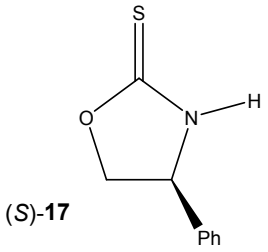
Supporting Information

Diastereoselective synthesis of the HIV protease inhibitor darunavir and related derivatives via a titanium tetrachloride mediated asymmetric glycolate aldol addition reaction

Jordan M. Witte, Emmanuel Ayim, Christopher J. Sams, Jasmine B. Service, Caitlyn C. Kant, Lillian Bambalas, Daniel Wright, Austin Carter, Kelly Moran, Isabella G. Rohrig, Gregory M. Ferrence, and Shawn R. Hitchcock*

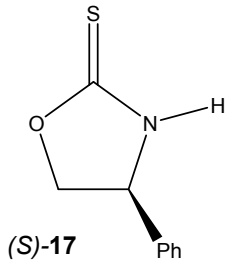
Department of Chemistry, Illinois State University, Normal, IL 617090-4160

Table of Contents



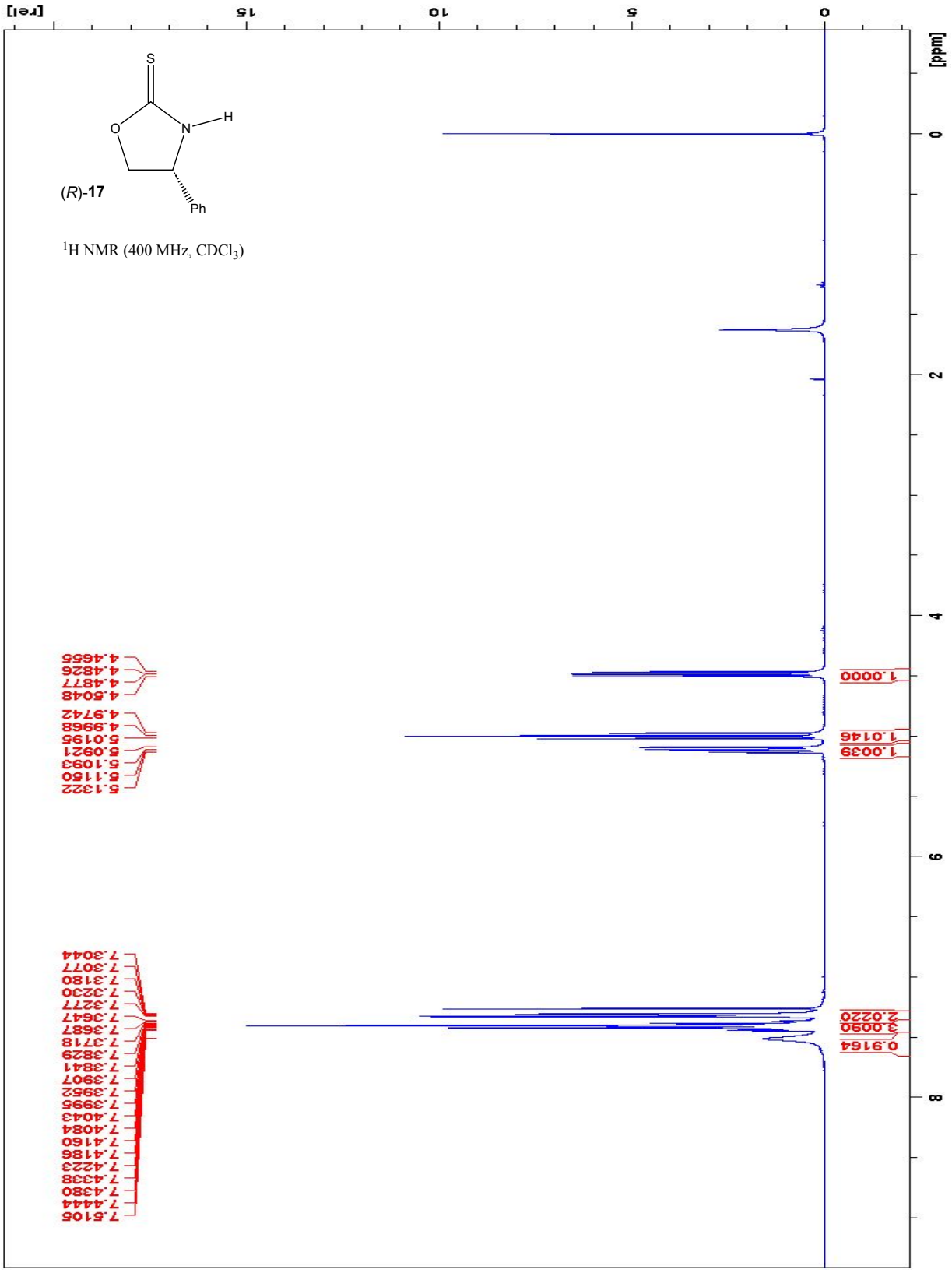
(*S*)-**17**

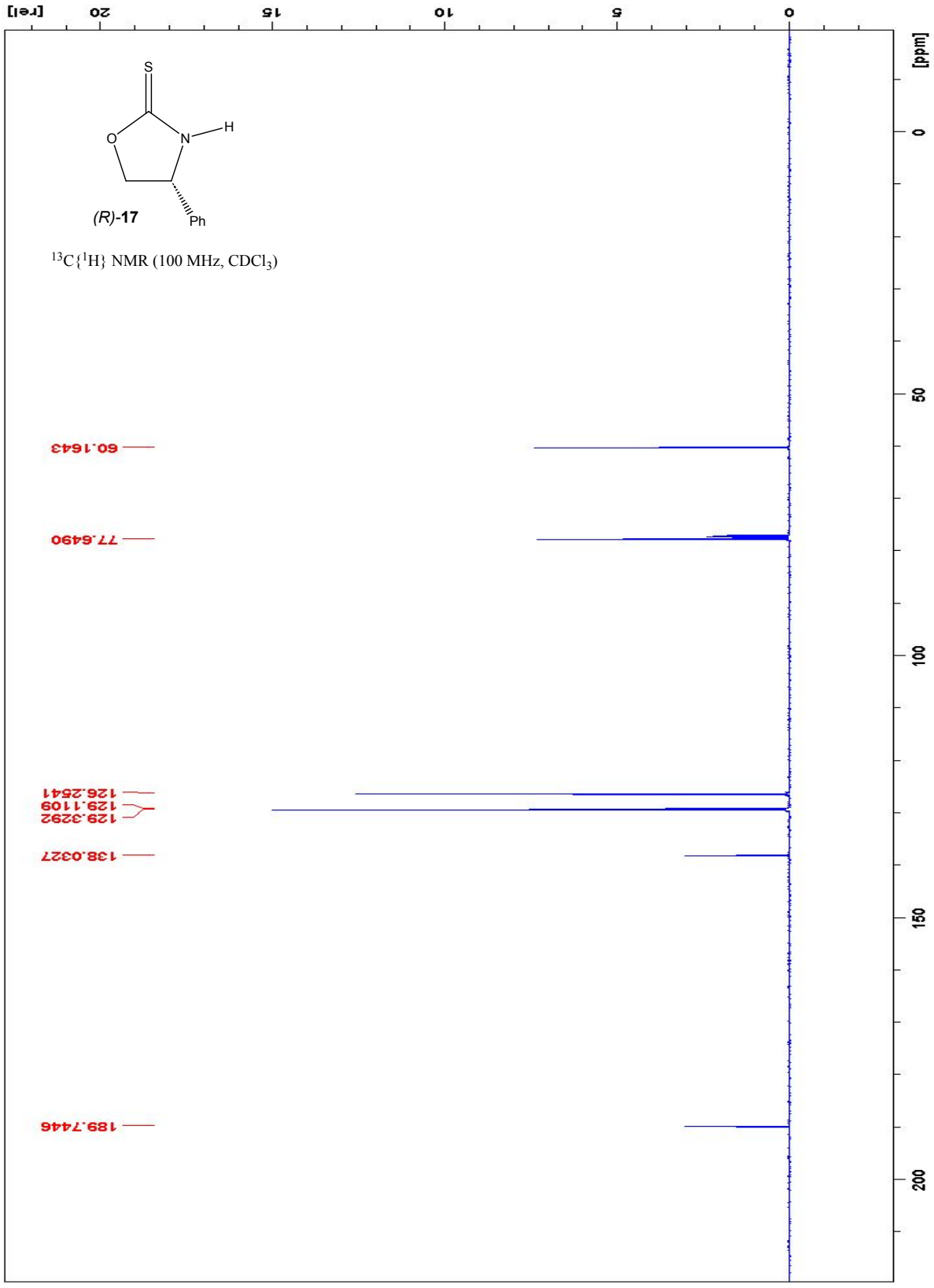
^{1}H NMR (500 MHz, $CDCl_3$)

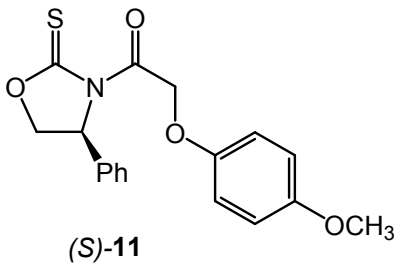


(S)-**17**

$^{13}C\{^{1}H\}$ NMR (125 MHz, $CDCl_3$)

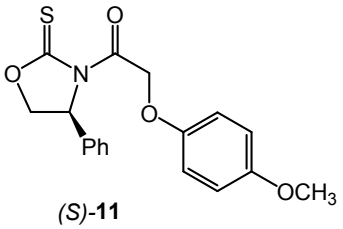

S
O
N
H
Ph
(R)-17
1H NMR (400 MHz, CDCl3)
[rel]
15
10
5
0
[ppm]
0
2
4
6
8
4.4655
4.4826
4.4877
4.5048
4.9742
4.9968
5.0195
5.0921
5.1093
5.1150
5.1322
7.3044
7.3077
7.3180
7.3230
7.3277
7.3647
7.3687
7.3718
7.3829
7.3841
7.3907
7.3962
7.3995
7.4043
7.4084
7.4160
7.4186
7.4223
7.4338
7.4380
7.4444
7.5105
1.0000
1.0146
1.0039
0.9164
3.0090
2.0220


S
O
N
H
Ph
(R)-17
13C{1H} NMR (100 MHz, CDCl3)
60.1643
77.6490
126.2541
129.1109
129.3292
138.0327
189.7446
[rel]
20
15
10
5
0
[ppm]
0
50
100
150
200



