APT of compound 17_a



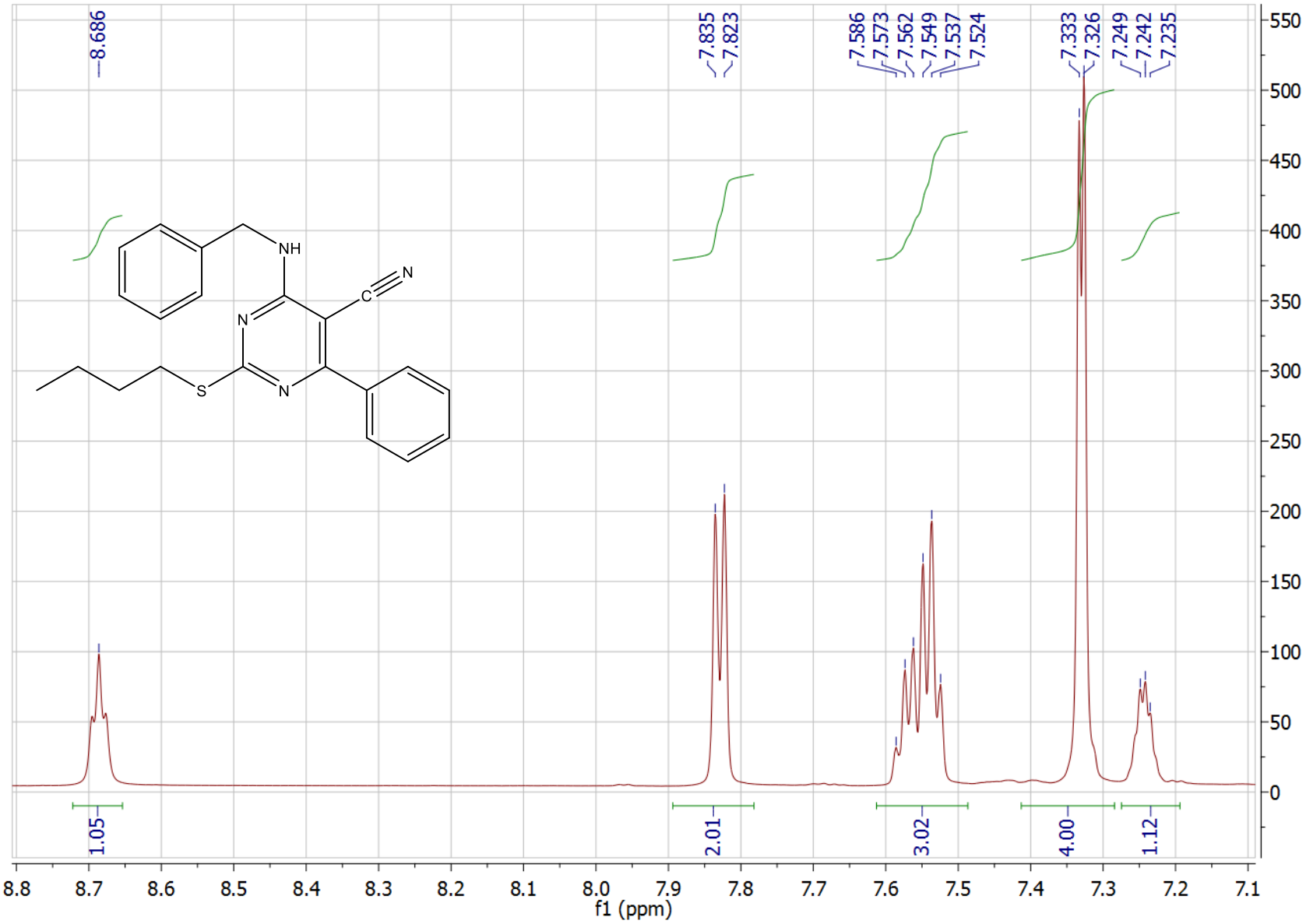
^{1}H NMR of compound 17_b



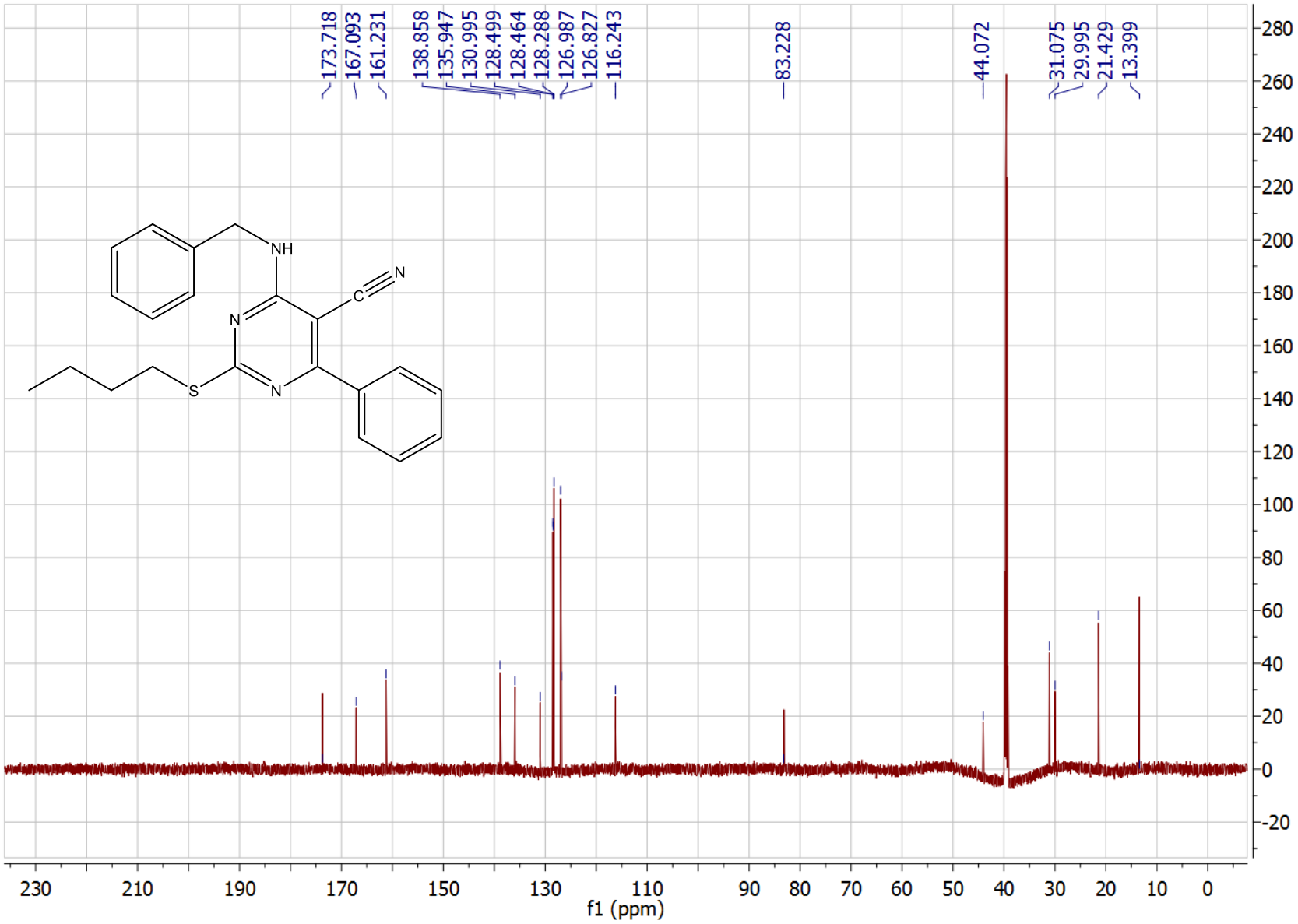
^{1}H NMR of compound 17_b



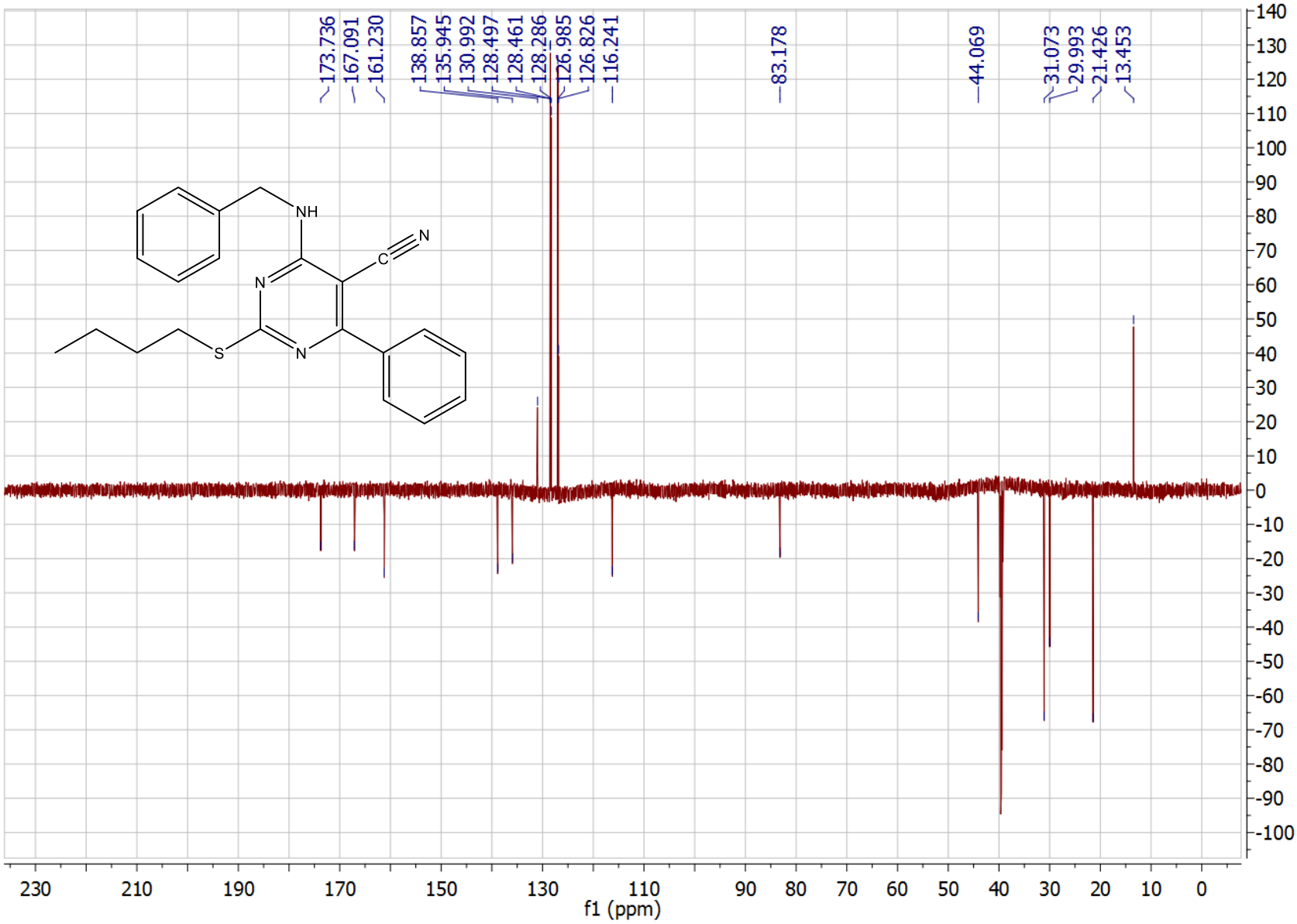
