Supporting Information

Development of *N*-(Functionalized Benzoyl)-homocycloleucyl-glycinonitriles as Potent Cathepsin K Inhibitors

Jure Borišek,[†] Matej Vizovišek,[‡] Piotr Sosnowski,[‡] Boris Turk,[‡,∥, §] Dušan Turk,*[,‡,§] Barbara Mohar,[†] and Marjana Novič*[,†]

[†]National Institute of Chemistry, Hajdrihova 19, SI-1001 Ljubljana, Slovenia
[‡]Department of Biochemistry, Molecular and Structural Biology, Jozef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[§]Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[I]Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia

Table of Contents

Table S1. Distribution of the various *N*-(functionalized benzoyl)-homocycloleucyl-glycinonitriles in the virtual combinatorial library.

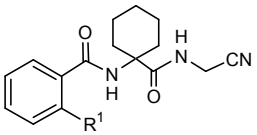
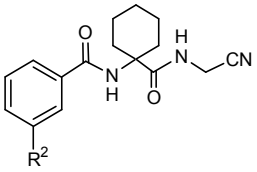
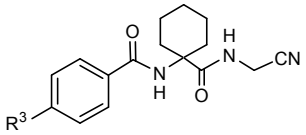
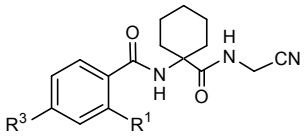
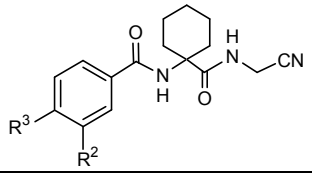
Virtual compound structure	Number of virtual compounds
	39
	46
	94
	18
	57
	Total: 254

Table S2. The post-docking analysis of the experimental and combinatorial datasets using the 4DMX binding site and the consensus results (merged post-docking and QSAR evaluation of the combinatorial dataset).

For this, see the appended Excel file.

Table S3. Molecular docking-calculated H-bond lengths (Å) between Cat K and compound **5**, **6**, or **9**.

Cat K amino acid residue	Compound atom	H-bond length with compound **5**	H-bond length with compound **6**	H-bond length with compound **9**
Cys25	C1	1.66	n.i.	1.77
Gln19	N1	3.02	n.i.	3.01
Asn161	N3	3.01	3.21	3.06
Gly66	O5	3.11	2.91	3.16
Gly66	N12	2.73	3.18	3.01
Gly64	N21	2.90	2.87	n.i.
Asn117[(b)]	N21[(a)]	2.53	3.20	n.i.
Asp61[(b)]	F22[(a)]	n.i.	n.i.	3.05
Glu118[(b)]	F22[(a)]	n.i.	n.i.	3.06

[a] First asymmetric unit of cathepsin K crystal structure.
[b] Second asymmetric unit of cathepsin K crystal structure.
n.i.: no interaction.

Table S4. Data and refinement statistics extracted by Phenix software.[1]

Cat K –"Ligand" complex	Compound 5	Compound 6	Compound 9
Wavelength (Å)	1.584	1.584	1.0
Resolution range (Å)	45.99 - 1.59 (1.642 - 1.585)	46.28-1.87 (1.937-1.87)	46.37 - 1.0 (1.036 - 1.0)
Space group	P 1 21 1	P 21 21 21	P 21 21 21
Unit cell	47.22 44.40 51.35 90.00 116.42 90.00	36.46 54.45 87.87 90.00 90.00 90.00	36.44 54.55 88.01 90.00 90.00 90.00
Total reflections	49603 (4338)	141950 (10892)	540950 (25908)
Unique reflections	25360 (2388)	14804 (1469)	92375 (6766)
Multiplicity	1.9 (1.8)	9.5 (7.4)	5.9 (3.8)
Completeness (%)	97.53 (90.01)	98.31 (916)	96.78 (71.72)
Mean I/sigma(I)	9.12 (1.17)	9.96 (3.78)	15.89 (1.79)
Wilson B-factor	13.81	8.42	8.34
R-merge	0.08534 (1.124)	0.2593 (0.9815)	0.05199 (0.5977)
R-work	0.1896 (0.3436)	0.1810 (0.2420)	0.1115 (0.2151)
R-free	0.2414 (0.3366)	0.2371 (0.3690)	0.1351 (0.2057)
Number of non-hydrogen atoms	2004	1909	2118
Protein residues	215	215	215
RMS (bonds)	0.019	0.019	0.022
RMS (angles)	1.91	1.76	2.10
Ramachandran favoured (%)	97	97	97
Ramachandran outliers (%)	0	0	0
Clashscore	3.12	3.31	3.01
Average B-factor	19	9	13.4
macromolecules	17.3	7.9	10.3
ligands	25.6	15.20	14.7
solvent	30.6	16.40	33.5

Statistics for the highest-resolution shell are shown in parentheses.

(1) (a) Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **2010**, *68*, 352–367.
(b) Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; McCoy, A.J.; Moriarty, N.W.; Oeffner, R.; Read, R.J.; Richardson, D.C.; Richardson, J.S.; Terwilliger, T.C.; Zwart, P.H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **2010**, *66*, 213–221.

Scheme S1. General synthetic pathway for the virtual combinatorial library of the *N*-(functionalized benzoyl)-homocycloleucyl-glycinonitriles.

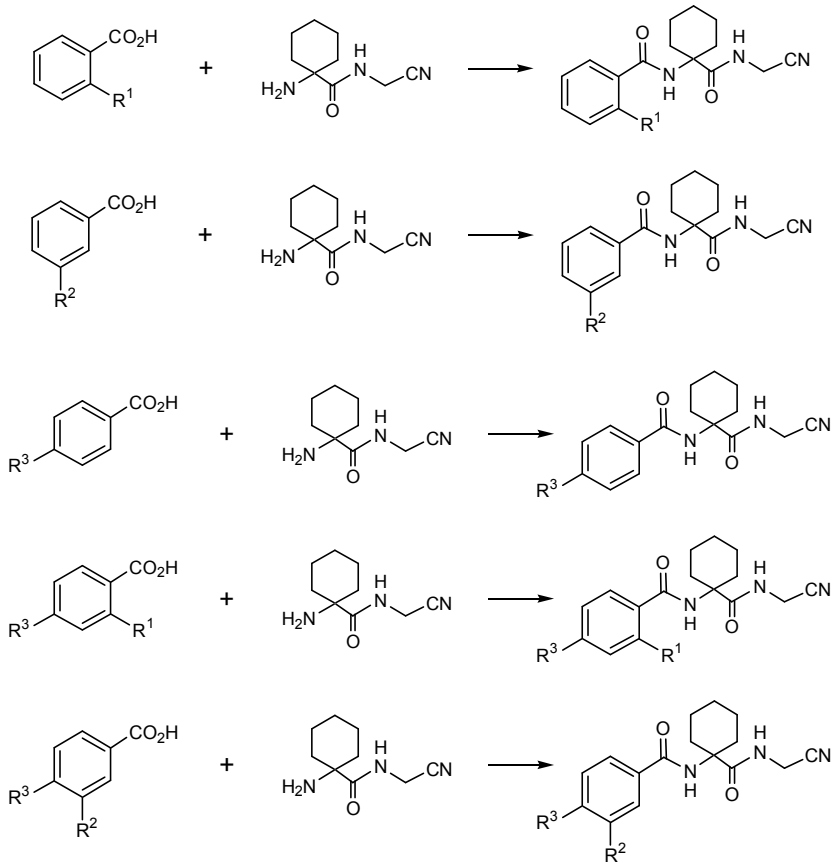
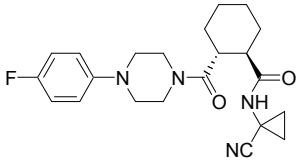


Figure S1. 0LB ligand.



(1*R*,2*R*)-*N*-(1-cyanocyclopropyl)-2-{[4-(4-fluorophenyl)piperazin-1-yl]carbonyl}cyclohexanecarboxamide

Figure S2. Reversibility of cathepsin K inhibition by compounds **5**, **6** and **9.** After 30 minutes of incubation of 1 μM cathepsin K with 5 μM compounds **5**, **6** or **9** (final concentration), the complex was rapidly diluted 100-fold in the reaction medium containing 50 μM Z-Gly-Pro-Arg-AMC, followed by monitoring substrate hydrolysis**.** E-64 was included as a control for irreversible inhibition. Substrate hydrolysis is plotted as relative fluorescence units (RFU) versus time. The solid curves represent the best fit to the equation $[P] = v_s t + (v_o - v_s)(1 - e^{-kt})/k$. All other experimental conditions were as described under Experimental section.

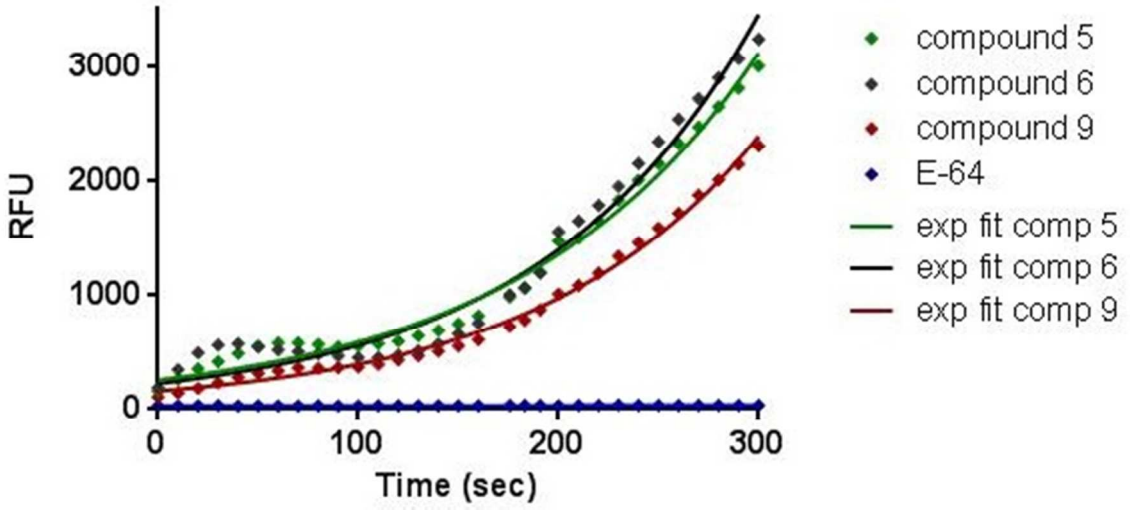


Figure S3. Progress curves for inhibition of cathepsin K by compounds **5**, **6** and **9**: (a) compound **5**; (b) compound **6**; (c) compound **9**. All experimental details are given in the Experimental Section (In vitro inhibition assay; Kinetics of Cat K inhibition by compounds **5**, **6** and **9**).

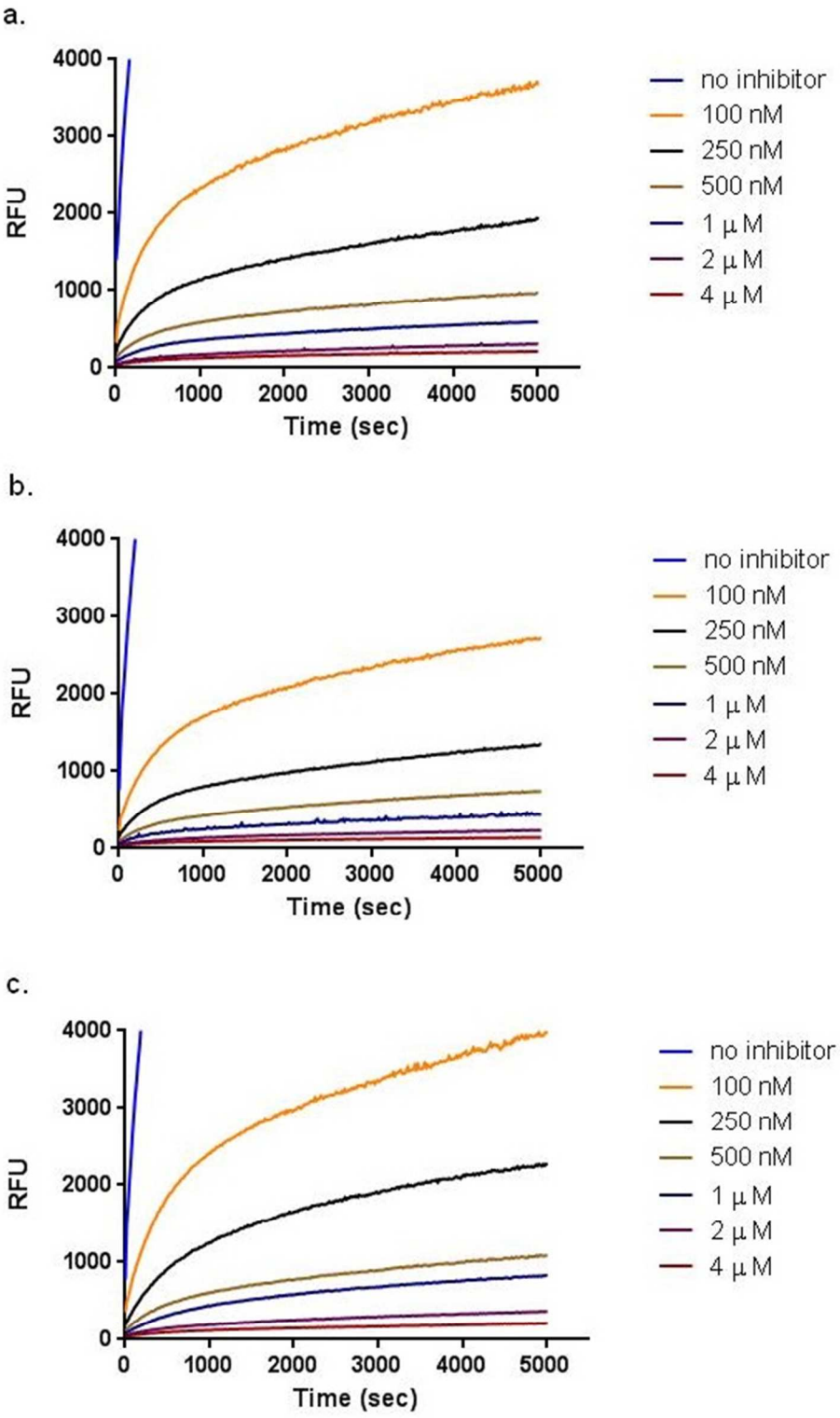
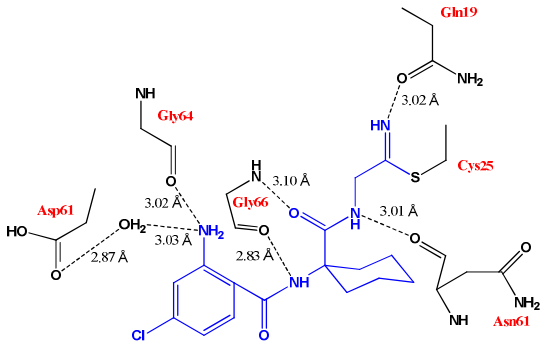


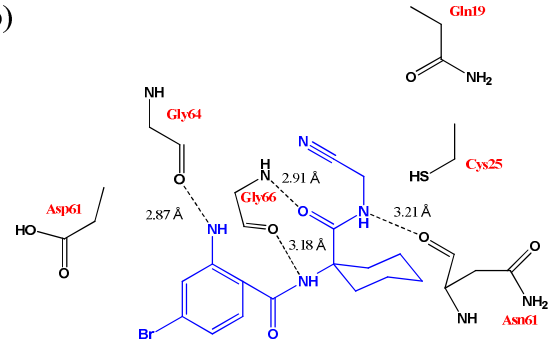
Figure S4. 2D representation of the key interactions (H-bonds) and distances in the crystal structures of cathepsin K: (a) for compound **5**, (b) for compound **6**, and (c) for compound **9**. Alternate conformation (noncovalent binding mode) of compound **9** and its interactions are colored in magenta.

a)