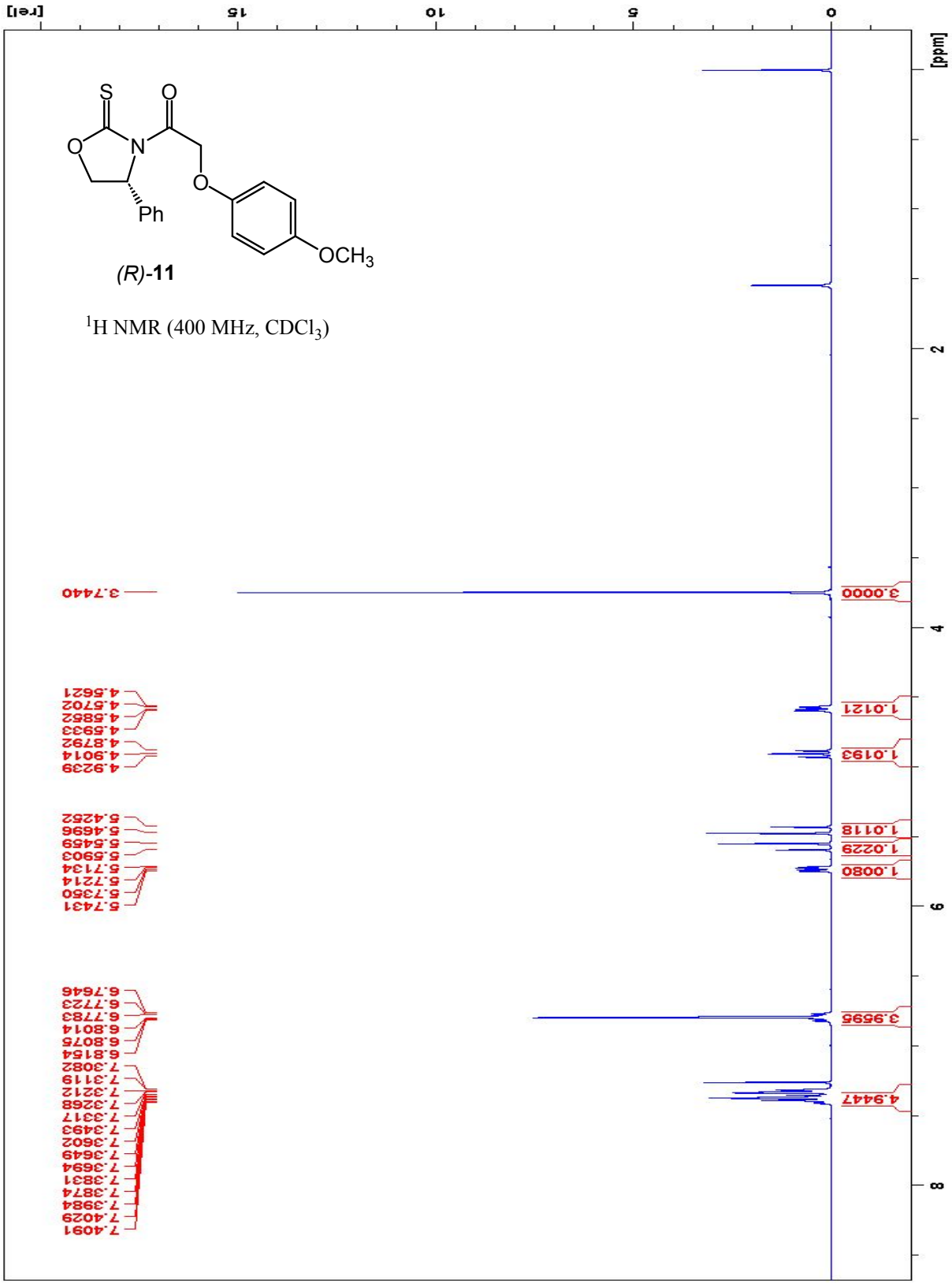
(S)-**11**

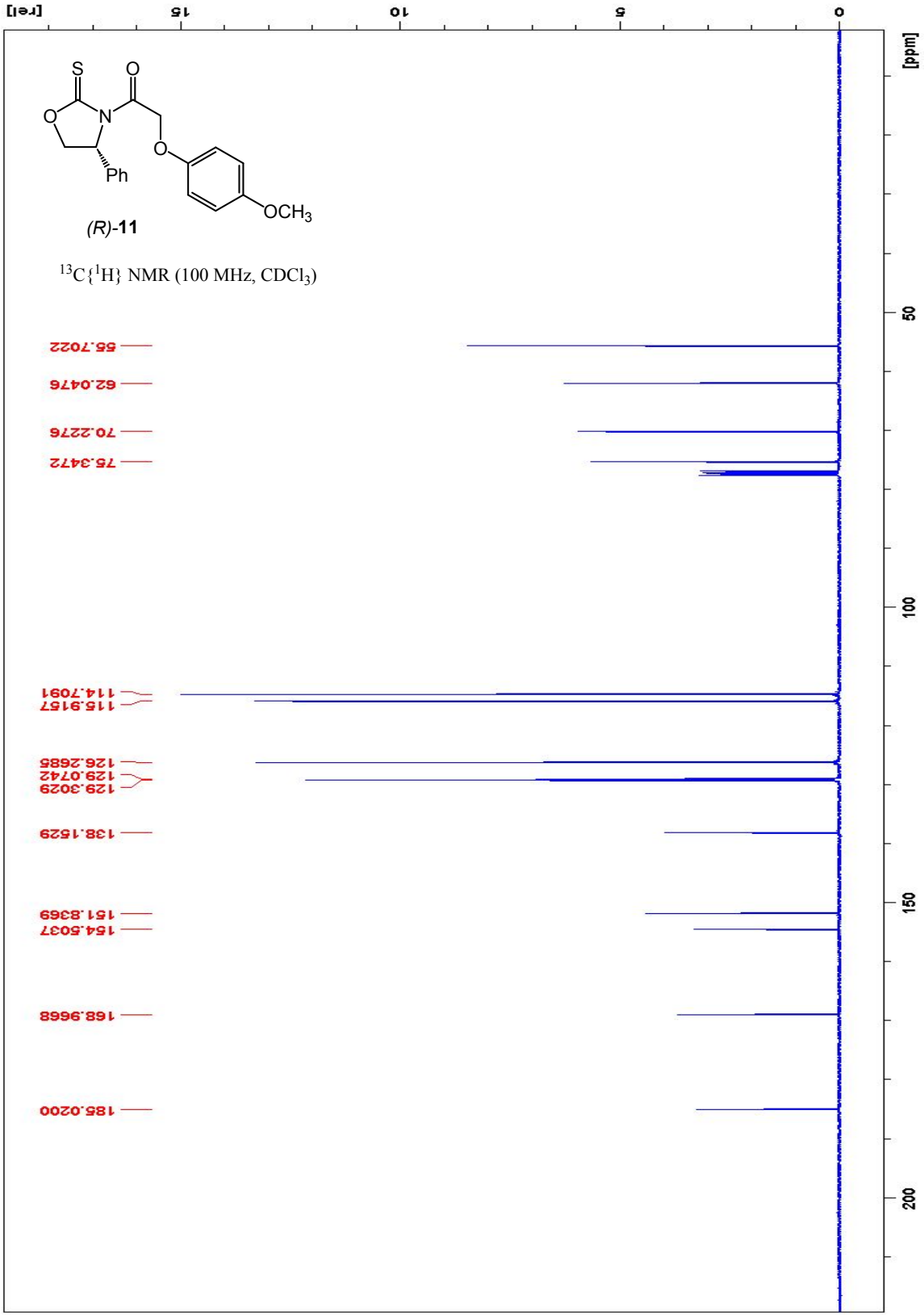
^{1}H NMR (400 MHz, $CDCl_3$)

