Electronic Supplementary Information (ESI) for

Exploring the role of the α-carboxyphosphonate moiety in the HIV-RT activity of α-carboxy nucleoside phosphonates

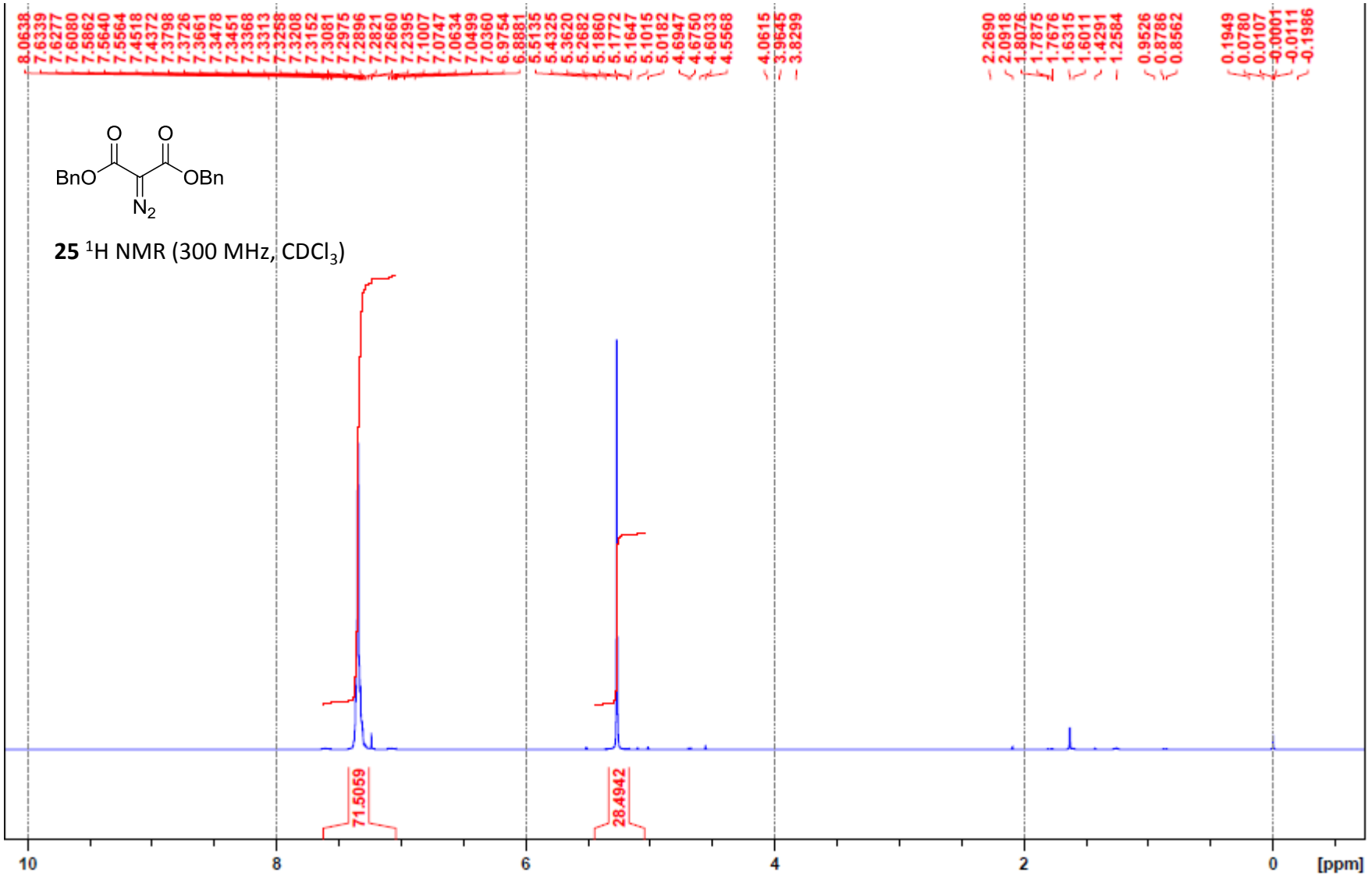
Nicholas D. Mullins,[†] Nuala M. Maguire,[†] Alan Ford,[†] Kalyan Das,[‖] Edward Arnold, [‖] Jan Balzarini[‡] and Anita R. Maguire[†§]*

[†]Department of Chemistry and [§]School of Pharmacy, Analytical and Biological Chemistry Research Facility, Synthesis and Solid State Pharmaceutical Centre, University College, Cork, Ireland

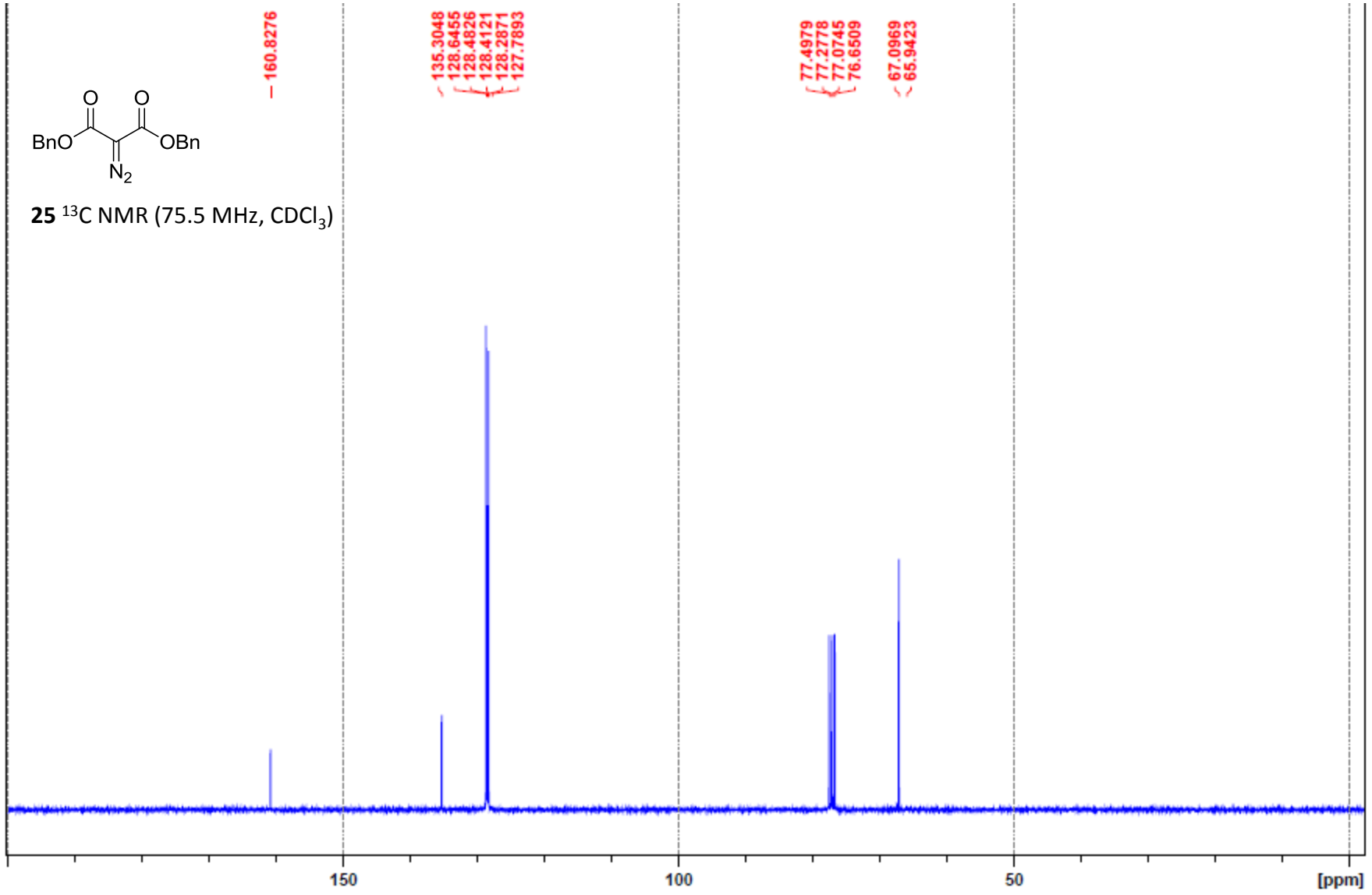
[‡]KU Leuven, Rega Institute for Medical Research, Minderbroedersstraat 10, B-3000 Leuven, Belgium

[‖]Center for Advanced Biotechnology and Medicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854-8020; USA

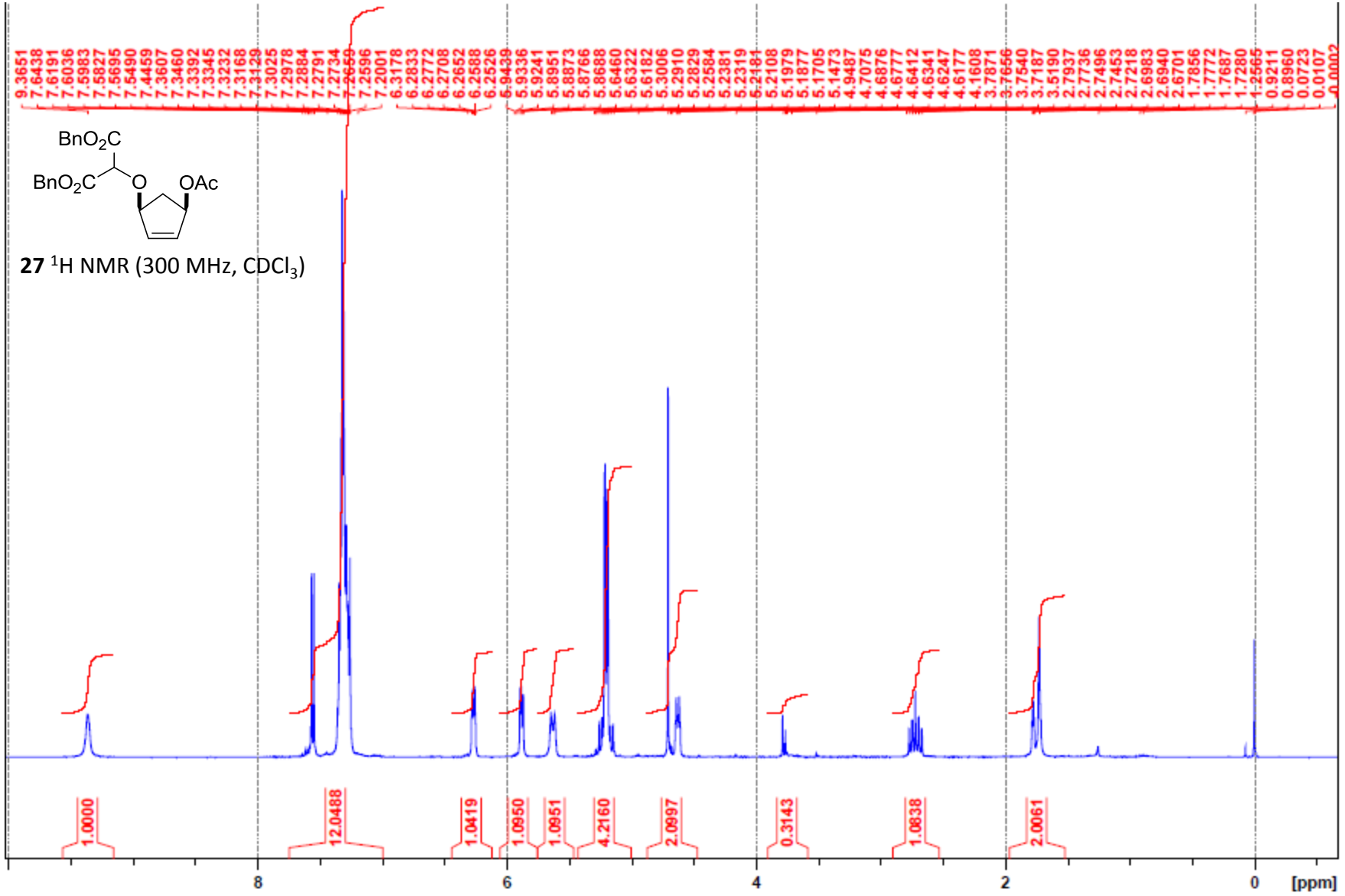
Table of Contents



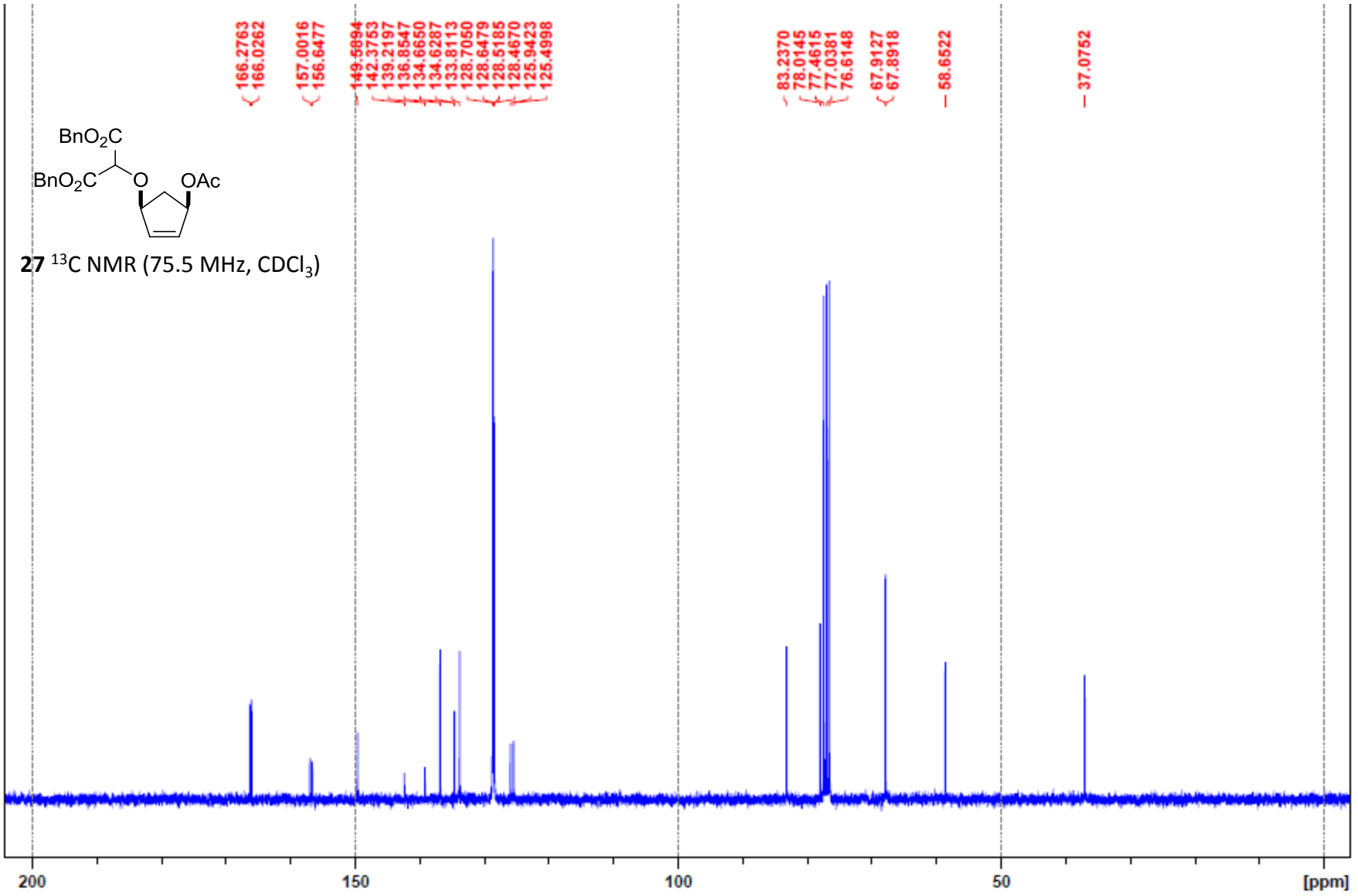
25 ^{1}H NMR (300 MHz, $CDCl_3$)



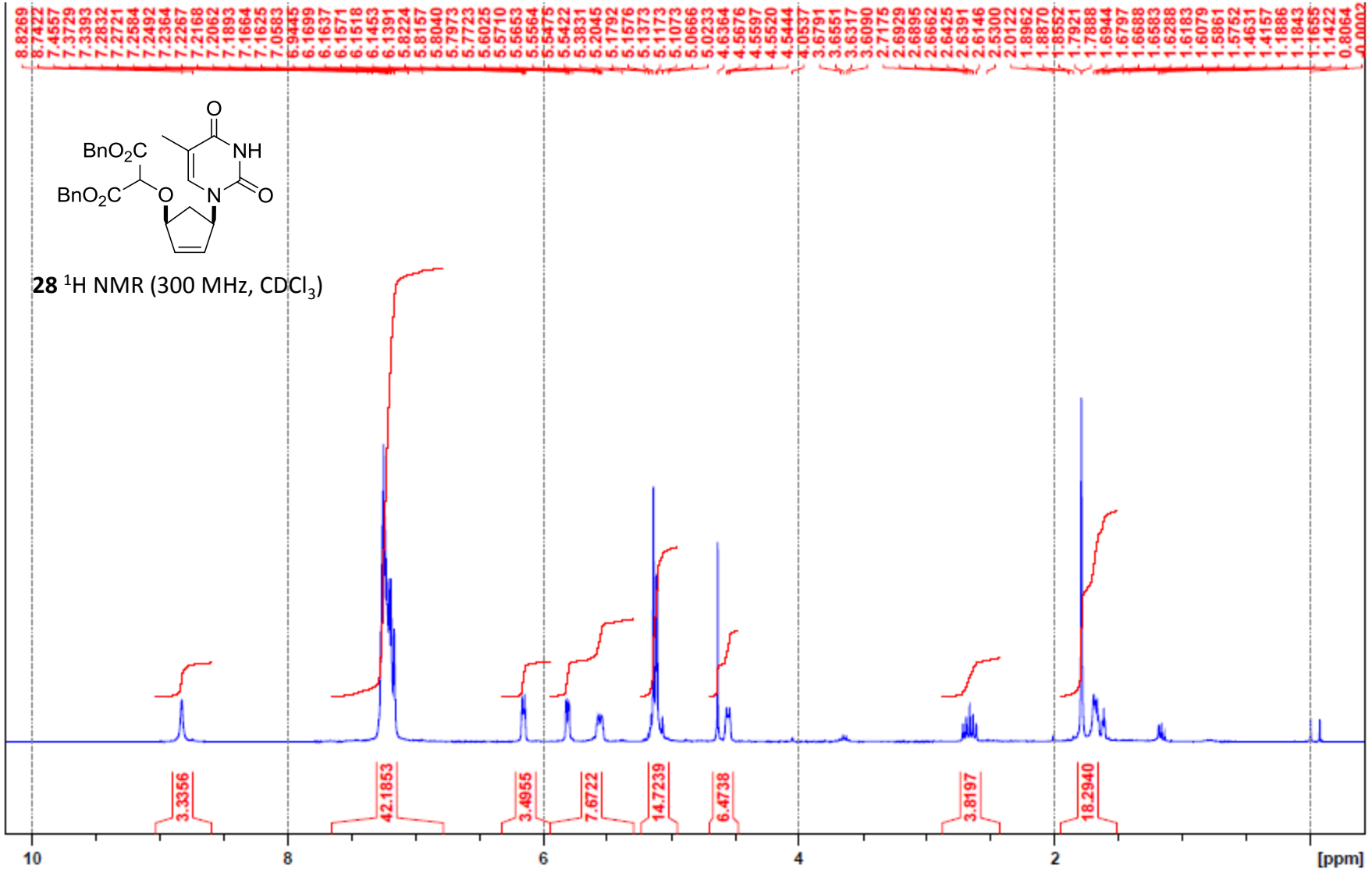
25 ^{13}C NMR (75.5 MHz, $CDCl_3$)