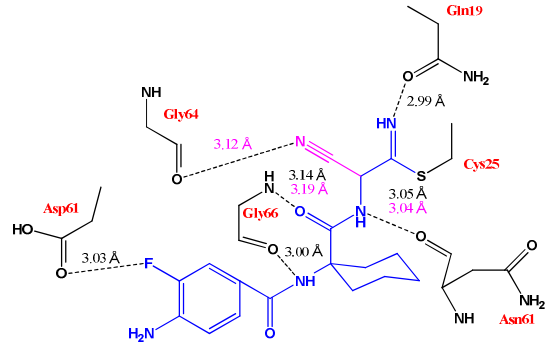
compound **5**

b)



compound **6**

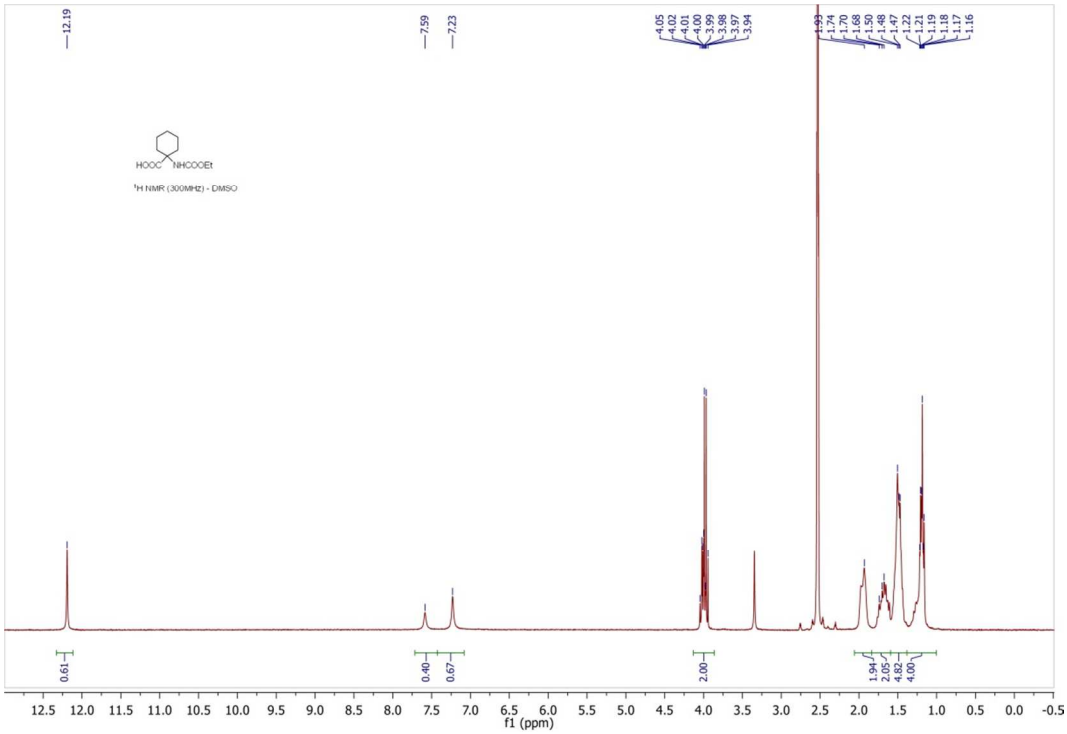
c)