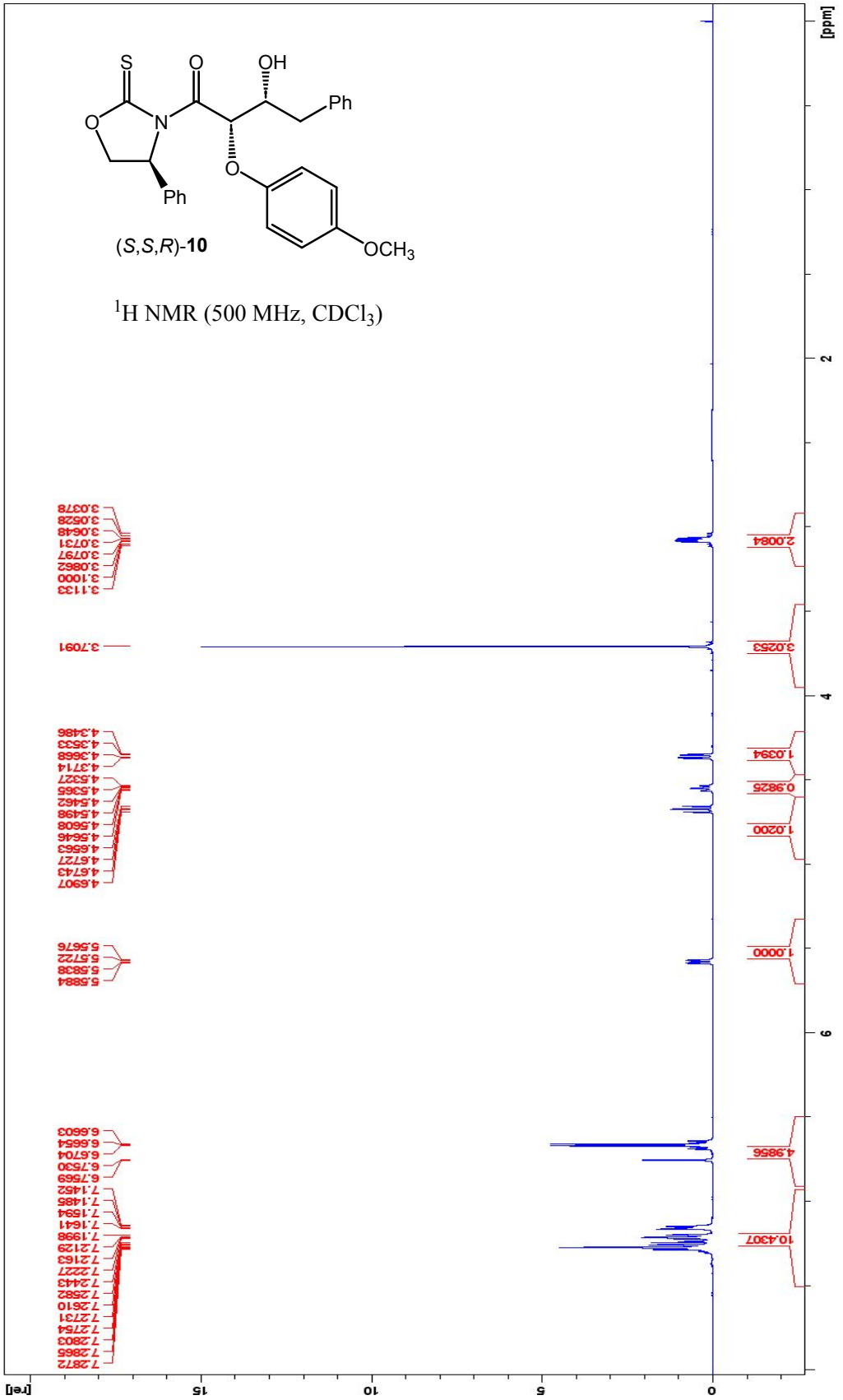


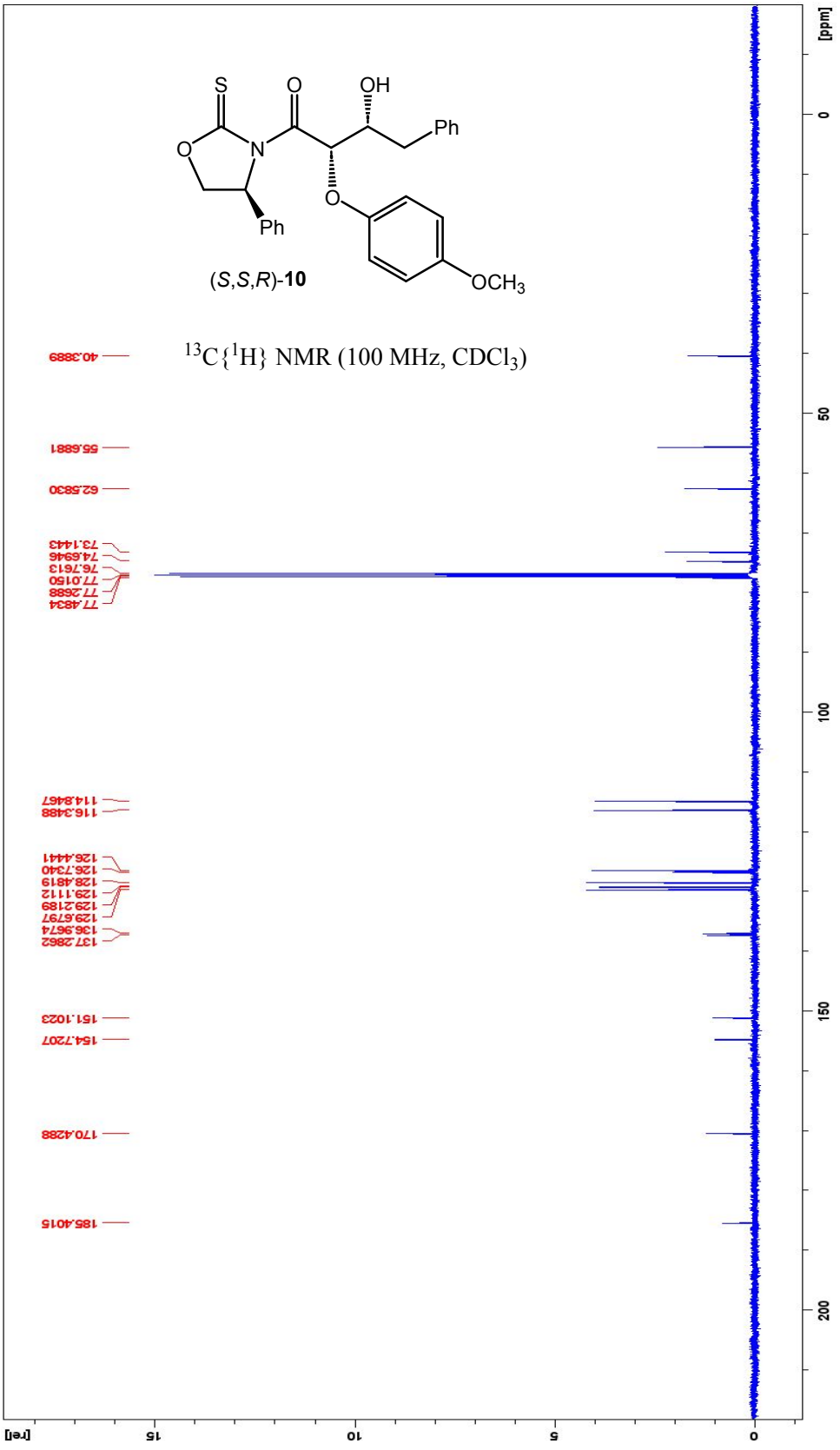
(S)-**11**

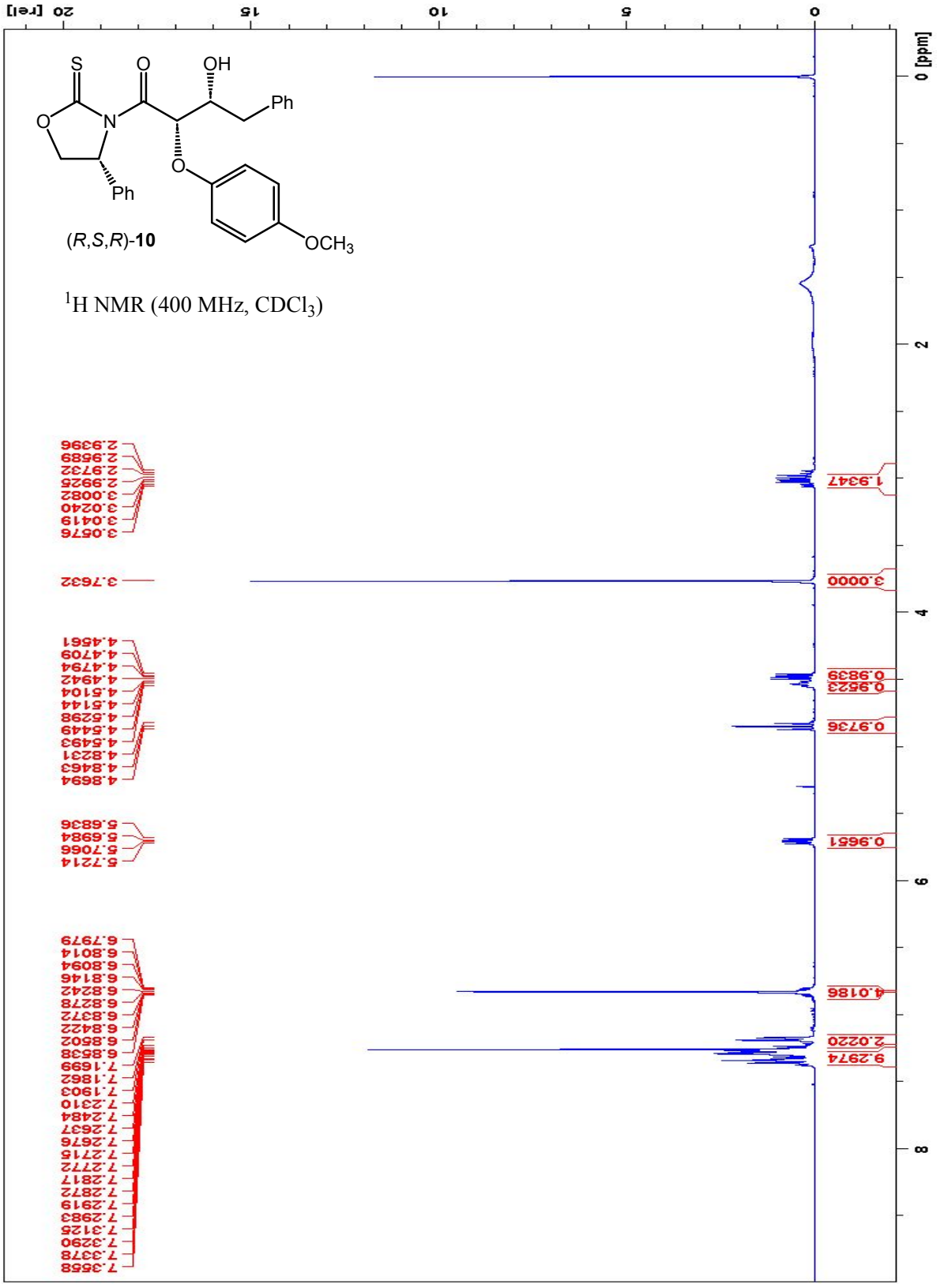
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$)