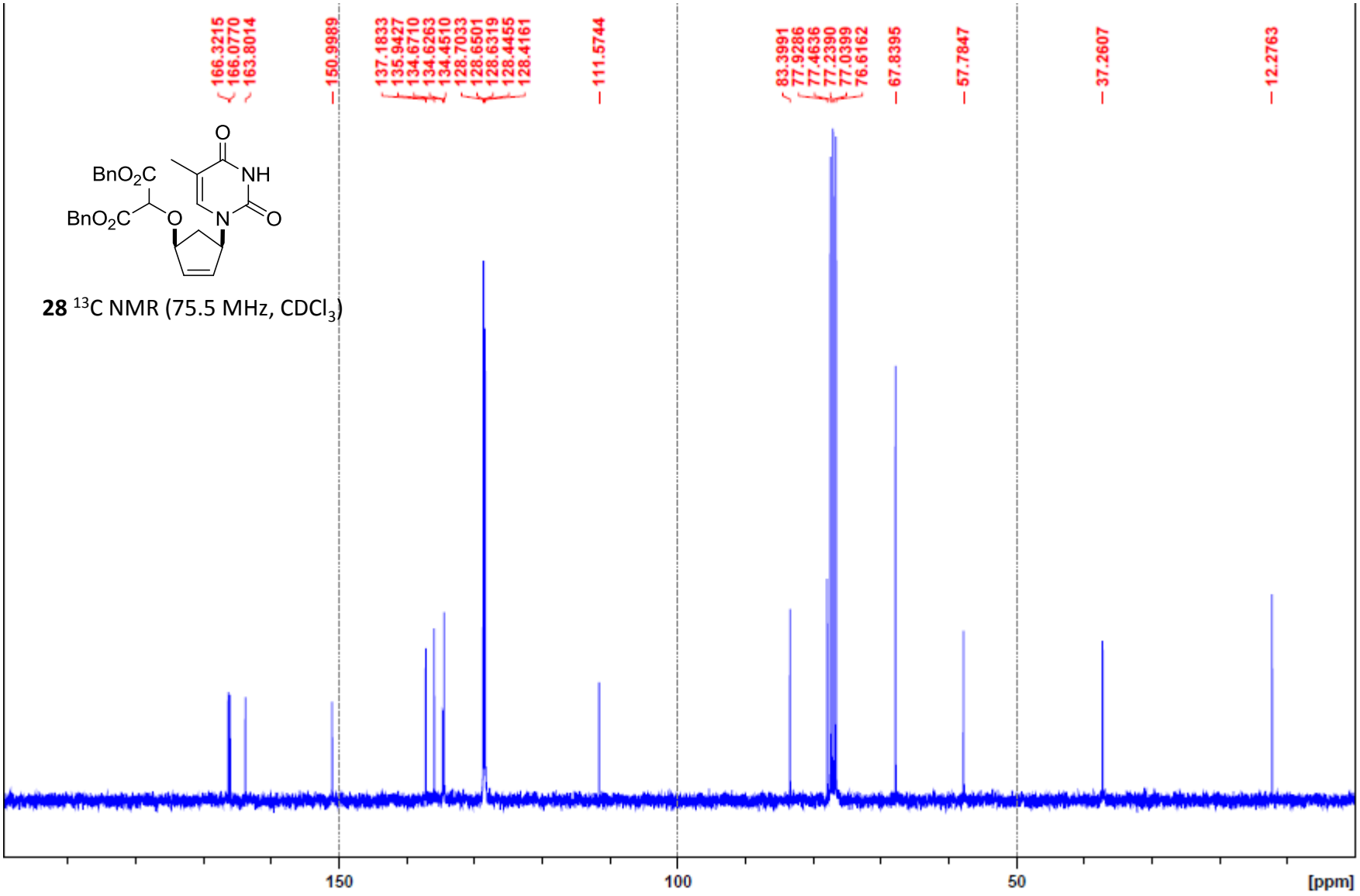


27 ^{1}H NMR (300 MHz, $CDCl_3$)

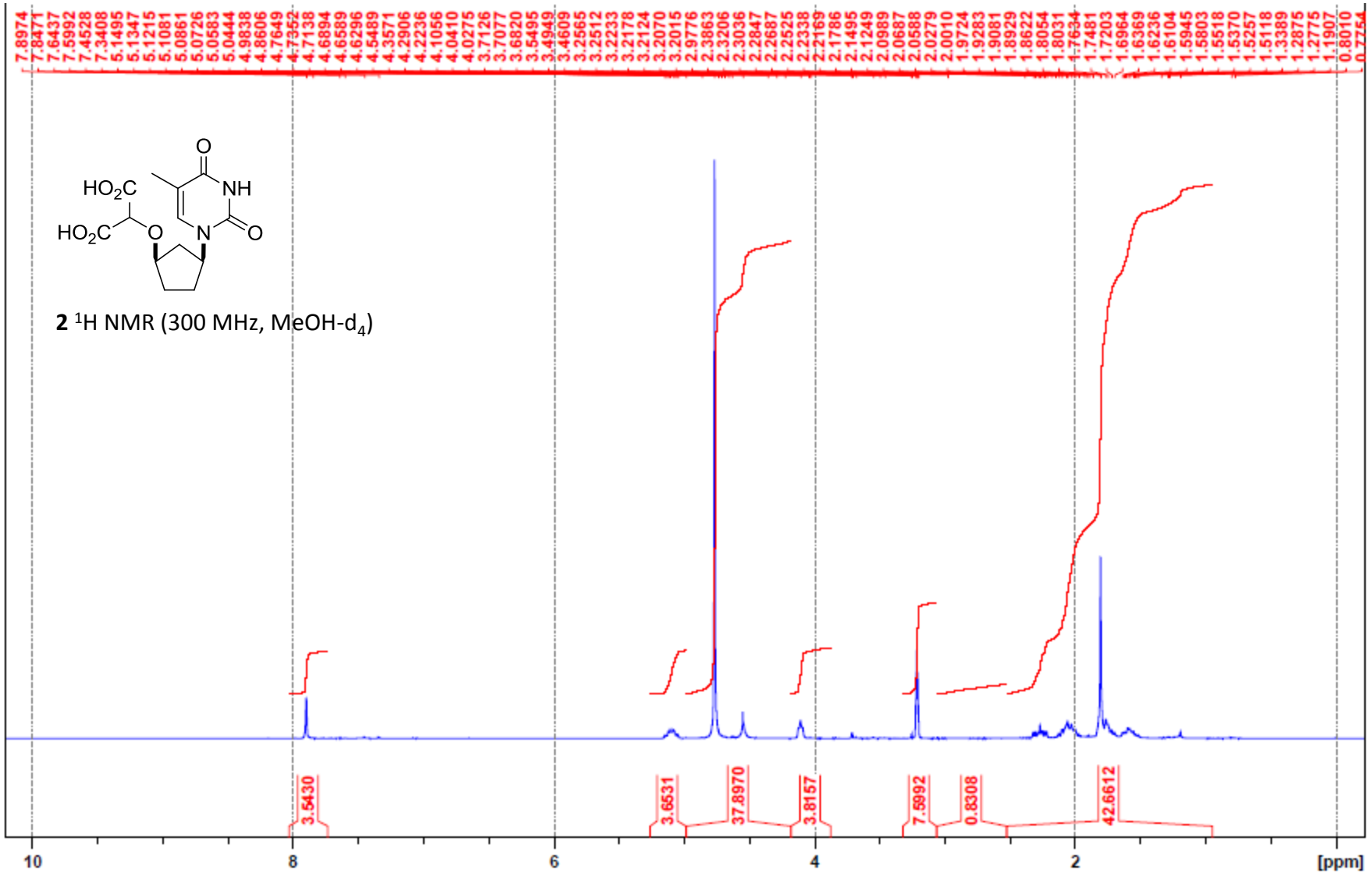


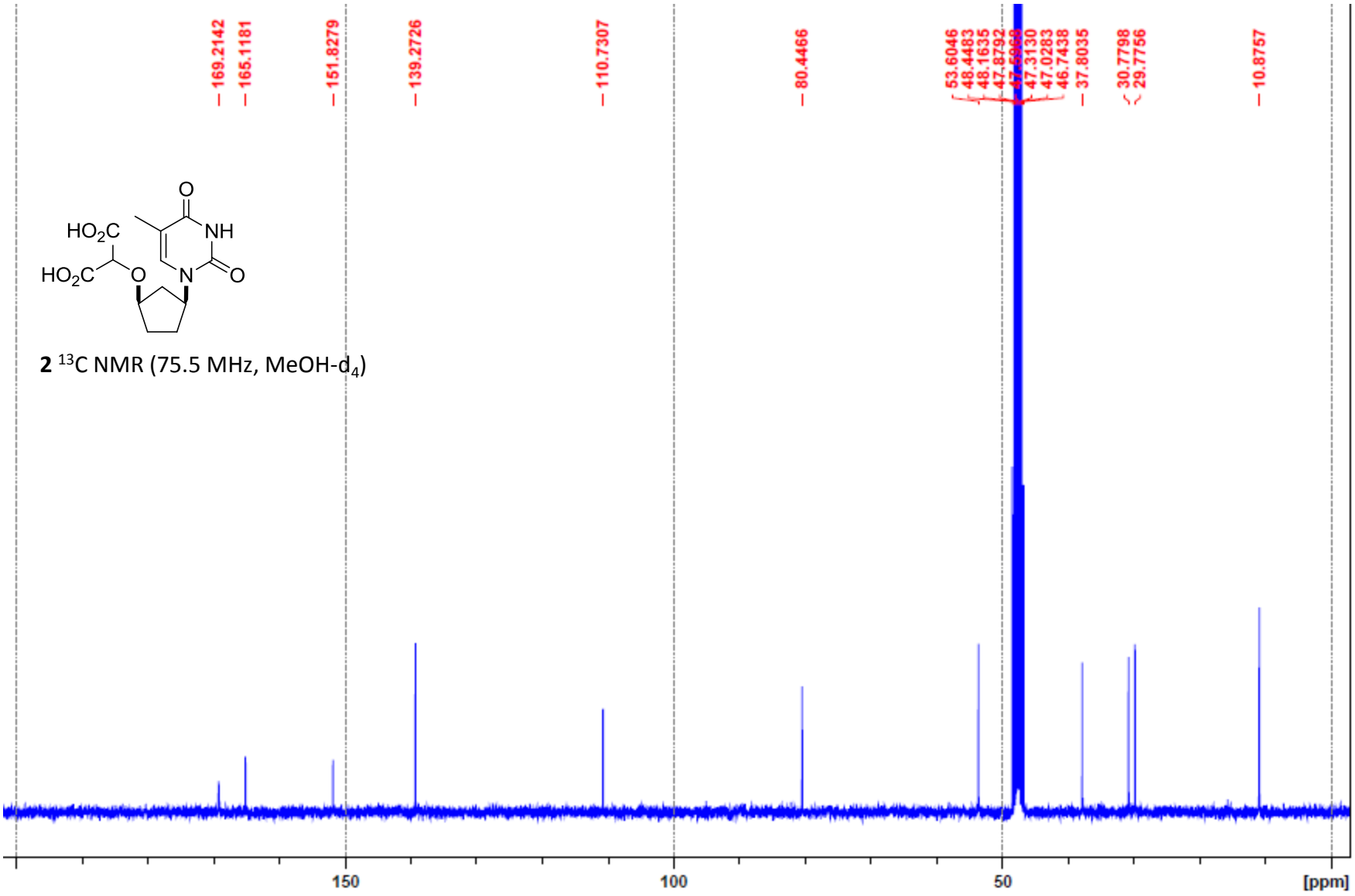
27 ^{13}C NMR (75.5 MHz, $CDCl_3$)


BnO2C
BnO2C
O
NH
O
N
O
28 1H NMR (300 MHz, CDCl3)
3.3356
42.1853
3.4955
7.6722
14.7239
6.4738
3.8197
18.2940
10
8
6
4
2
[ppm]

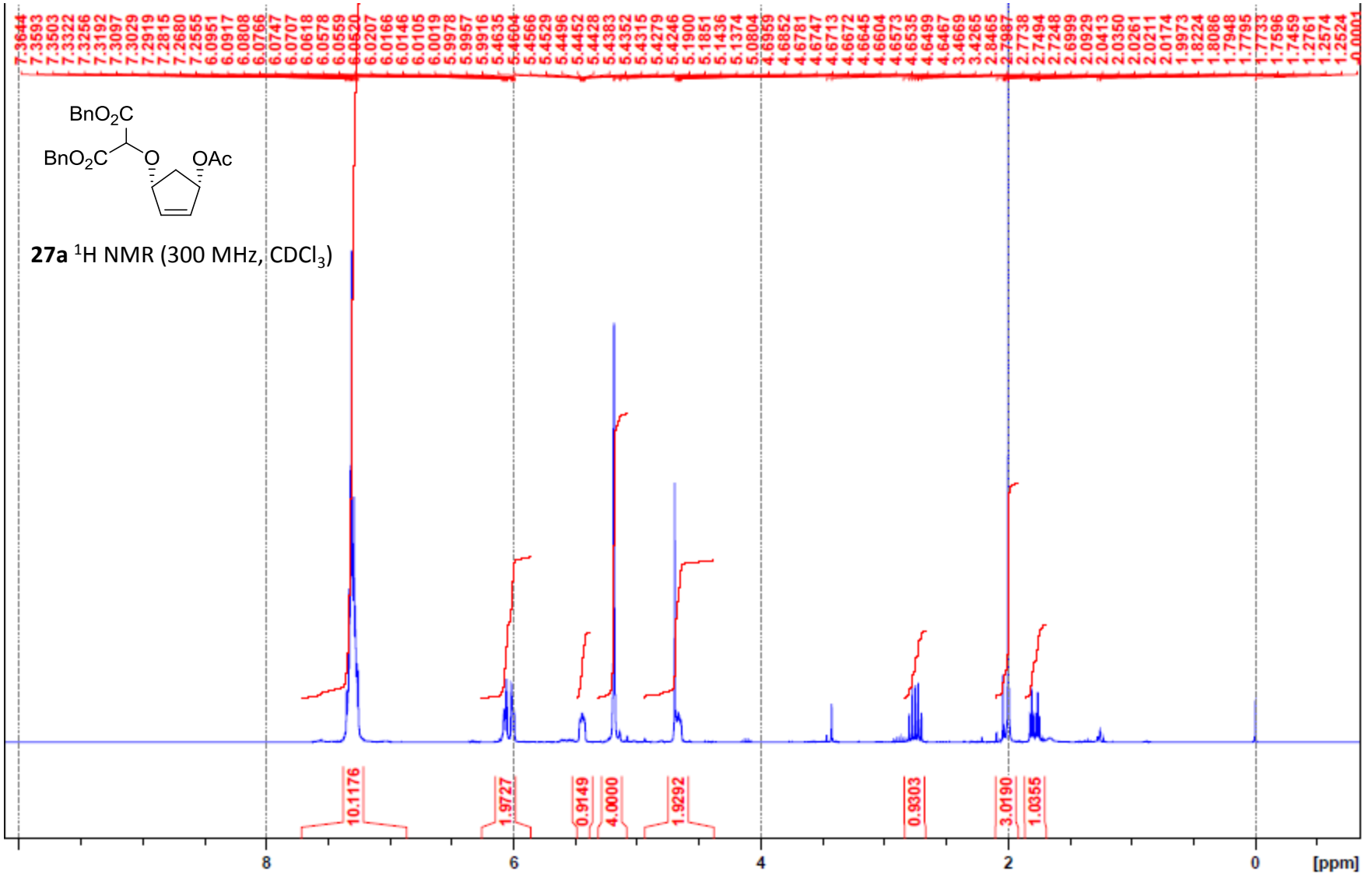


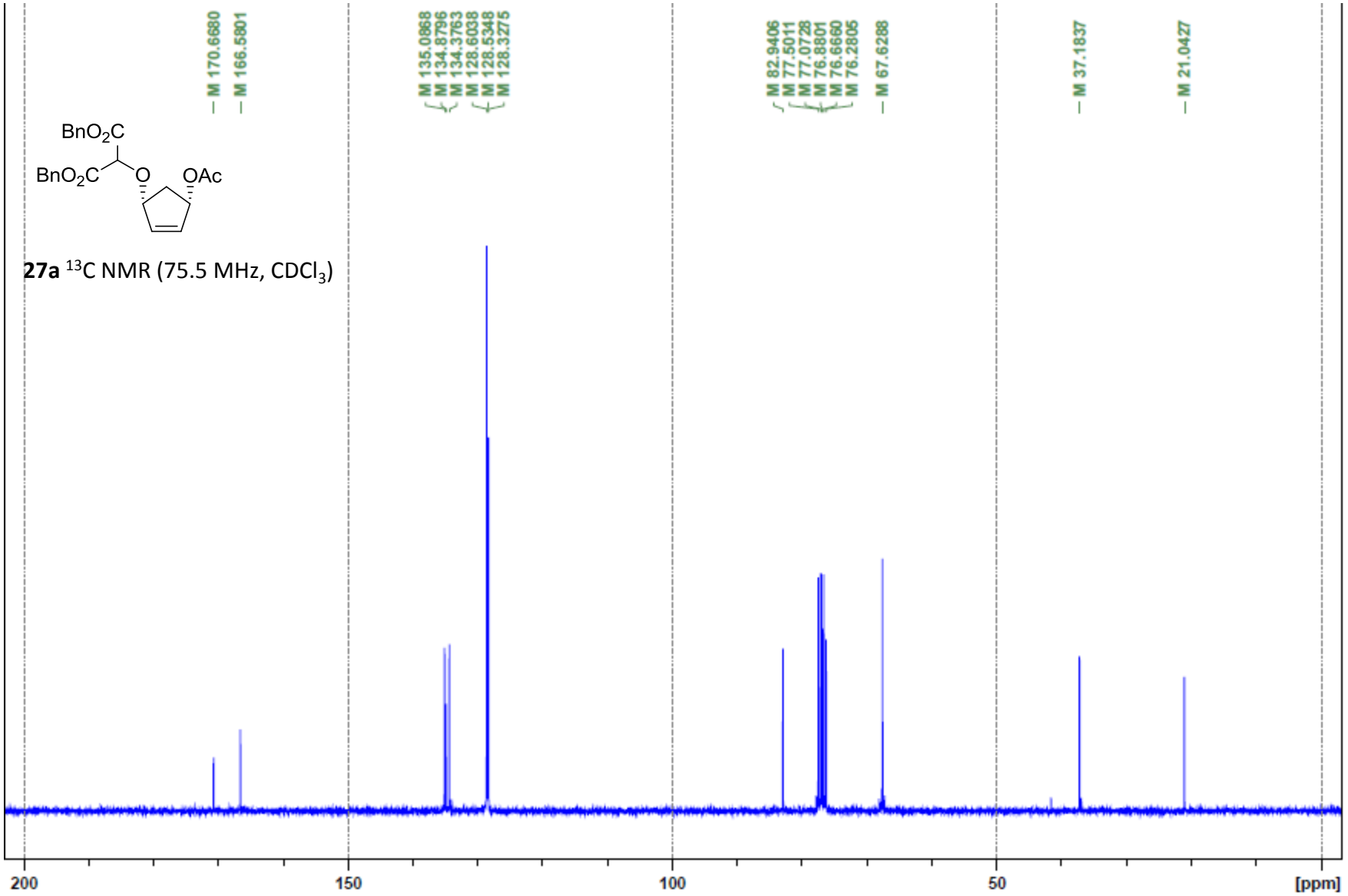
28 ^{13}C NMR (75.5 MHz, $CDCl_3$)