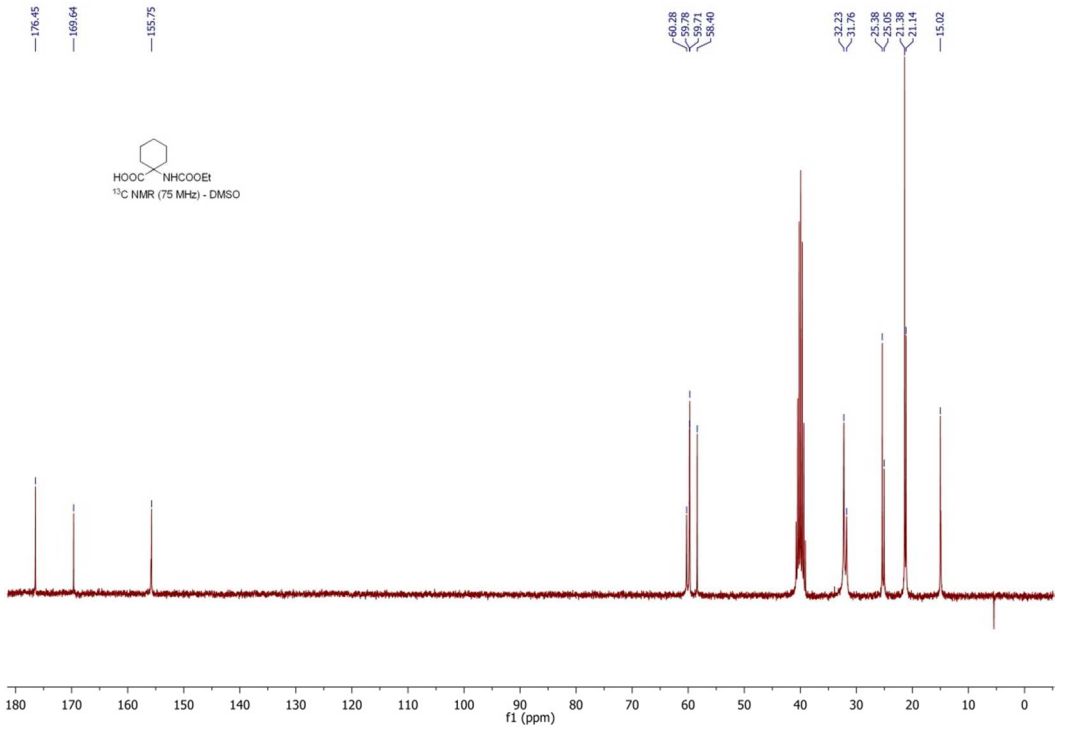


compound **9**

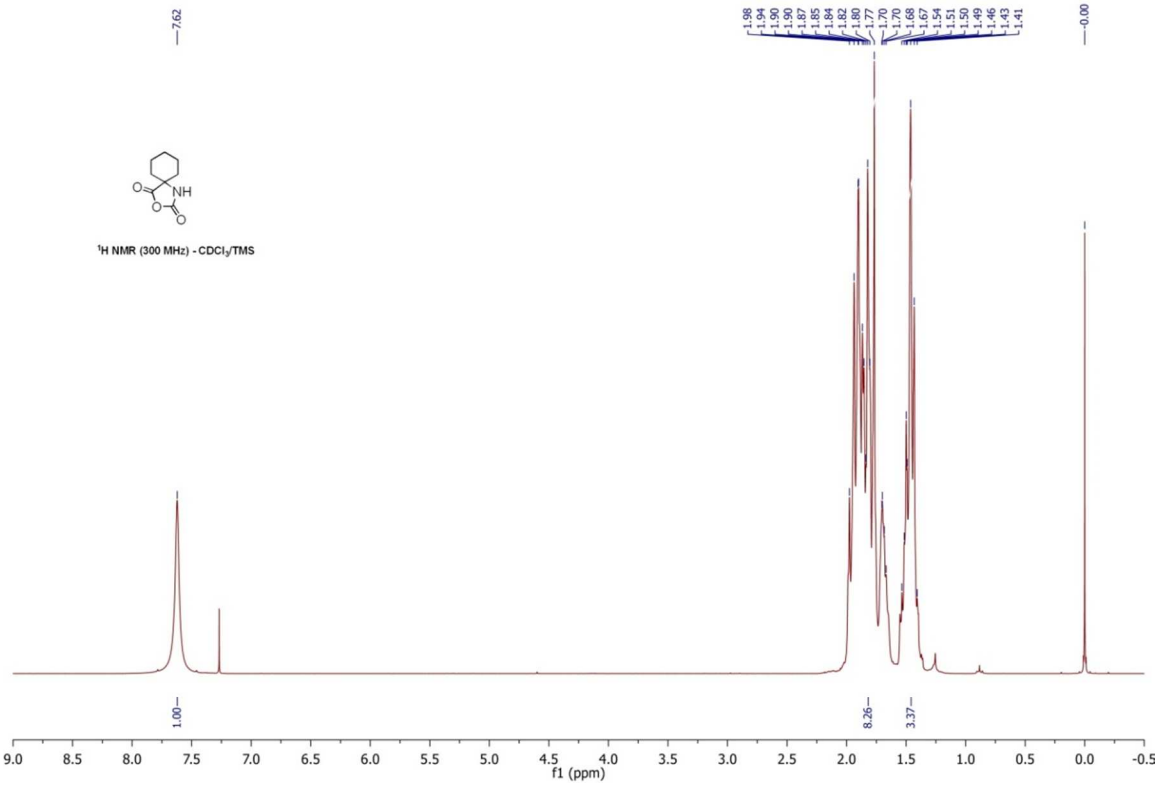
NMR spectra

Compound **1**

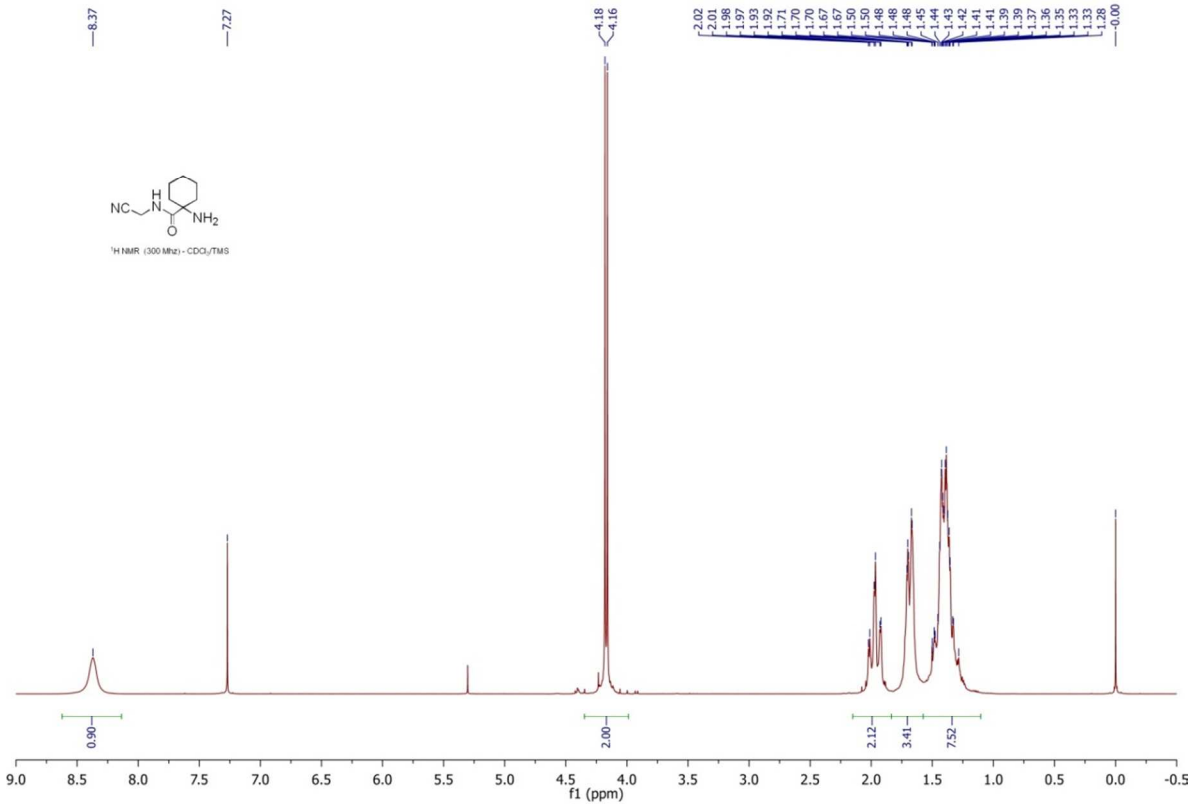




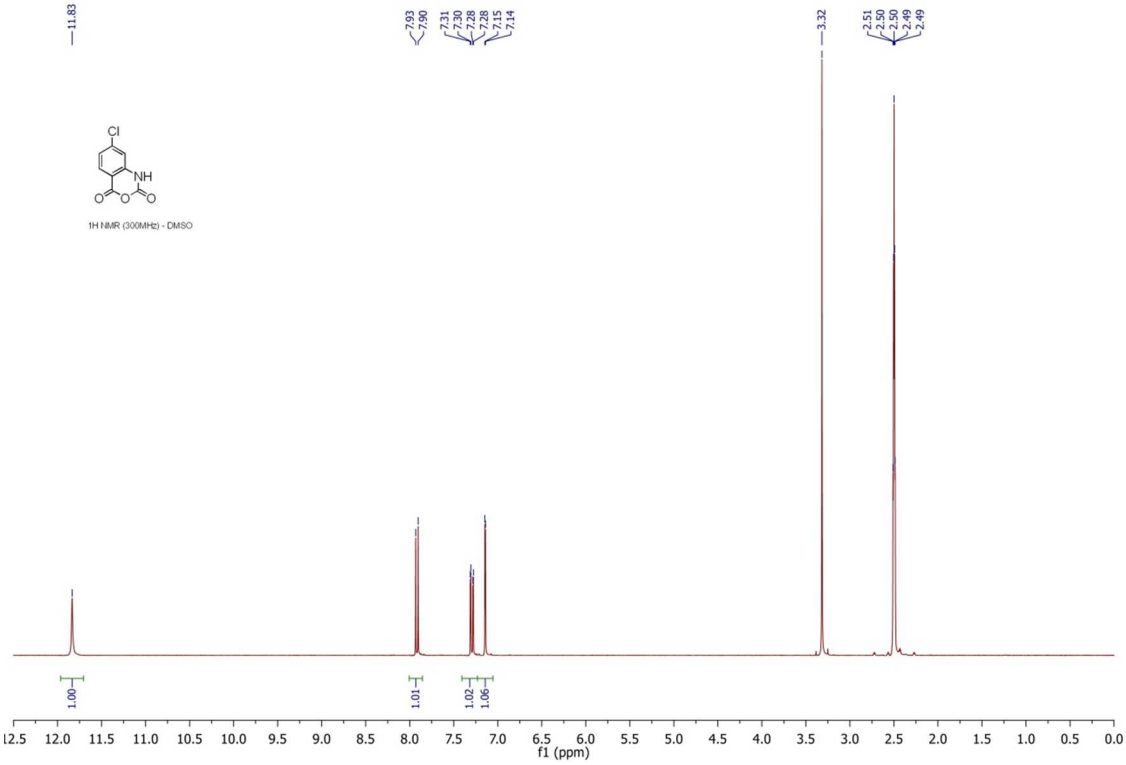
Compound **2**



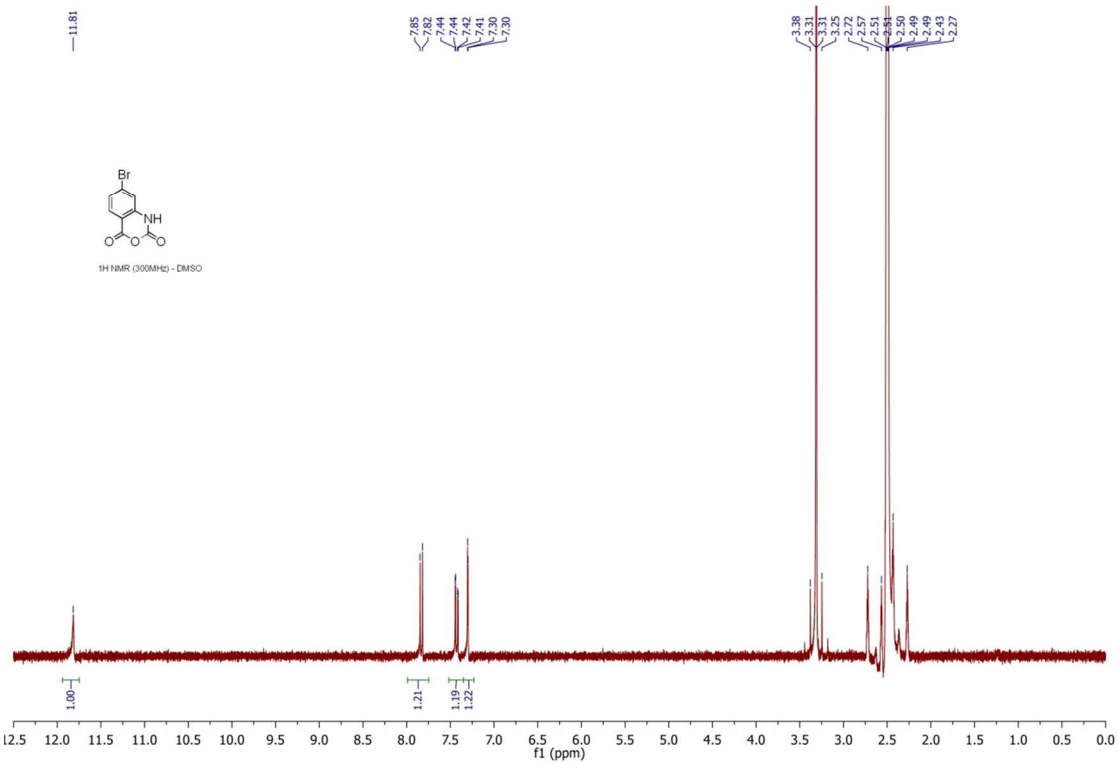
Compound **3**



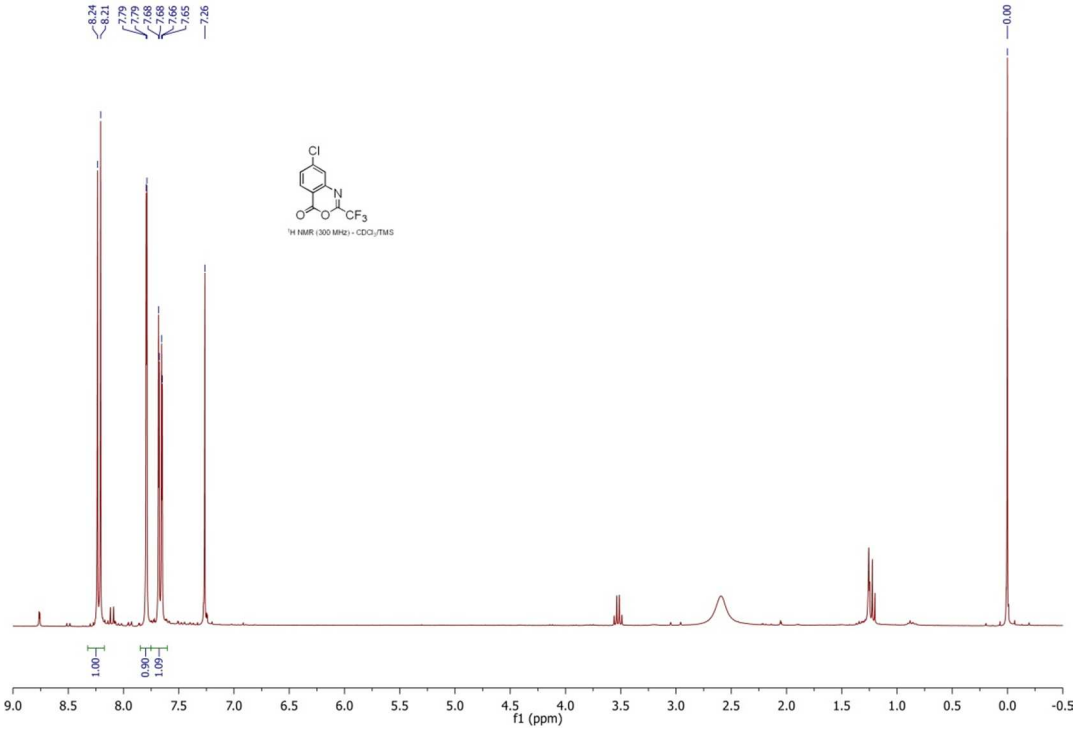
4-Chloroisatoic anhydride



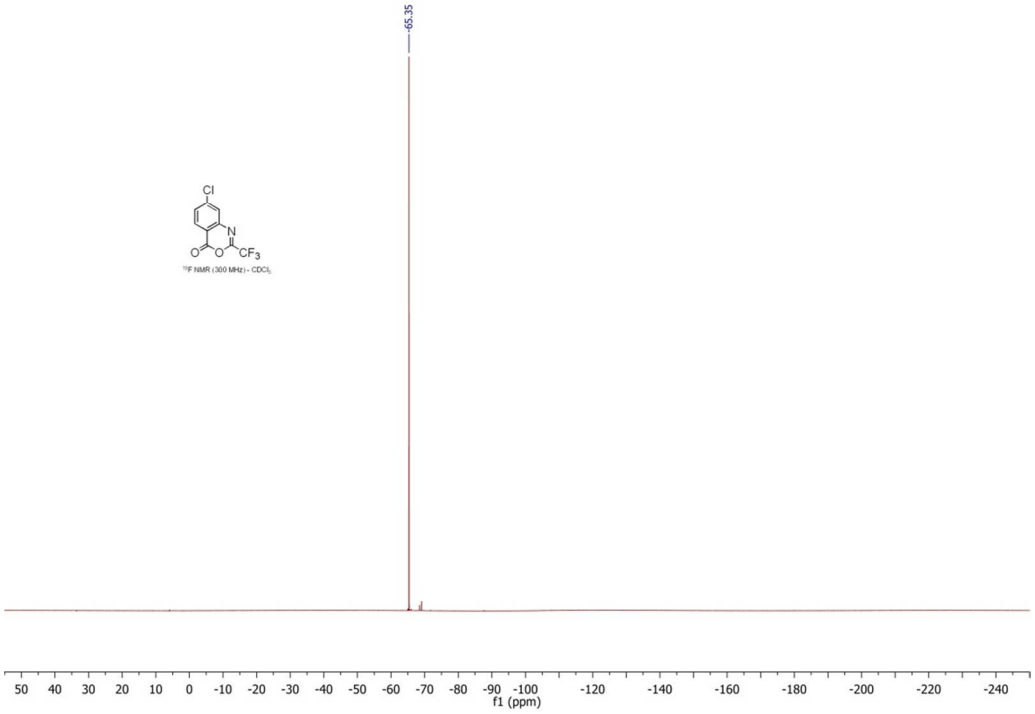
4-Bromoisatoic anhydride



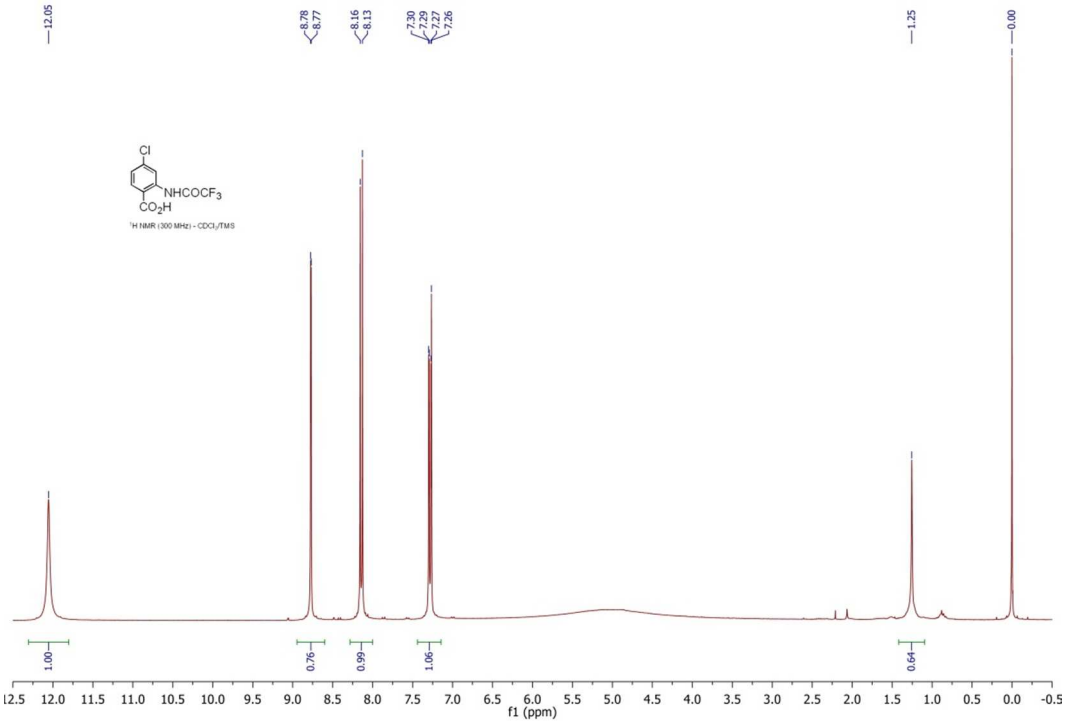
7-Chloro-2-trifluoromethyl-4*H*-3,1-benzoxazin-4-one

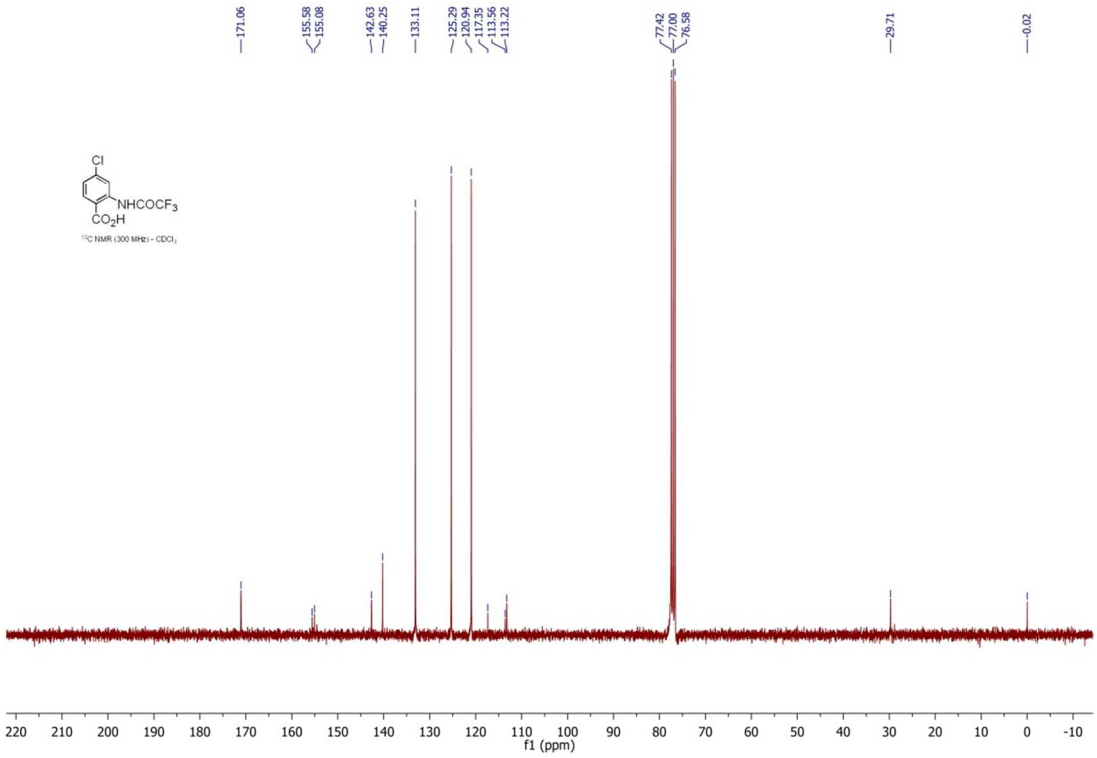


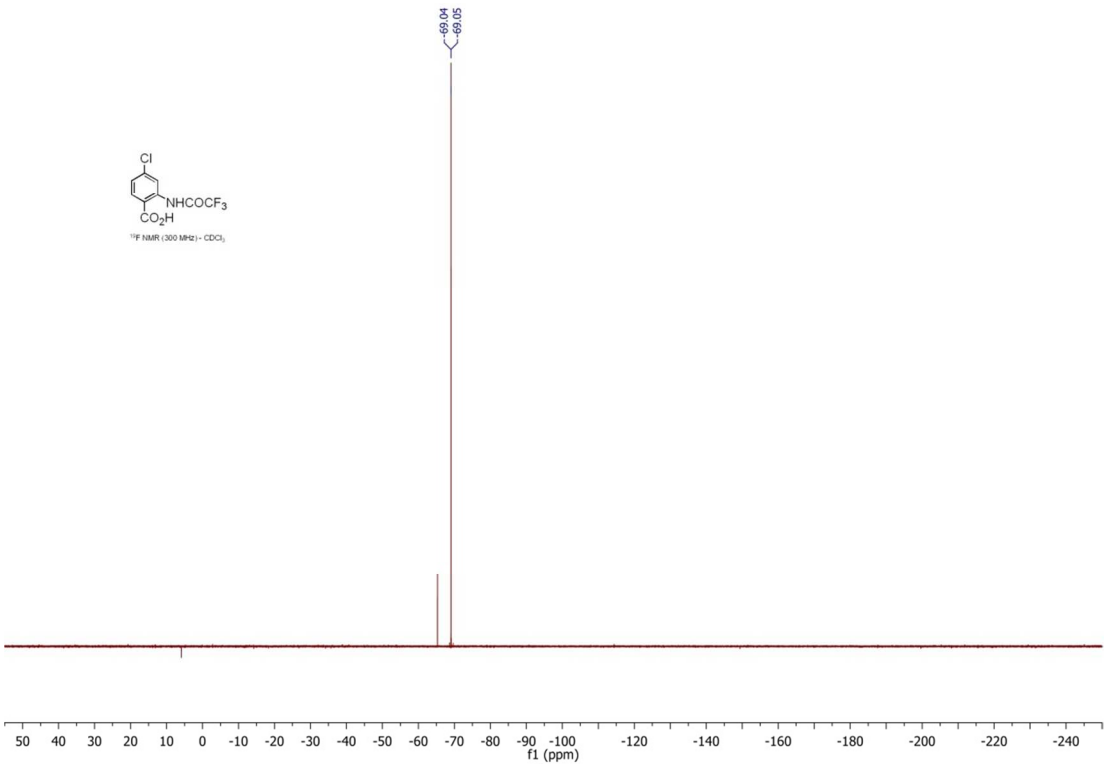
7-Chloro-2-trifluoromethyl-4*H*-3,1-benzoxazin-4-one



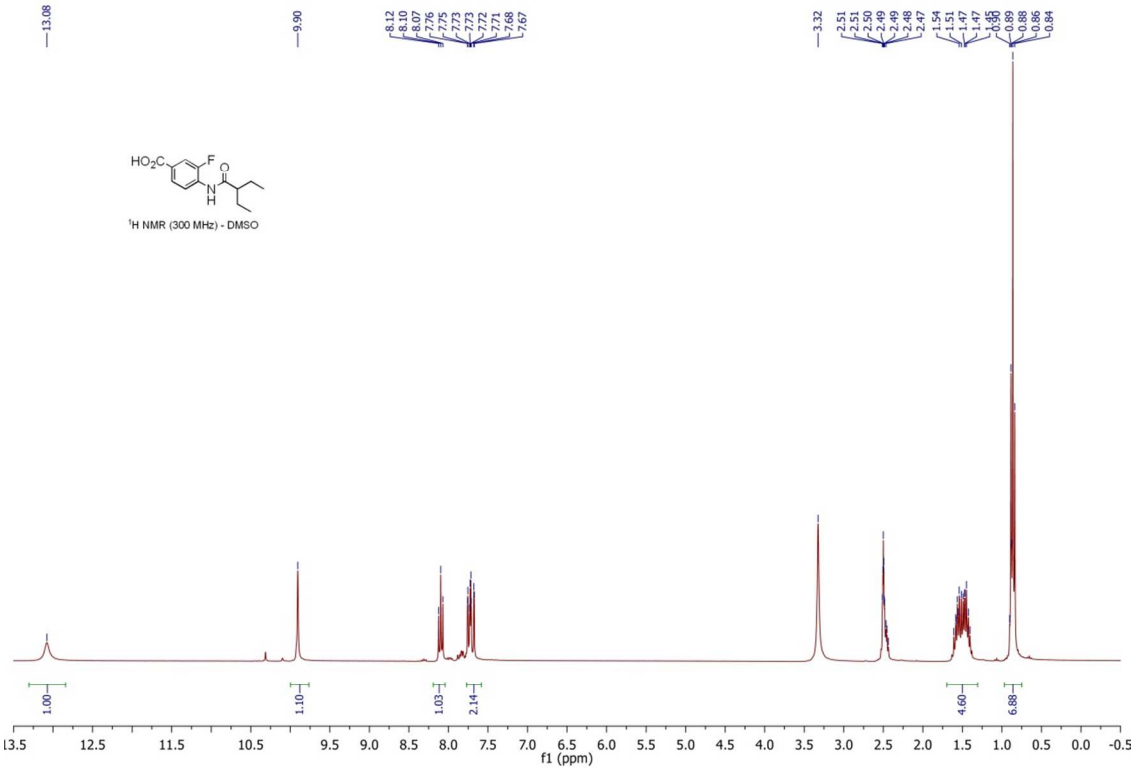
N-Trifluoroacetyl-4-chloroanthranilic acid