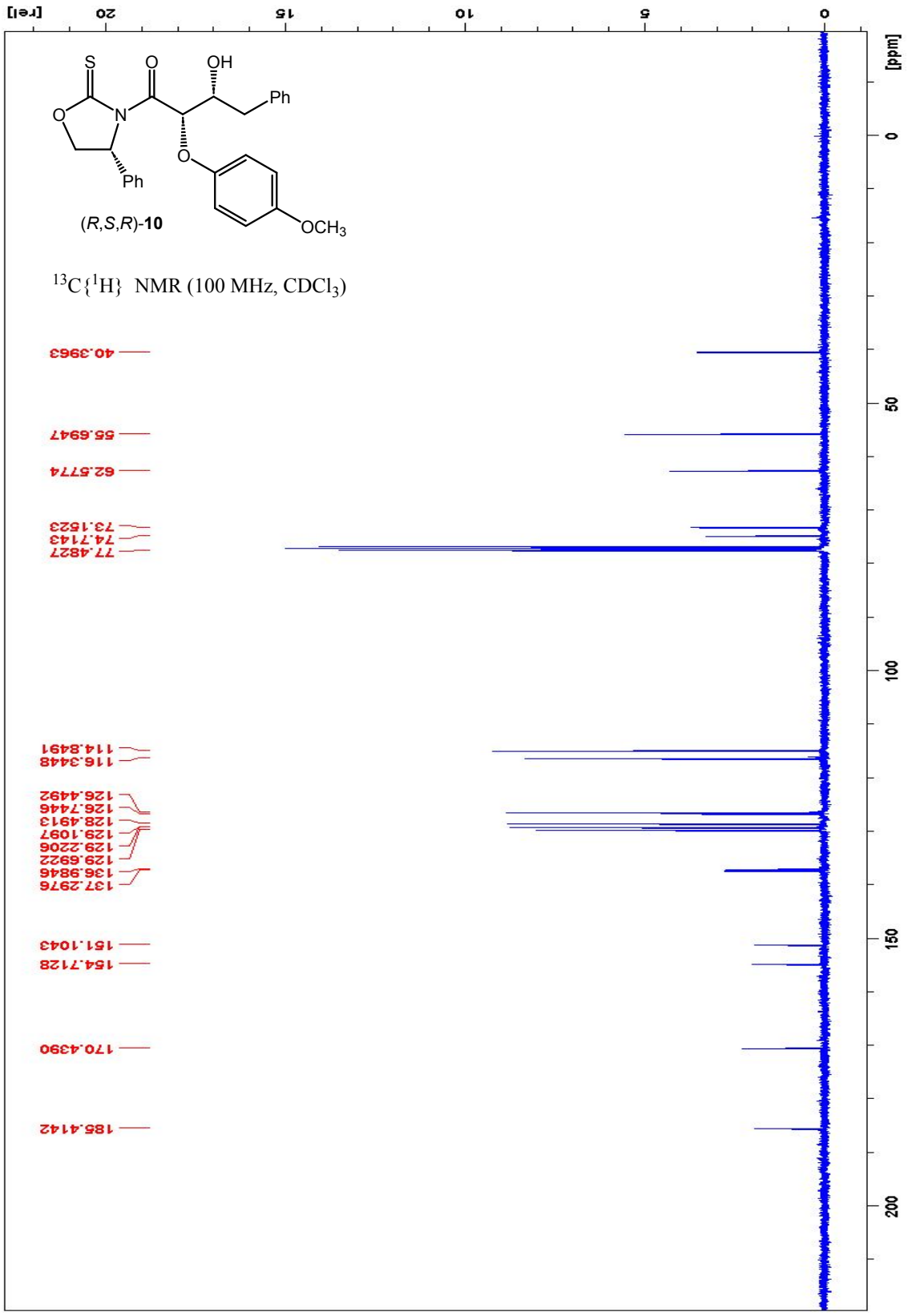

S
O
O
N
O
Ph
OCH3
(R)-11
1H NMR (400 MHz, CDCl3)
[rel]
15
10
5
0
[ppm]
2
4
6
8
3.7440
4.5621
4.5702
4.5862
4.5933
4.8792
4.9014
4.9239
5.4252
5.4696
5.5459
5.5903
5.7134
5.7214
5.7350
5.7431
6.7646
6.7723
6.7783
6.8014
6.8075
6.8154
7.3082
7.3119
7.3212
7.3268
7.3317
7.3493
7.3602
7.3649
7.3694
7.3831
7.3874
7.3984
7.4029
7.4091
3.0000
1.0121
1.0193
1.0118
1.0229
1.0080
3.9595
4.9447

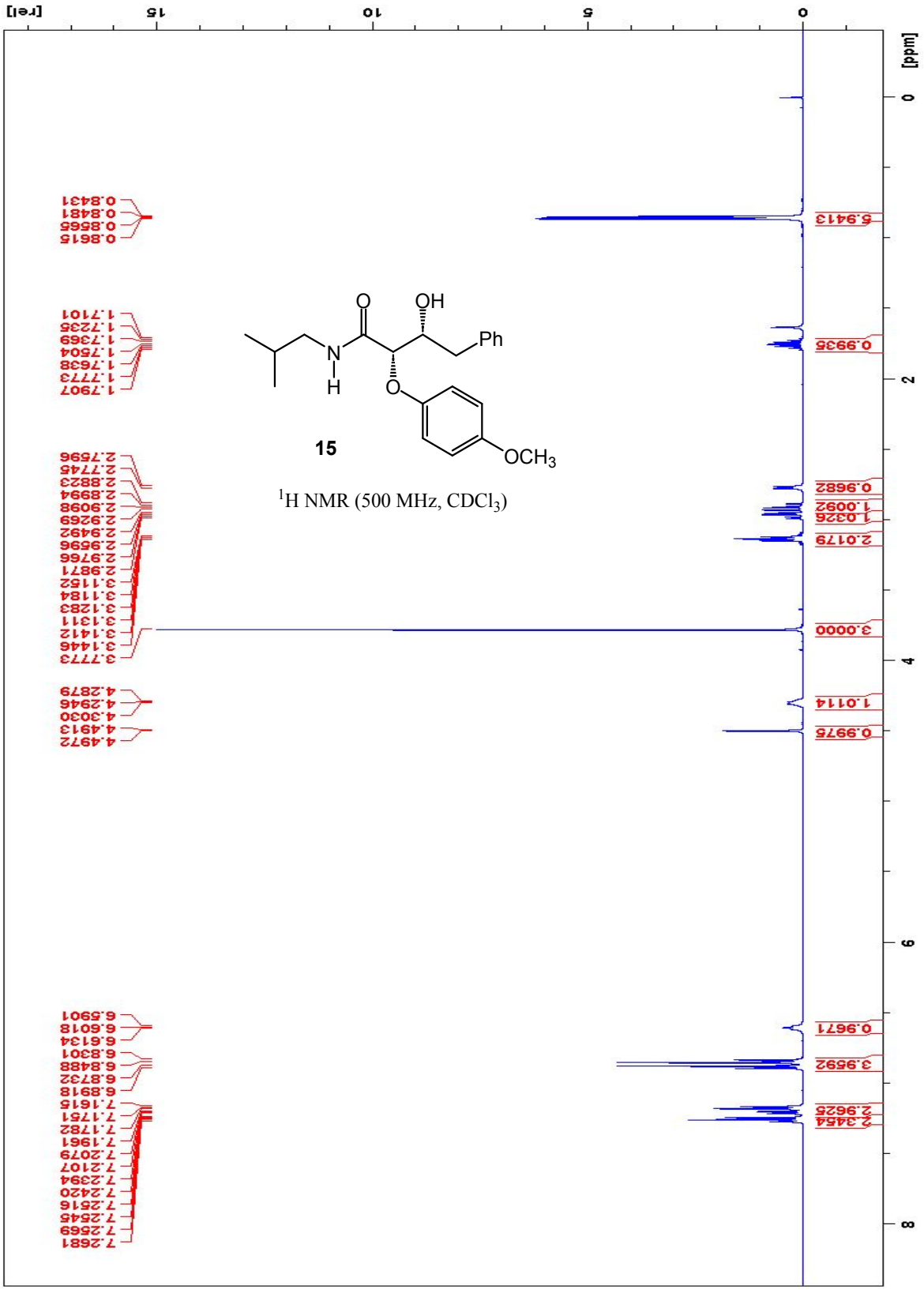

S
O
O
N
O
Ph
OCH3
(R)-11
13C{1H} NMR (100 MHz, CDCl3)
55.7022
62.0476
70.2276
75.3472
114.7091
115.9157
126.2685
129.0742
129.3029
138.1529
151.8369
154.5037
168.9668
185.0200
[rel]
15
10
5
0
[ppm]
50
100
150
200


S
O
OH
Ph
O
N
O
Ph
OCH3
(S,S,R)-10
1H NMR (500 MHz, CDCl3)
[ppm]
[rel]

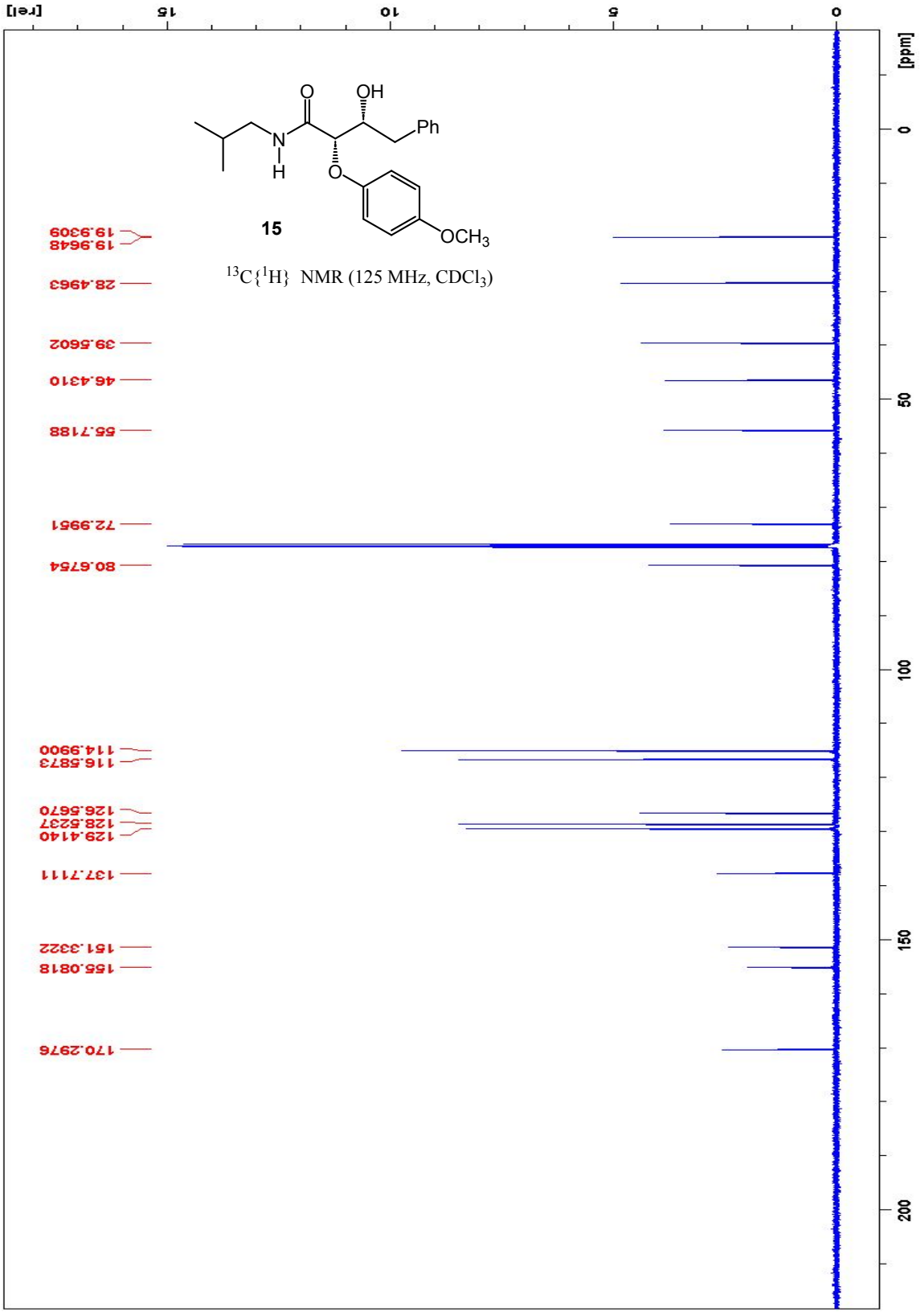

S
O
OH
Ph
O
N
O
Ph
OCH3
(S,S,R)-10
13C{1H} NMR (100 MHz, CDCl3)
40.3889
55.6881
62.5830
73.1443
74.6946
76.7613
77.0150
77.2688
77.4834
114.8467
116.3488
126.4441
126.7340
128.4819
129.1112
129.2189
129.6797
136.9674
137.2862
151.1023
154.7207
170.4288
185.4015
[rel]
15
10
5
0
[ppm]
0
50
100
150
200