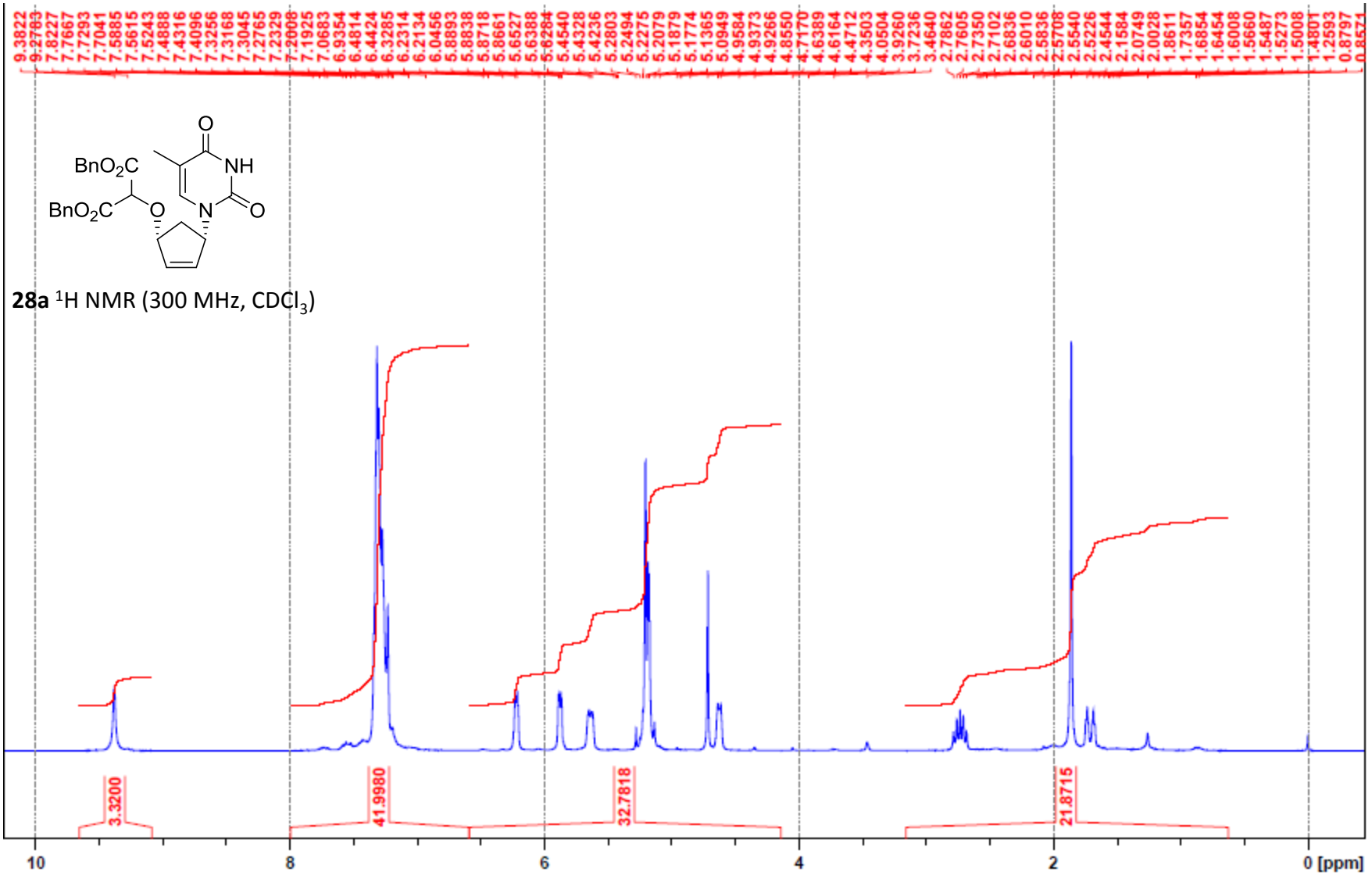

HO2C
HO2C
O
NH
N
O
O
2 1H NMR (300 MHz, MeOH-d4)
3.5430
3.6531
37.8970
3.8157
7.5992
0.8308
42.6612
10
8
6
4
2
[ppm]

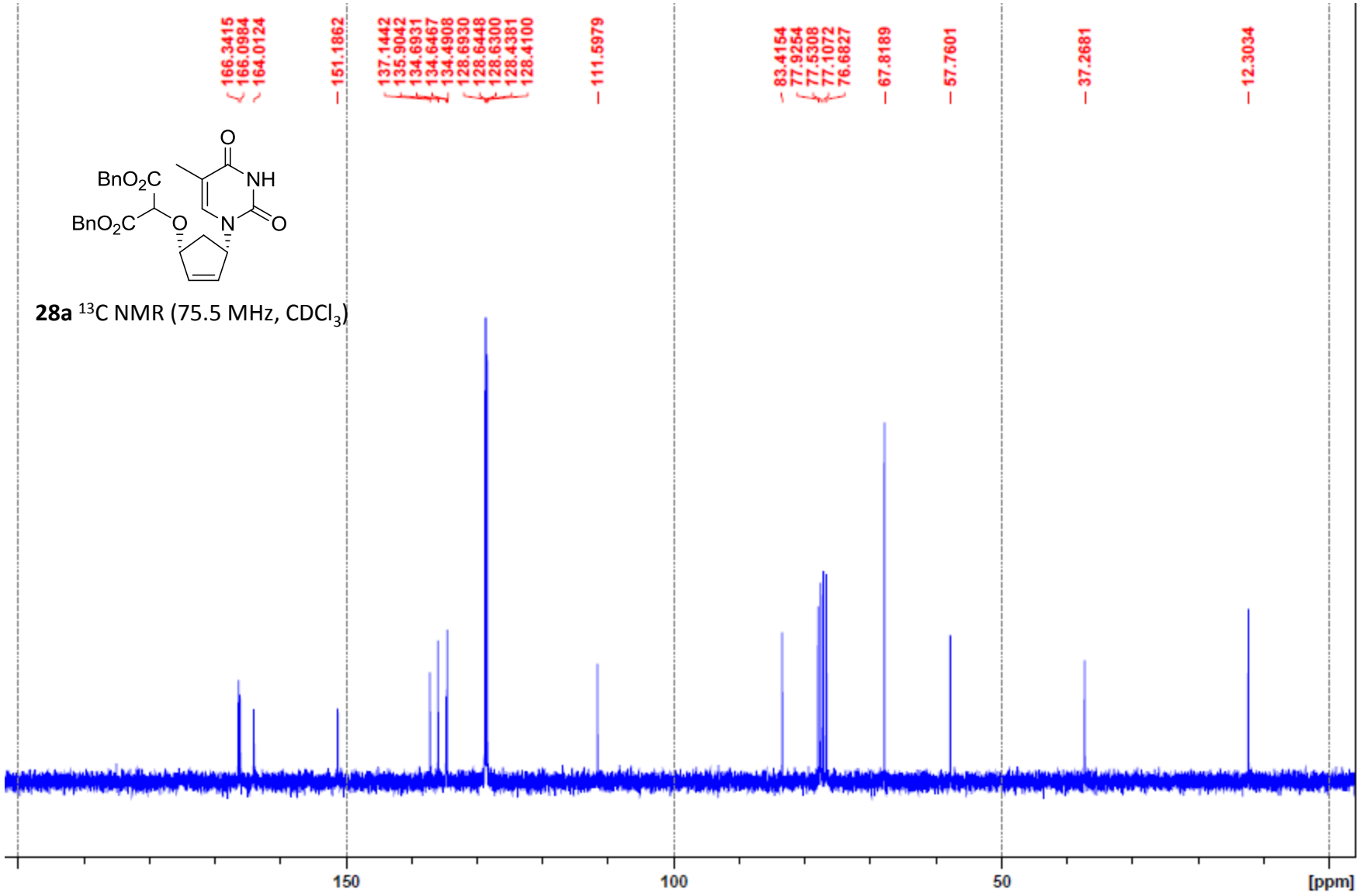


2 ^{13}C NMR (75.5 MHz, MeOH-d_4)

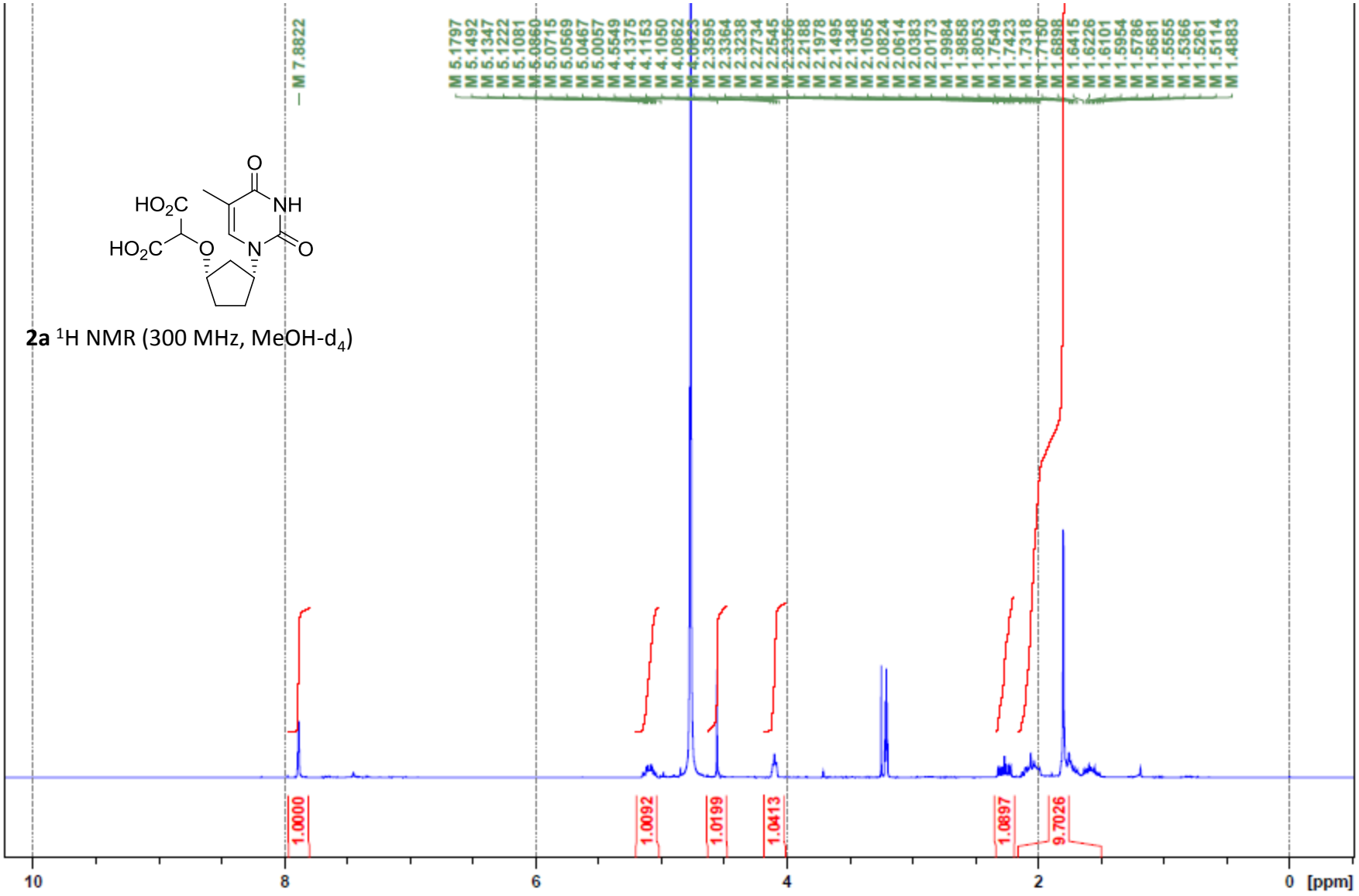

BnO2C
BnO2C
O
OAc
27a 1H NMR (300 MHz, CDCl3)
10.1176
1.9727
0.9149
4.0000
1.9292
0.9303
3.0190
1.0355
8
6
4
2
0
[ppm]


M 170.6680
M 166.5801
M 135.0868
M 134.8796
M 134.3763
M 128.6038
M 128.5348
M 128.3275
M 82.9406
M 77.5011
M 77.0728
M 76.8801
M 76.6660
M 76.2805
M 67.6288
M 37.1837
M 21.0427
BnO2C
BnO2C
O
OAc
27a 13C NMR (75.5 MHz, CDCl3)
200
150
100
50
[ppm]

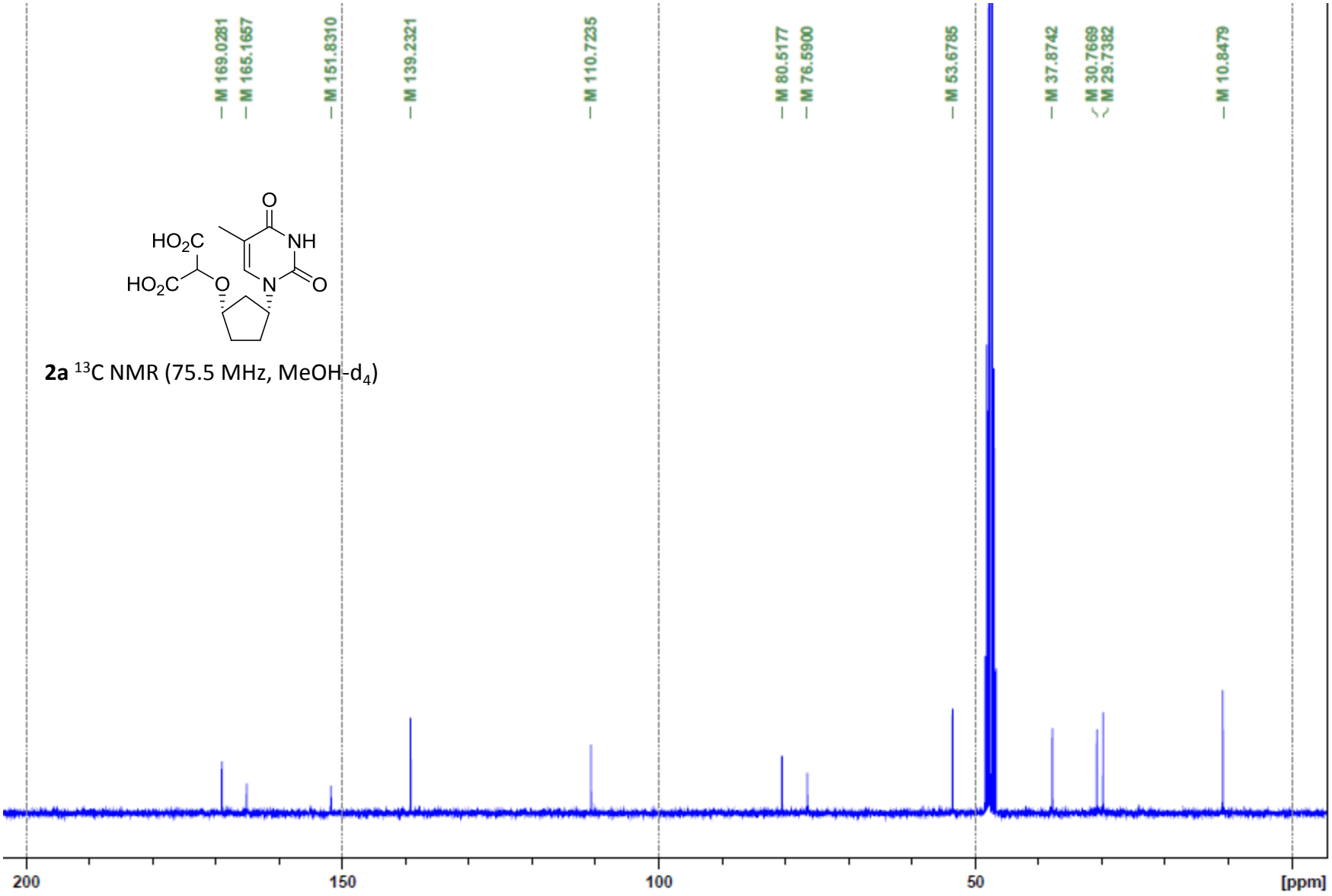

9.3822
9.2783
7.8227
7.7667
7.7293
7.7041
7.5885
7.5615
7.5243
7.4888
7.4316
7.4096
7.3256
7.3168
7.3045
7.2765
7.2329
7.2008
7.1925
7.0683
6.9354
6.4814
6.4424
6.3285
6.2314
6.2134
6.0456
5.8893
5.8838
5.8718
5.8661
5.6527
5.6388
5.6284
5.4540
5.4328
5.4236
5.2803
5.2494
5.2275
5.2079
5.1879
5.1774
5.1365
5.0949
4.9584
4.9373
4.9266
4.8550
4.7170
4.6389
4.6164
4.4712
4.3503
4.0504
3.9260
3.7236
3.4640
2.7862
2.7605
2.7350
2.7102
2.6836
2.6010
2.5836
2.5708
2.5540
2.5226
2.4544
2.1584
2.0749
2.0028
1.8611
1.7357
1.6854
1.6454
1.6008
1.5660
1.5487
1.5273
1.5008
1.4801
1.2593
0.8797
0.8571
BnO2C
BnO2C
O
NH
N
O
O
28a 1H NMR (300 MHz, CDCl3)
3.3200
41.9980
32.7818
21.8715
10
8
6
4
2
0 [ppm]