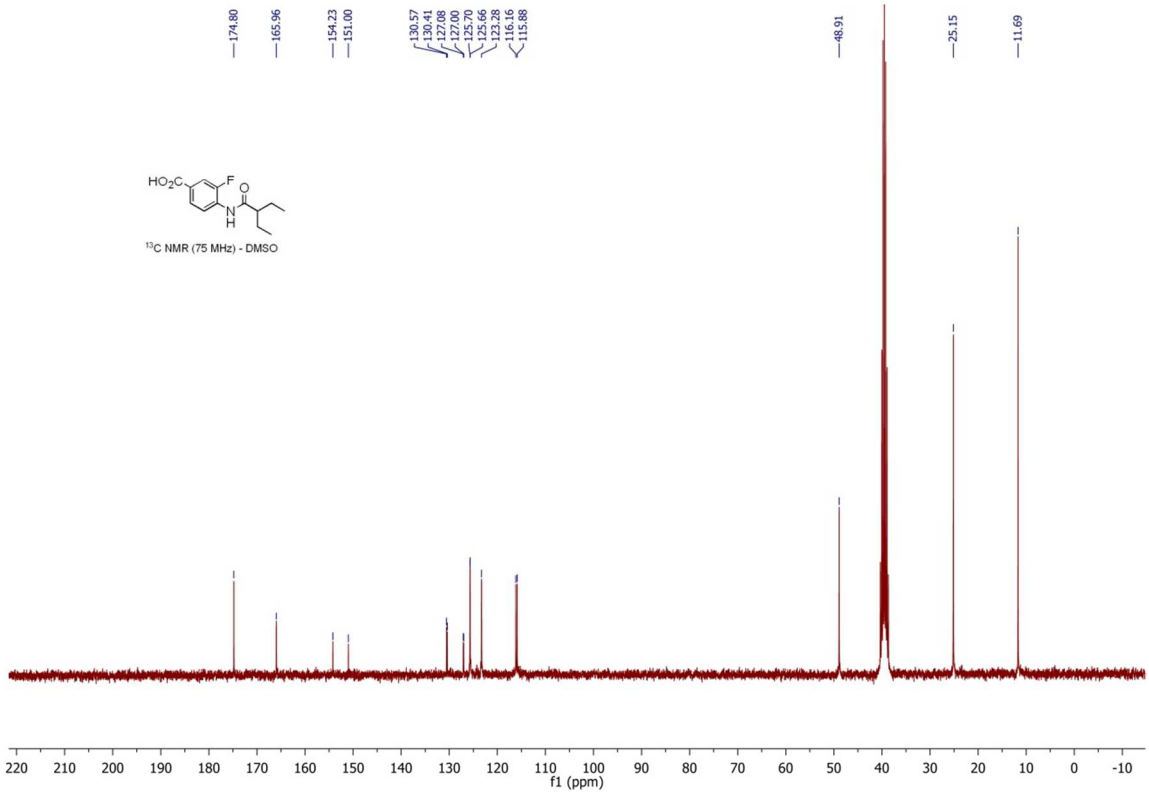


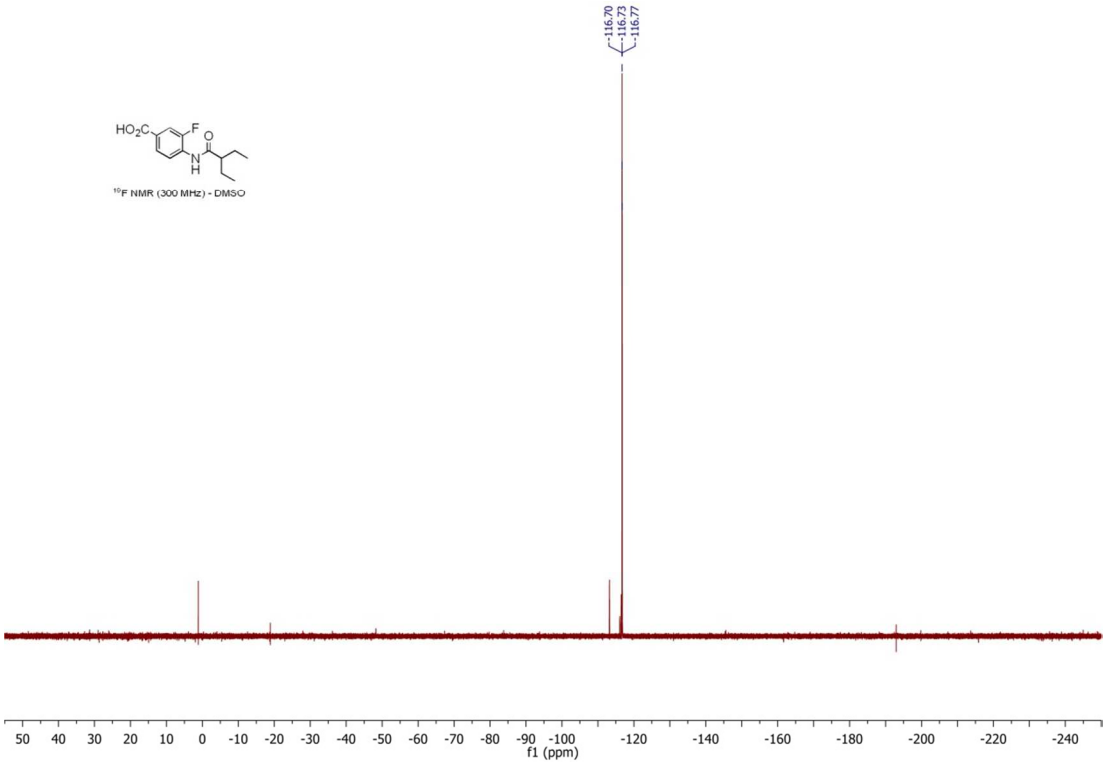



-69.04
-69.05
Cl
NHCOCF3
CO2H
19F NMR (300 MHz) - CDCl3
50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -120 -140 -160 -180 -200 -220 -240
f1 (ppm)

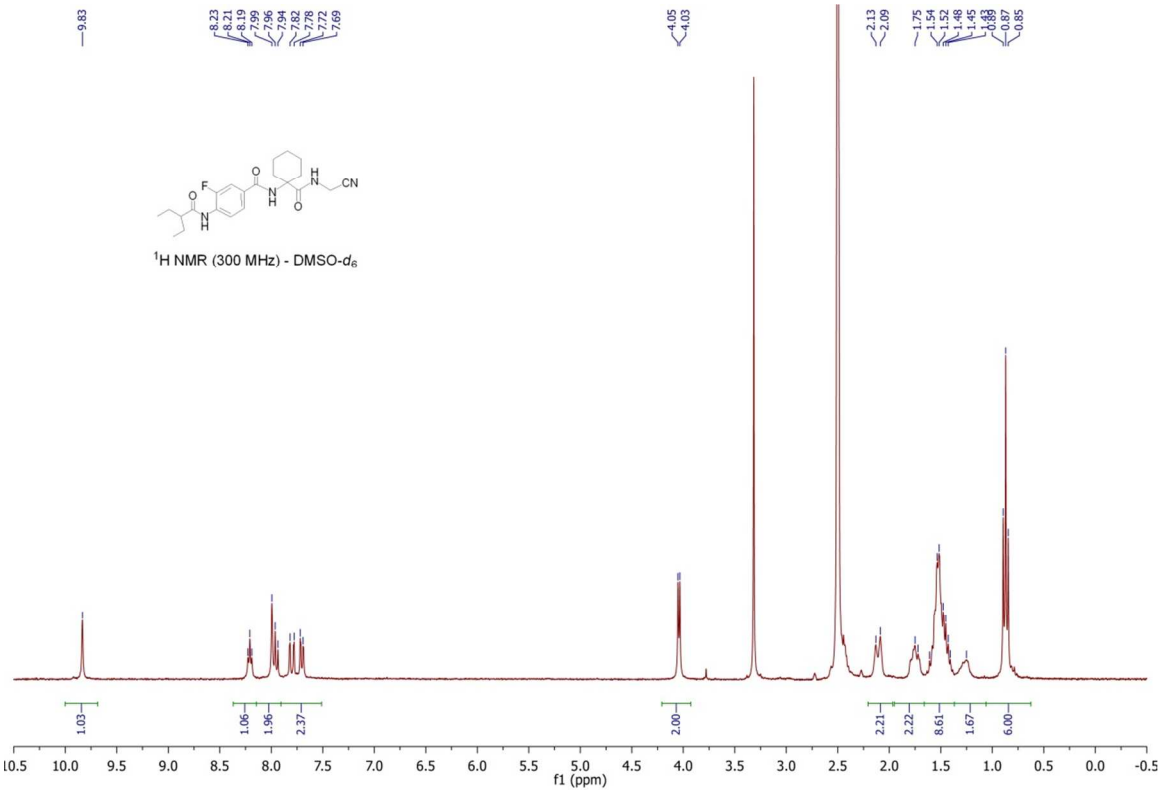
4-(2-Ethylbutyrylamino)-3-fluorobenzoic acid

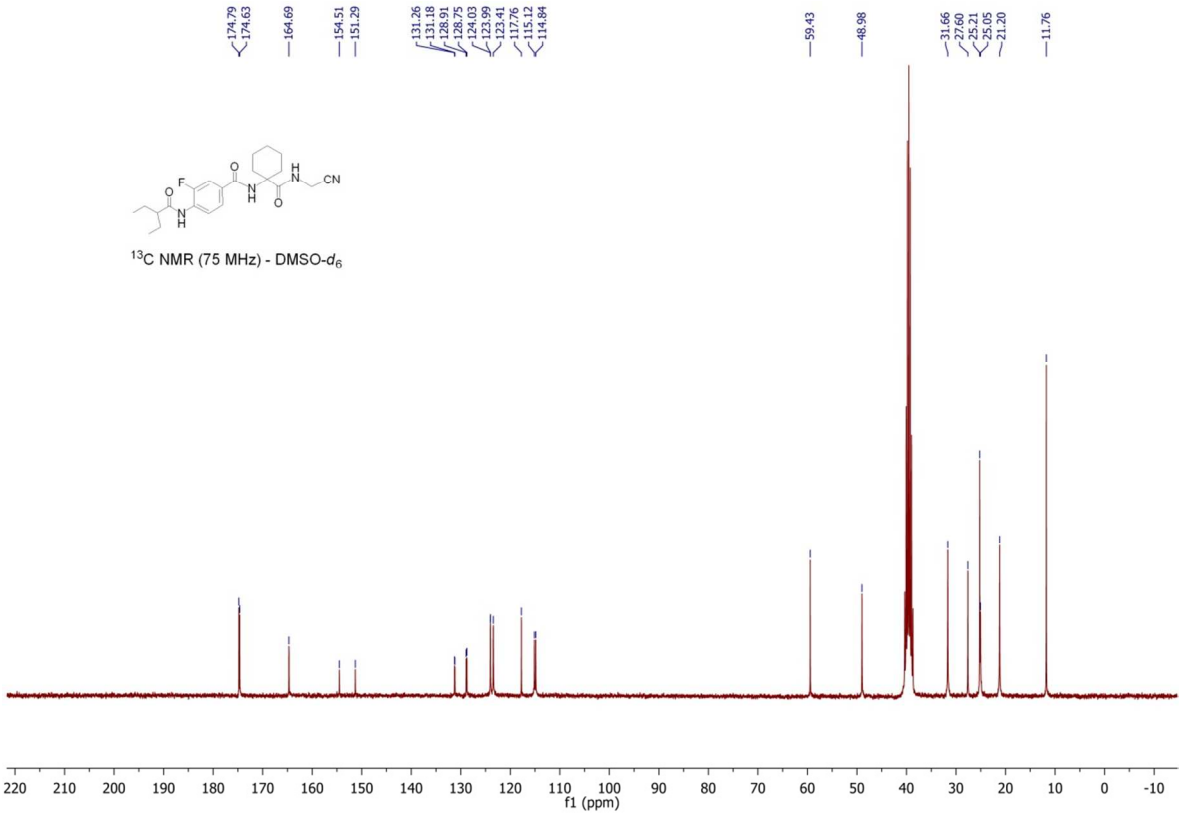




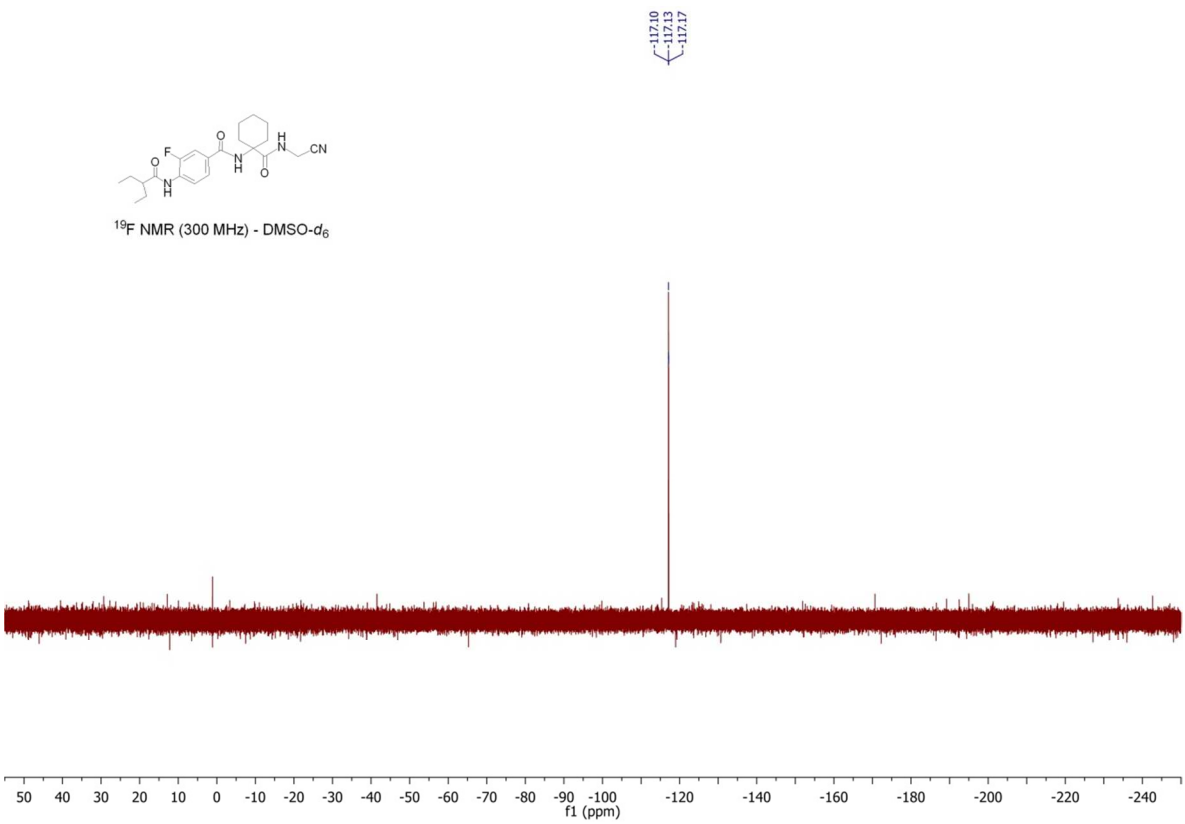

116.70
116.73
116.77
HO2C
F
O
N
H
19F NMR (300 MHz) - DMSO
50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -120 -140 -160 -180 -200 -220 -240
f1 (ppm)