^{1}H NMR of compound 17_b



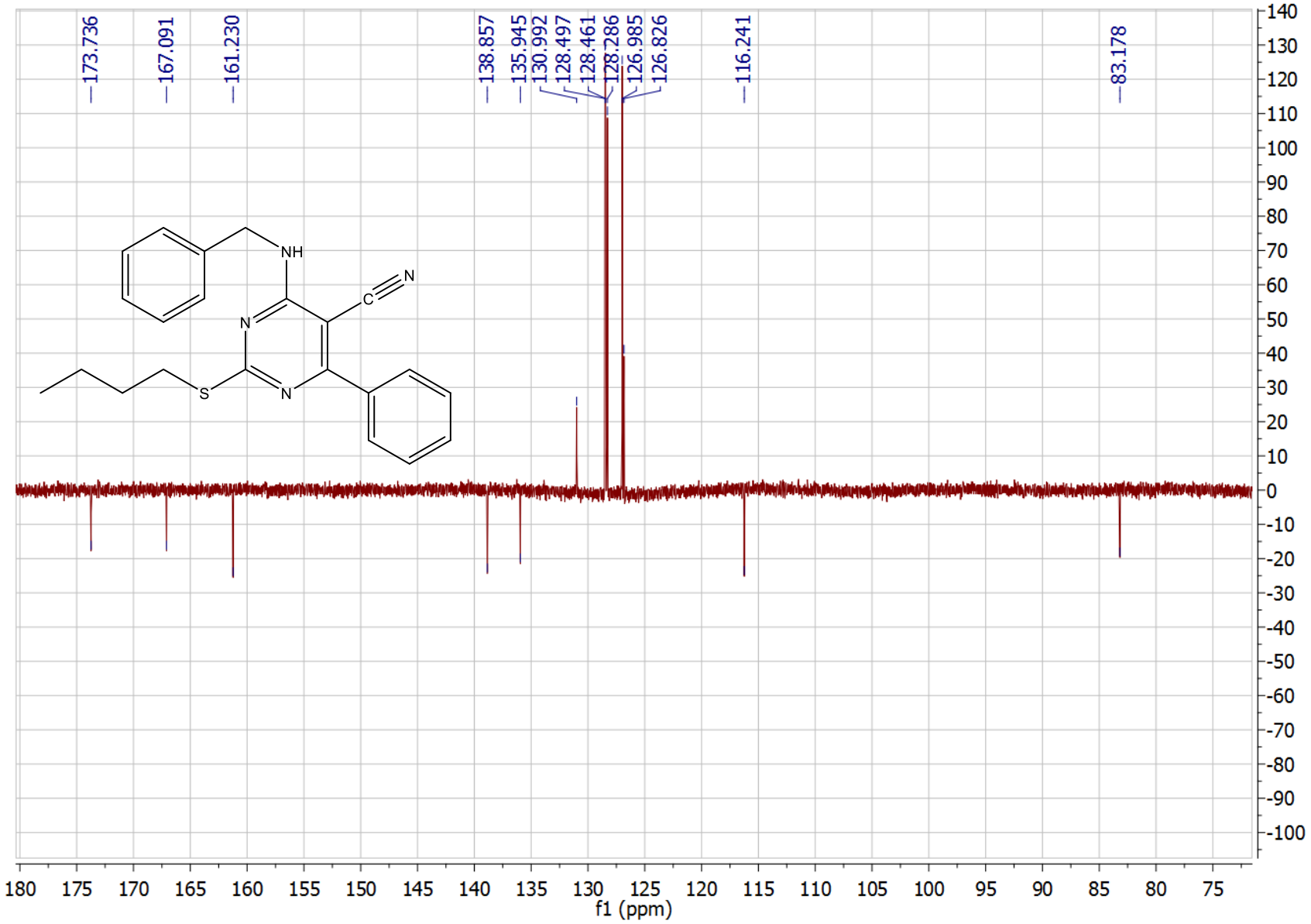
^{13}C NMR of compound 17_b



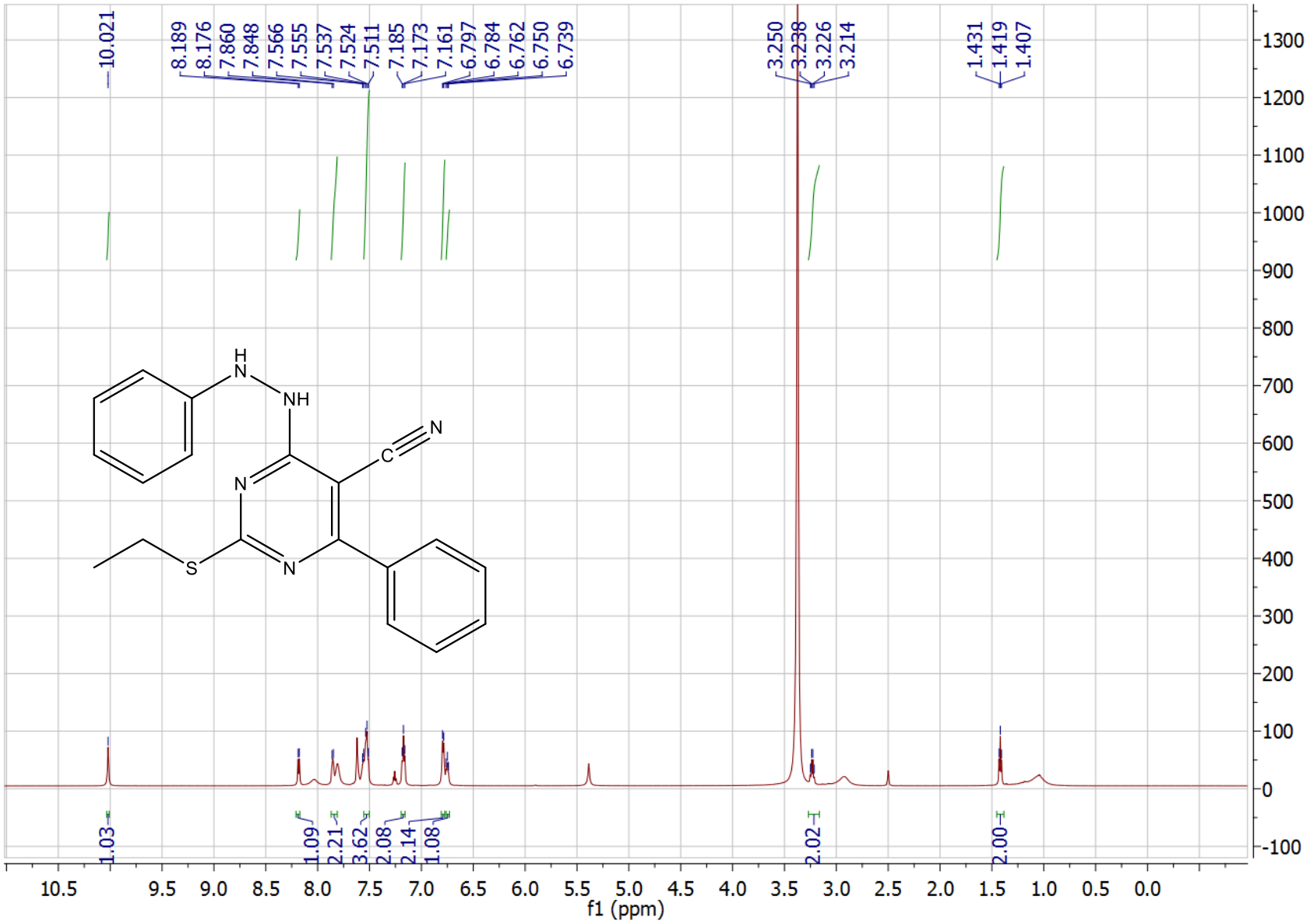
APT of compound 17_b



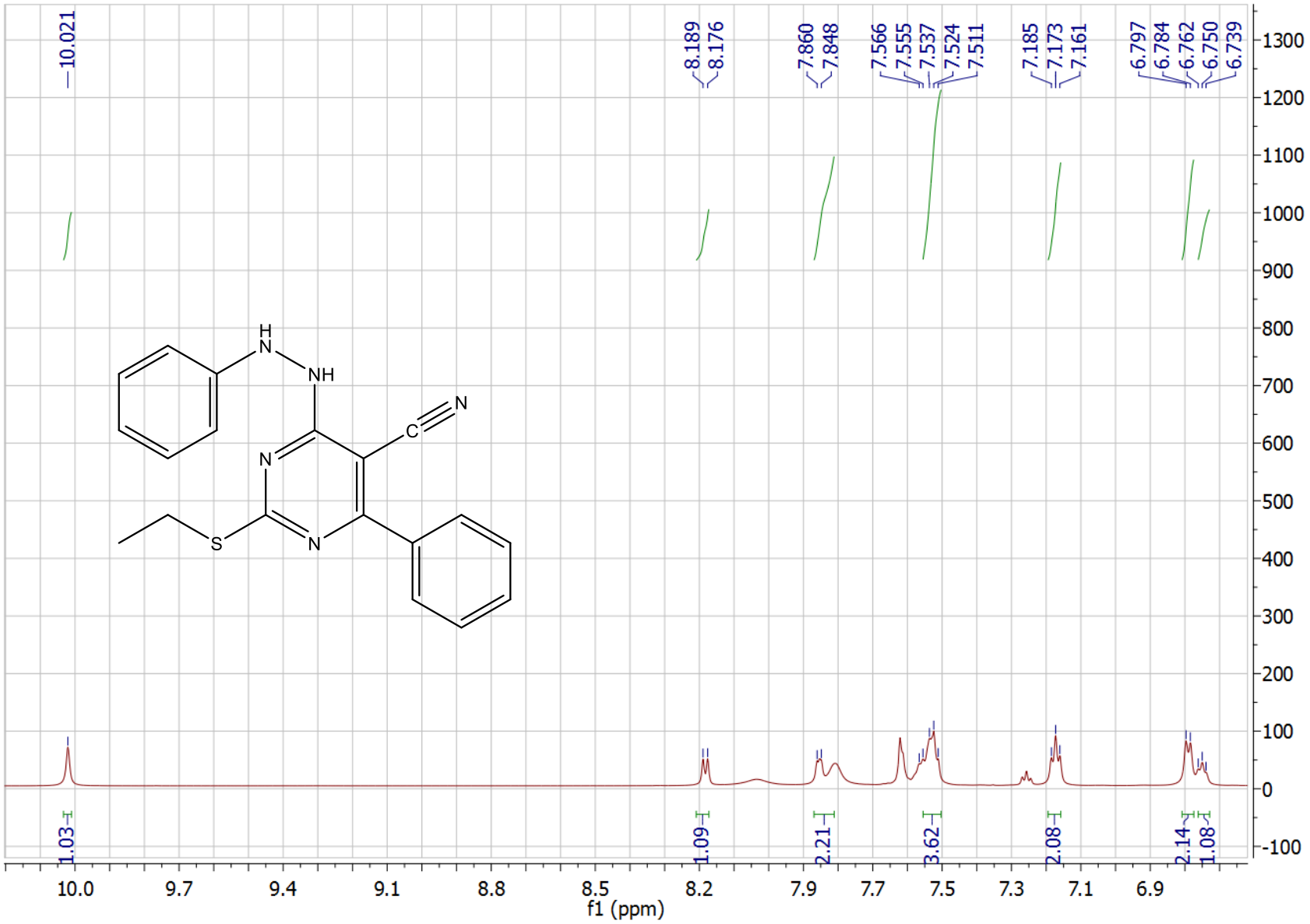