S
O
OH
Ph
O
N
O
Ph
OCH3
(R,S,R)-10
1H NMR (400 MHz, CDCl3)
[rel]
0 [ppm]
1.9347
3.0000
0.9839
0.9823
0.9736
0.9651
4.0186
2.0220
9.2974

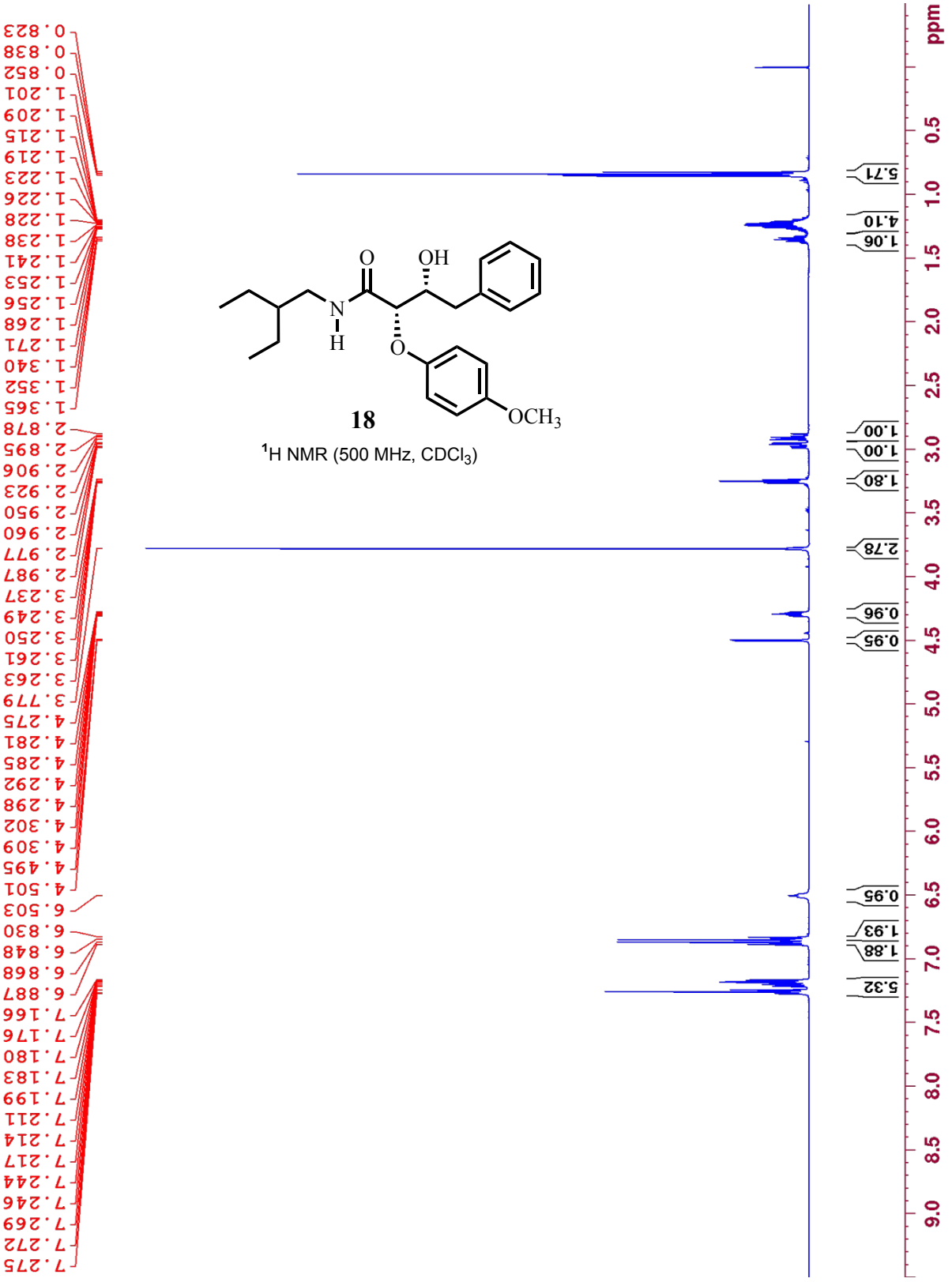

S
O
OH
Ph
O
N
O
Ph
OCH3
(R,S,R)-10
13C{1H} NMR (100 MHz, CDCl3)
40.3963
55.6947
62.5774
73.1623
74.7143
77.4827
114.8491
116.3448
126.4492
126.7446
128.4913
129.1097
129.2206
129.6922
136.9846
137.2976
151.1043
154.7128
170.4390
185.4142
[rel]
[ppm]

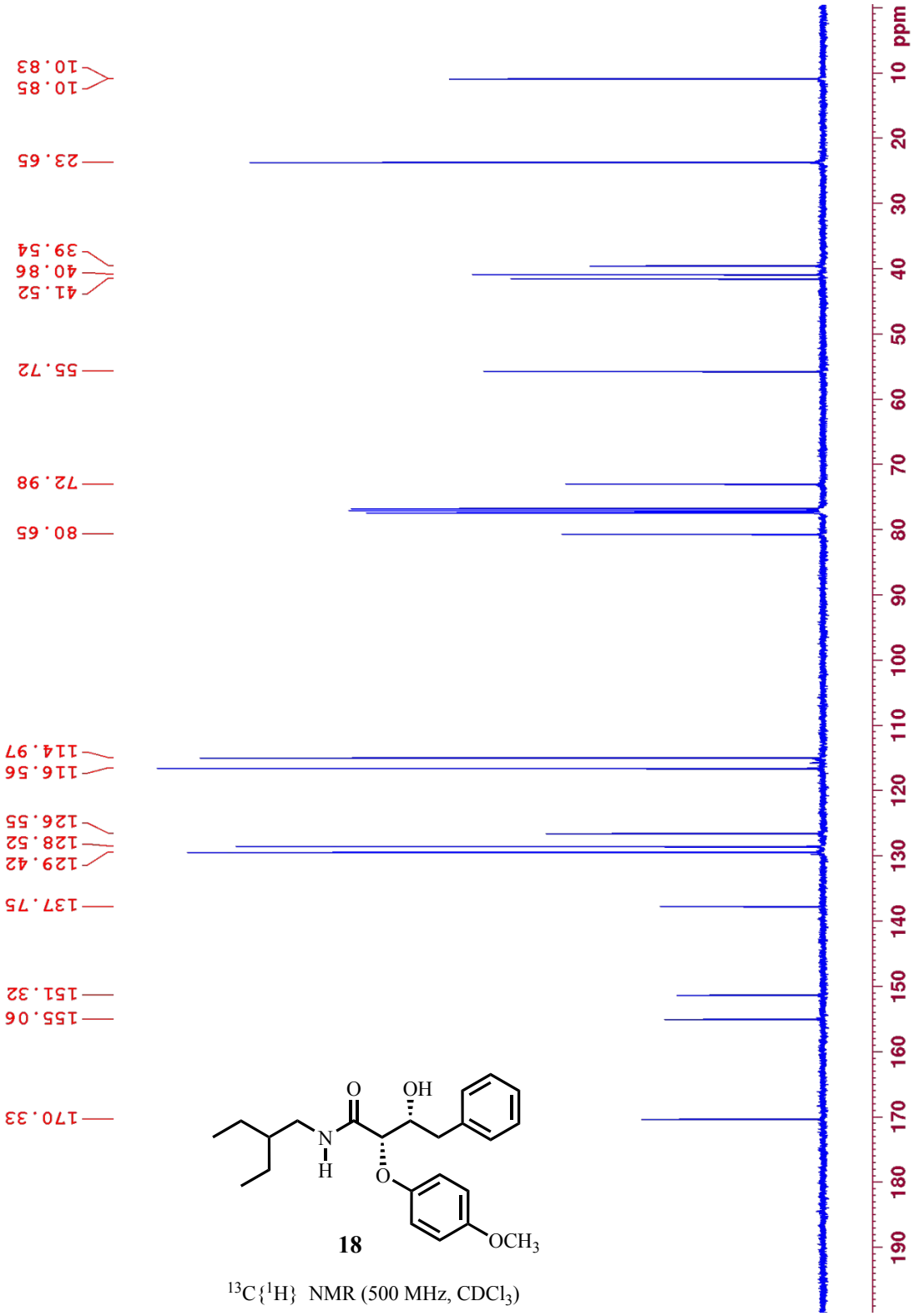


^{1}H NMR (500 MHz, $CDCl_3$)