Compound **4**


9.83
8.23
8.21
8.19
7.99
7.96
7.94
7.82
7.78
7.72
7.69
4.05
4.03
2.13
2.09
1.75
1.54
1.52
1.48
1.45
0.87
0.85
O
F
N
H
O
N
H
CN
O
1H NMR (300 MHz) - DMSO-d6
1.03
1.06
1.96
2.37
2.00
2.21
2.22
8.61
1.67
6.00
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

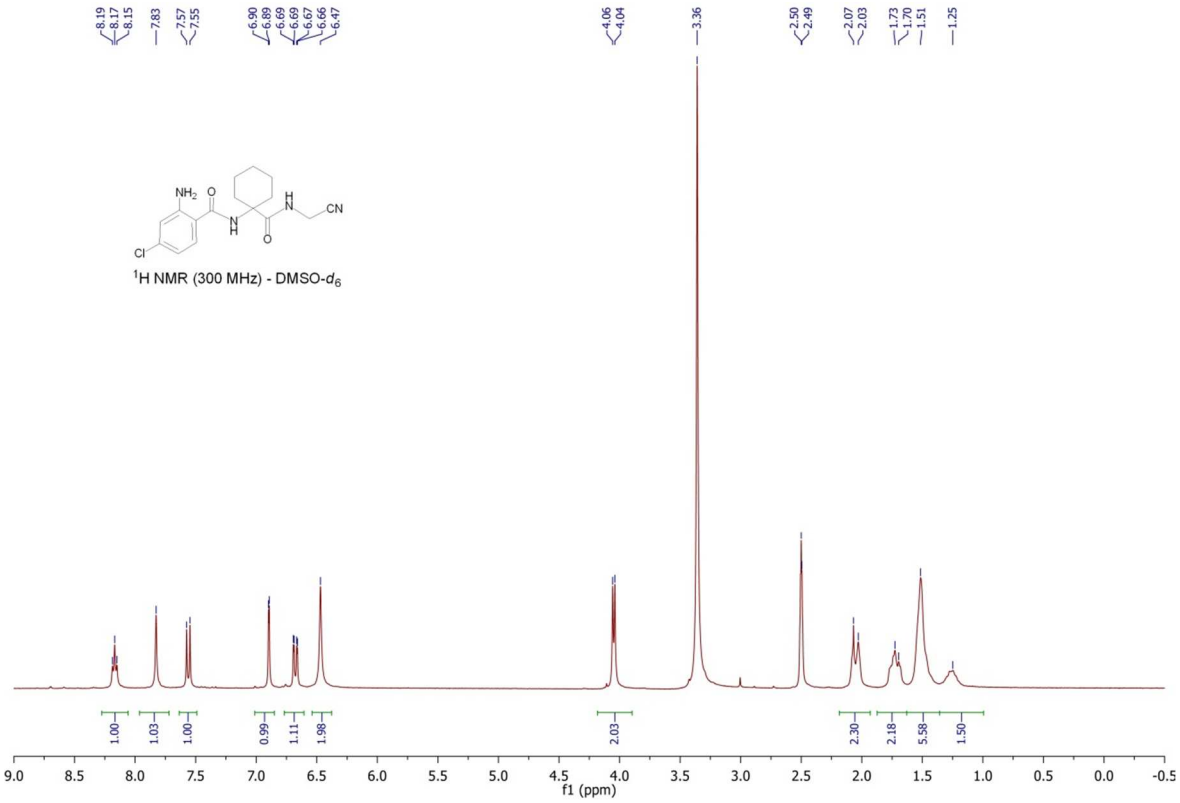


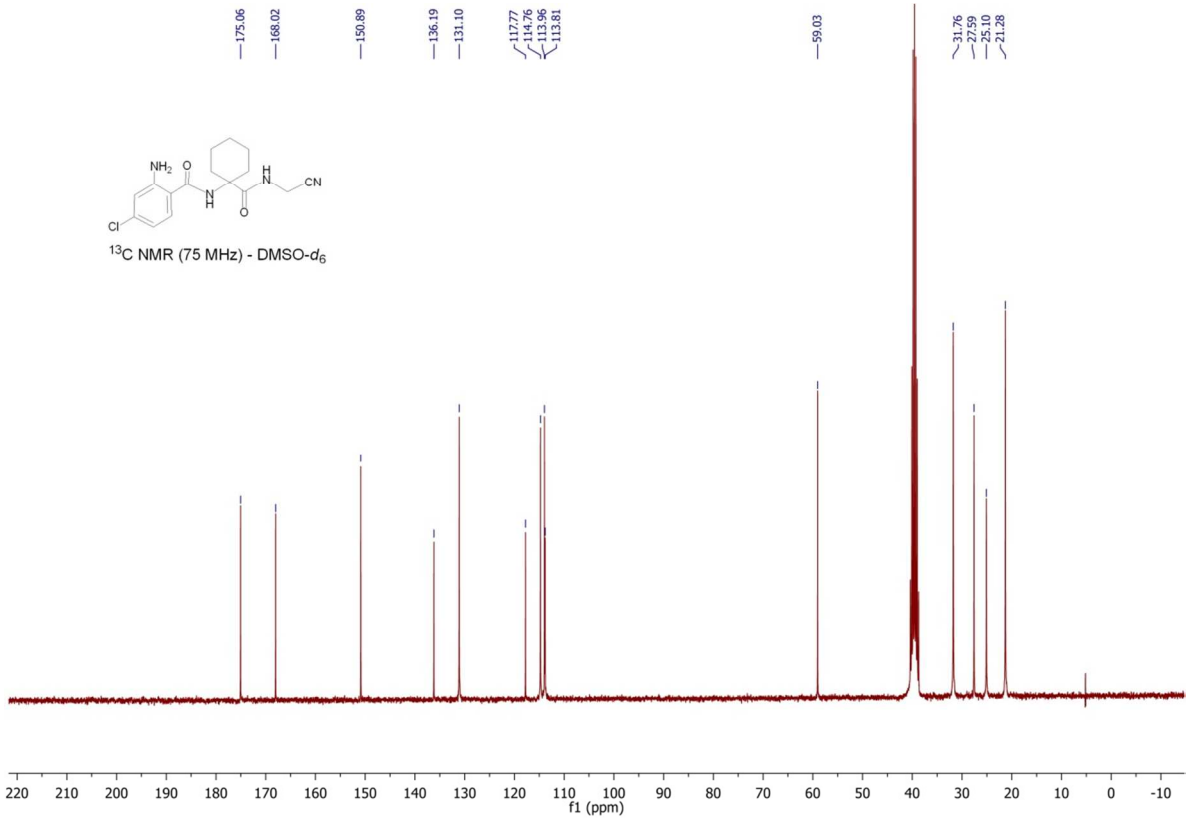
^{13}C NMR (75 MHz) - DMSO-d_6



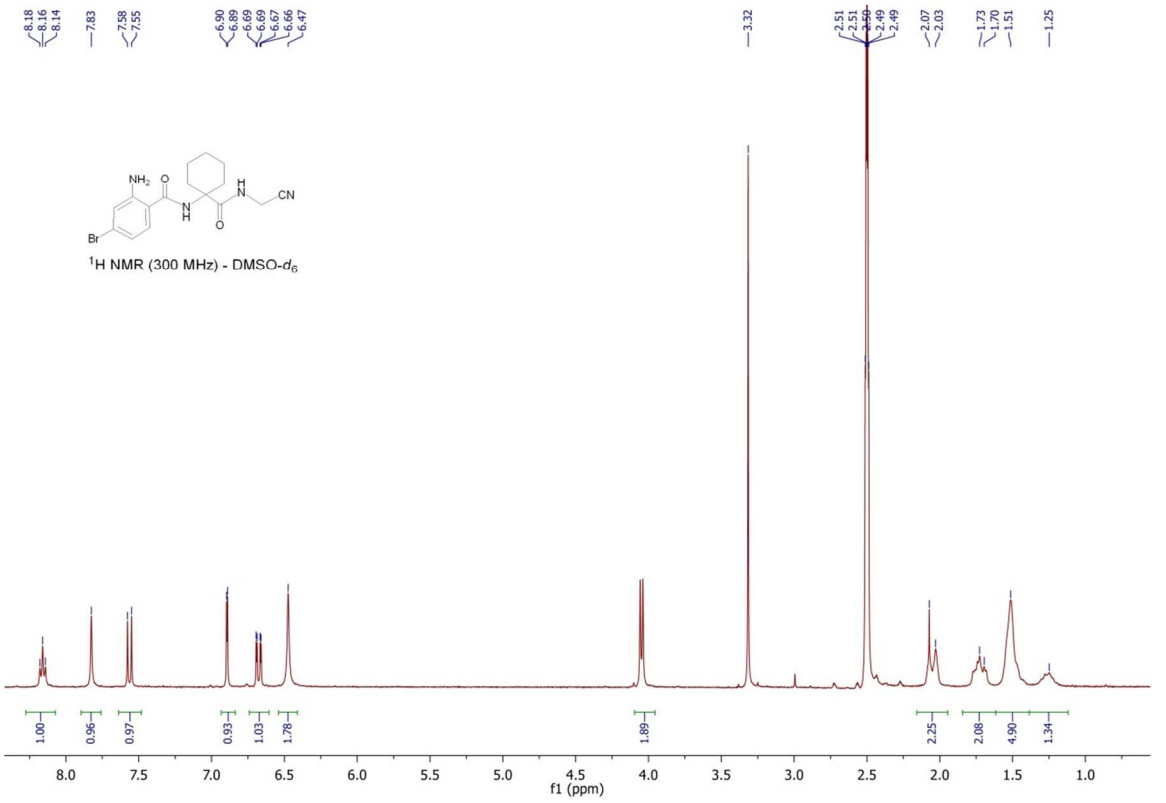
^{19}F NMR (300 MHz) - DMSO-d_6

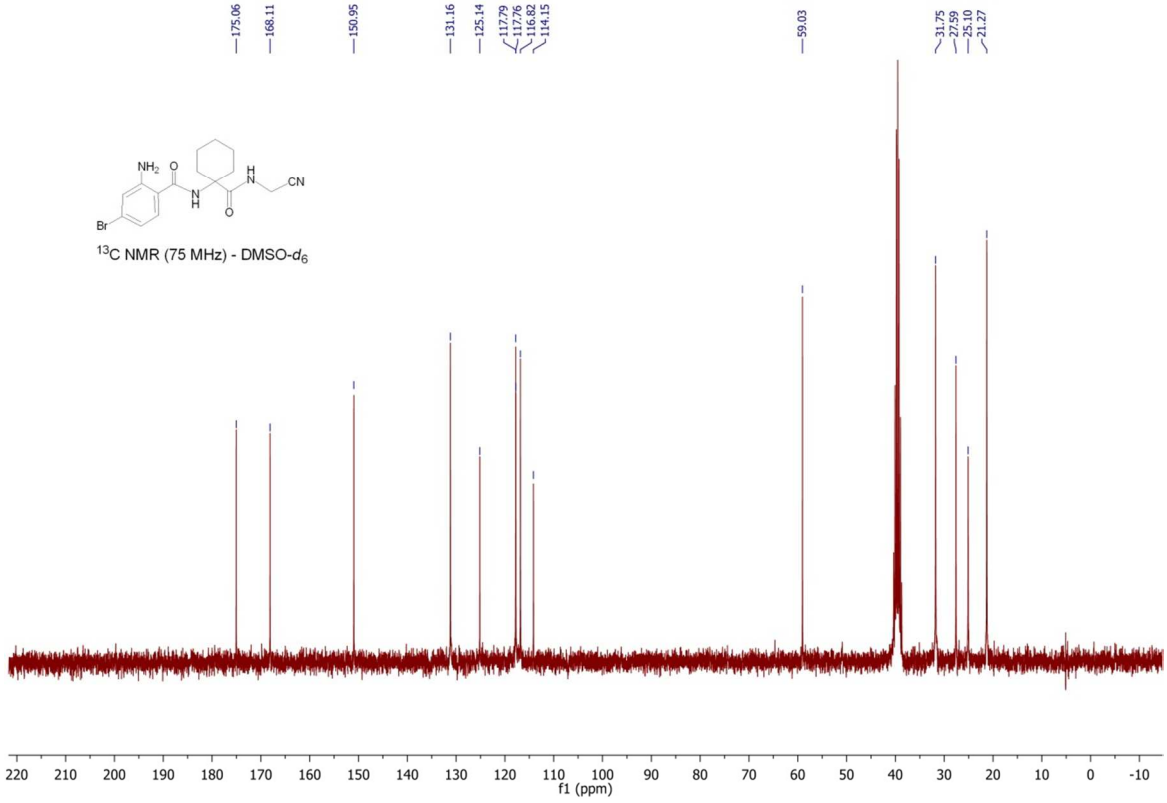
Compound **5**


8.19
8.17
8.15
7.83
7.57
7.55
6.90
6.89
6.69
6.69
6.67
6.66
6.47
4.06
4.04
3.36
2.50
2.49
2.07
2.03
1.73
1.70
1.51
1.25
NH2
O
N
H
H
N
CN
O
Cl
1H NMR (300 MHz) - DMSO-d6
1.00
1.03
1.00
0.99
1.11
1.98
2.03
2.30
2.18
5.58
1.50
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)

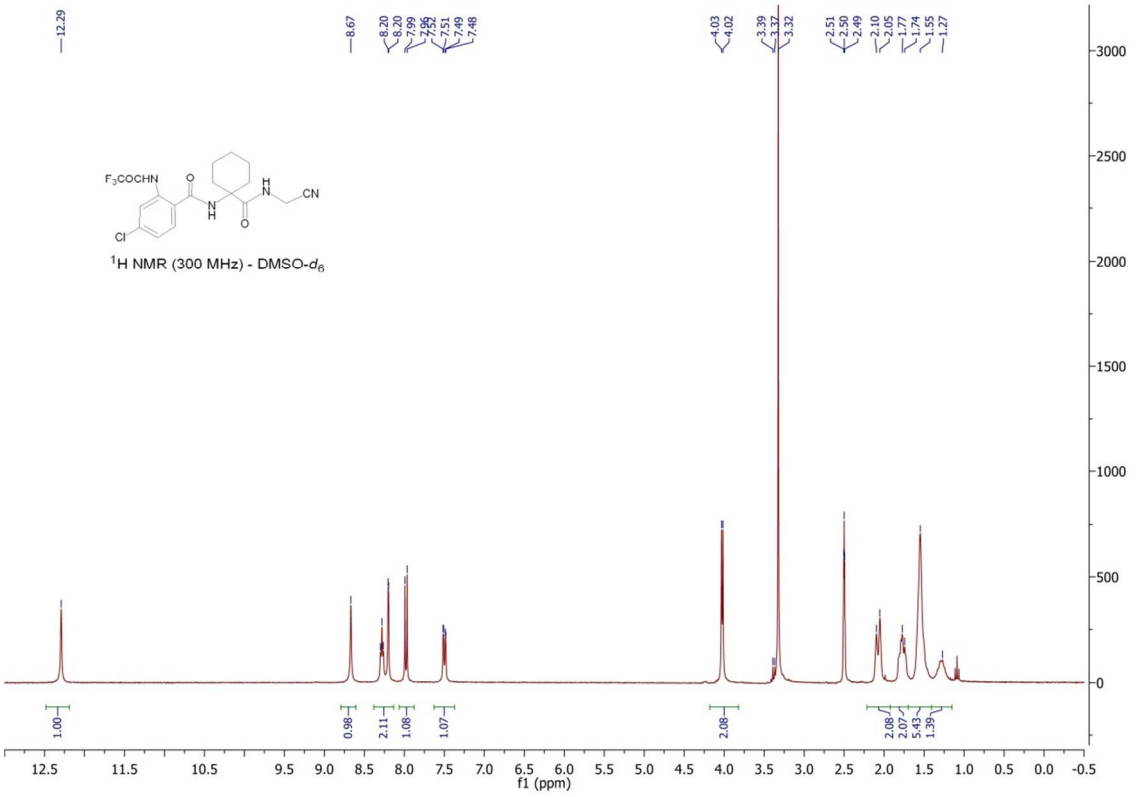

175.06
168.02
150.89
136.19
131.10
117.77
114.76
113.96
113.81
59.03
31.76
27.59
25.10
21.28
NH2
O
N
H
H
N
CN
O
Cl
13C NMR (75 MHz) - DMSO-d6
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