APT of compound 17_b



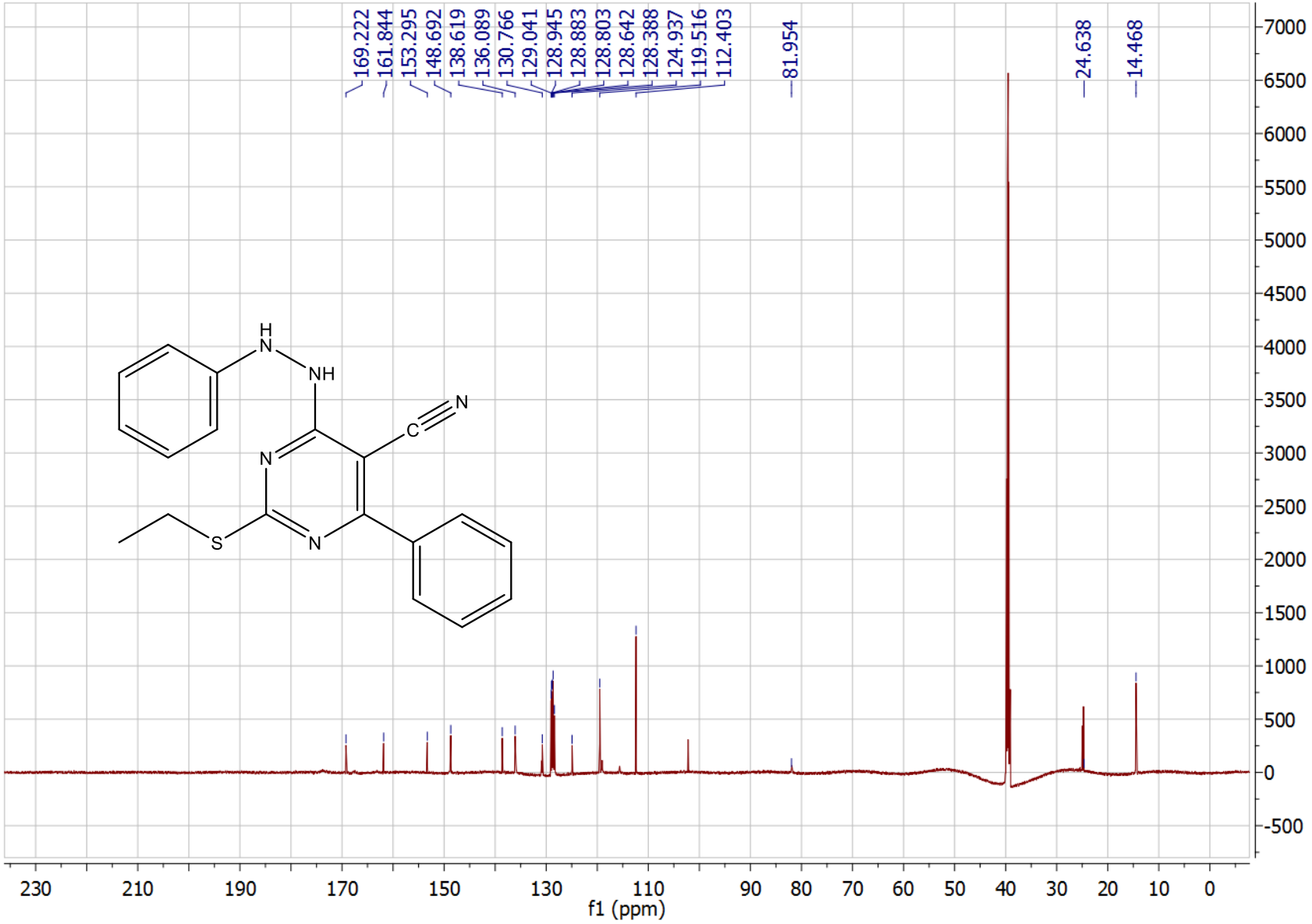
^{1}H NMR of compound 18$_a$



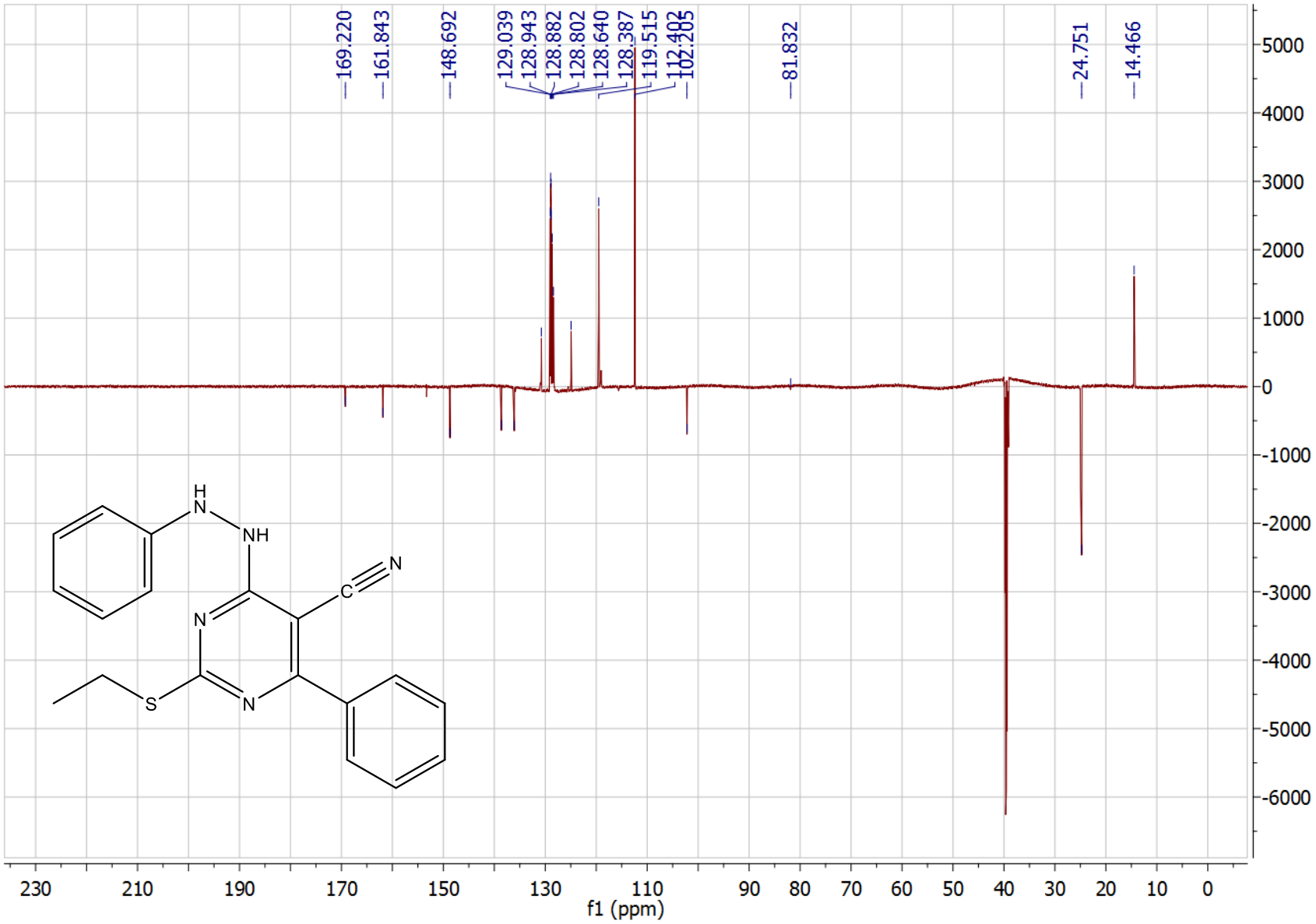
^{1}H NMR of compound 18$_a$



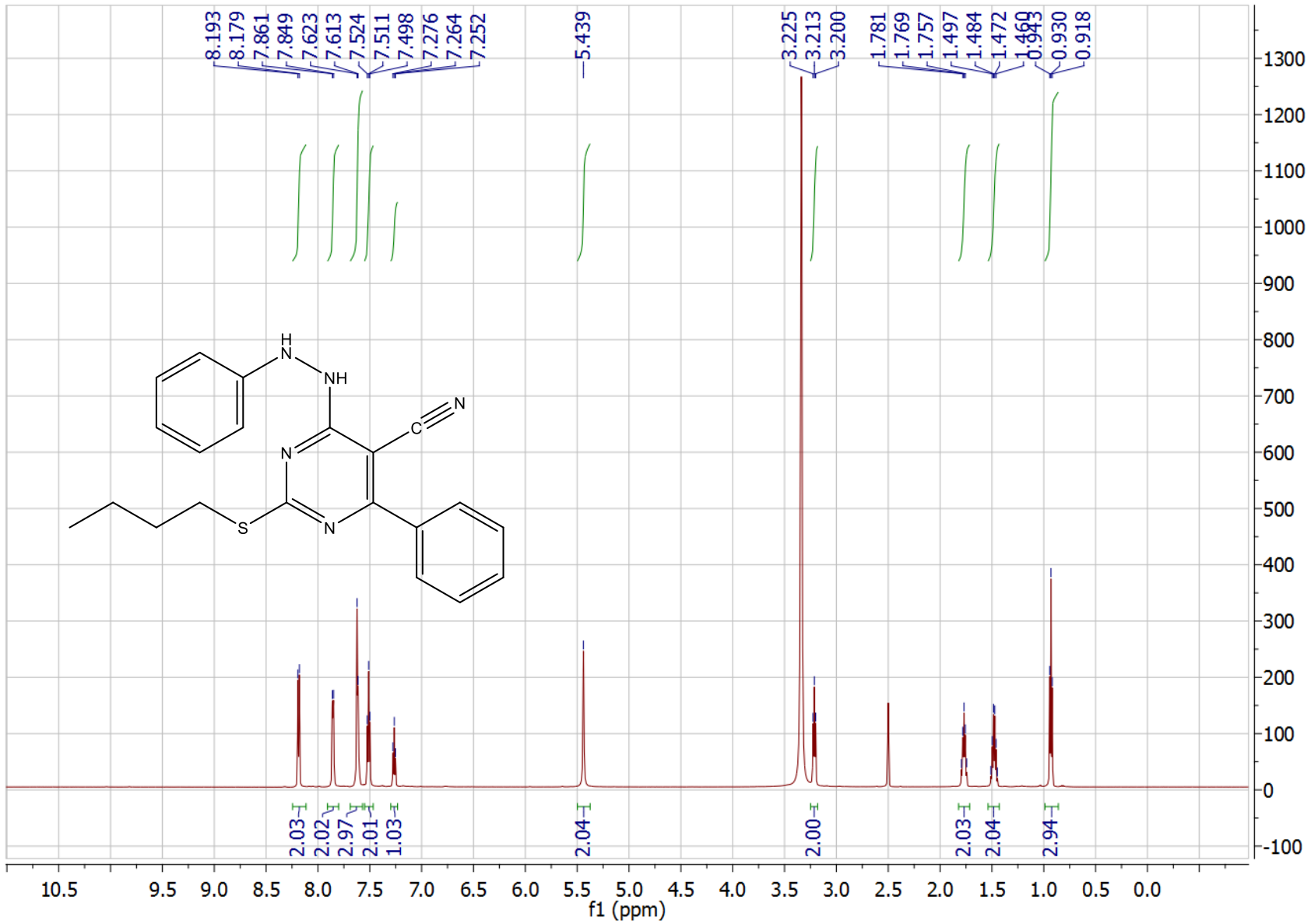