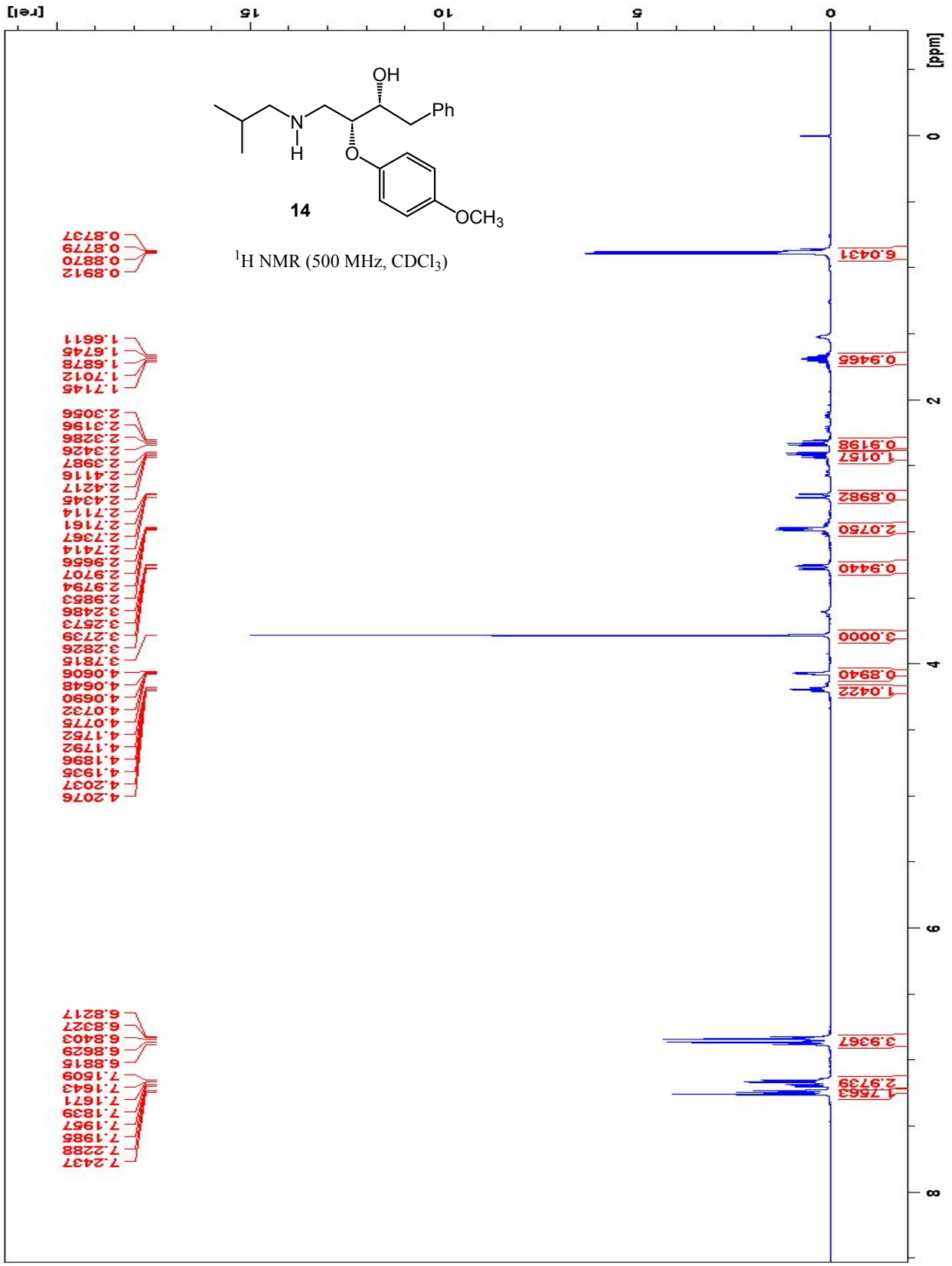
$^{13}C\{^{1}H\}$ NMR (125 MHz, $CDCl_3$)


0.823
0.838
0.852
1.201
1.209
1.215
1.219
1.223
1.226
1.228
1.238
1.241
1.253
1.256
1.268
1.271
1.340
1.352
1.365
2.878
2.895
2.906
2.923
2.950
2.960
2.977
2.987
3.237
3.249
3.250
3.261
3.263
3.779
4.275
4.281
4.285
4.292
4.298
4.302
4.309
4.495
4.501
6.503
6.830
6.848
6.868
6.887
7.166
7.176
7.180
7.183
7.199
7.211
7.214
7.217
7.244
7.246
7.269
7.272
7.275
O
OH
N
H
O
OCH3
18
1H NMR (500 MHz, CDCl3)
5.71
4.10
1.06
1.00
1.00
1.80
2.78
0.96
0.95
0.95
1.93
1.88
5.32
ppm
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0

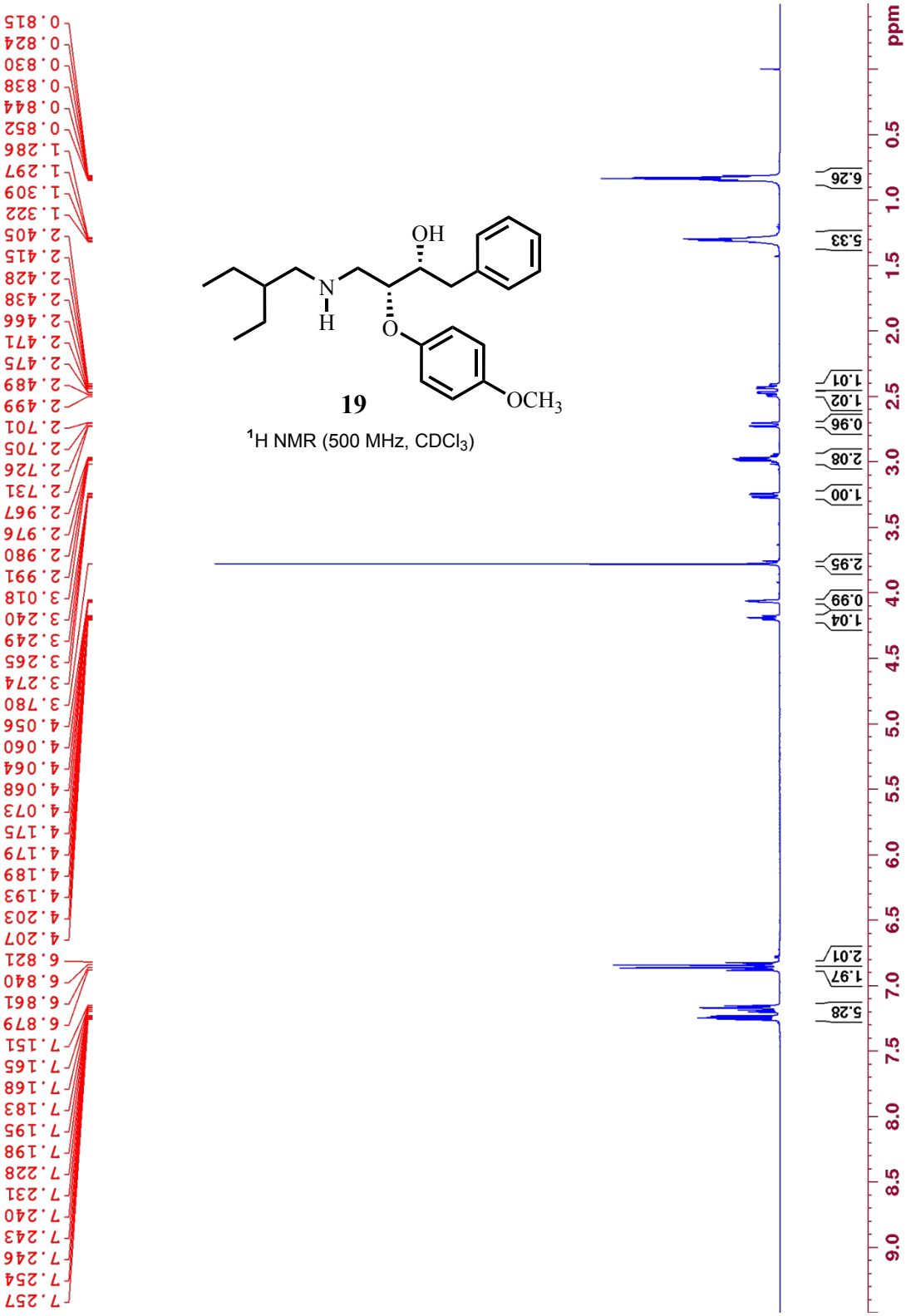


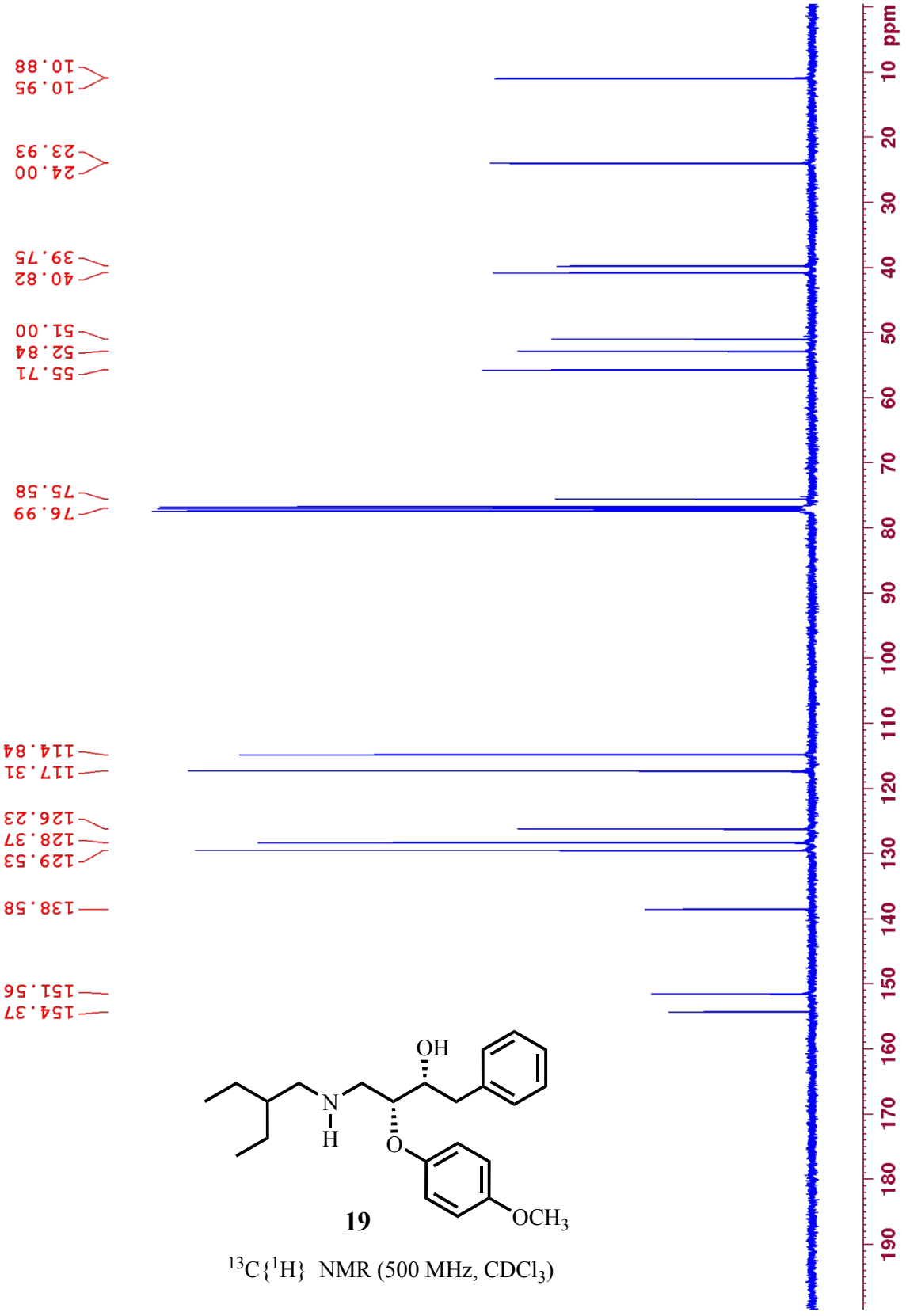
18

$^{13}C\{^{1}H\}$ NMR (500 MHz, $CDCl_3$)