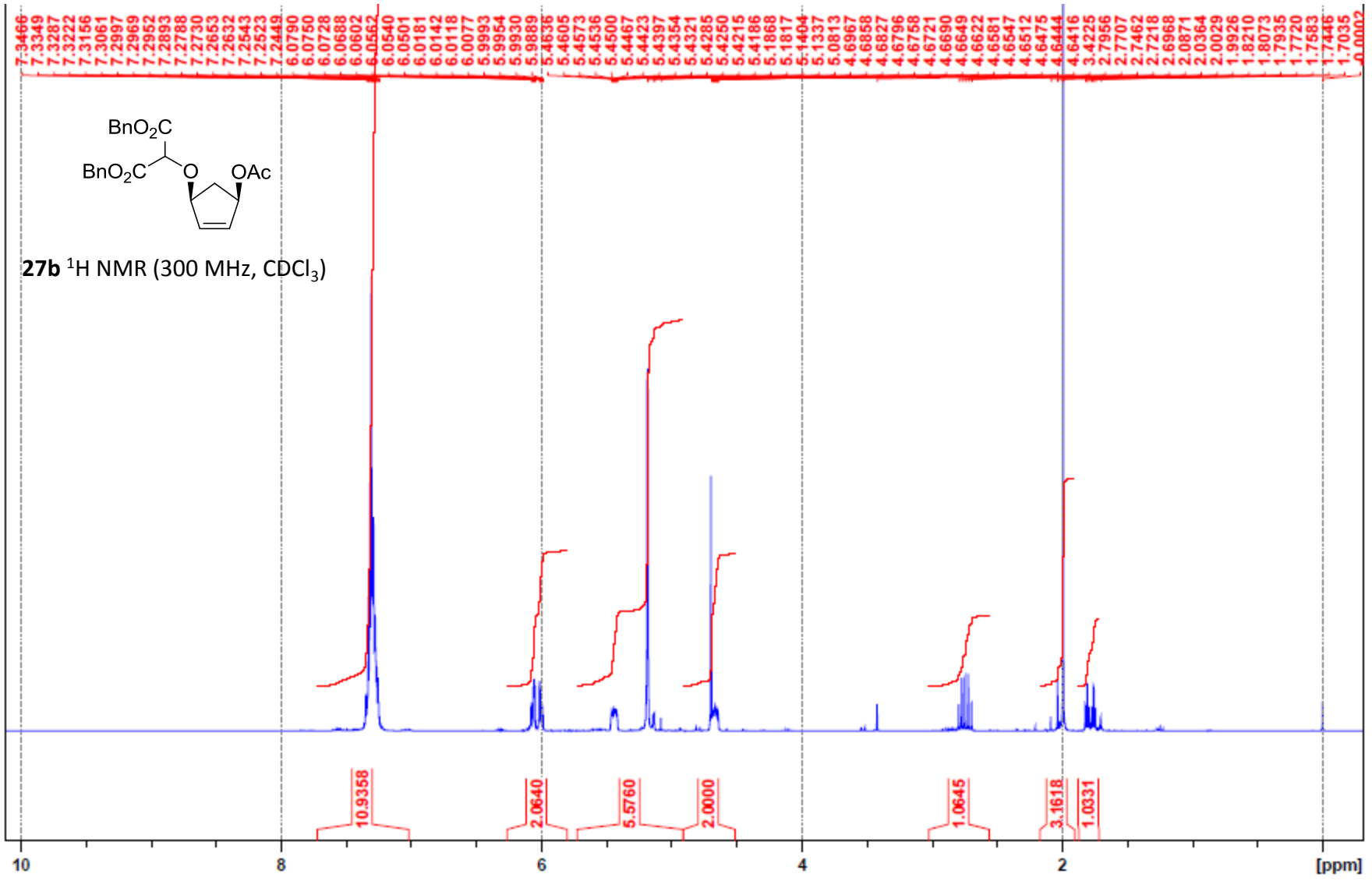
28a ^{13}C NMR (75.5 MHz, $CDCl_3$)

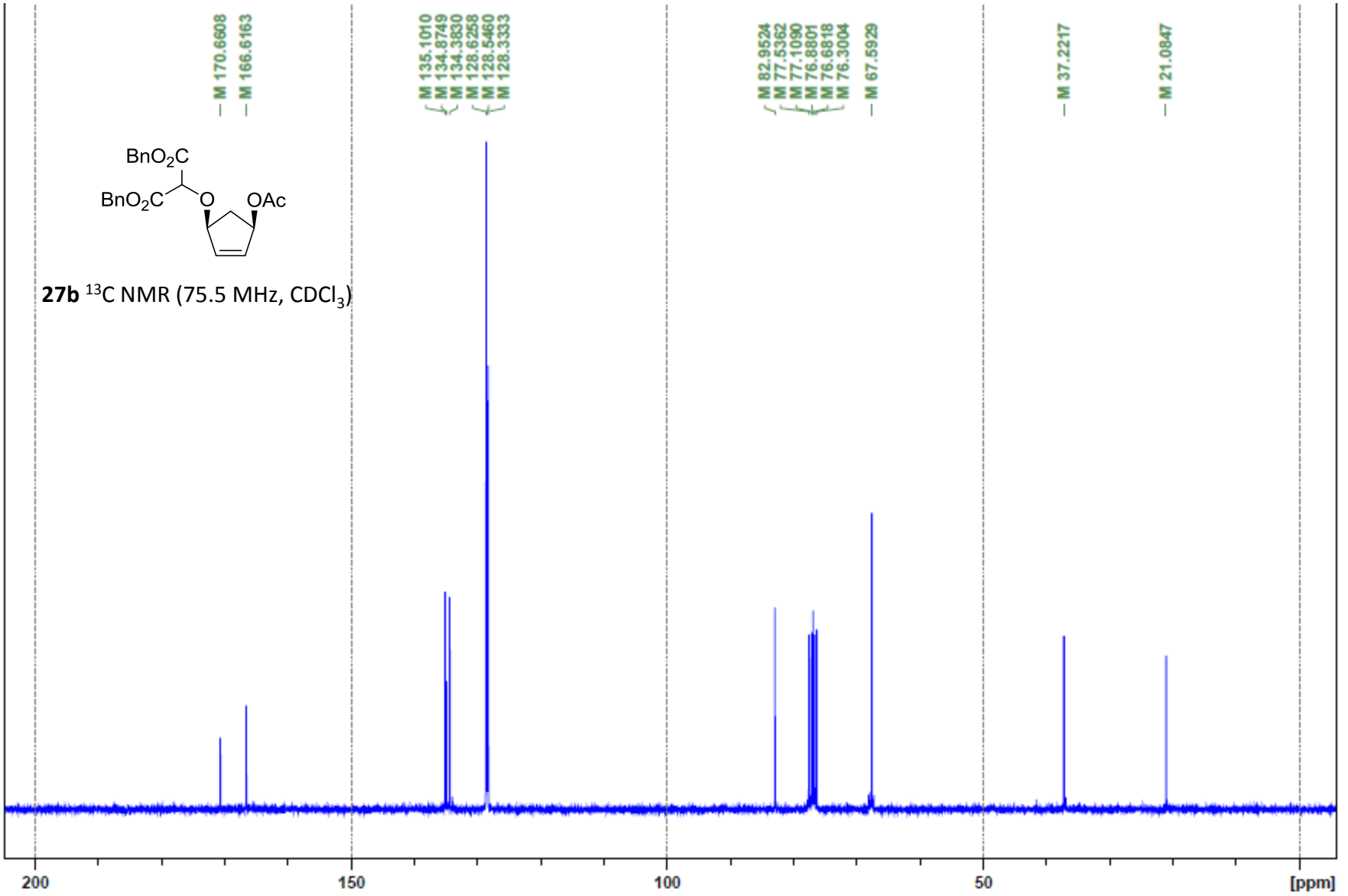


2a ^{1}H NMR (300 MHz, MeOH-d_4)

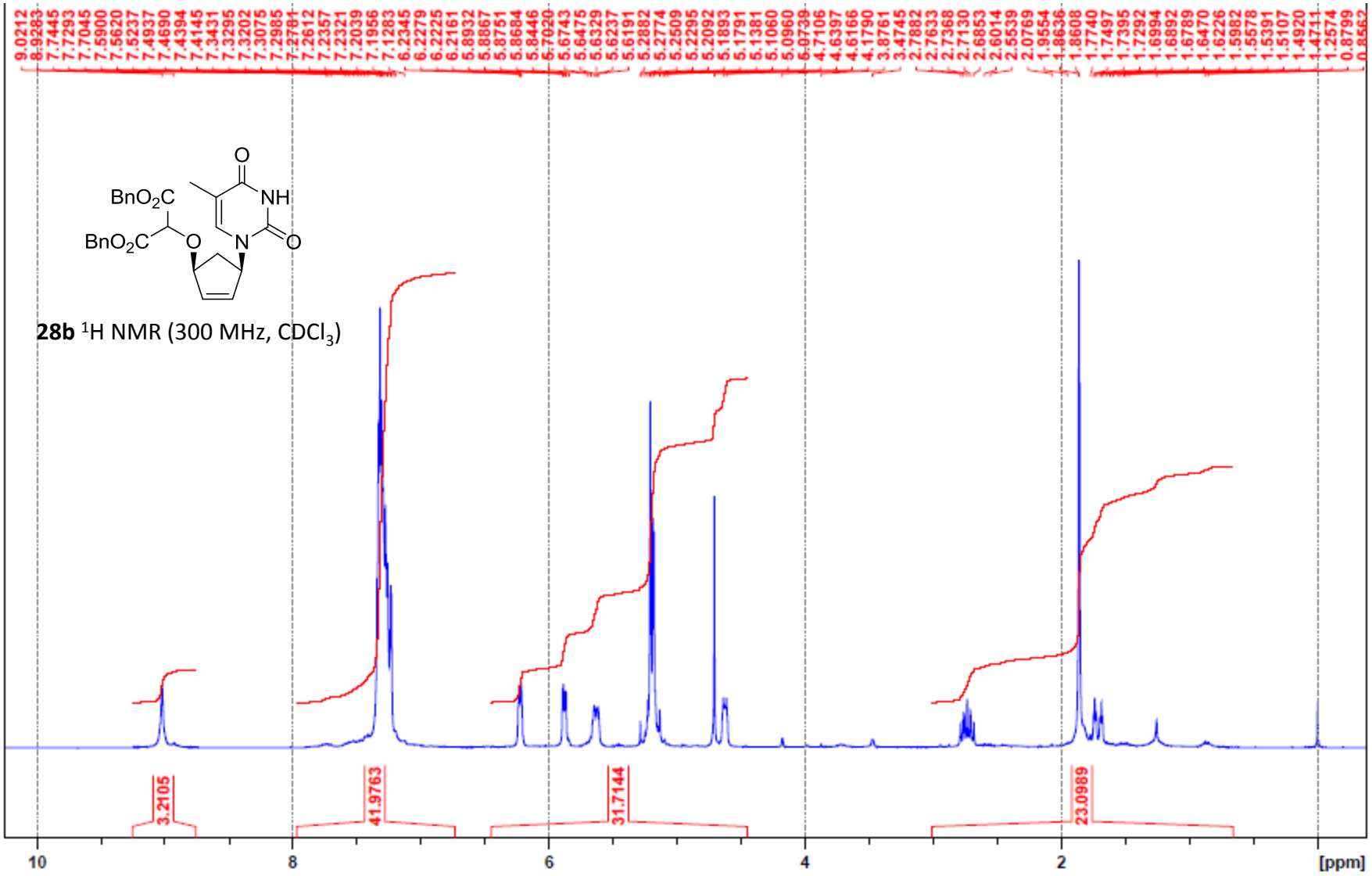


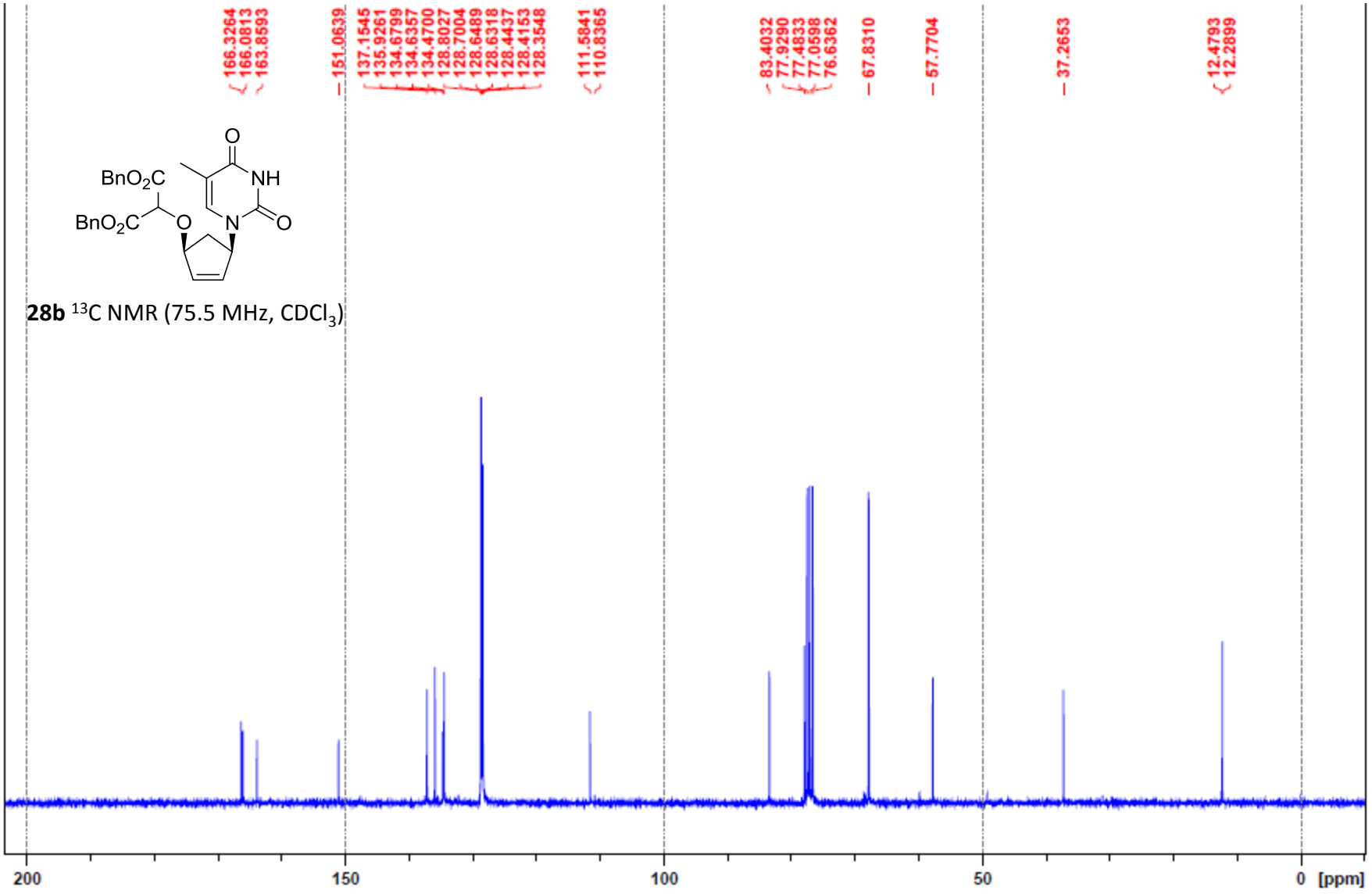
2a ^{13}C NMR (75.5 MHz, MeOH-d_4)


BnO2C
BnO2C
O
OAc
27b 1H NMR (300 MHz, CDCl3)
10.9358
2.0640
5.5760
2.0000
1.0645
3.1618
1.0331
10
8
6
4
2
[ppm]