^{13}C NMR of compound 18_a



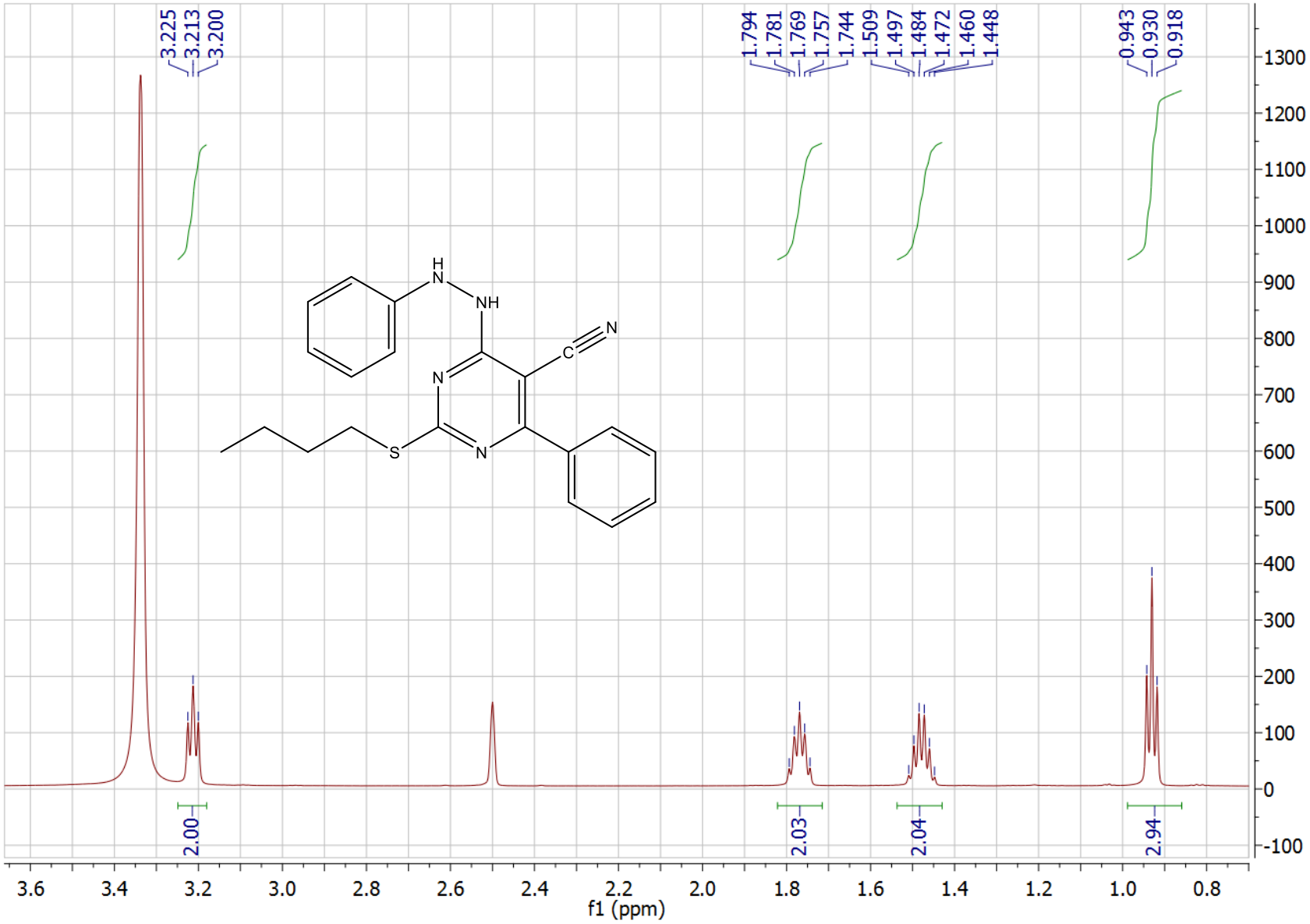
APT of compound 18_a



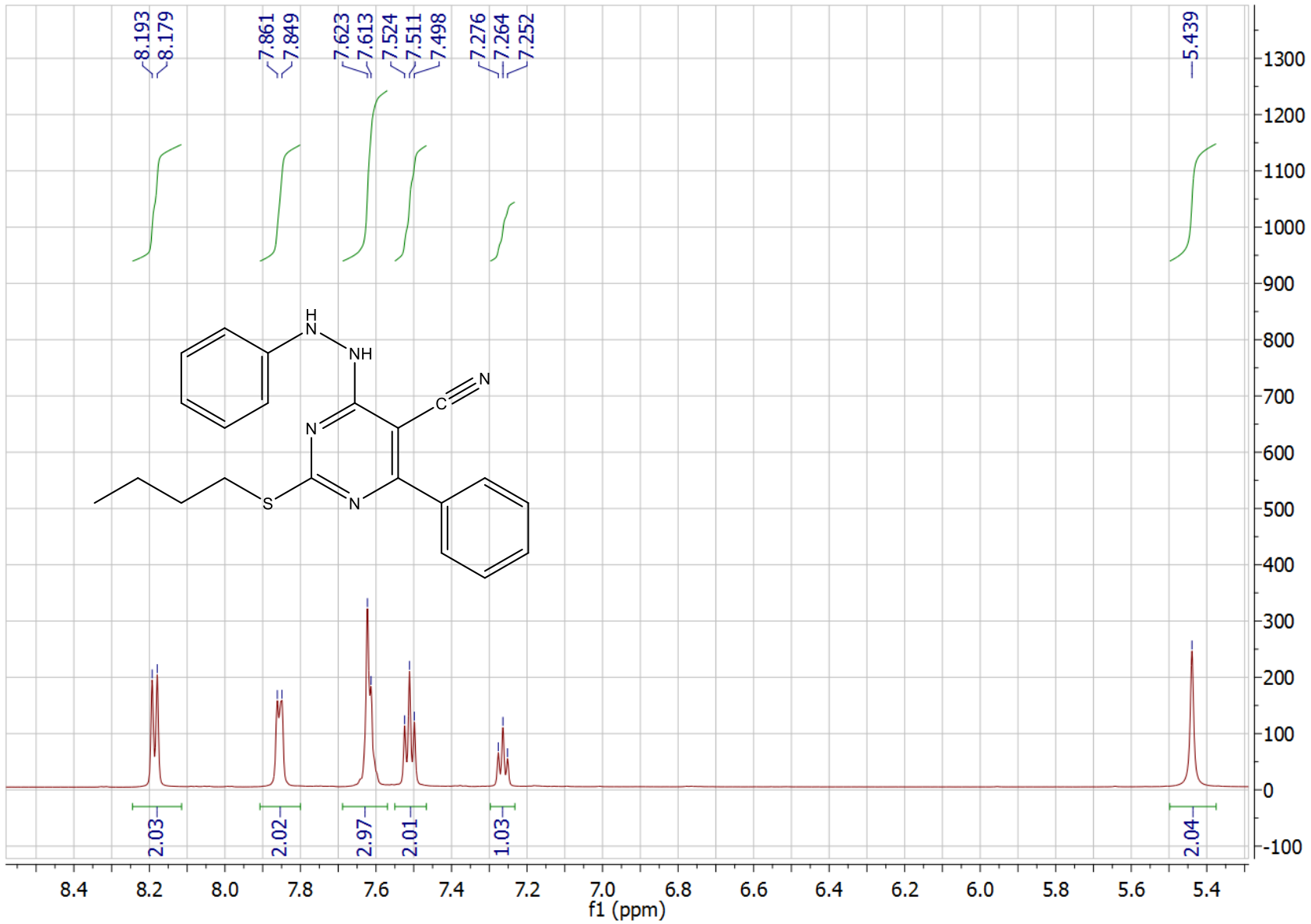
^{1}H NMR of compound 18_b



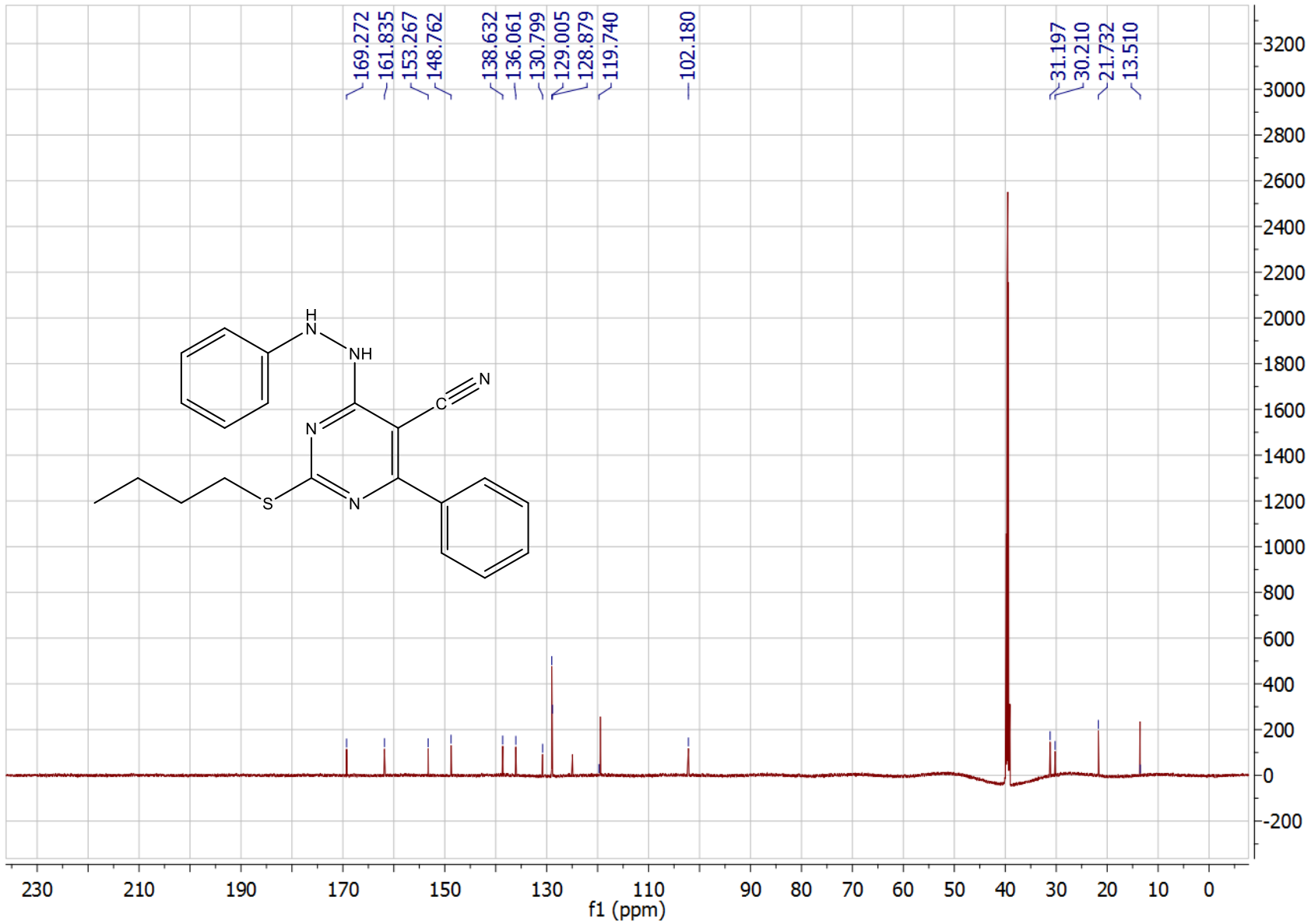