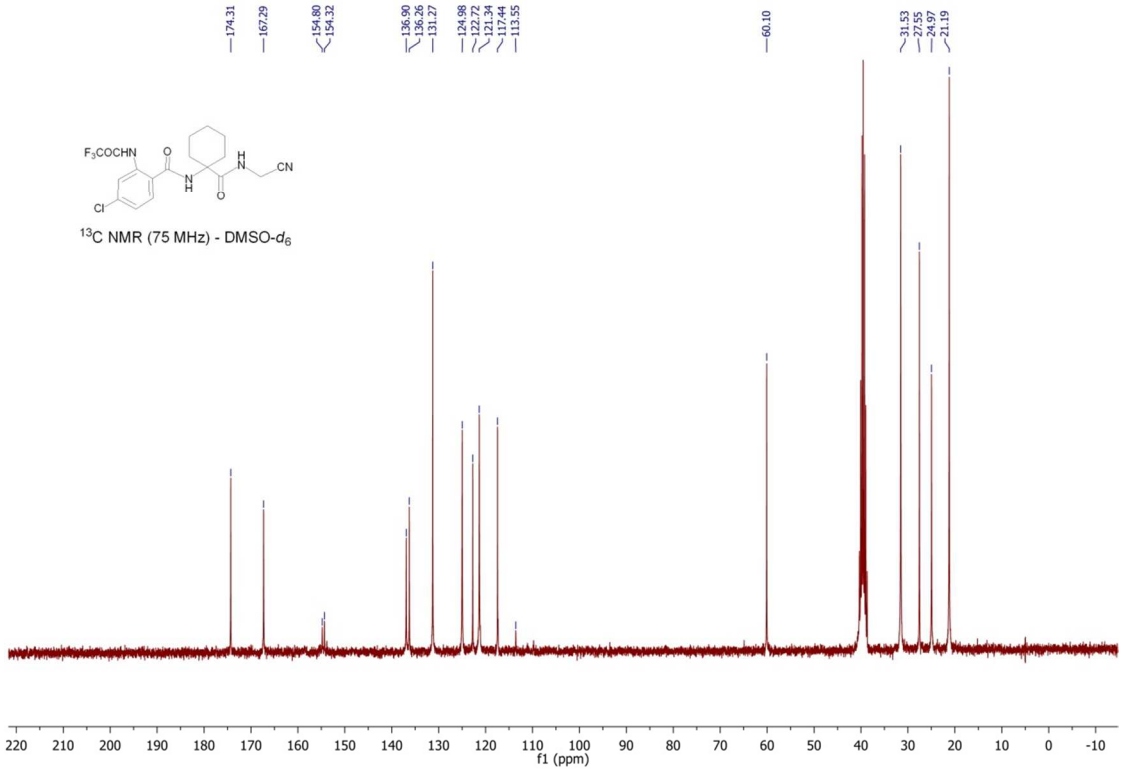
Compound **6**

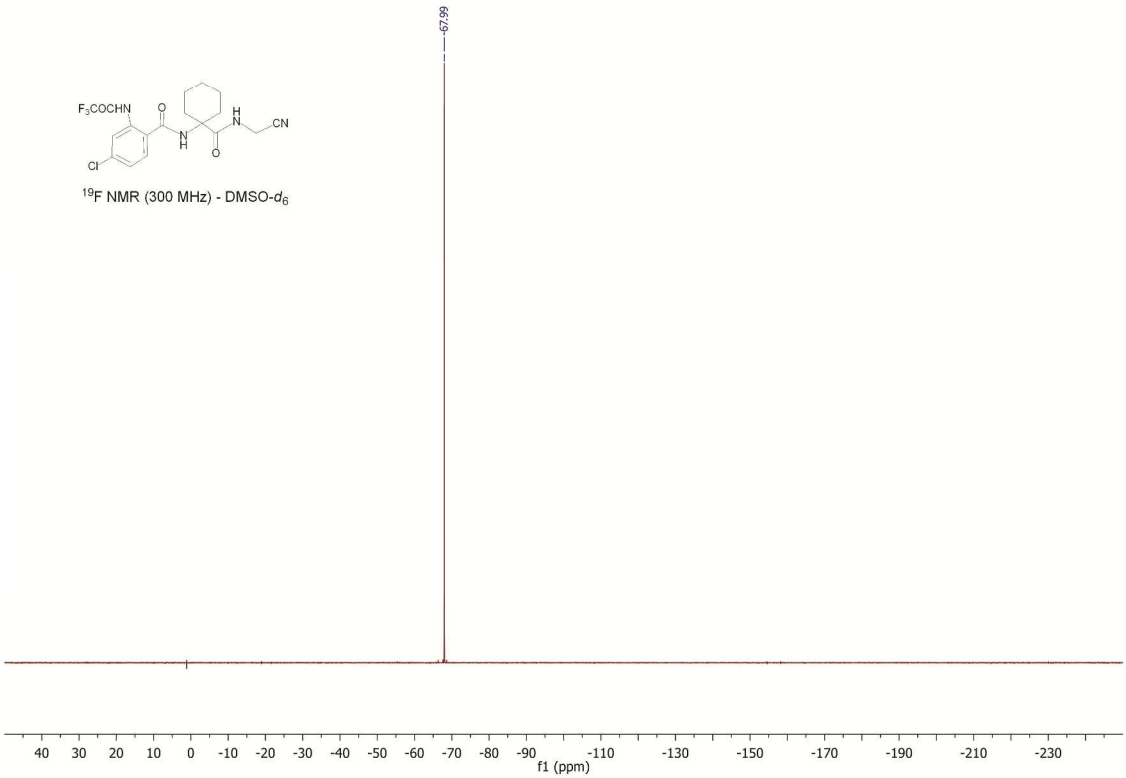

8.18
8.16
8.14
7.83
7.58
7.55
6.90
6.89
6.69
6.69
6.67
6.66
6.47
3.32
2.51
2.51
2.49
2.49
2.07
2.03
1.73
1.70
1.51
1.25
NH2
O
N
H
N
H
O
CN
Br
1H NMR (300 MHz) - DMSO-d6
1.00
0.96
0.97
0.93
1.03
1.78
1.89
2.25
2.08
4.90
1.34
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
f1 (ppm)


175.06
168.11
150.95
131.16
125.14
117.79
117.76
116.82
114.15
59.03
31.75
27.59
25.10
21.27
NH2
O
N
H
N
H
O
CN
Br
13C NMR (75 MHz) - DMSO-d6
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

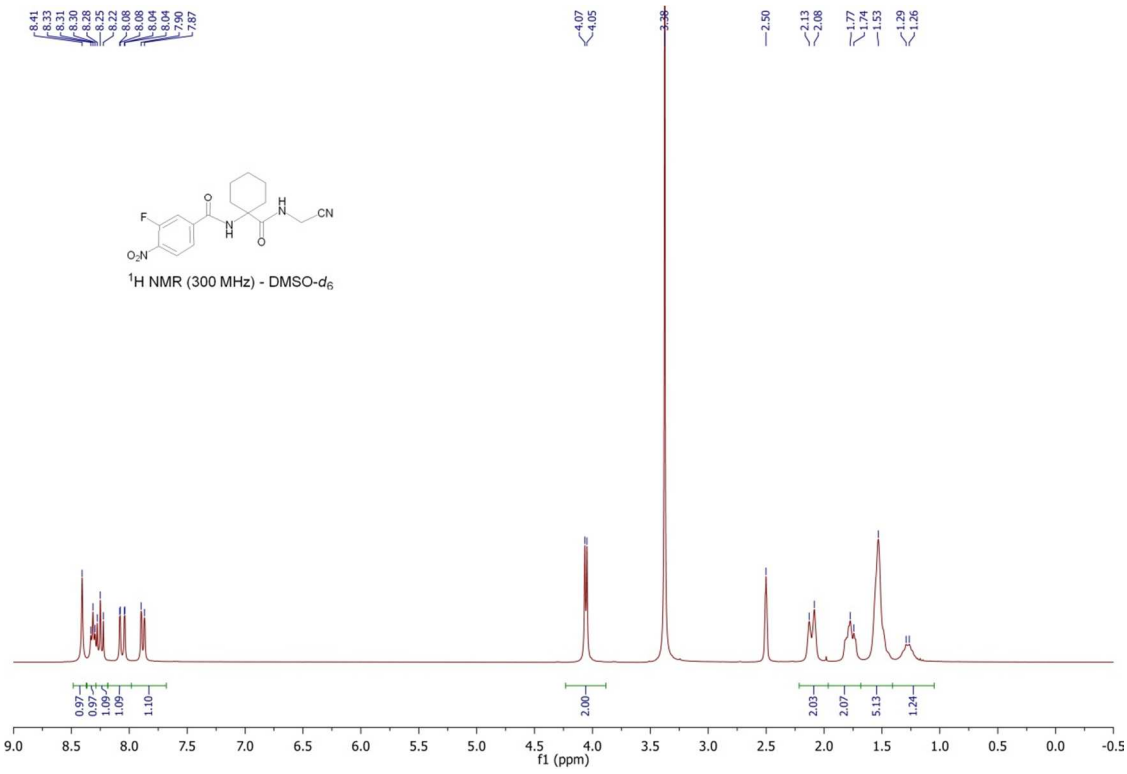
Compound **7**

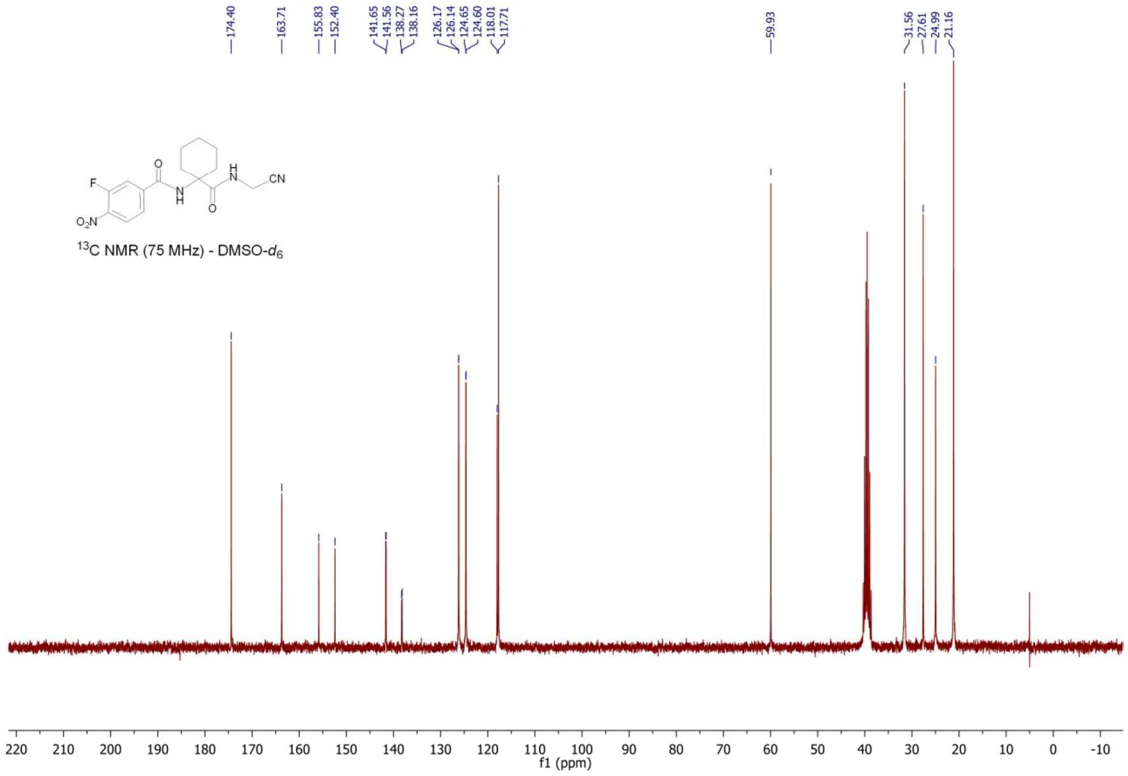

12.29
8.67
8.20
8.20
7.99
7.51
7.49
7.48
4.03
4.02
3.39
3.37
3.32
2.51
2.50
2.49
2.10
2.05
1.77
1.74
1.55
1.27
F3COCHN
CN
Cl
1H NMR (300 MHz) - DMSO-d6
1.00
0.98
2.11
1.08
1.07
2.08
2.08
2.07
5.43
1.39
f1 (ppm)

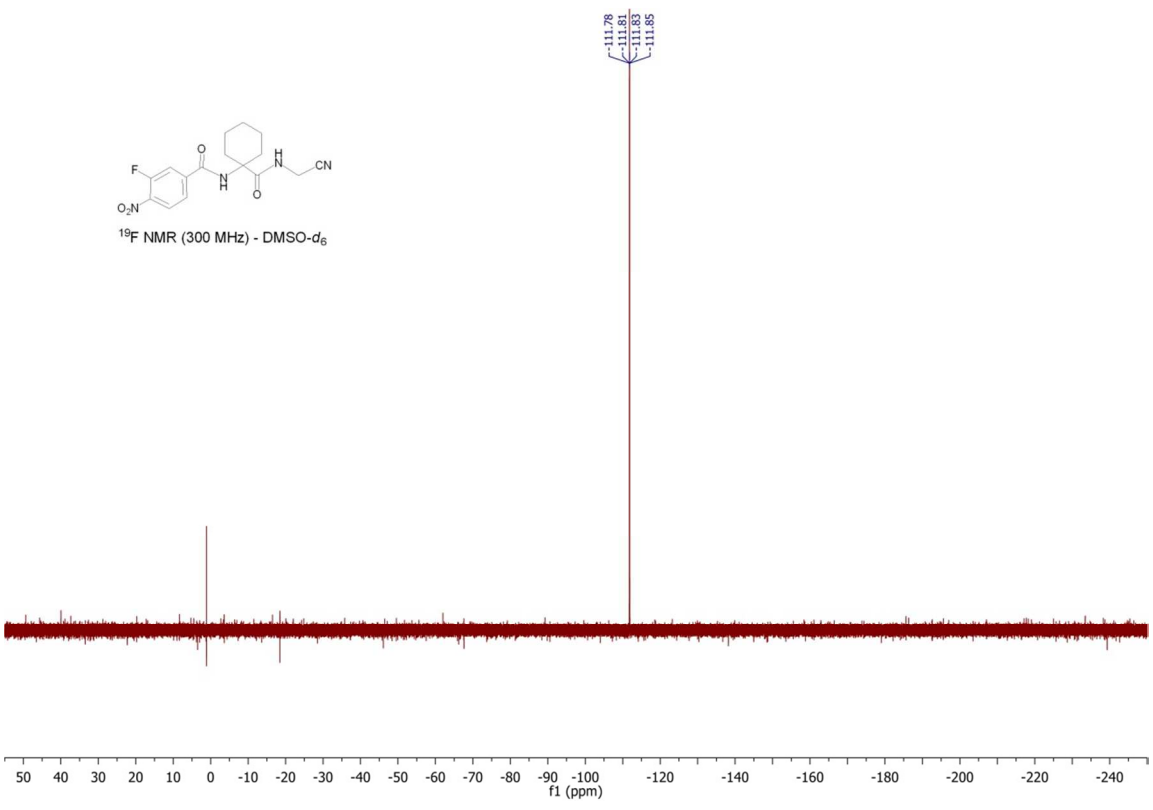

174.31
167.29
154.80
154.32
136.90
136.26
131.27
124.98
122.72
121.34
117.44
113.55
60.10
31.53
27.55
24.97
21.19
F3COCHN
CN
Cl
13C NMR (75 MHz) - DMSO-d6
f1 (ppm)