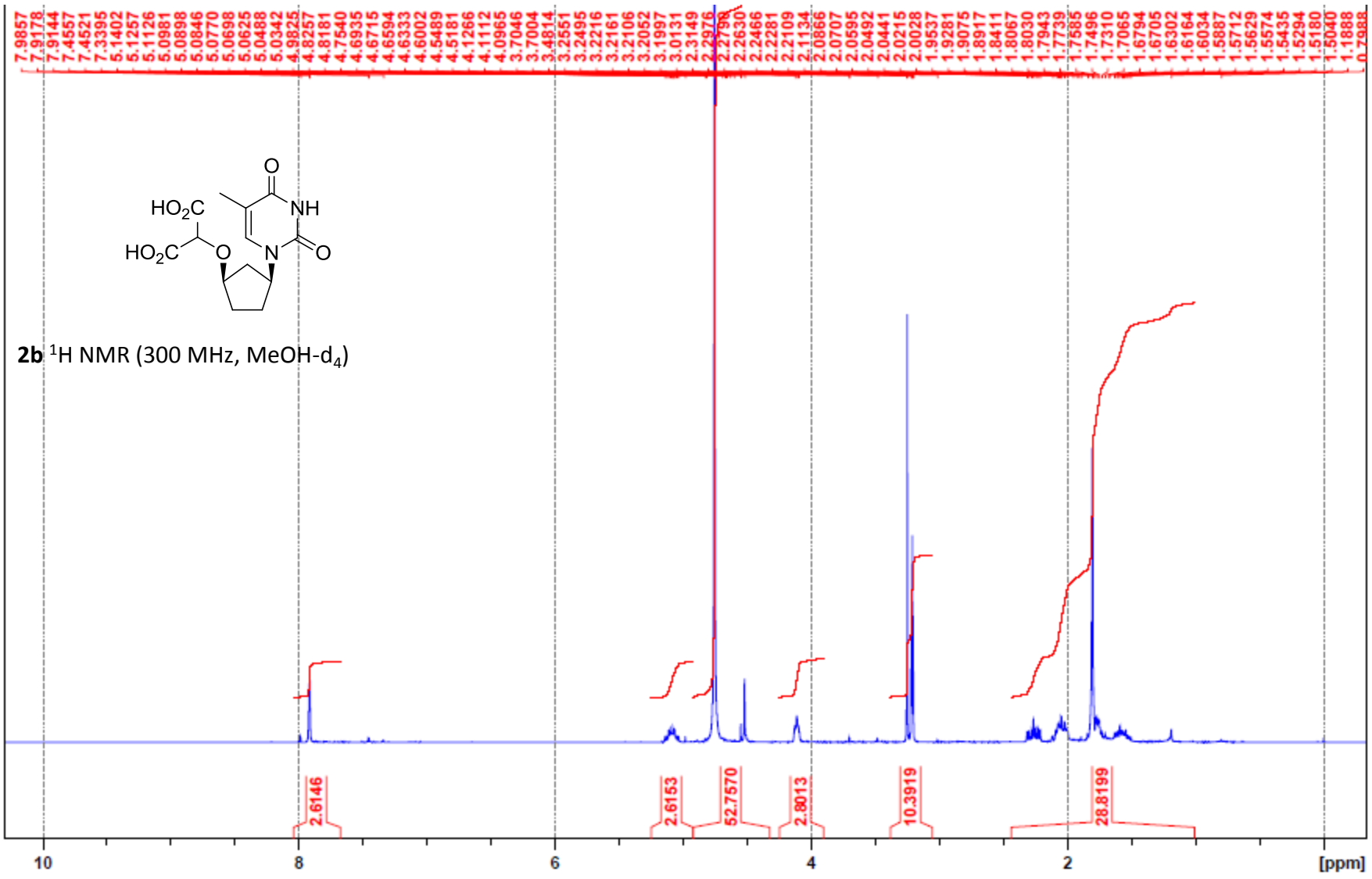


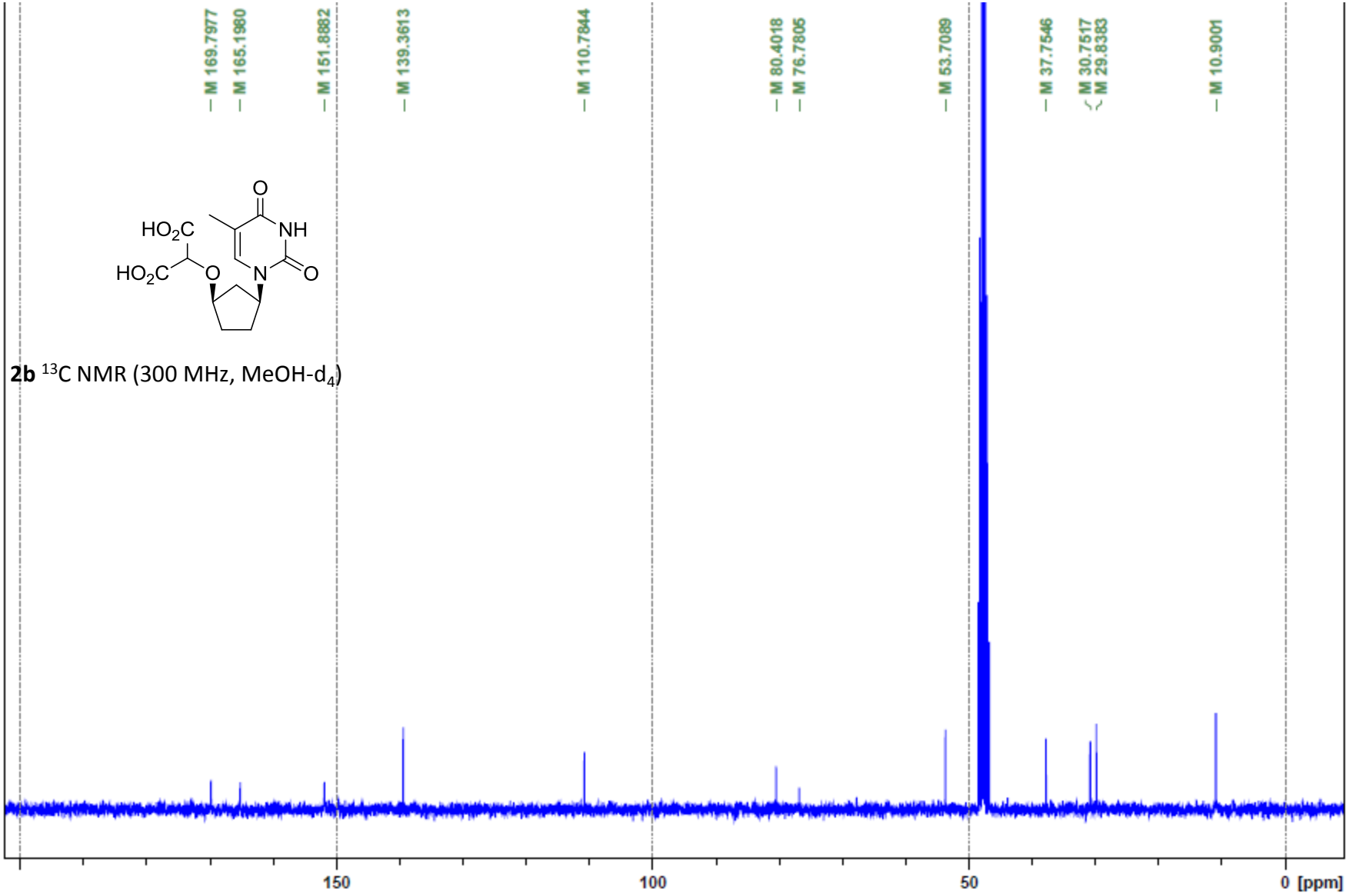
27b ^{13}C NMR (75.5 MHz, $CDCl_3$)


9.0212
8.9283
7.7445
7.7293
7.7045
7.5900
7.5620
7.5237
7.4937
7.4690
7.4394
7.4145
7.3431
7.3295
7.3202
7.3075
7.2985
7.2781
7.2612
7.2357
7.2321
7.2039
7.1956
7.1283
6.2345
6.2279
6.2225
6.2161
5.8932
5.8867
5.8751
5.8684
5.8446
5.7020
5.6743
5.6475
5.6329
5.6237
5.6191
5.2882
5.2774
5.2509
5.2295
5.2092
5.1893
5.1791
5.1381
5.1060
5.0960
4.7106
4.6397
4.6166
4.1790
3.8761
3.4745
2.7882
2.7633
2.7368
2.7130
2.6853
2.6014
2.5539
2.0769
1.9554
1.8636
1.8608
1.7740
1.7497
1.7395
1.7292
1.6994
1.6892
1.6789
1.6470
1.6226
1.5982
1.5578
1.5391
1.5107
1.4920
1.4711
1.2574
0.8799
0.8562
BnO2C
BnO2C
O
NH
N
O
O
28b 1H NMR (300 MHz, CDCl3)
3.2105
41.9763
31.7144
23.0989
10
8
6
4
2
[ppm]

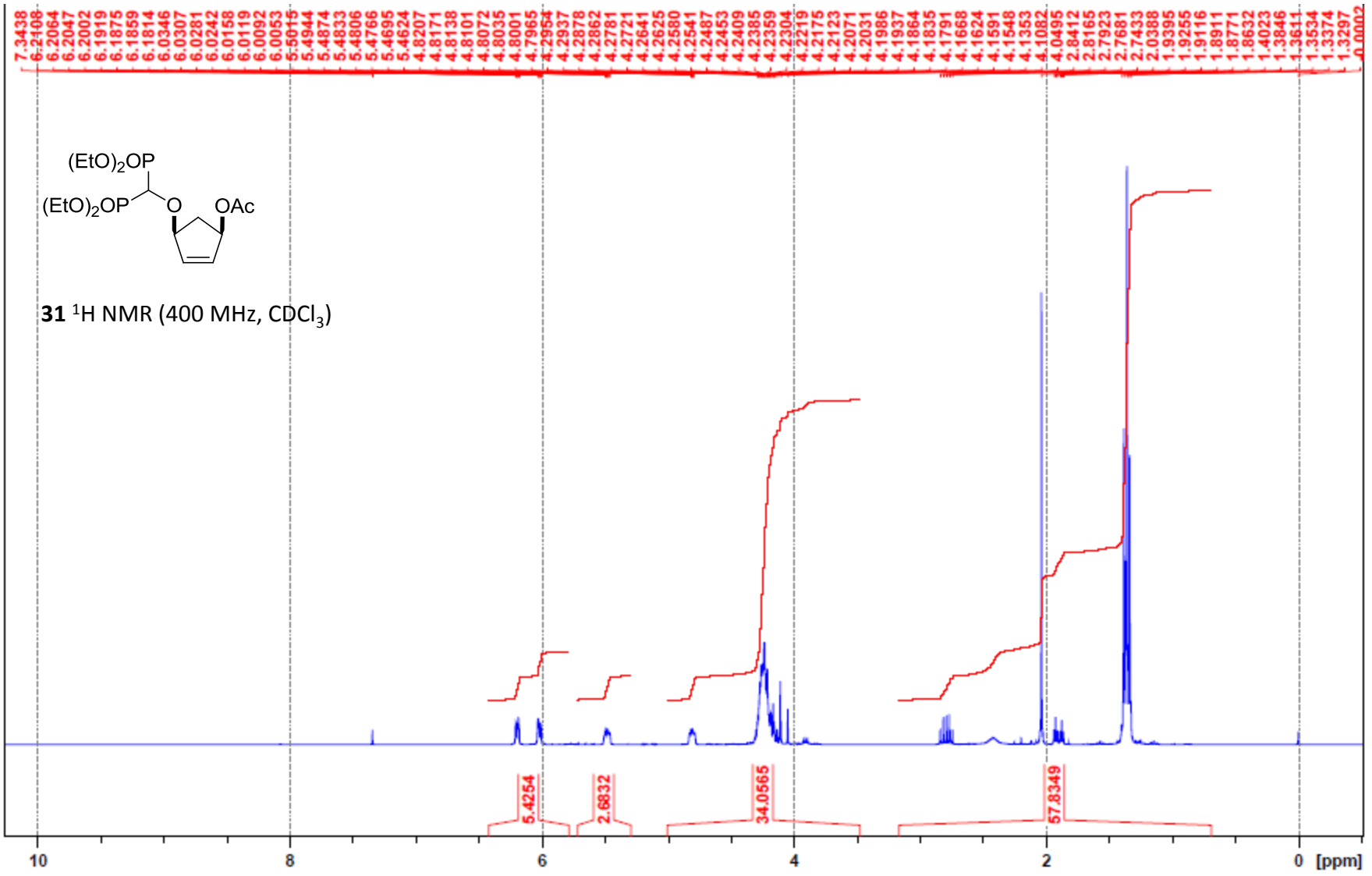


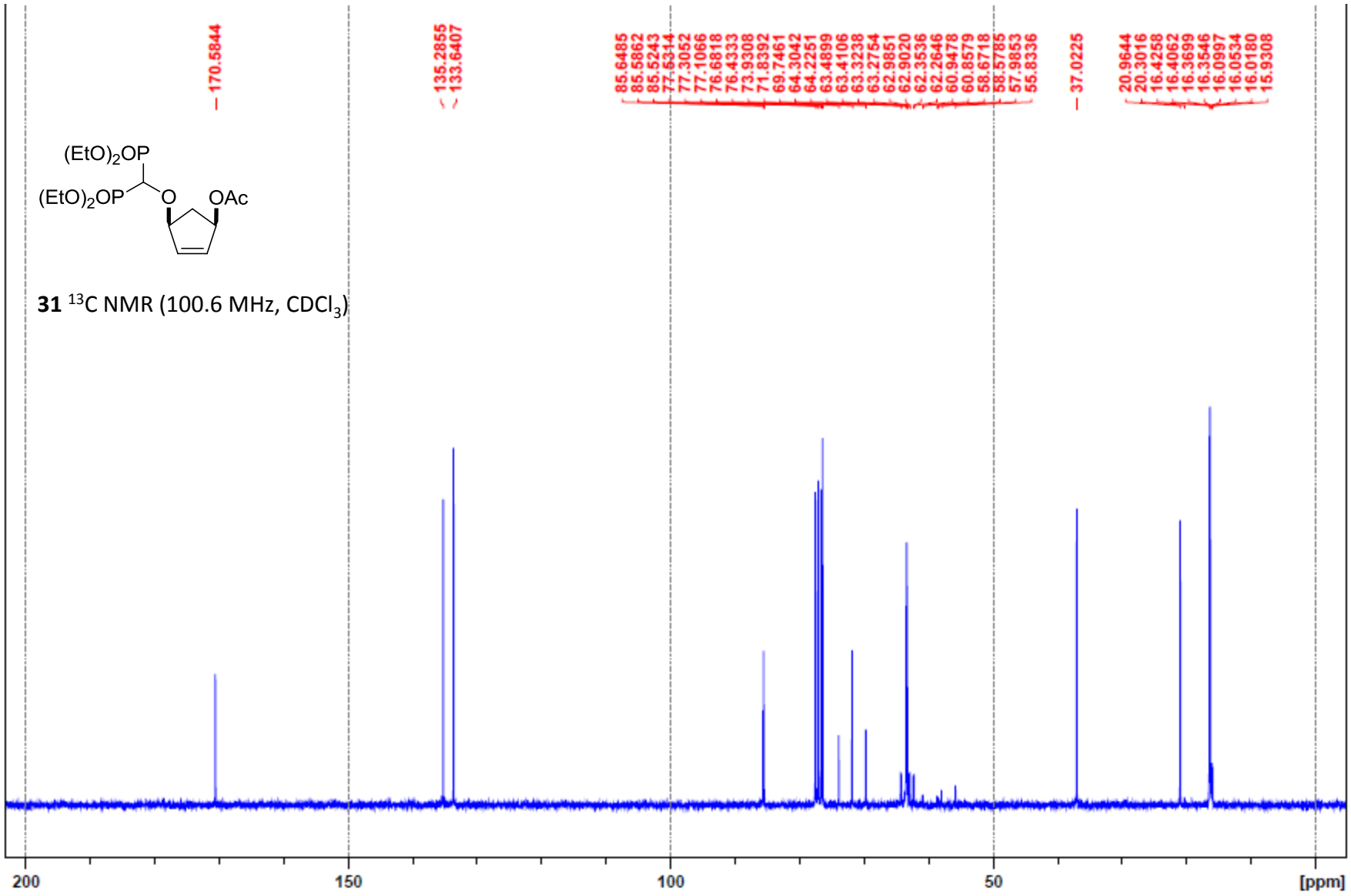
28b ^{13}C NMR (75.5 MHz, $CDCl_3$)


HO2C
HO2C
O
NH
O
N
O
2b 1H NMR (300 MHz, MeOH-d4)
2.6146
2.6153
52.7570
2.8013
10.3919
28.8199
10
8
6
4
2
[ppm]

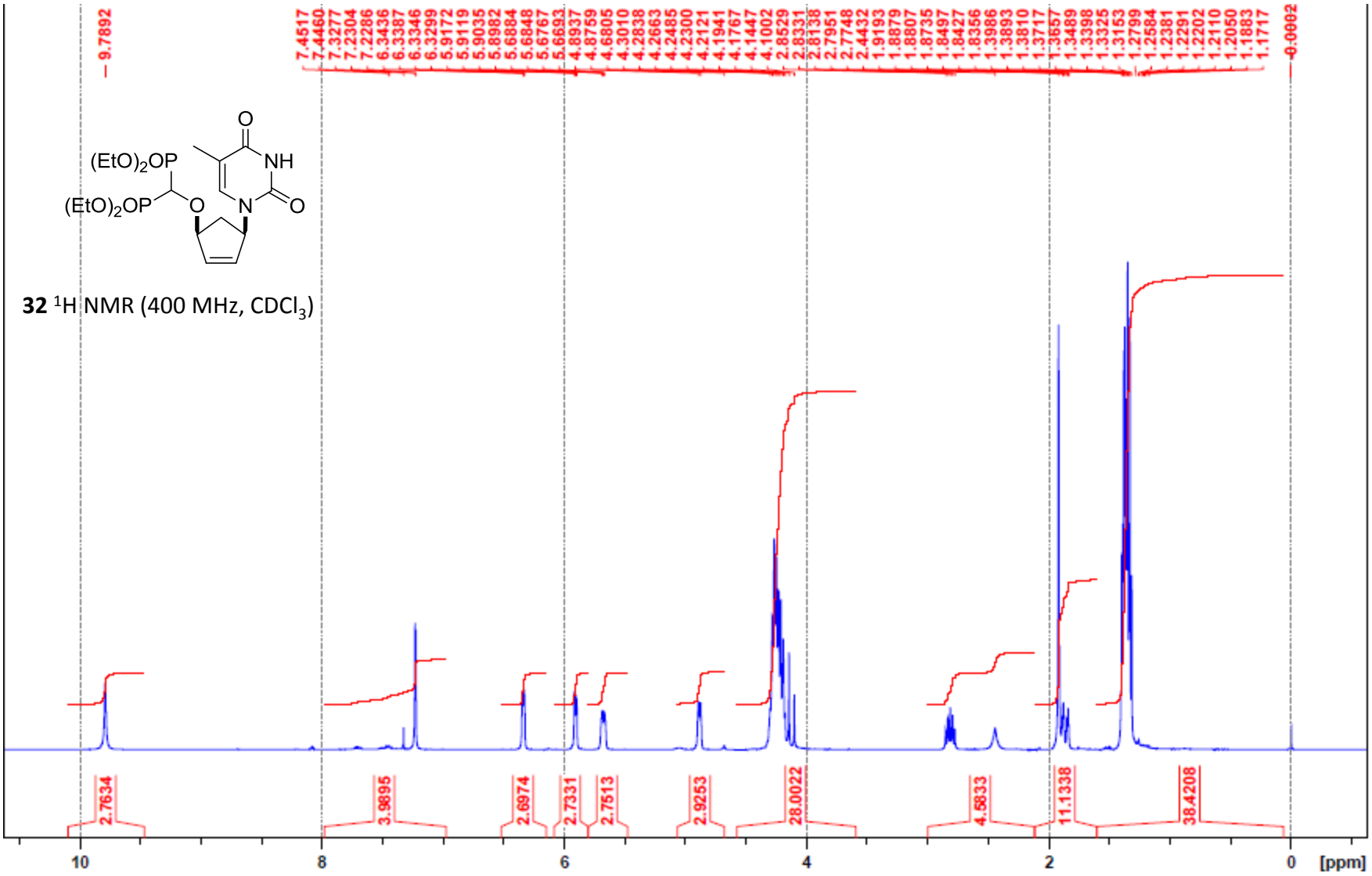


2b ^{13}C NMR (300 MHz, MeOH-d_4)

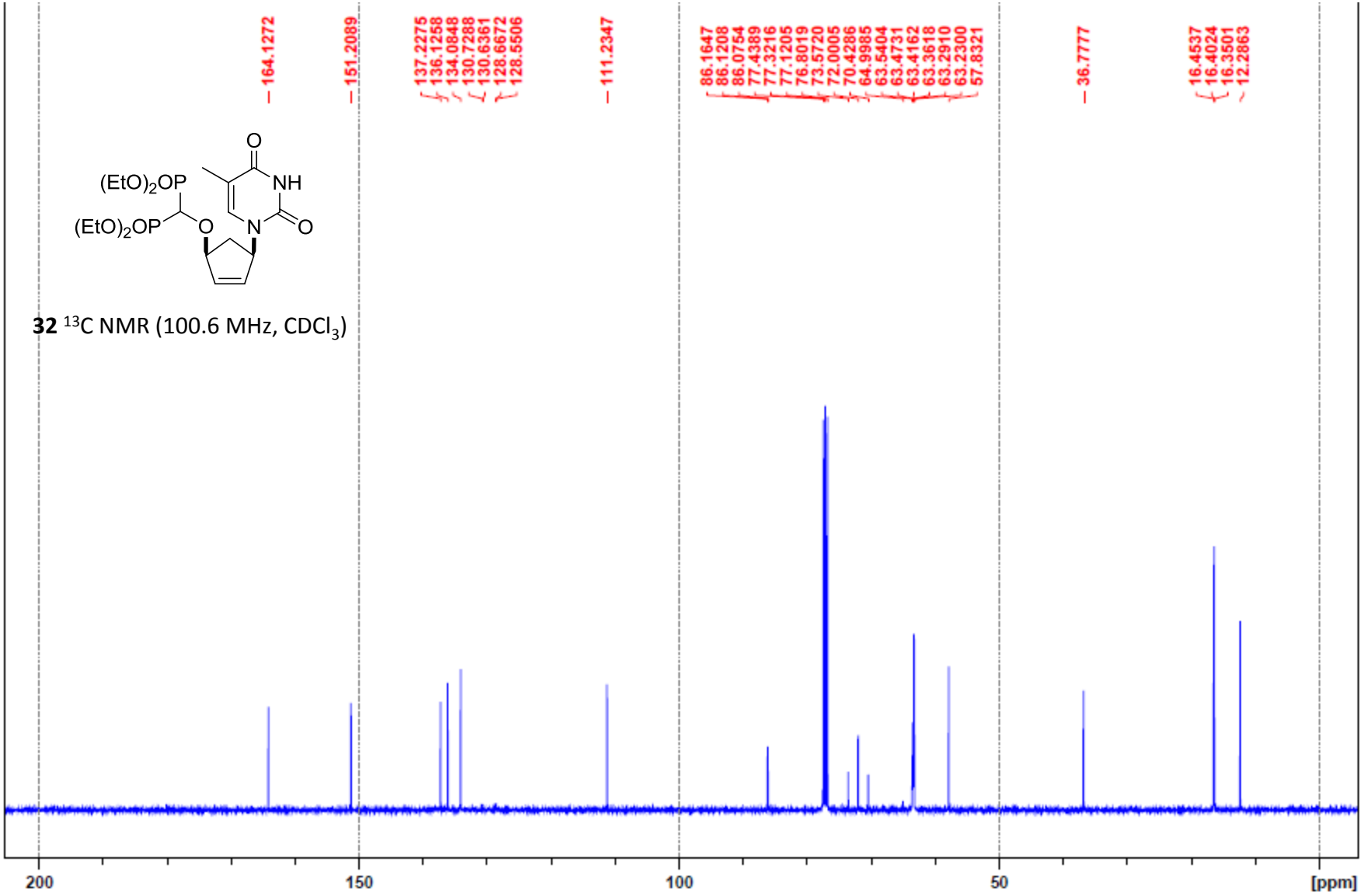

(EtO)2OP
(EtO)2OP
O
OAc
31 1H NMR (400 MHz, CDCl3)
5.4254
2.6832
34.0565
57.8349
10
8
6
4
2
0
[ppm]