^{1}H NMR of compound 18_b



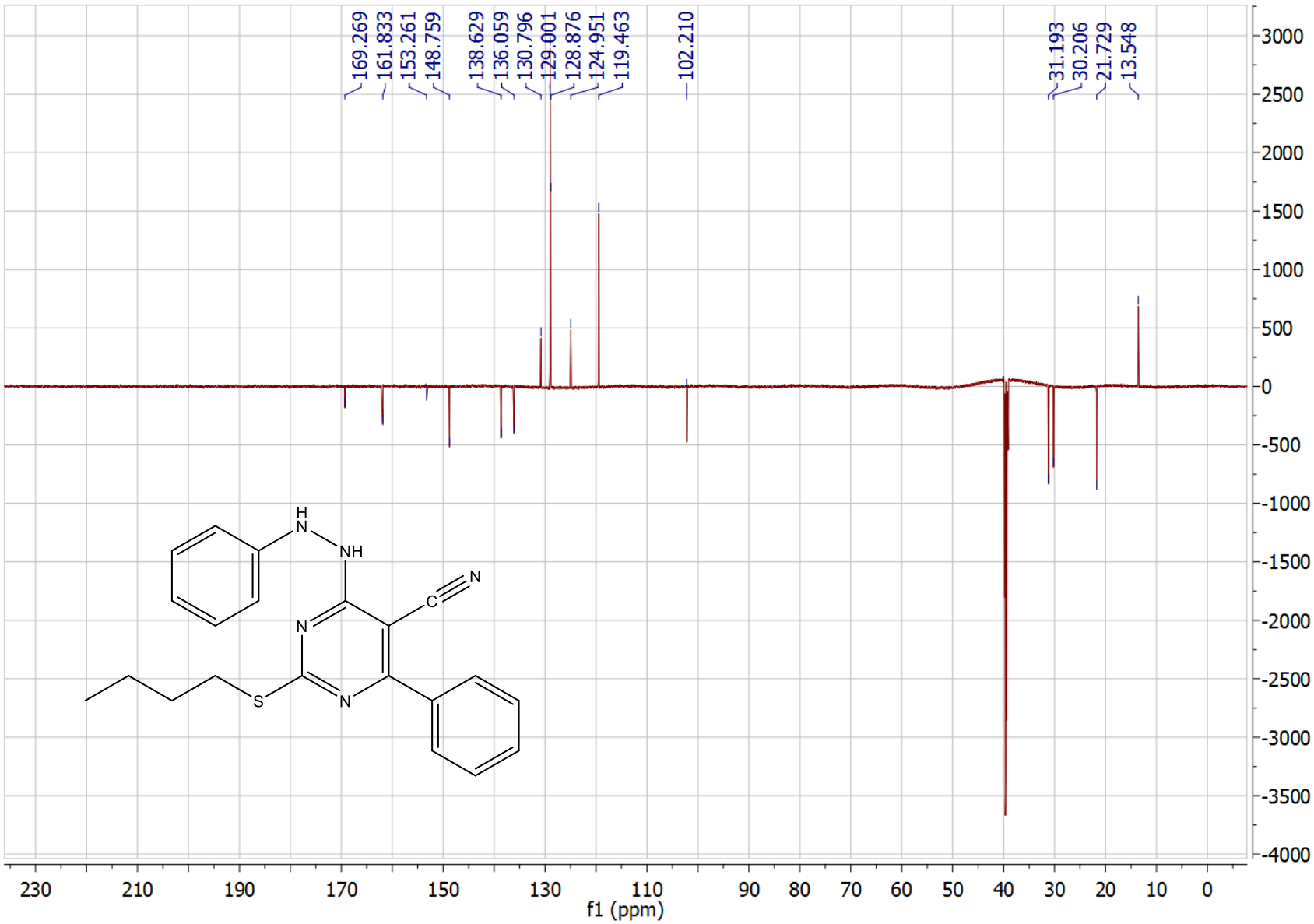
^{1}H NMR of compound 18_b



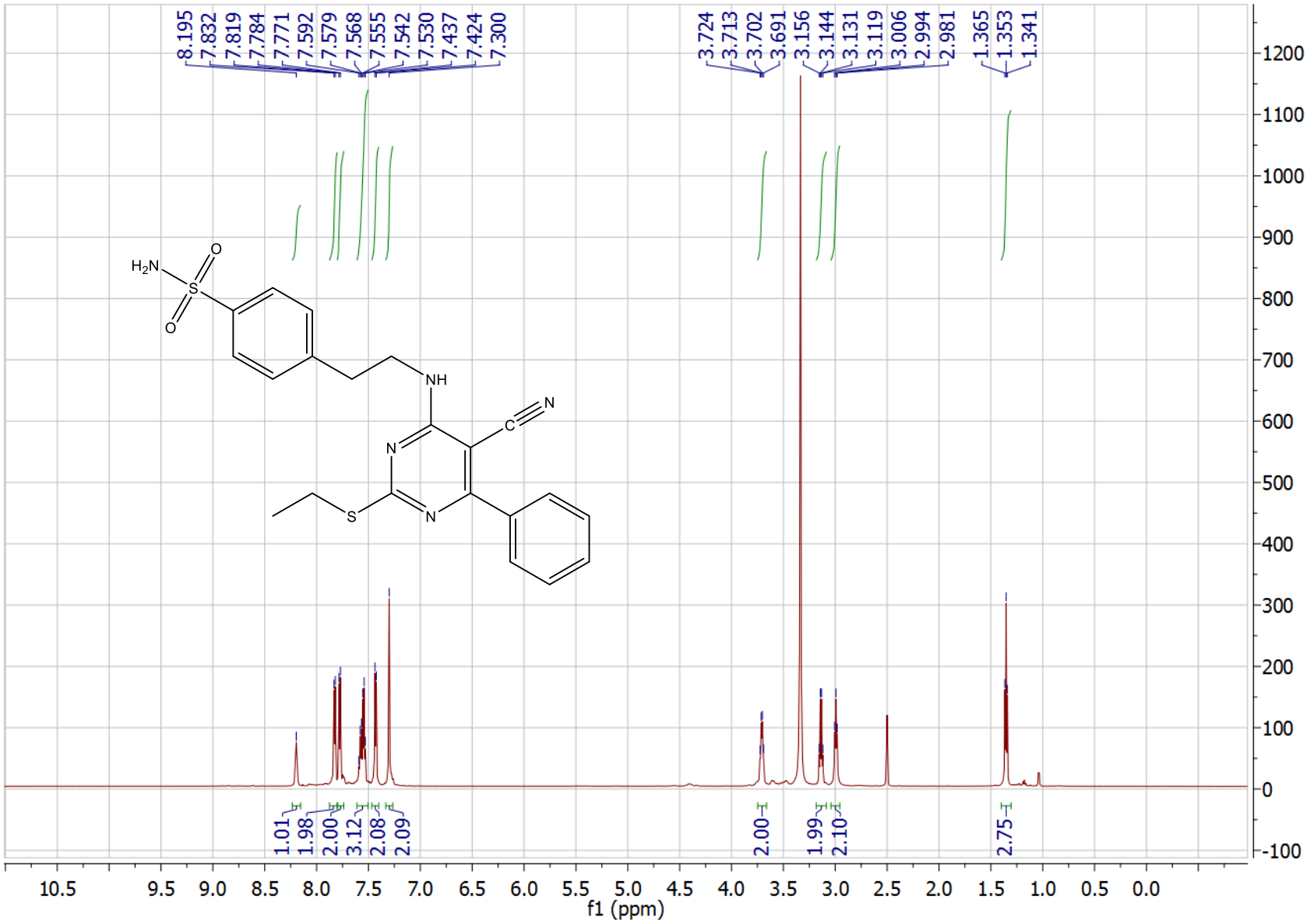
^{13}C NMR of compound 18$_b$



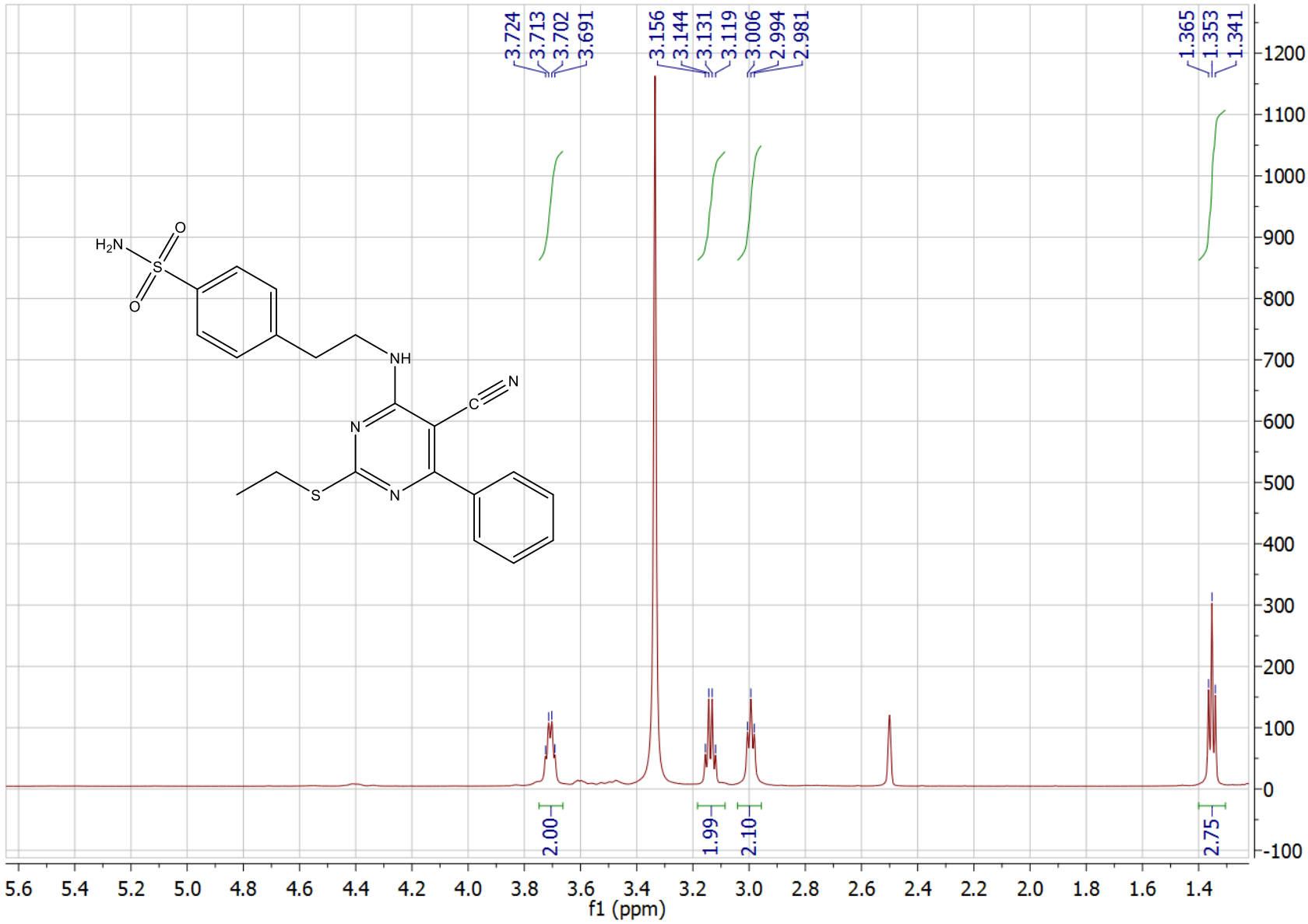