F3COCHN
O
N
H
H
N
CN
O
Cl
19F NMR (300 MHz) - DMSO-d6
-67.99
40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -110 -130 -150 -170 -190 -210 -230
f1 (ppm)

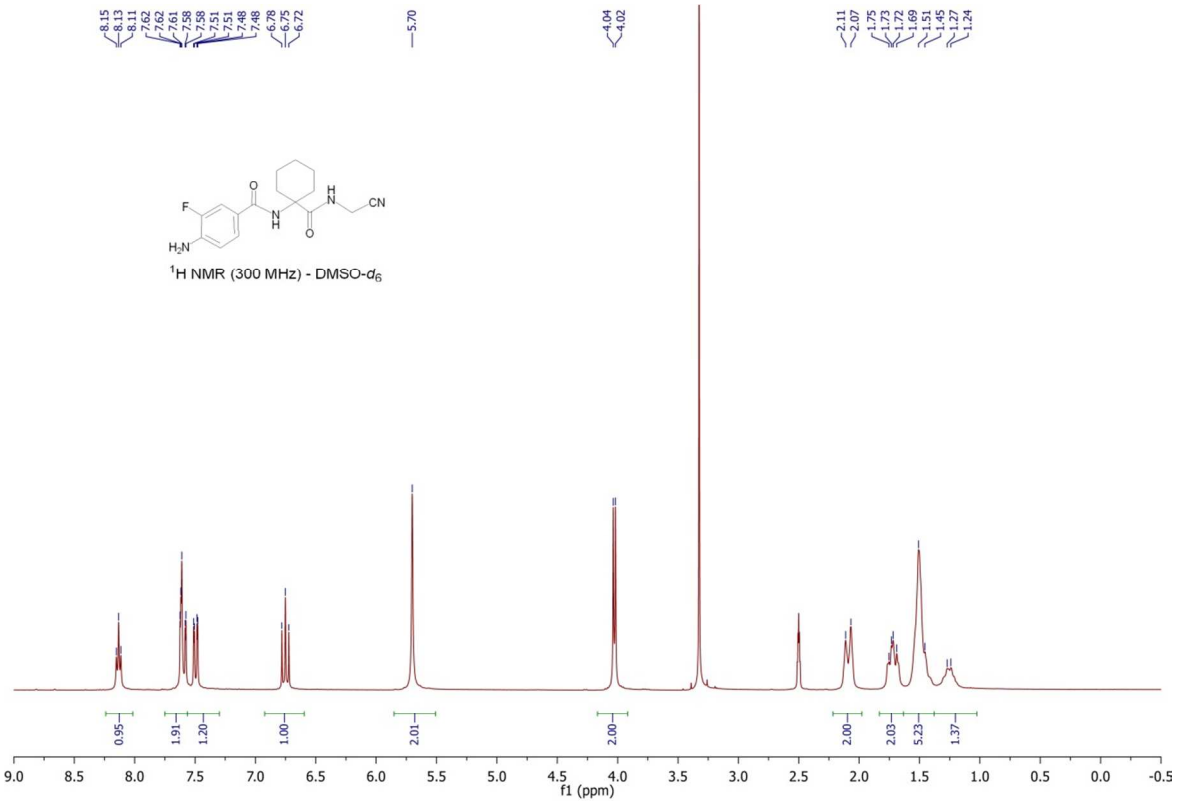
Compound **8**

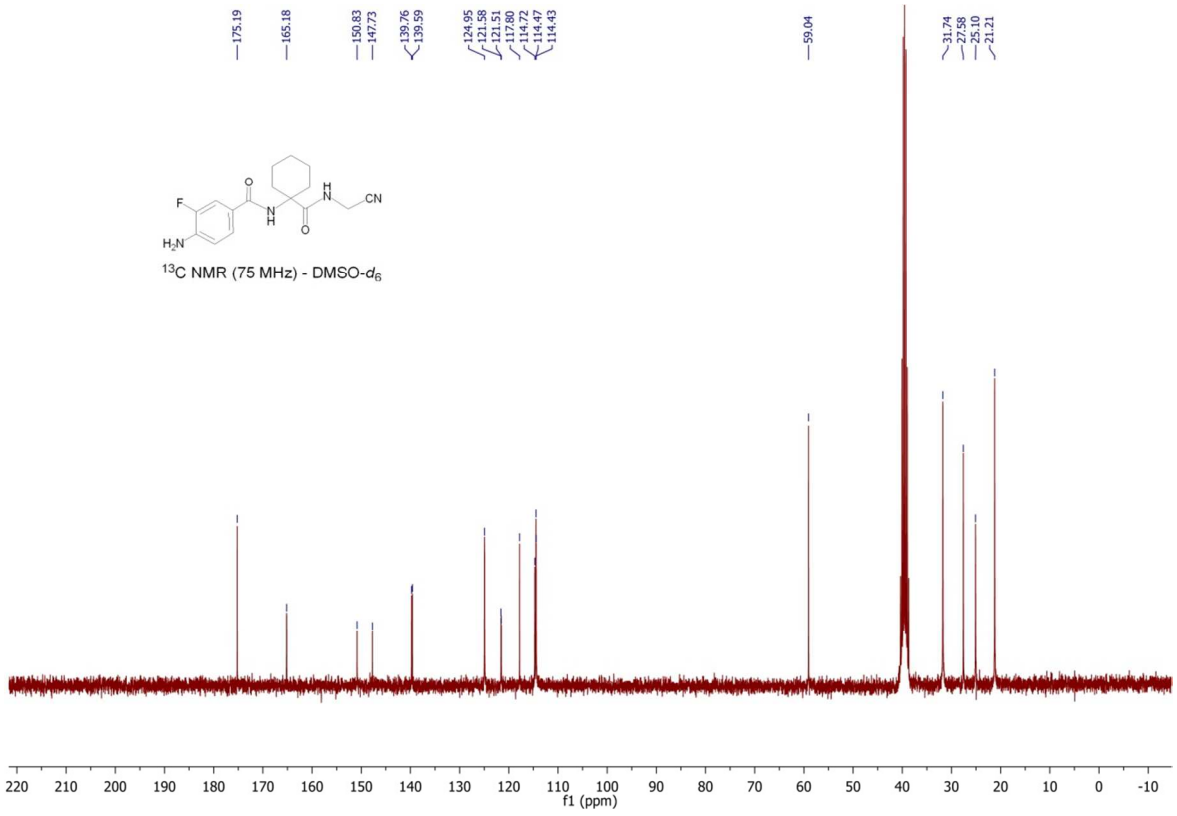

8.41
8.33
8.31
8.30
8.28
8.25
8.22
8.08
8.08
8.04
8.04
7.90
7.87
4.07
4.05
3.38
2.50
2.13
2.08
1.77
1.74
1.53
1.29
1.26
F
O
N
H
H
N
CN
O
O2N
1H NMR (300 MHz) - DMSO-d6
0.97
0.97
1.09
1.09
1.10
2.00
2.03
2.07
5.13
1.24
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

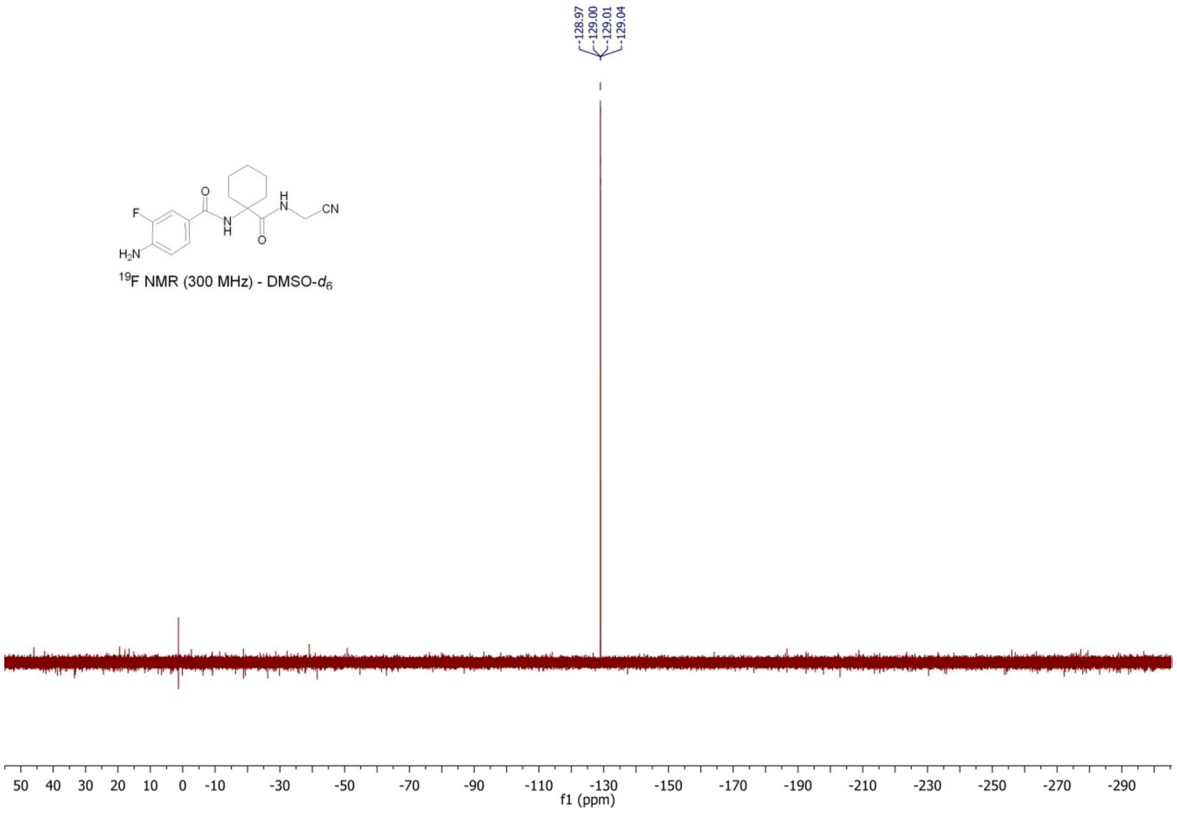

174.40
163.71
155.83
152.40
141.65
141.56
138.27
138.16
126.17
126.14
124.65
124.60
118.01
117.71
59.93
31.56
27.61
24.99
21.16
F
O2N
O
N
H
H
N
CN
O
13C NMR (75 MHz) - DMSO-d6
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)


-111.78
-111.81
-111.83
-111.85
F
O2N
O
N
H
H
N
CN
O
19F NMR (300 MHz) - DMSO-d6
50 40 30 20 10 0 -10 -20 -30 -40 -50 -60 -70 -80 -90 -100 -120 -140 -160 -180 -200 -220 -240
f1 (ppm)

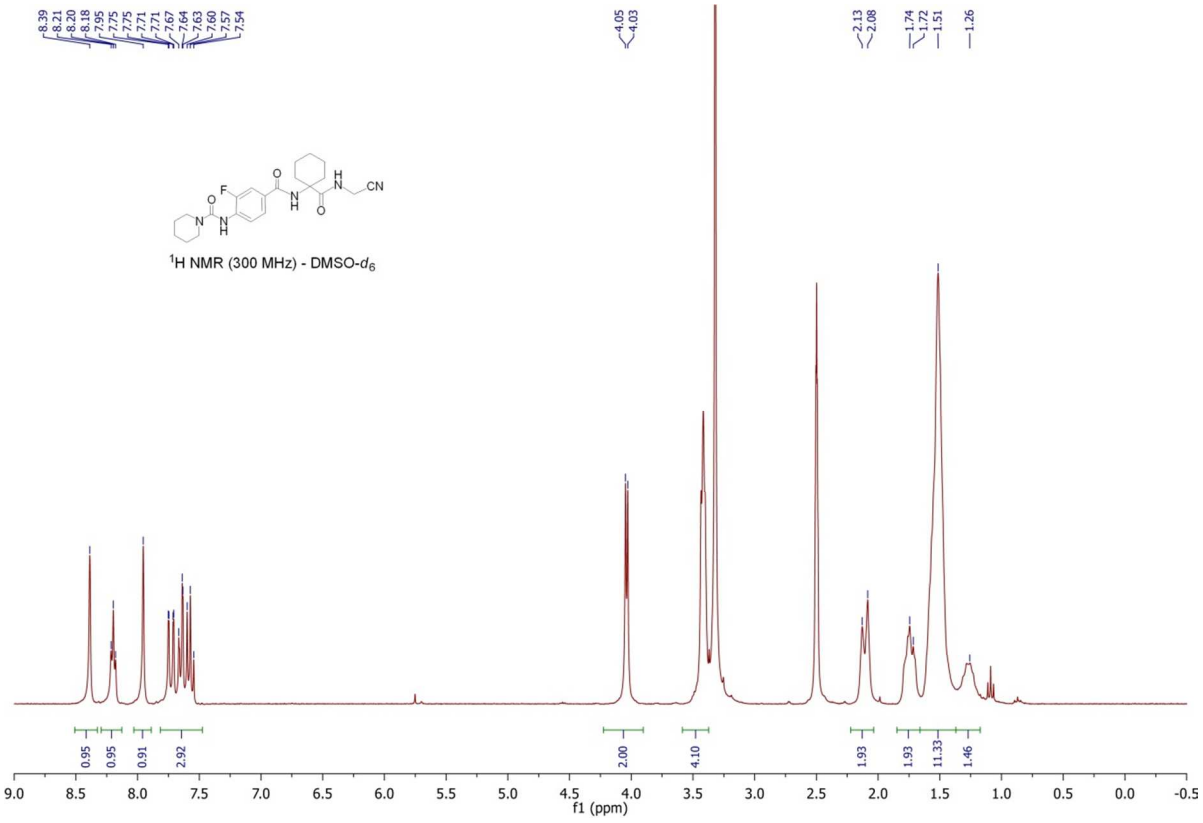
Compound **9**

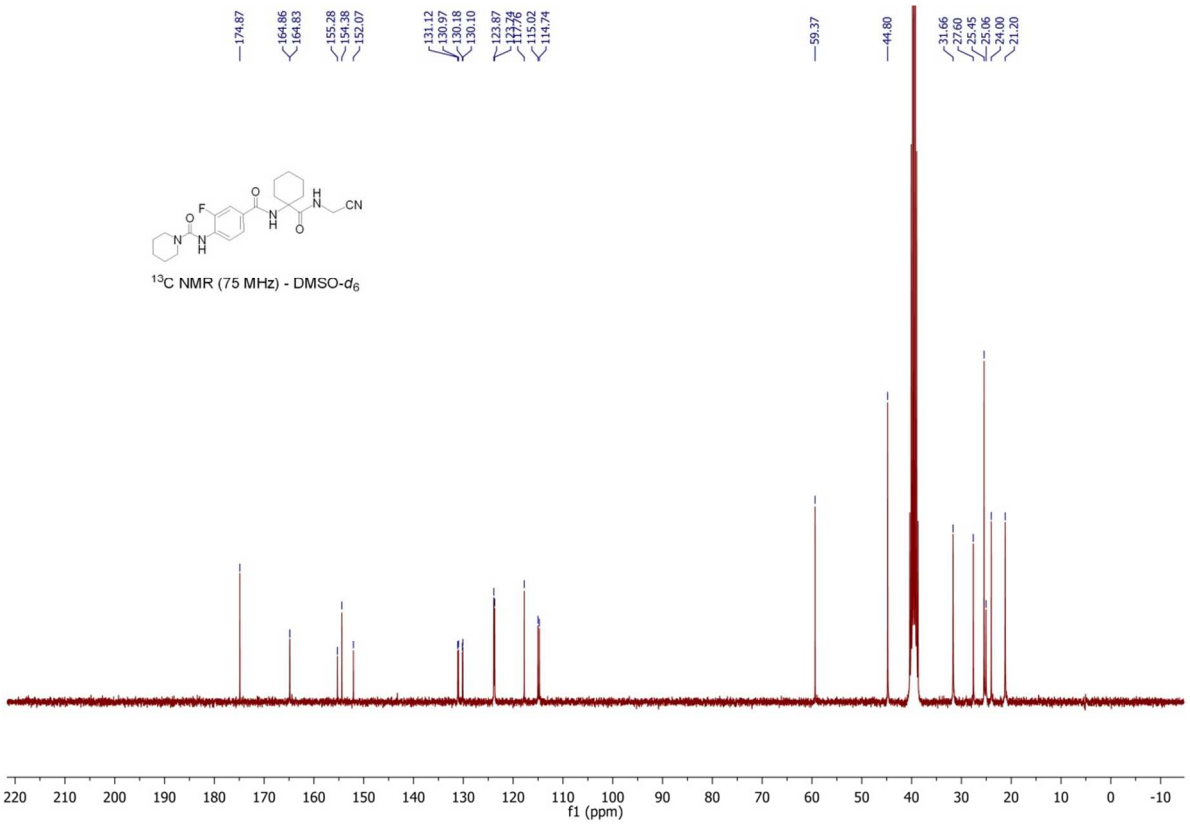

8.15
8.13
8.11
7.62
7.62
7.61
7.58
7.58
7.51
7.51
7.48
7.48
6.78
6.75
6.72
5.70
4.04
4.02
2.11
2.07
1.75
1.73
1.72
1.69
1.51
1.45
1.27
1.24
F
O
N
H
H
N
CN
O
H2N
1H NMR (300 MHz) - DMSO-d6
0.95
1.91
1.20
1.00
2.01
2.00
2.00
2.03
5.23
1.37
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)