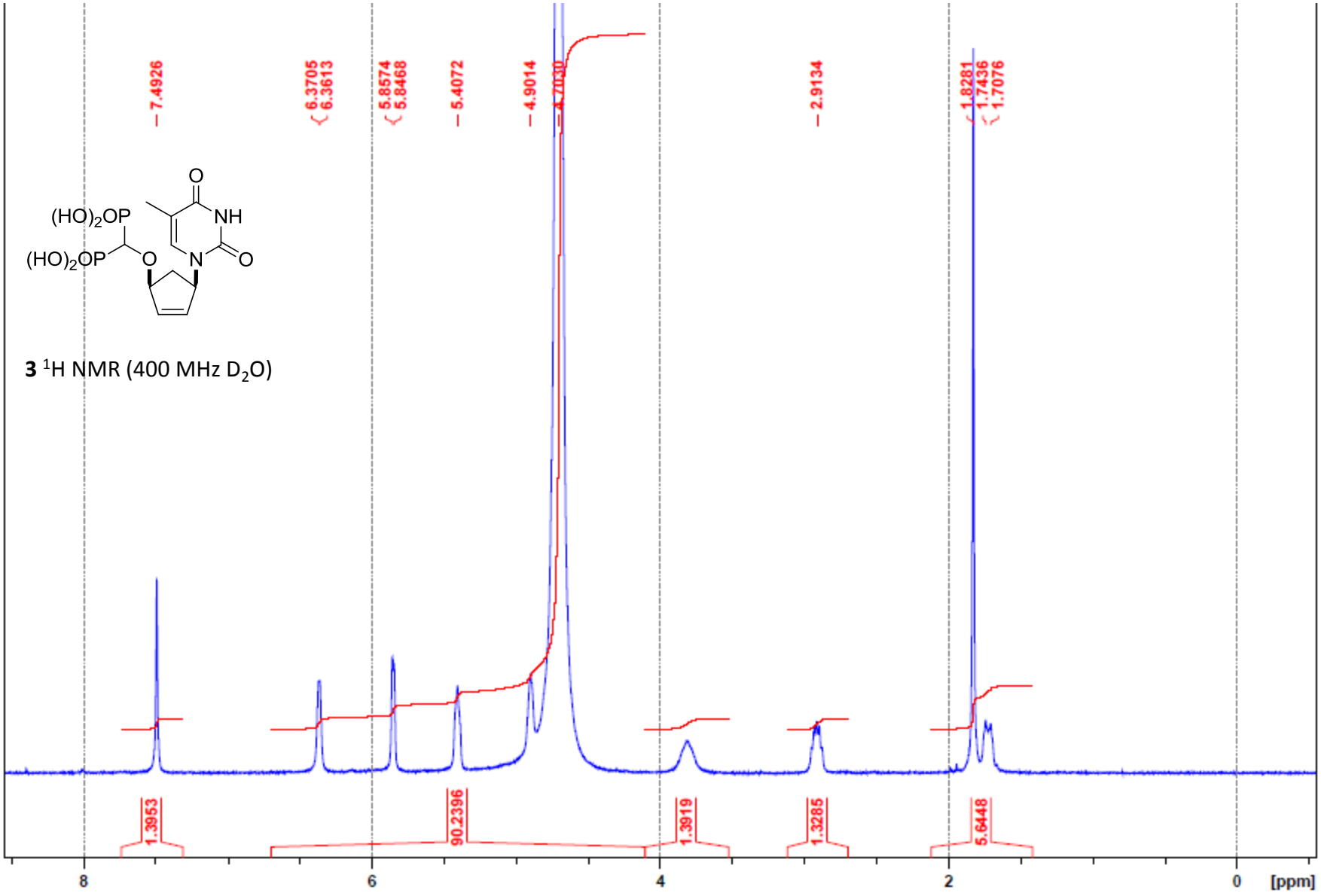
31 ^{13}C NMR (100.6 MHz, $CDCl_3$)



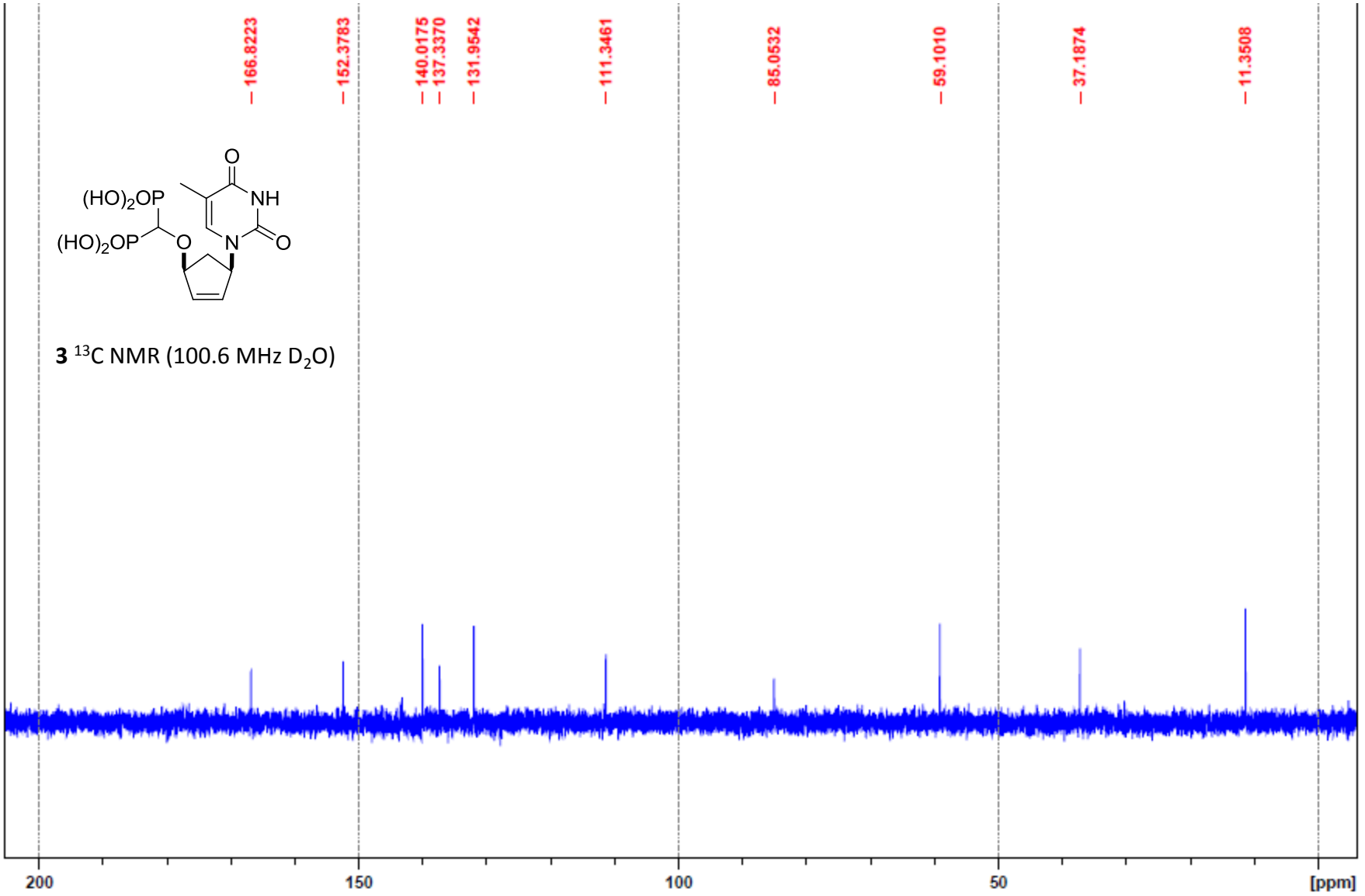
32 ^{1}H NMR (400 MHz, $CDCl_3$)



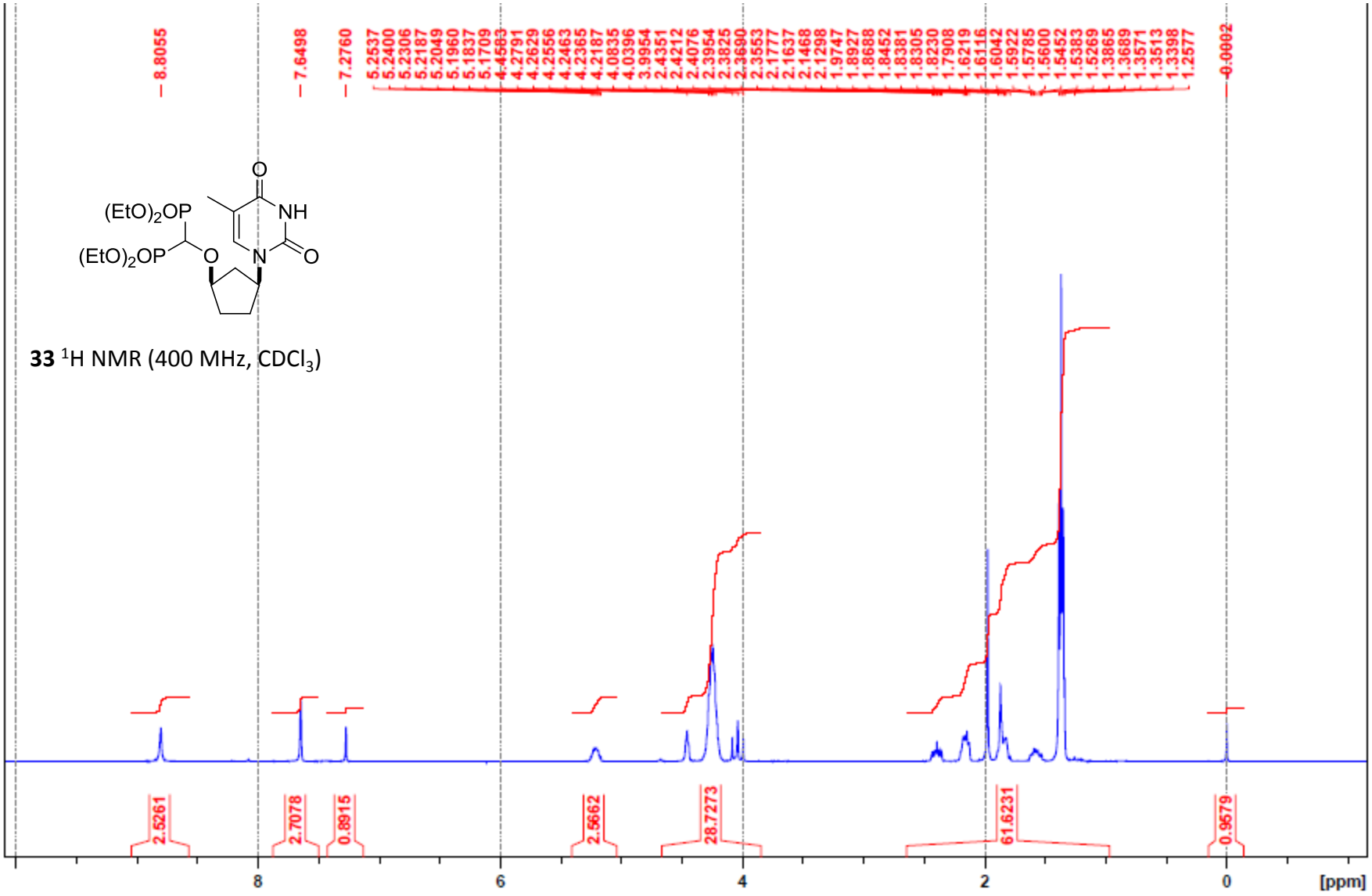
32 ^{13}C NMR (100.6 MHz, $CDCl_3$)



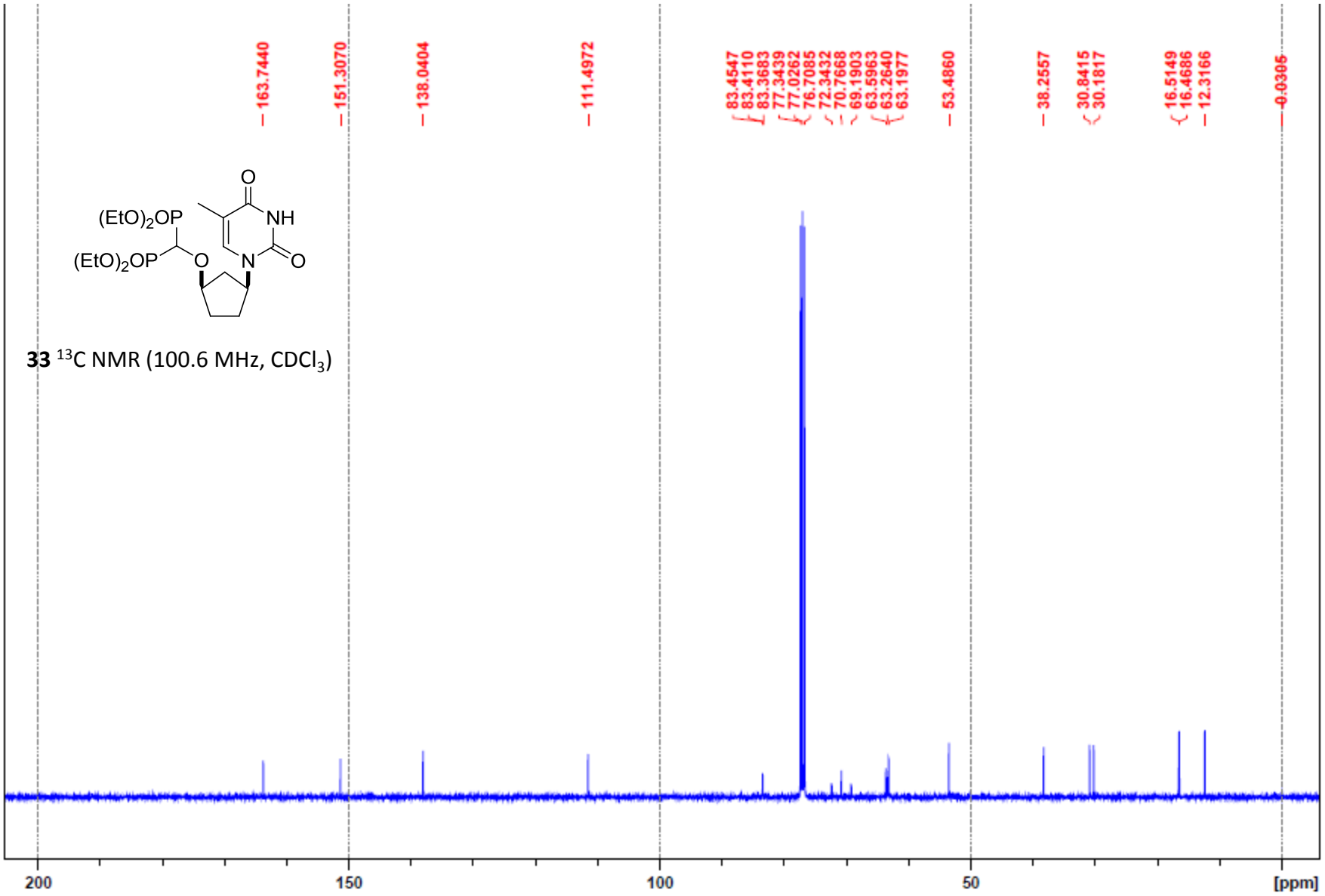
3 ^{1}H NMR (400 MHz D_2O)



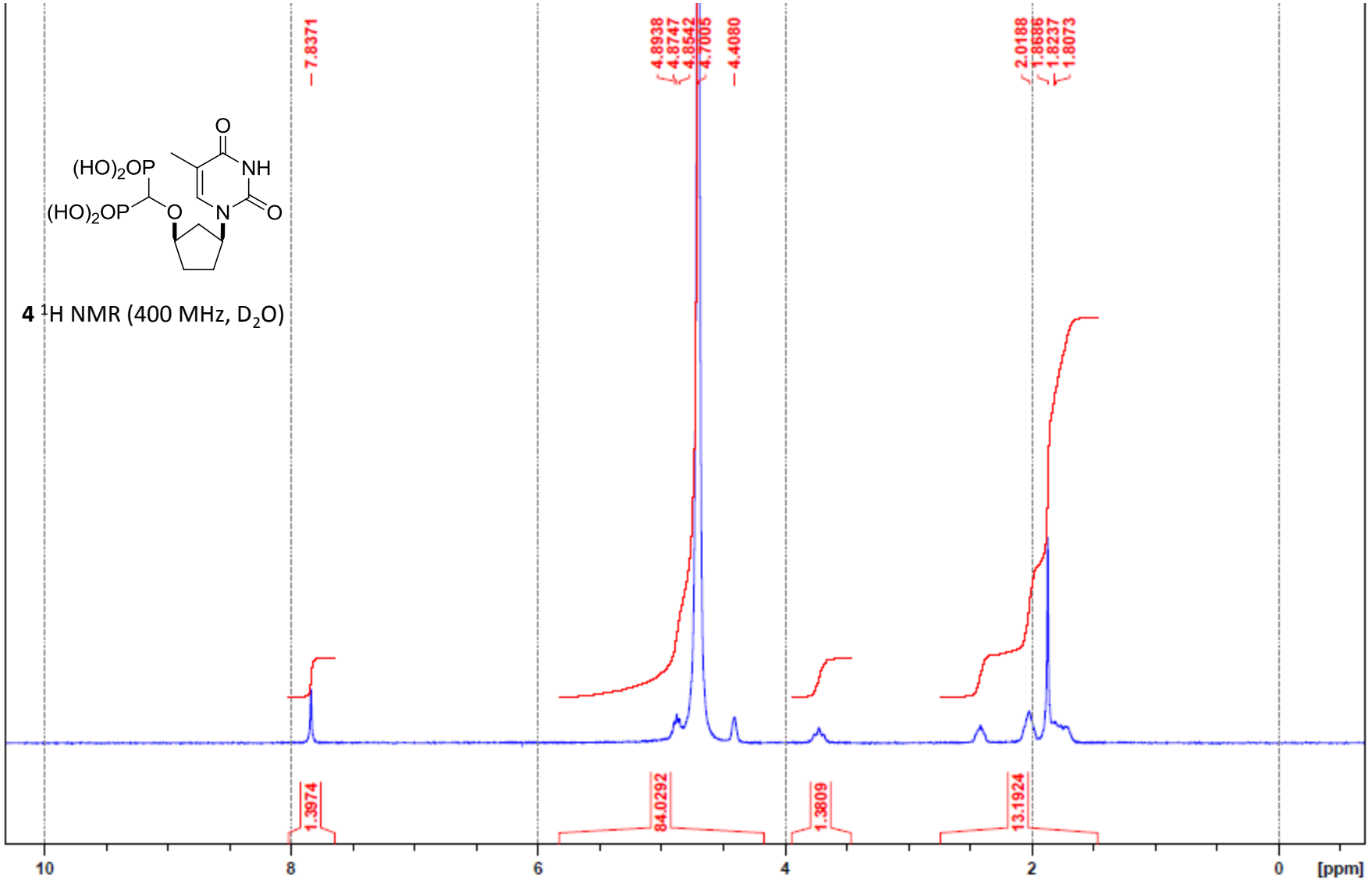
3 ^{13}C NMR (100.6 MHz D_2O)