APT of compound 18_b



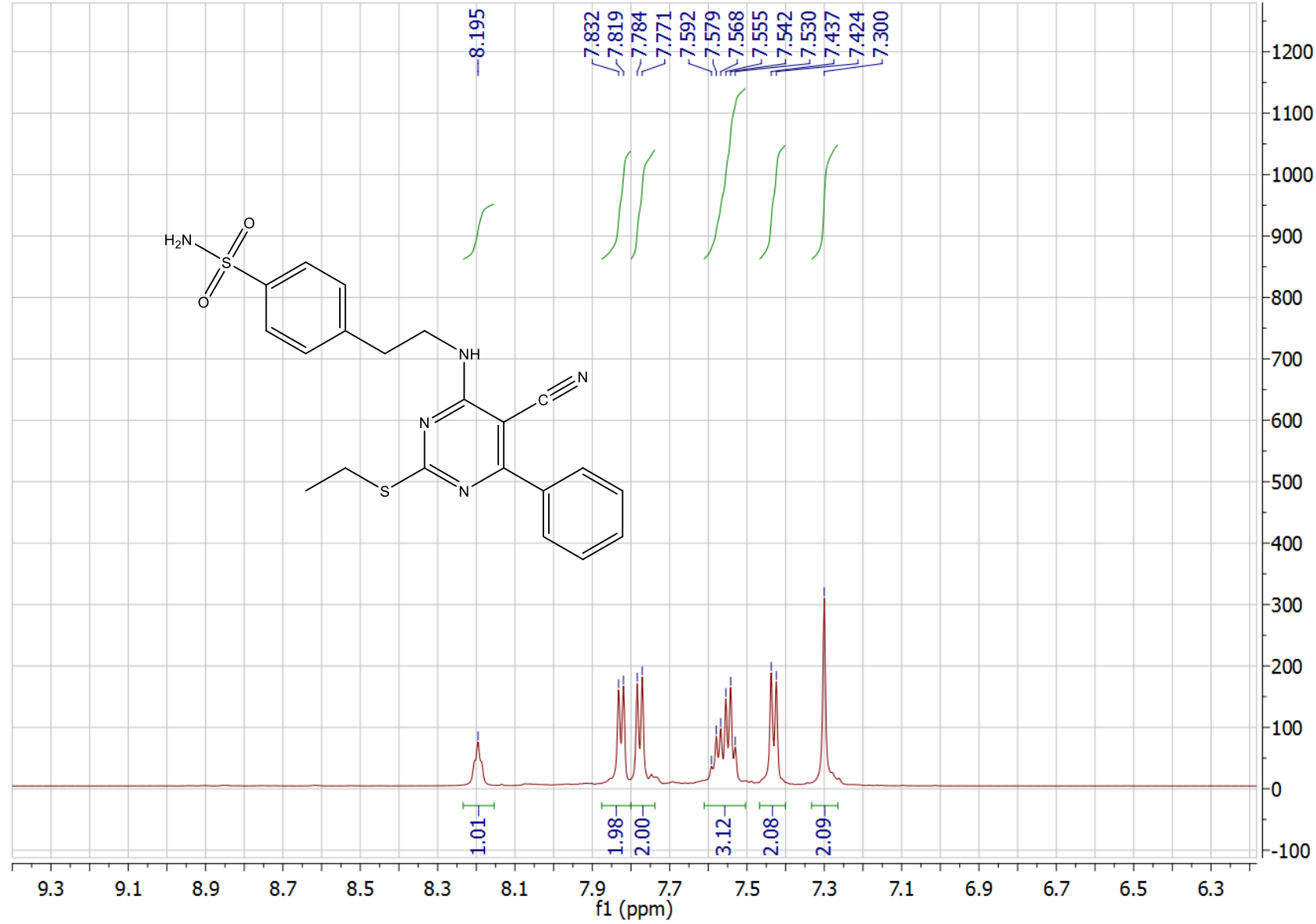
^{1}H NMR of compound 19$_a$



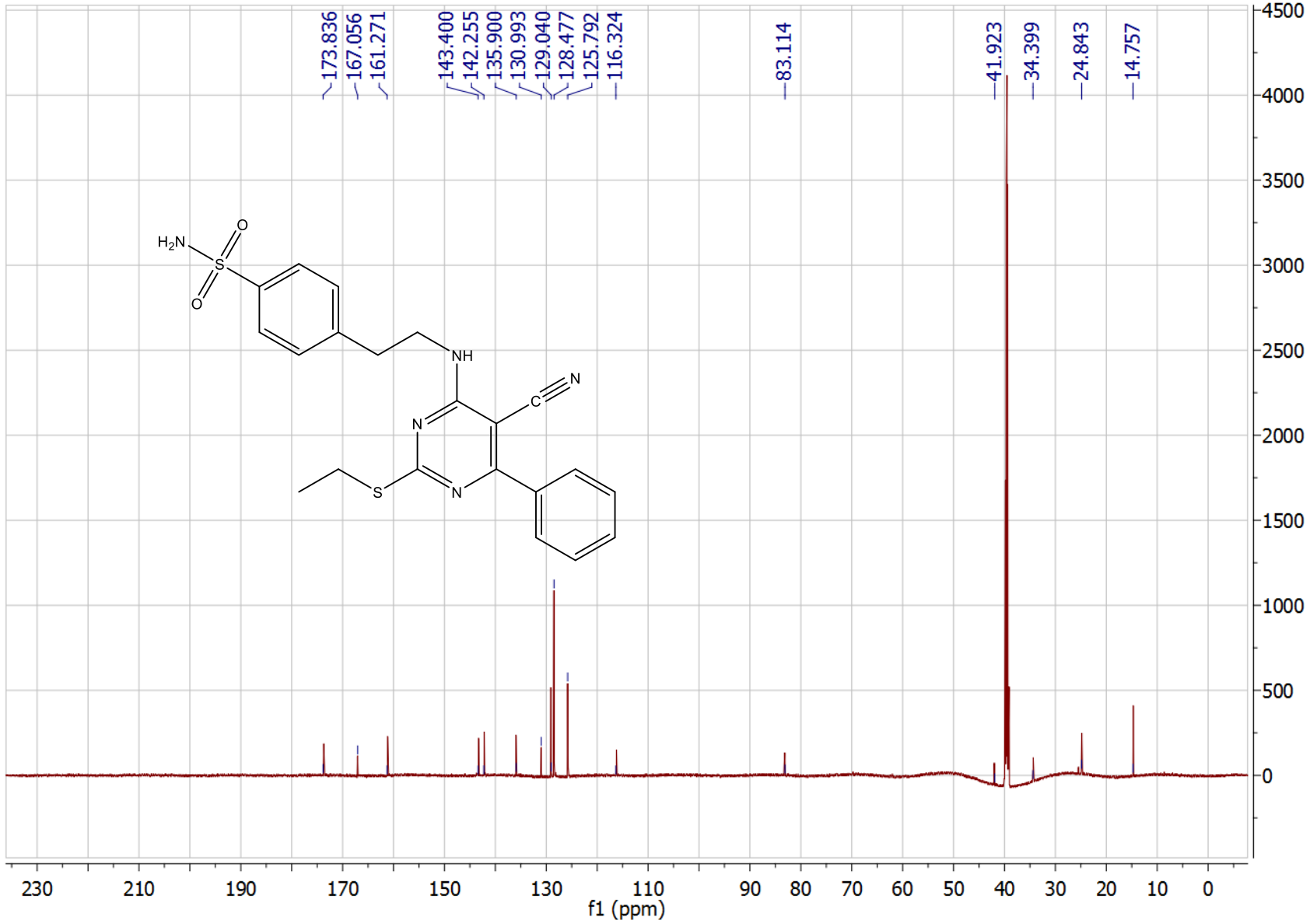
^{1}H NMR of compound 19$_{a}$



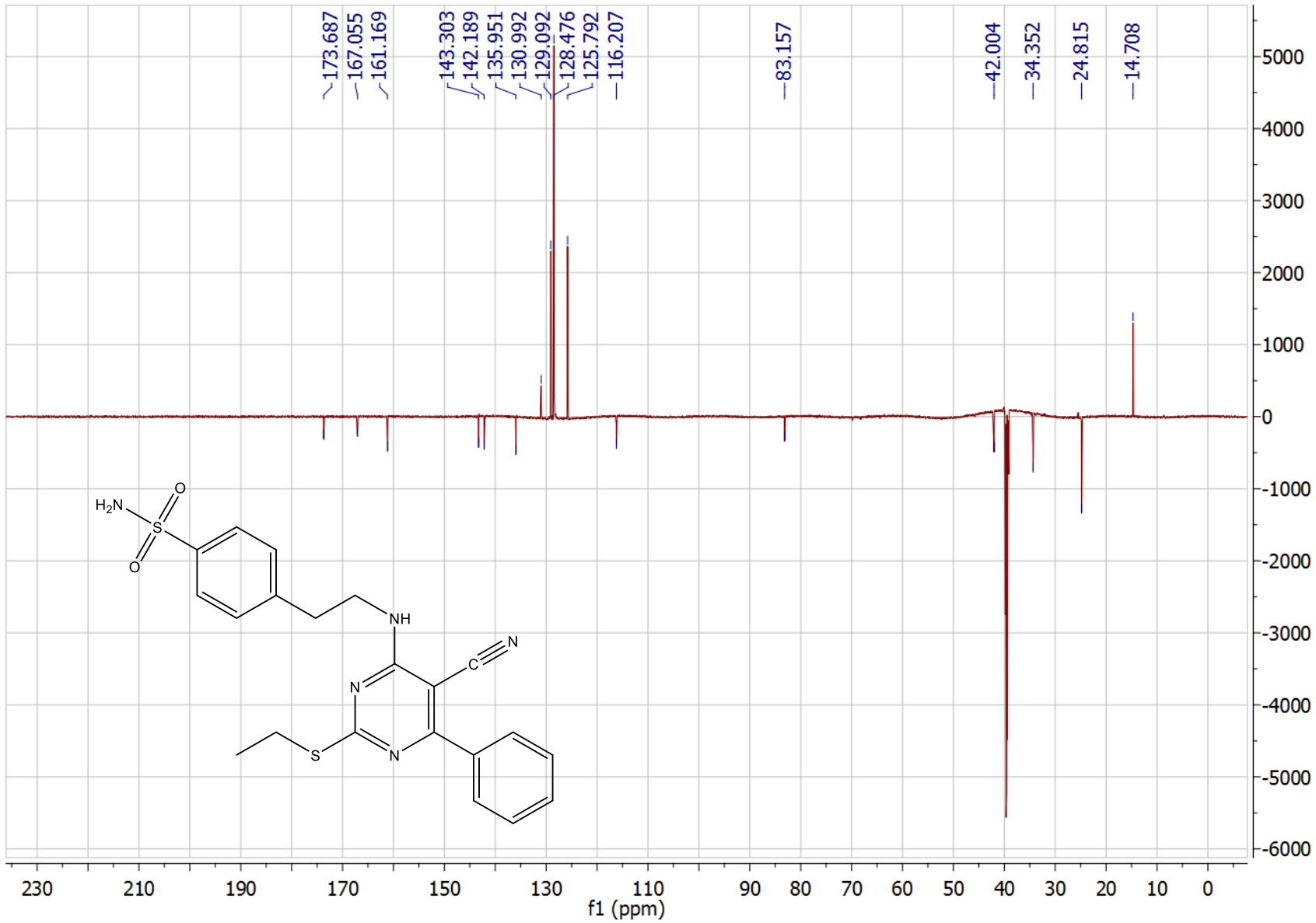