175.19
165.18
150.83
147.73
139.76
139.59
124.95
121.58
121.51
117.80
114.72
114.47
114.43
59.04
31.74
27.58
25.10
21.21
F
O
N
H
H
N
CN
O
H2N
13C NMR (75 MHz) - DMSO-d6
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)

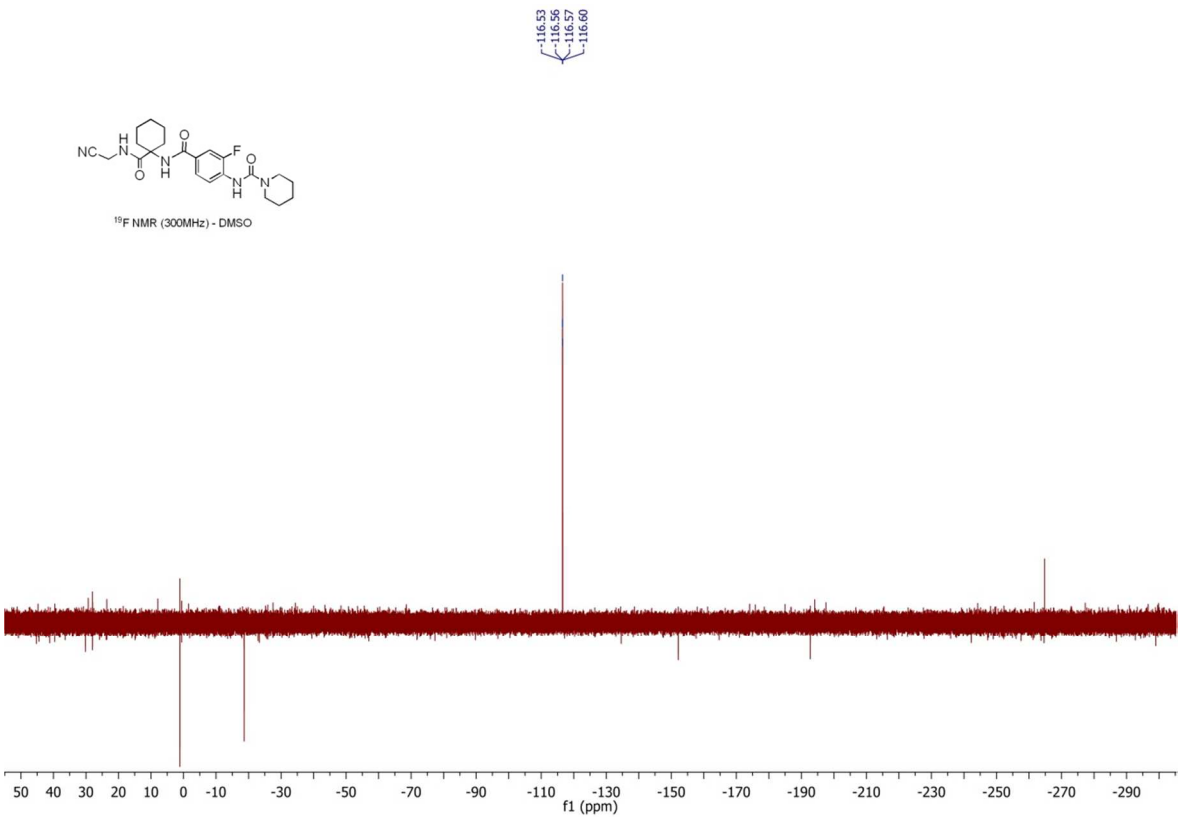

-128.97
-129.00
-129.01
-129.04
F
O
N
H
H
N
CN
O
H_2N
^{19}F NMR (300 MHz) - DMSO-d_6
50 40 30 20 10 0 -10 -30 -50 -70 -90 -110 -130 -150 -170 -190 -210 -230 -250 -270 -290
f1 (ppm)

Compound **10**


8.39
8.21
8.20
8.18
7.95
7.75
7.75
7.71
7.71
7.67
7.64
7.63
7.60
7.57
7.54
4.05
4.03
2.13
2.08
1.74
1.72
1.51
1.26
F
O
N
H
H
N
CN
O
O
N
H
^{1}H NMR (300 MHz) - DMSO-d_6
0.95
0.95
0.91
2.92
2.00
4.10
1.93
1.93
11.33
1.46
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 -0.5
f1 (ppm)

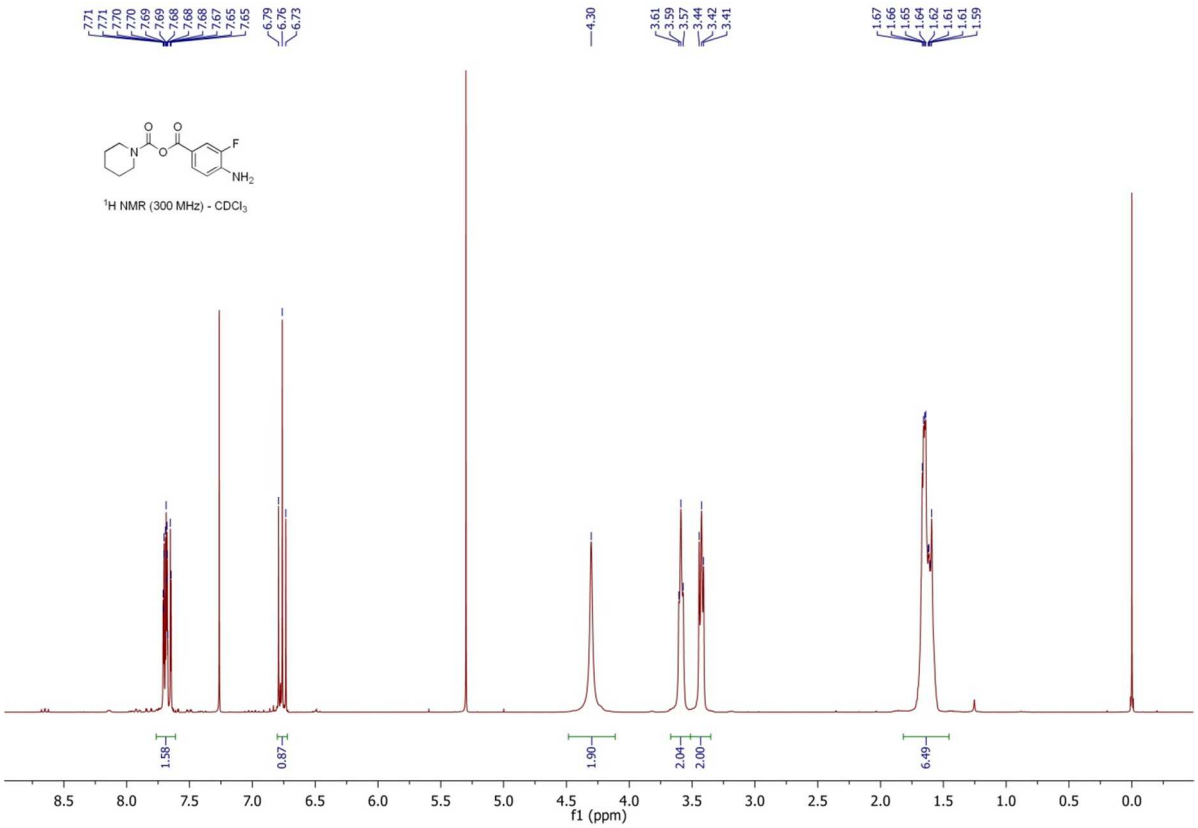


^{13}C NMR (75 MHz) - DMSO-d_6

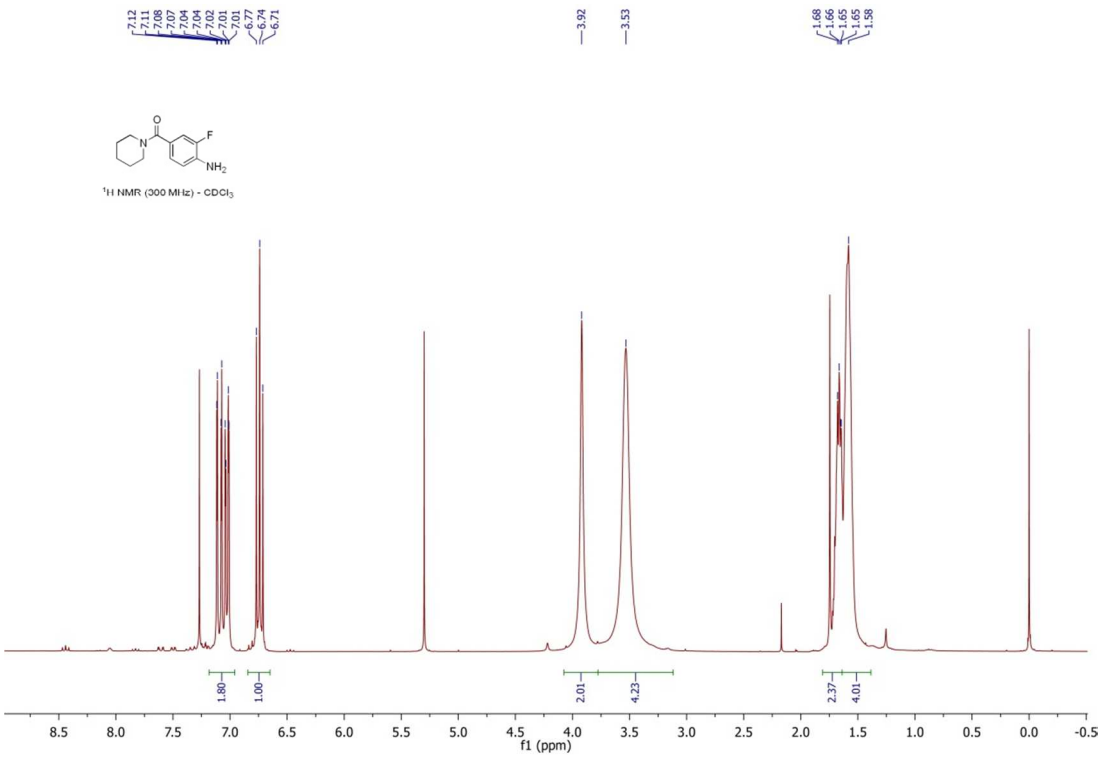


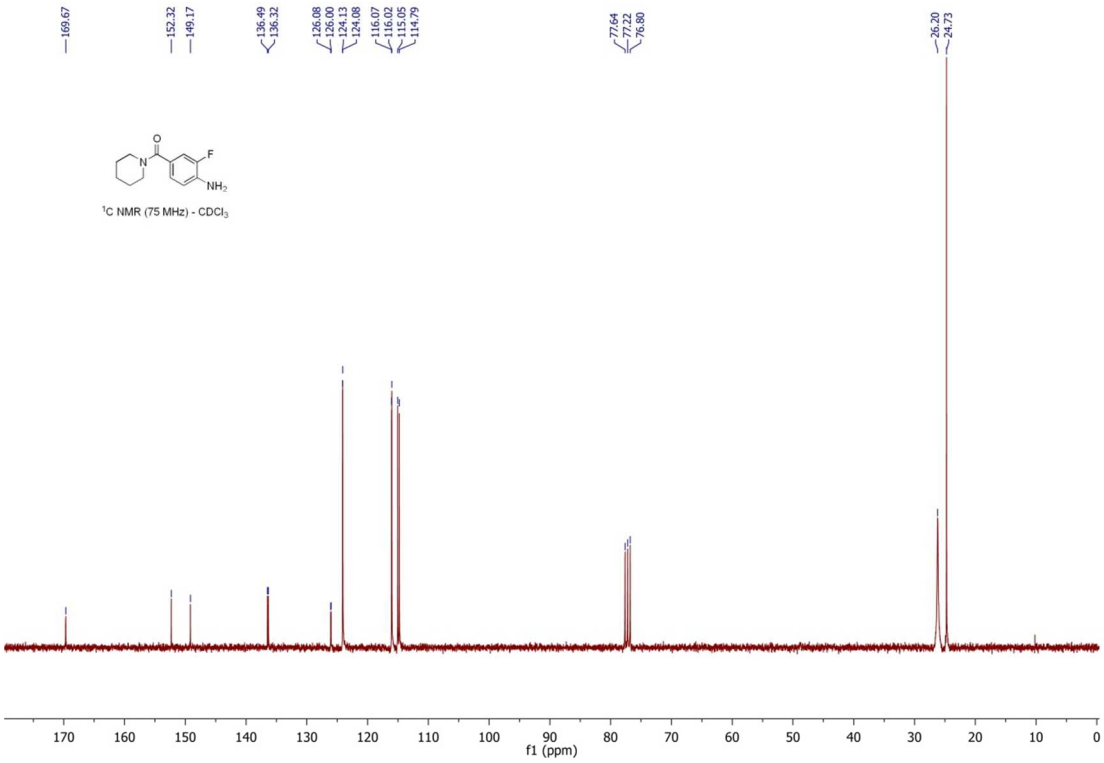
^{19}F NMR (300MHz) - DMSO

Compound **12**


7.71
7.71
7.70
7.70
7.69
7.69
7.68
7.68
7.68
7.67
7.65
7.65
6.79
6.76
6.73
4.30
3.61
3.59
3.57
3.44
3.42
3.41
1.67
1.66
1.65
1.64
1.62
1.61
1.61
1.59
O
O
N
O
F
NH2
1H NMR (300 MHz) - CDCl3
1.58
0.87
1.90
2.04
2.00
6.49
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

Compound **13**


7.12
7.11
7.08
7.07
7.04
7.04
7.02
7.01
7.01
6.77
6.74
6.71
3.92
3.53
1.68
1.66
1.65
1.65
1.58
F
NH2
1H NMR (300 MHz) - CDCl3
1.80
1.00
2.01
4.23
2.37
4.01
f1 (ppm)


169.67
152.32
149.17
136.49
136.32
126.08
126.00
124.13
124.08
116.07
116.02
115.05
114.79
77.64
77.22
76.80
26.20
24.73
F
NH2
13C NMR (75 MHz) - CDCl3
f1 (ppm)