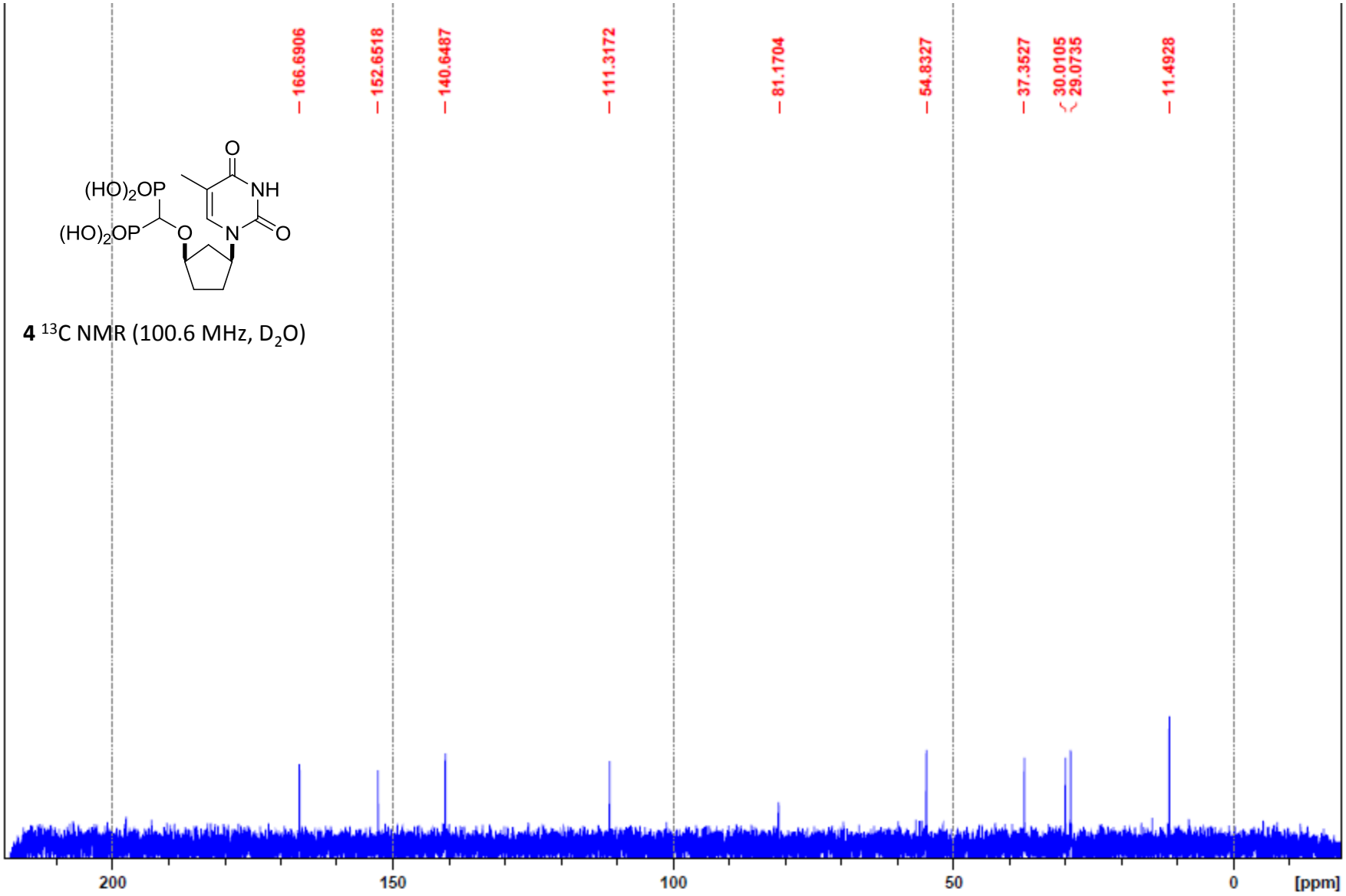


33 ^{1}H NMR (400 MHz, $CDCl_3$)

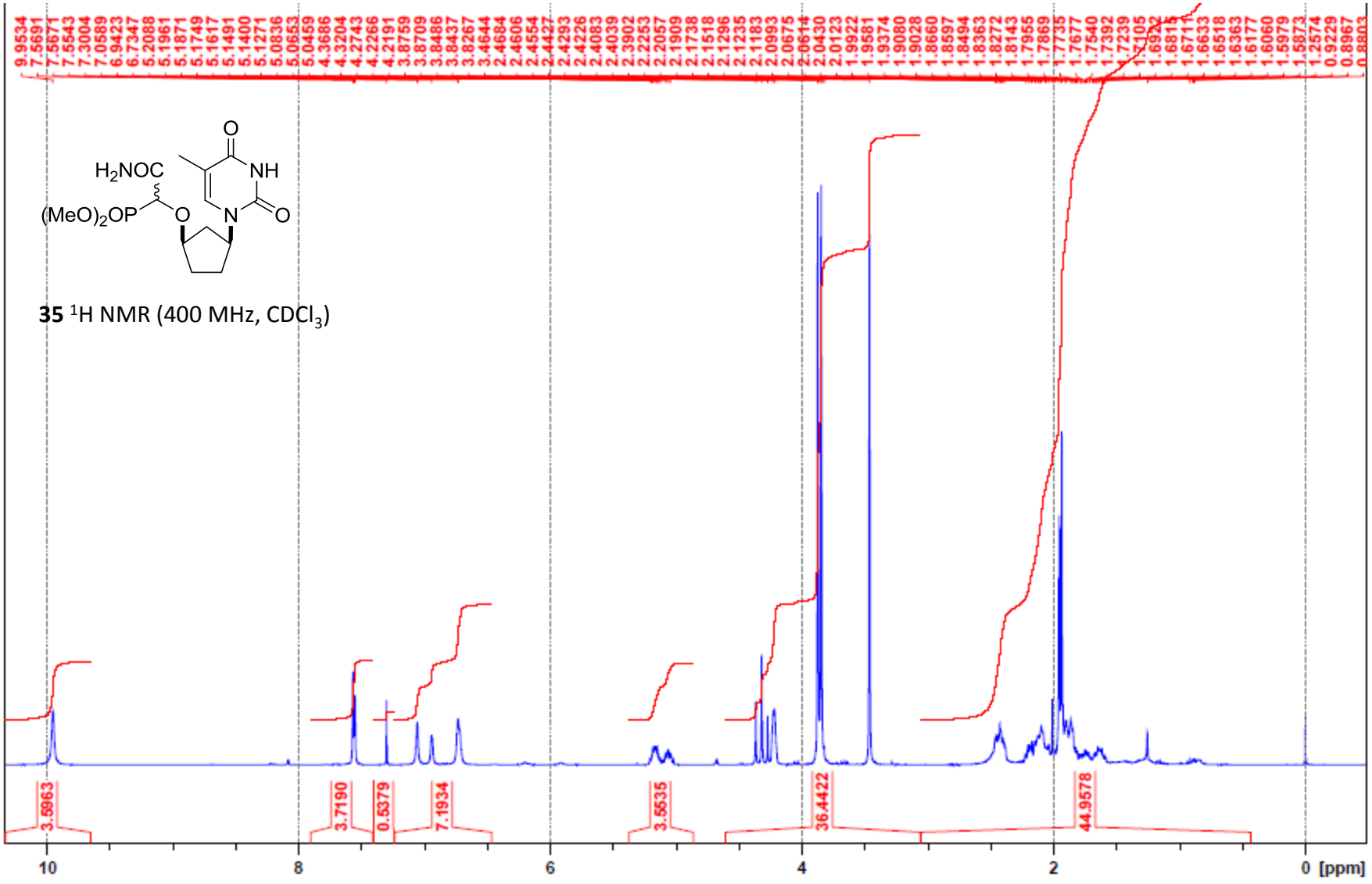


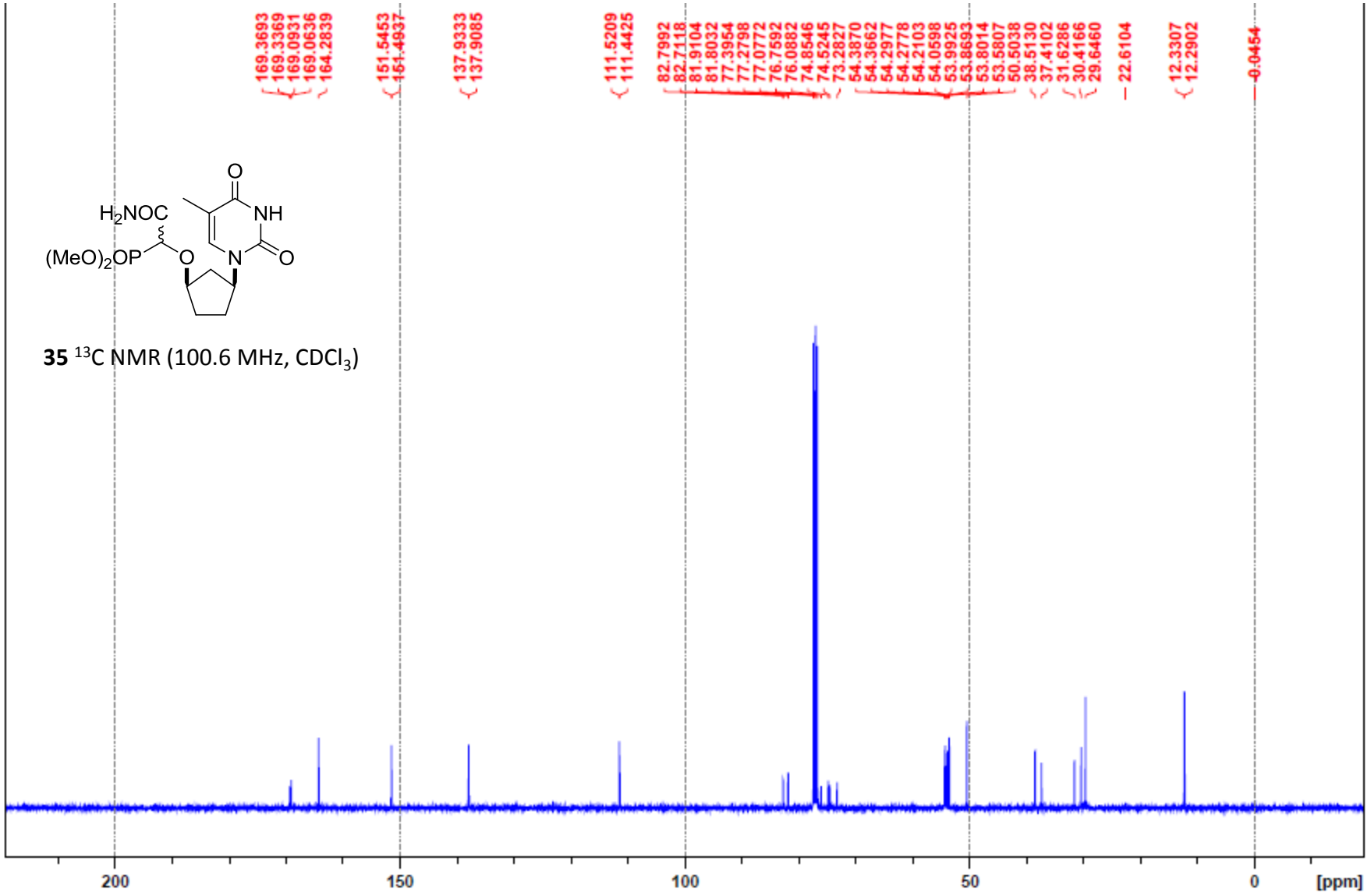
33 ^{13}C NMR (100.6 MHz, $CDCl_3$)


(HO)2OP
(HO)2OP
O
NH
N
O
O
4 1H NMR (400 MHz, D2O)
7.8371
4.8938
4.8747
4.8542
4.7005
4.4080
2.0188
1.8686
1.8237
1.8073
1.3974
84.0292
1.3809
13.1924
10
8
6
4
2
0
[ppm]

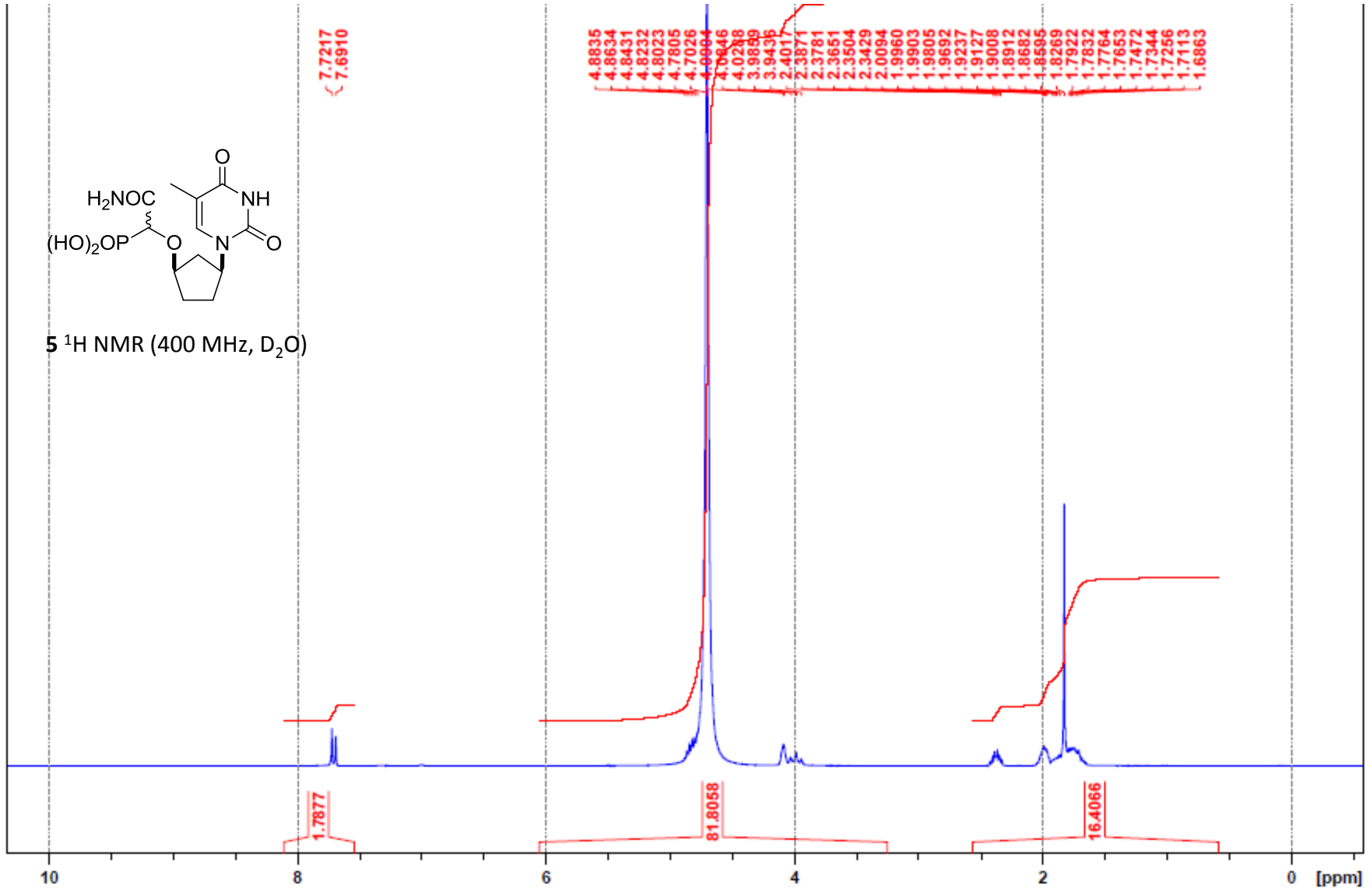


4 ^{13}C NMR (100.6 MHz, D_2O)