^{1}H NMR of compound 19$_{a}$



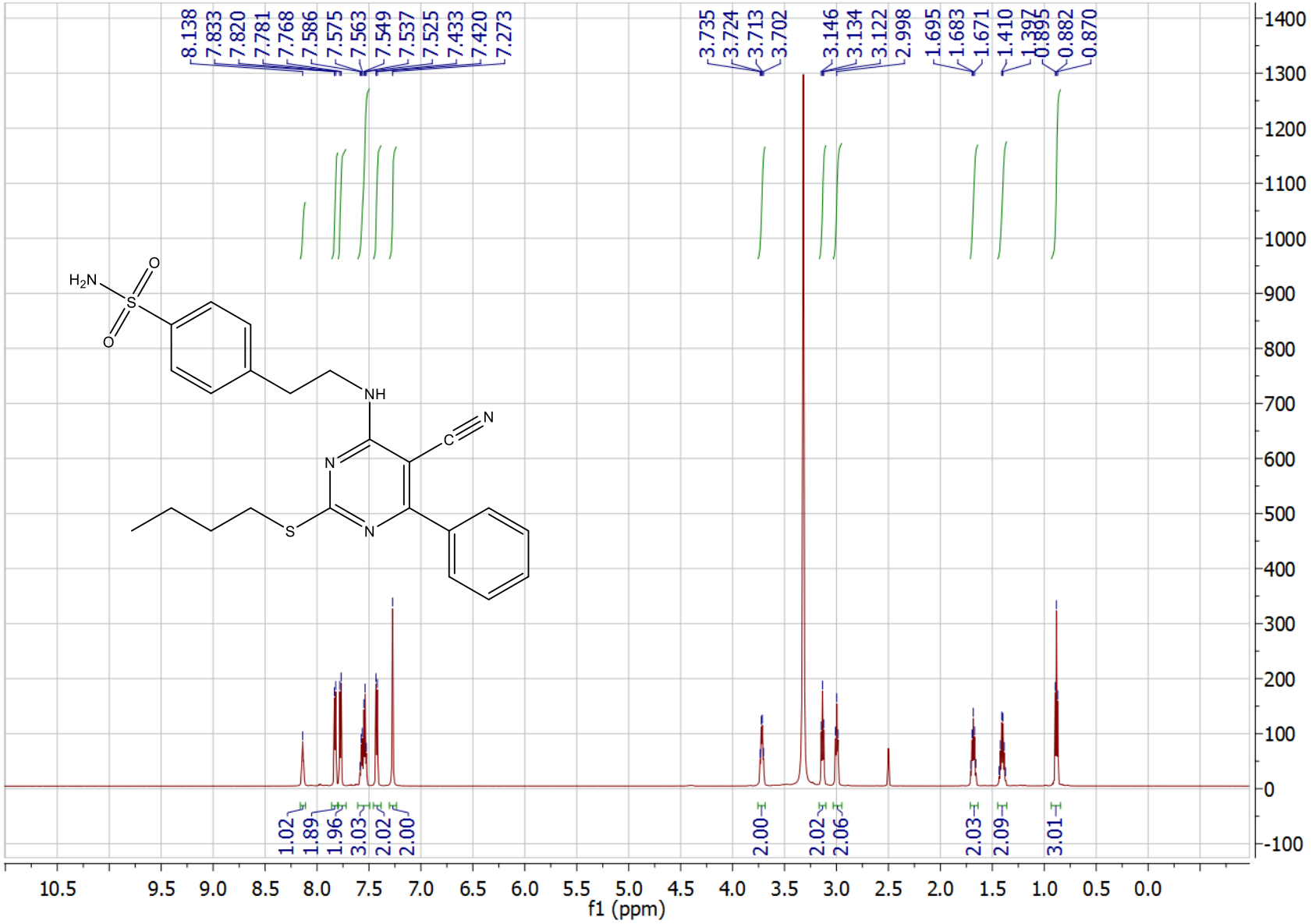
^{13}C NMR of compound 19_a



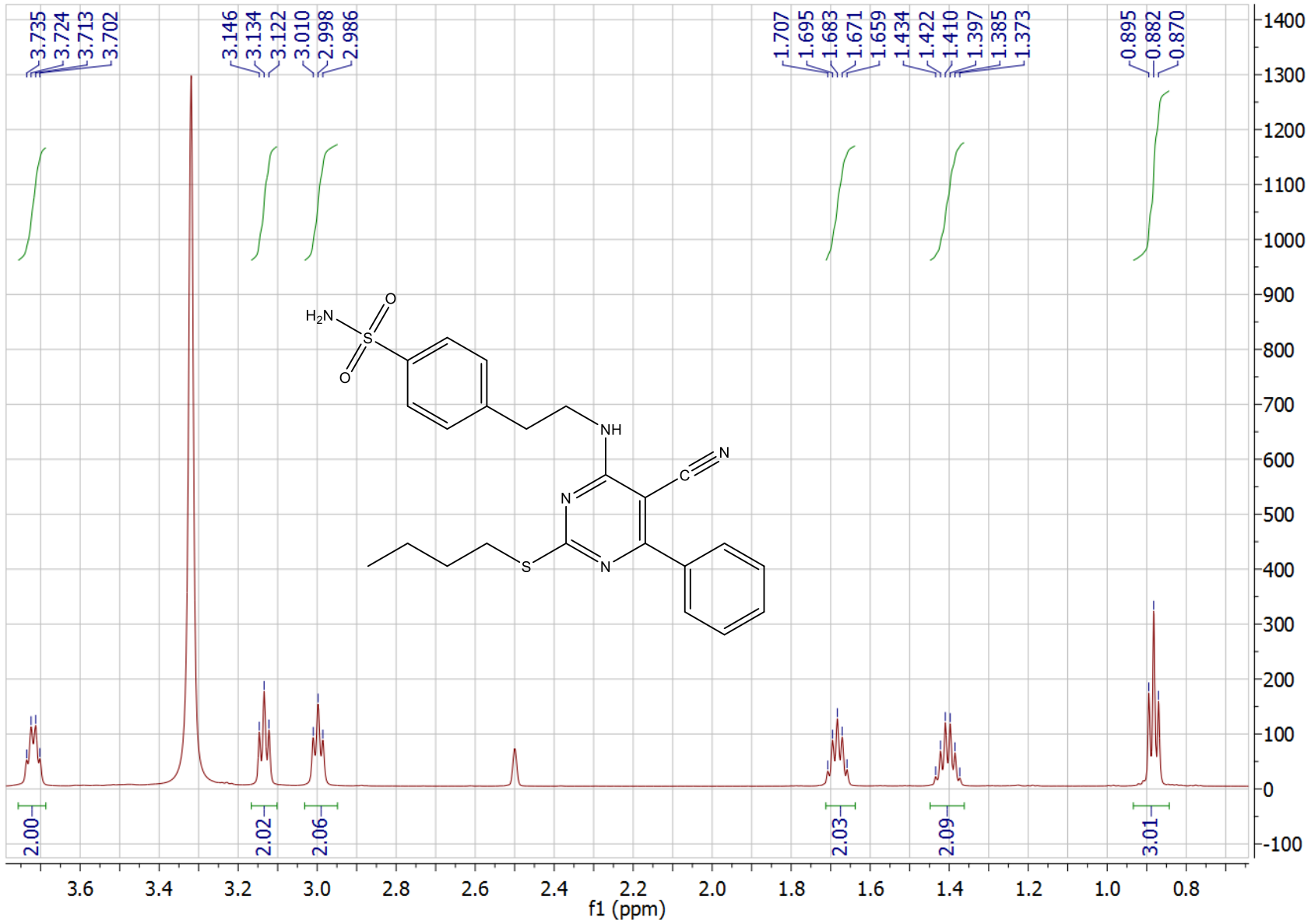
APT of compound 19_a