OH
Ph
N
H
O
OCH3
14
[1H NMR (500 MHz, CDCl3)]
[rel]
[ppm]

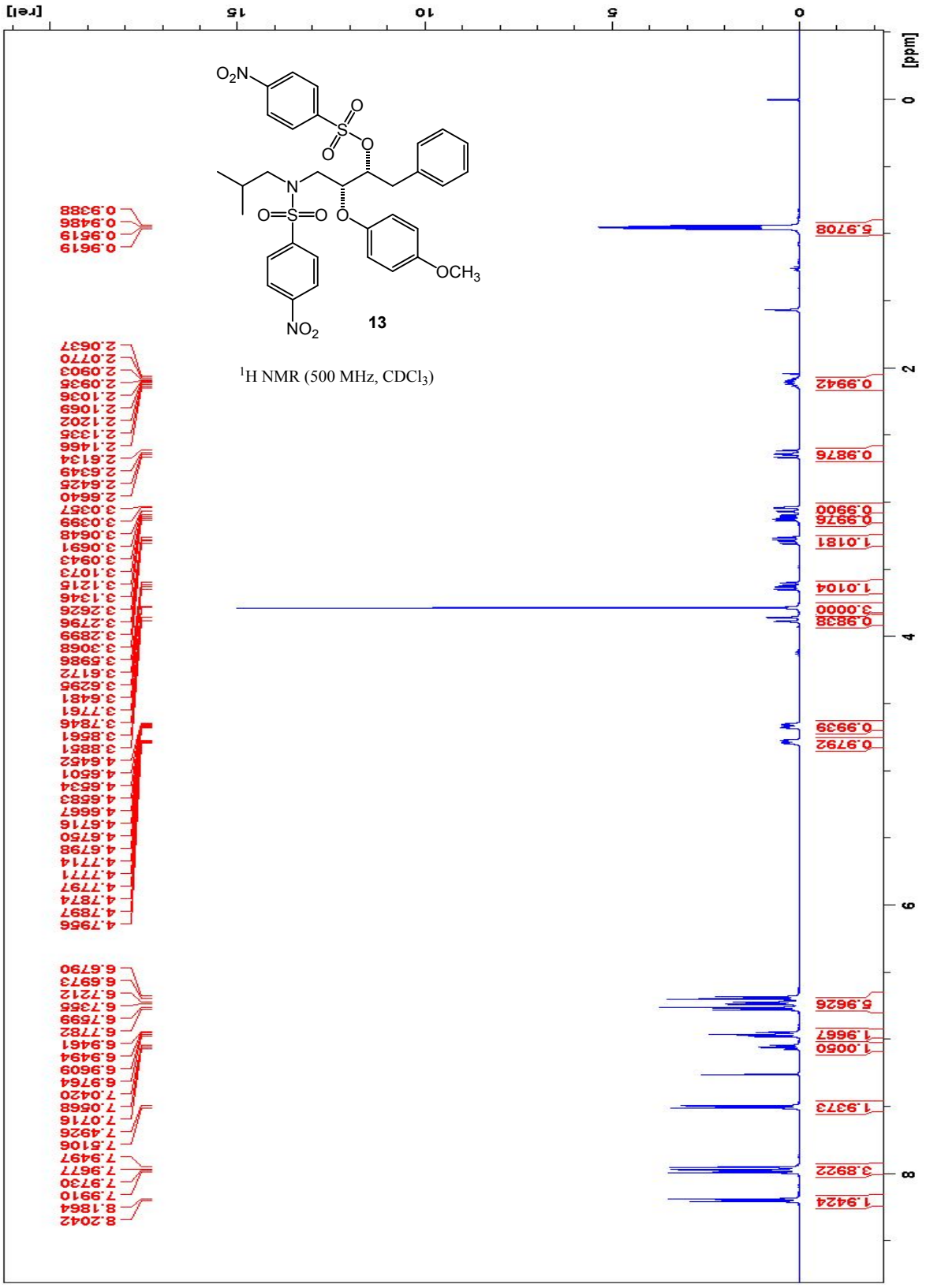
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$)