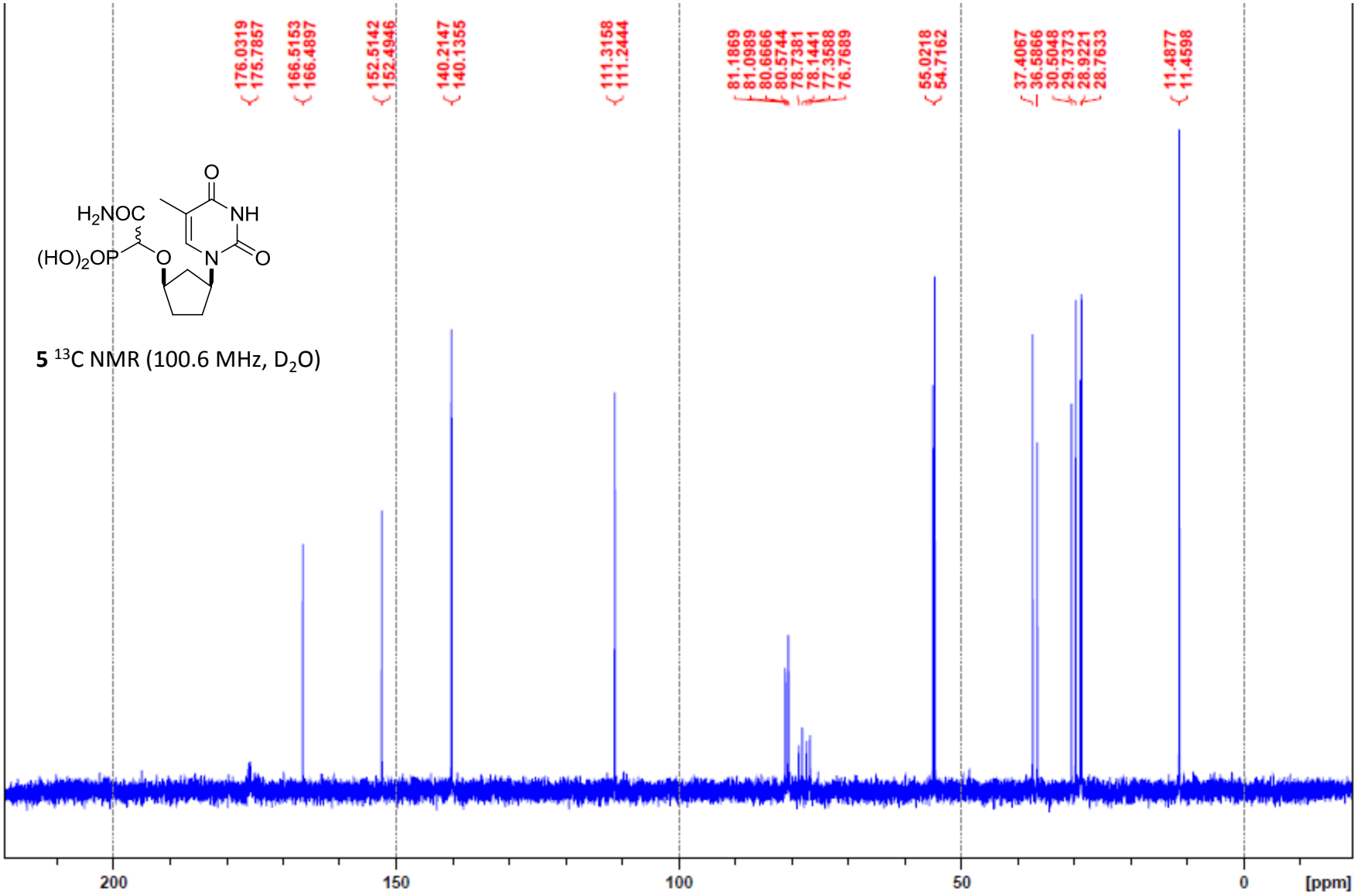

H2NOC
(MeO)2OP
O
NH
N
O
35 1H NMR (400 MHz, CDCl3)
3.5963
3.7190
0.5379
7.1934
3.5535
36.4422
44.9578
10
8
6
4
2
0 [ppm]



35 ^{13}C NMR (100.6 MHz, $CDCl_3$)



5 ^{1}H NMR (400 MHz, D_2O)



5 ^{13}C NMR (100.6 MHz, D_2O)

All tested compounds had a purity of > 95%

Conditions

Column: Agilent Zorbax C8-RX 5μm 4.6 × 250mm

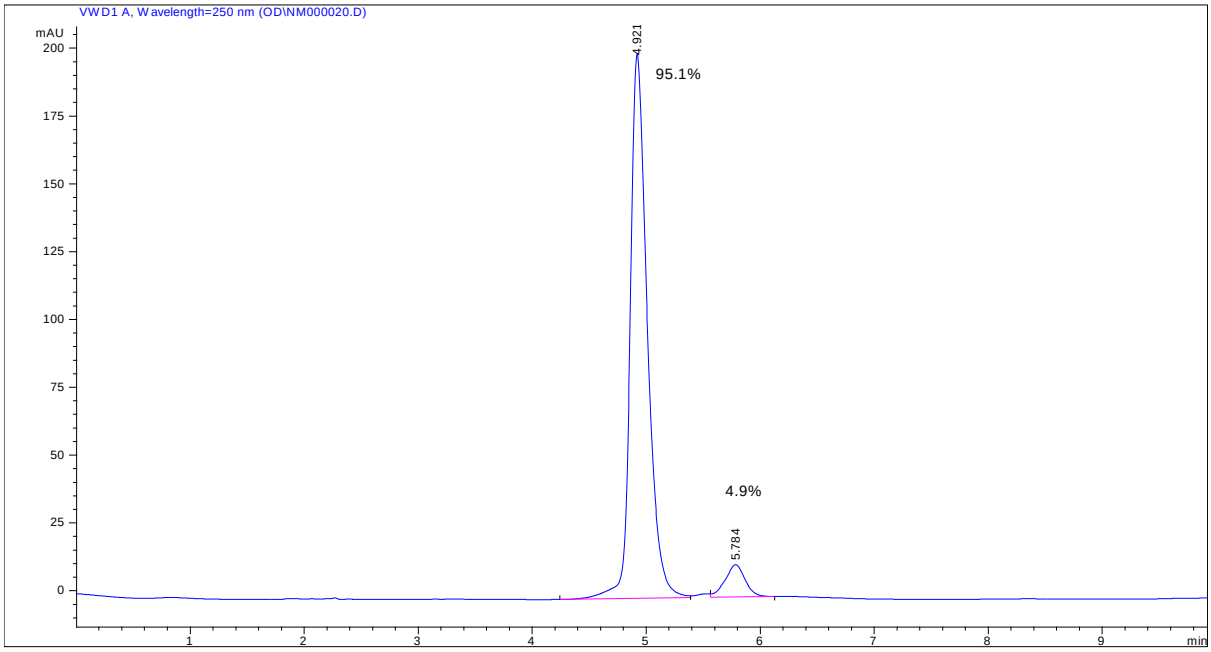
Mobile Phase: 10mmol triethylammonium formate : MeOH 98:2

Wavelength: 254 nm

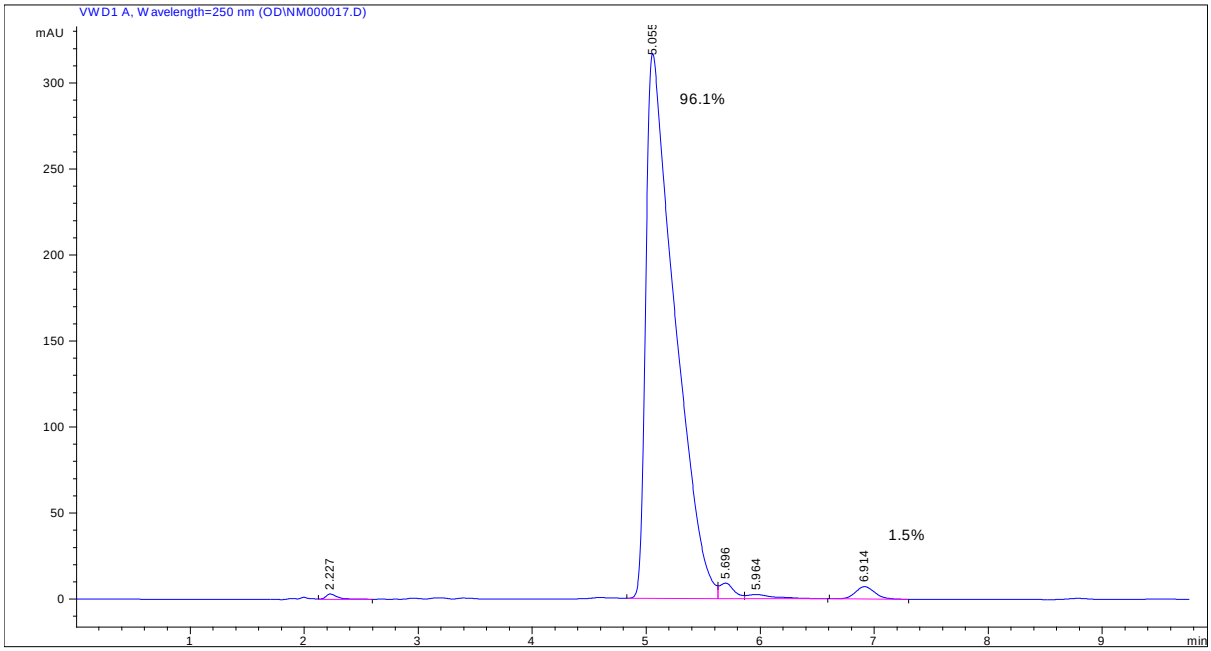
Injection volume: 10μL

Temperature: 25°C

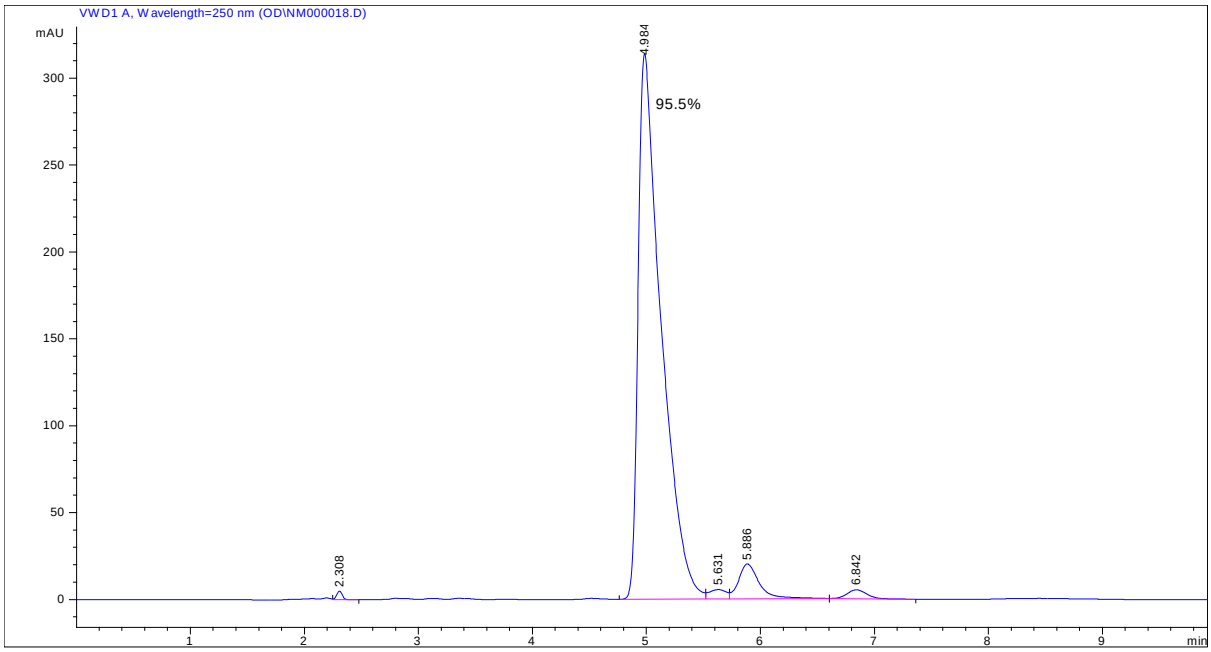
Compound **2**: 95.1%



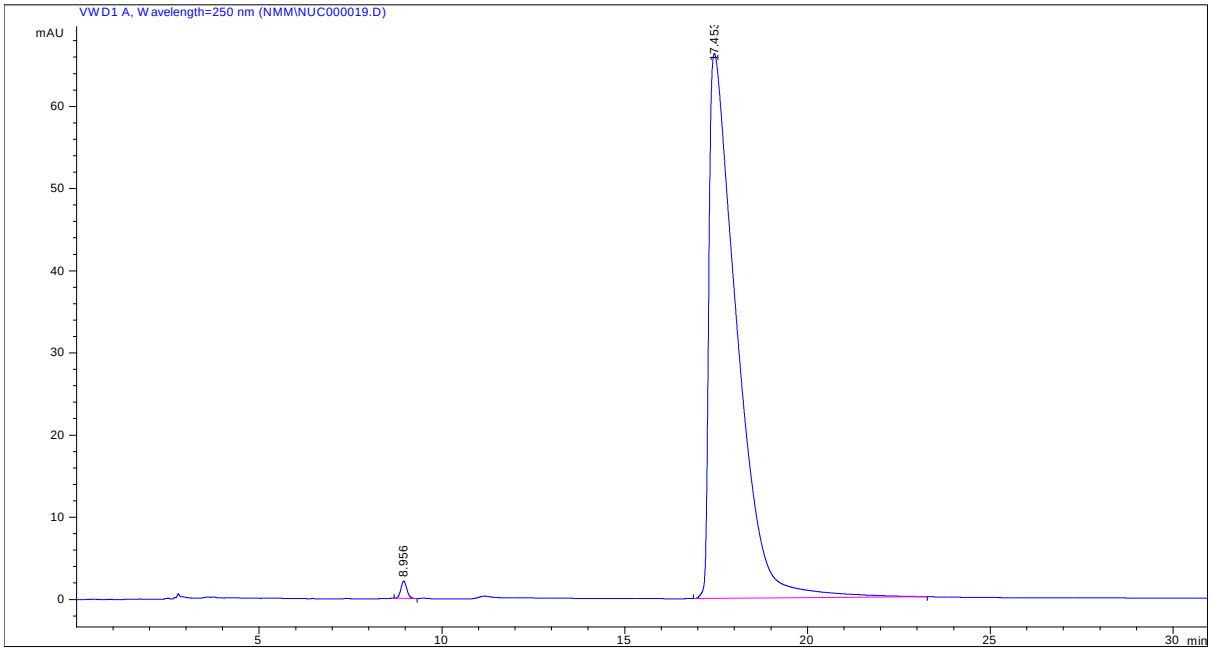
Compound **2a**: 96.1%



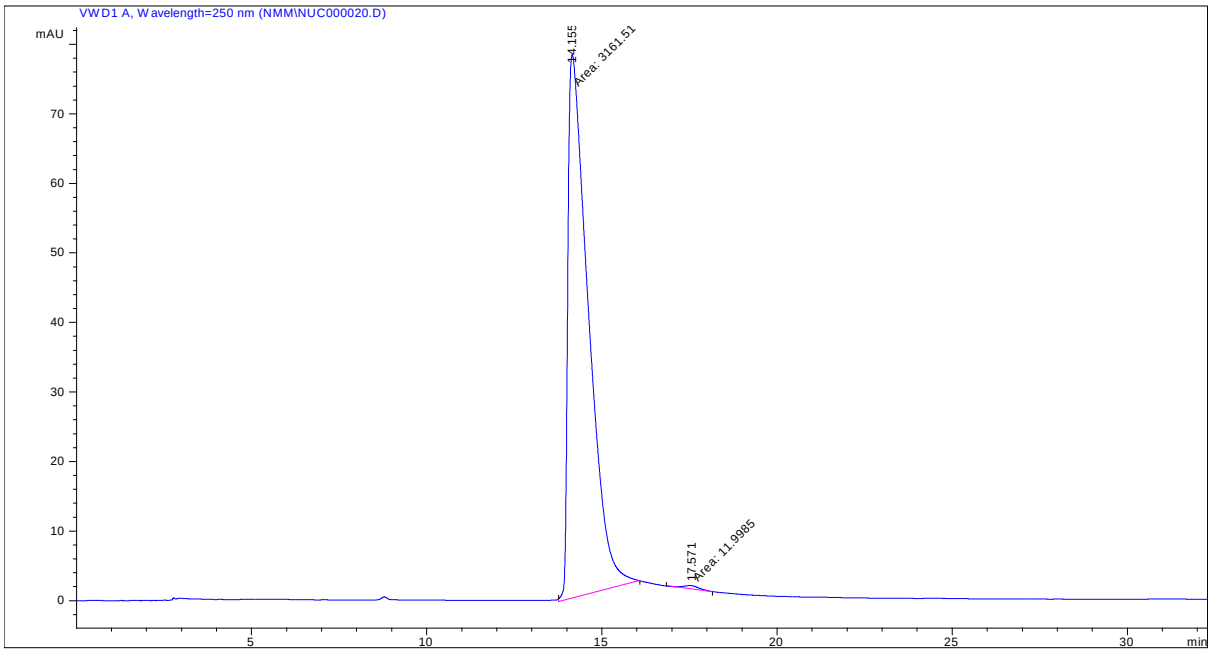
Compound **2b**: 95.5%



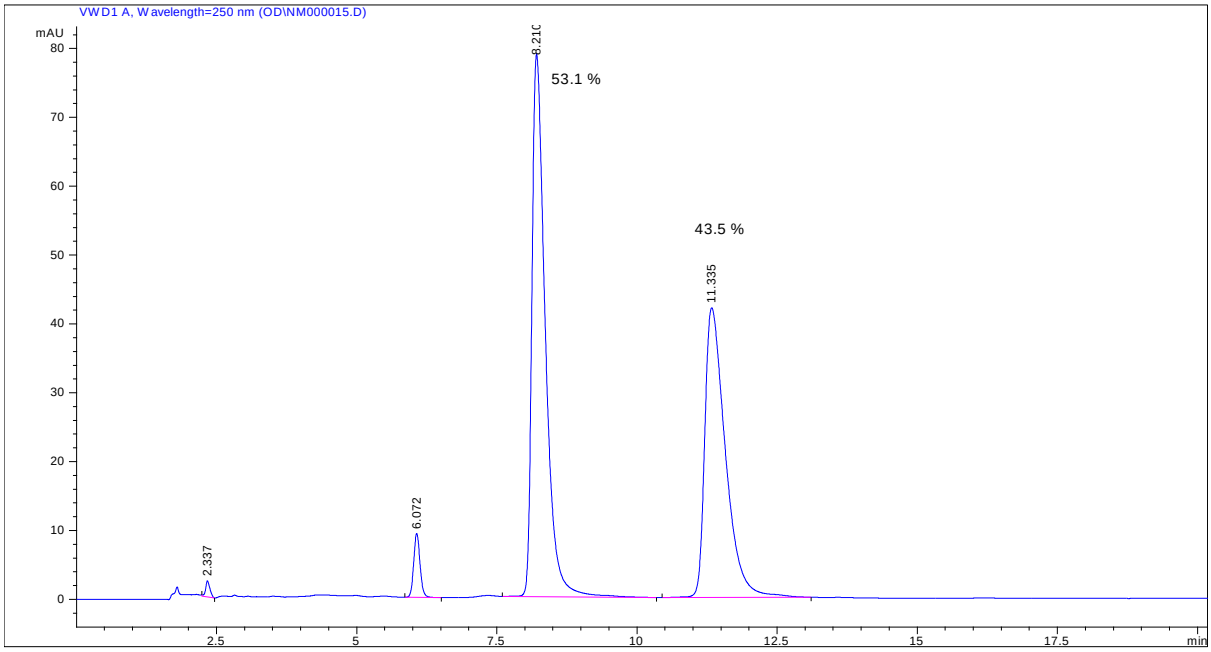
Compound **3**: 99.3%



Compound **4**: 99.6%



Compound **5**: Sum of diastereomers: 96.6%



SI Table 1: Crystallographic data processing and refinement statistics.

	RT/38-Apt DNA/compound **2a**
PDB Id	5HLF
Data collection Source	APS 23ID-D
Space Group	$P2_1$
Cell Parameters	90.06, 128.56, 132.41 Å; 90, 101.46, 90 º
Resolution (Highest Resolution Cell) Å	43 - 2.95(3.04 - 2.95)
R_{merge}	0.059 (0.884)
Number of Unique Reflections	52,130 (4,646)
Completeness	84.8 (87.5)
Multiplicity	2.1 (2.1)
I/sig(I)	9.1 (1.2)
Wilson B ($Å^2$)	84.9
Refinement Statistics	
Cutoff criteria	$F \leq 0$
Number of reflections (R-free set)	52,070 (2,565)
R_{work}/R_{free}	0.208/0.257
Number of atoms refined	17,442
Stereochemistry (RMSDs)	
Bond lengths (Å)	0.011
Bond angles (º)	1.434