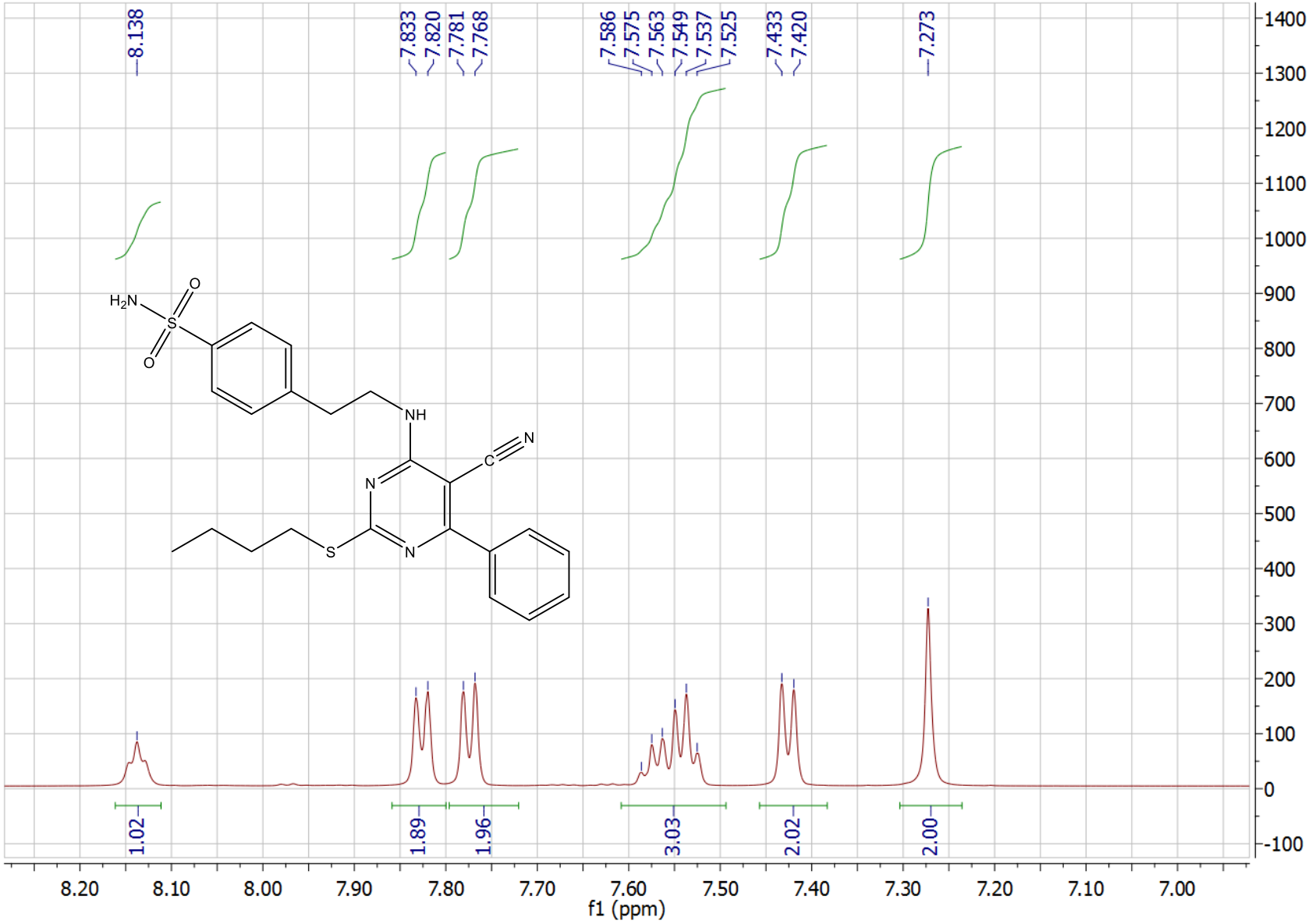
^{1}H NMR of compound 19_b



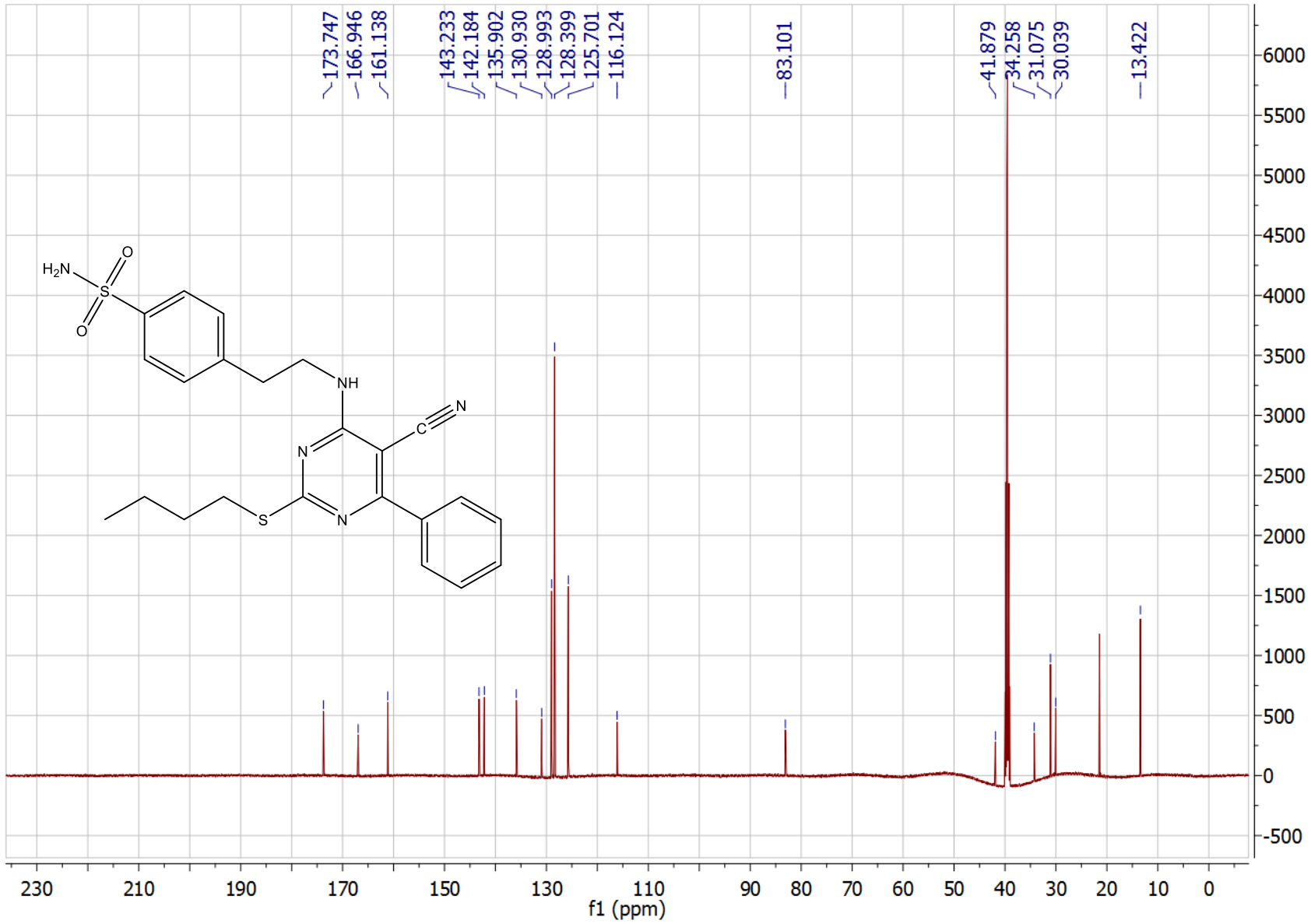
^{1}H NMR of compound 19_b



^{1}H NMR of compound 19$_{b}$



^{13}C NMR of compound 19_b



APT of compound 19_b