OH
N
H
O
OCH3
19
1H NMR (500 MHz, CDCl3)
ppm

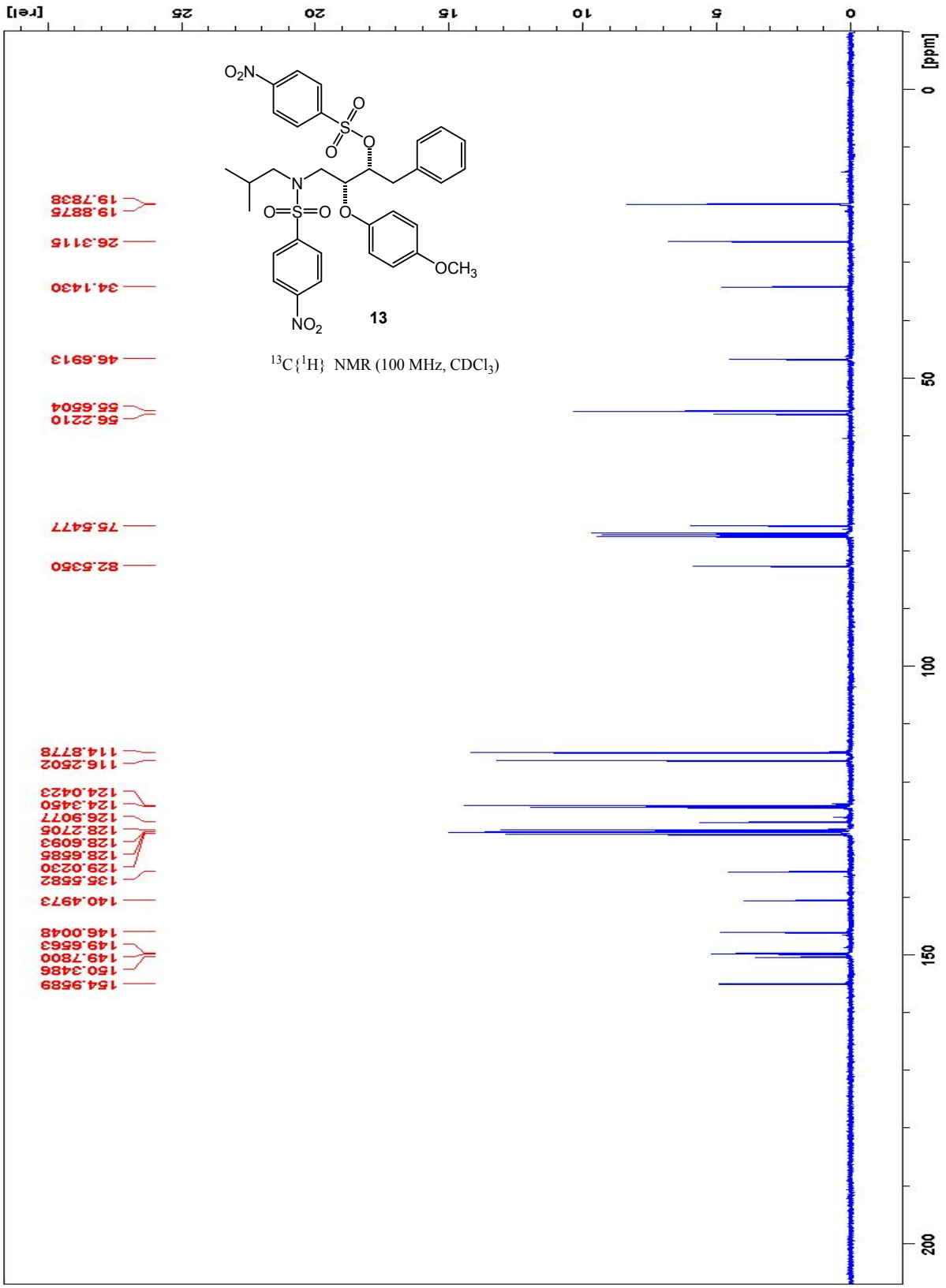


19

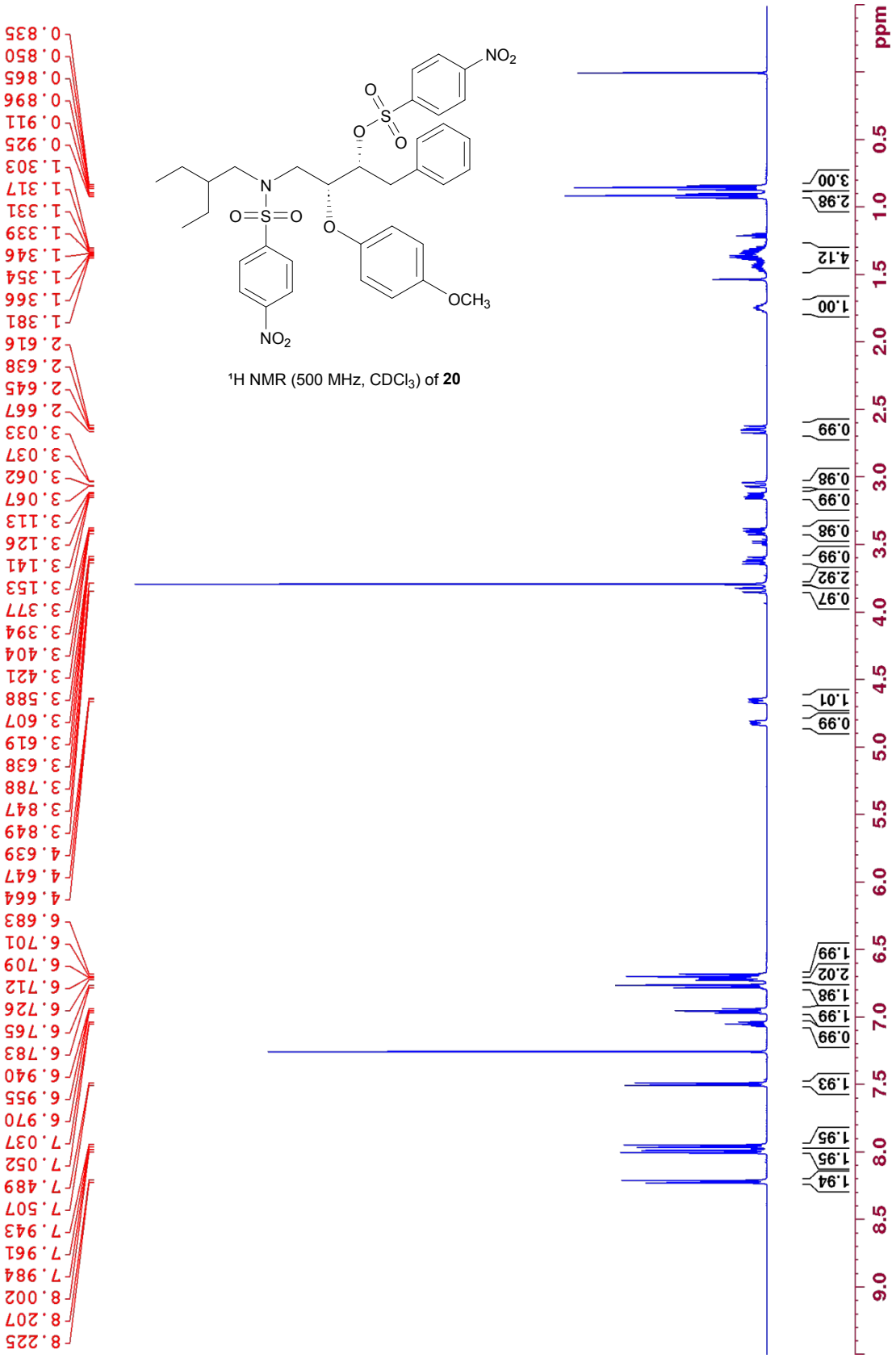
$^{13}C\{^{1}H\}$ NMR (500 MHz, $CDCl_3$)



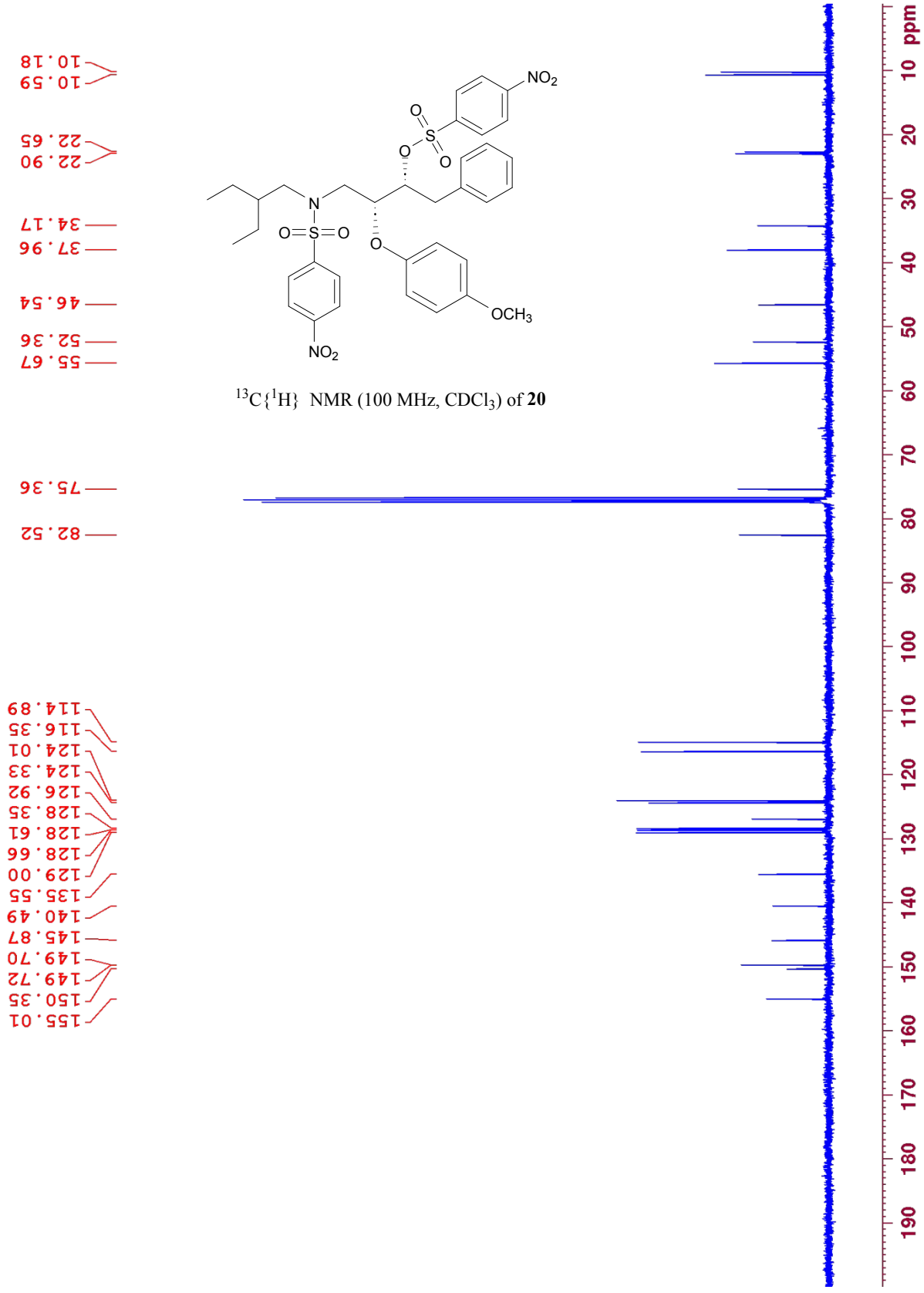
^{1}H NMR (500 MHz, $CDCl_3$)



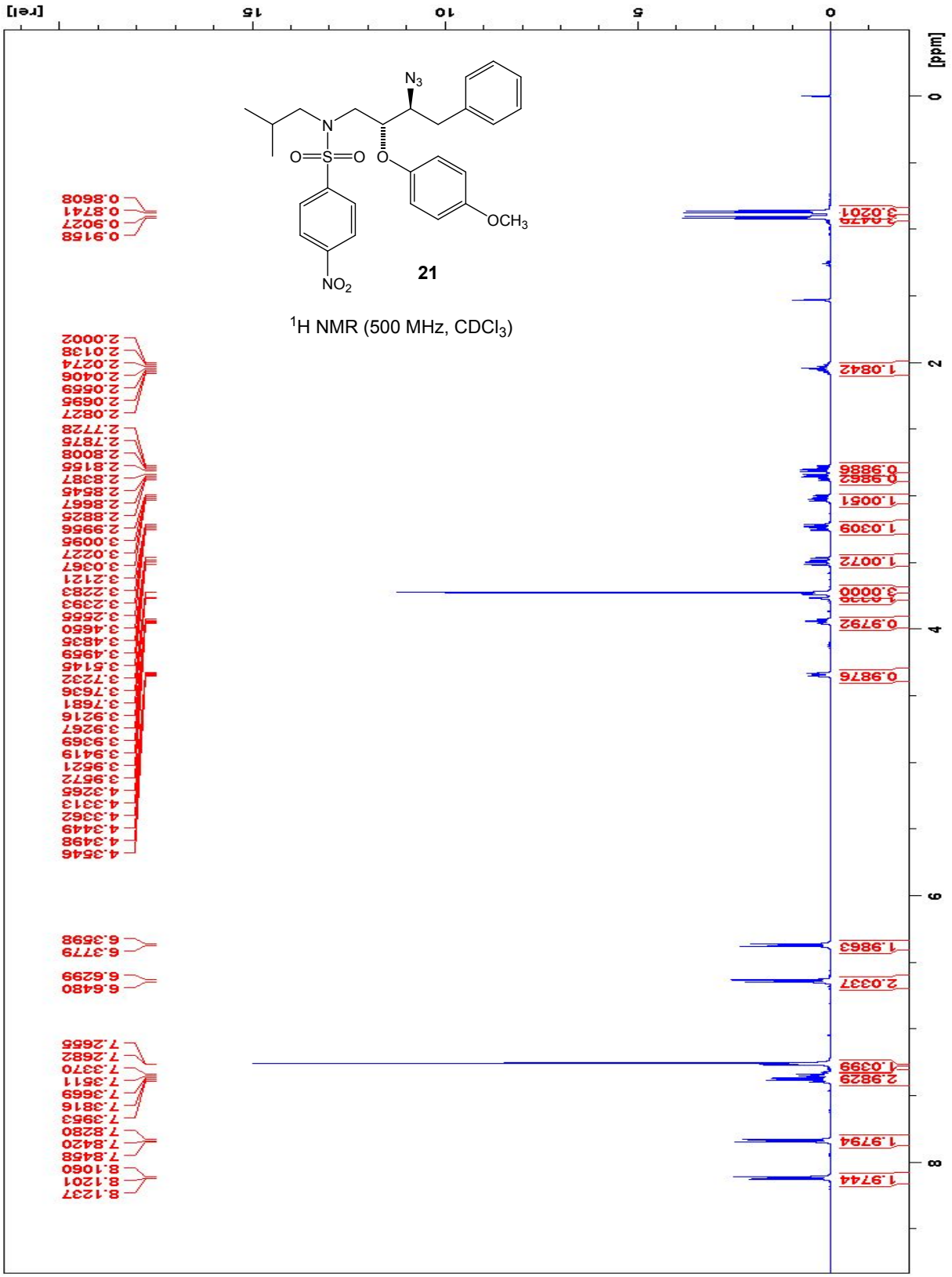
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$)

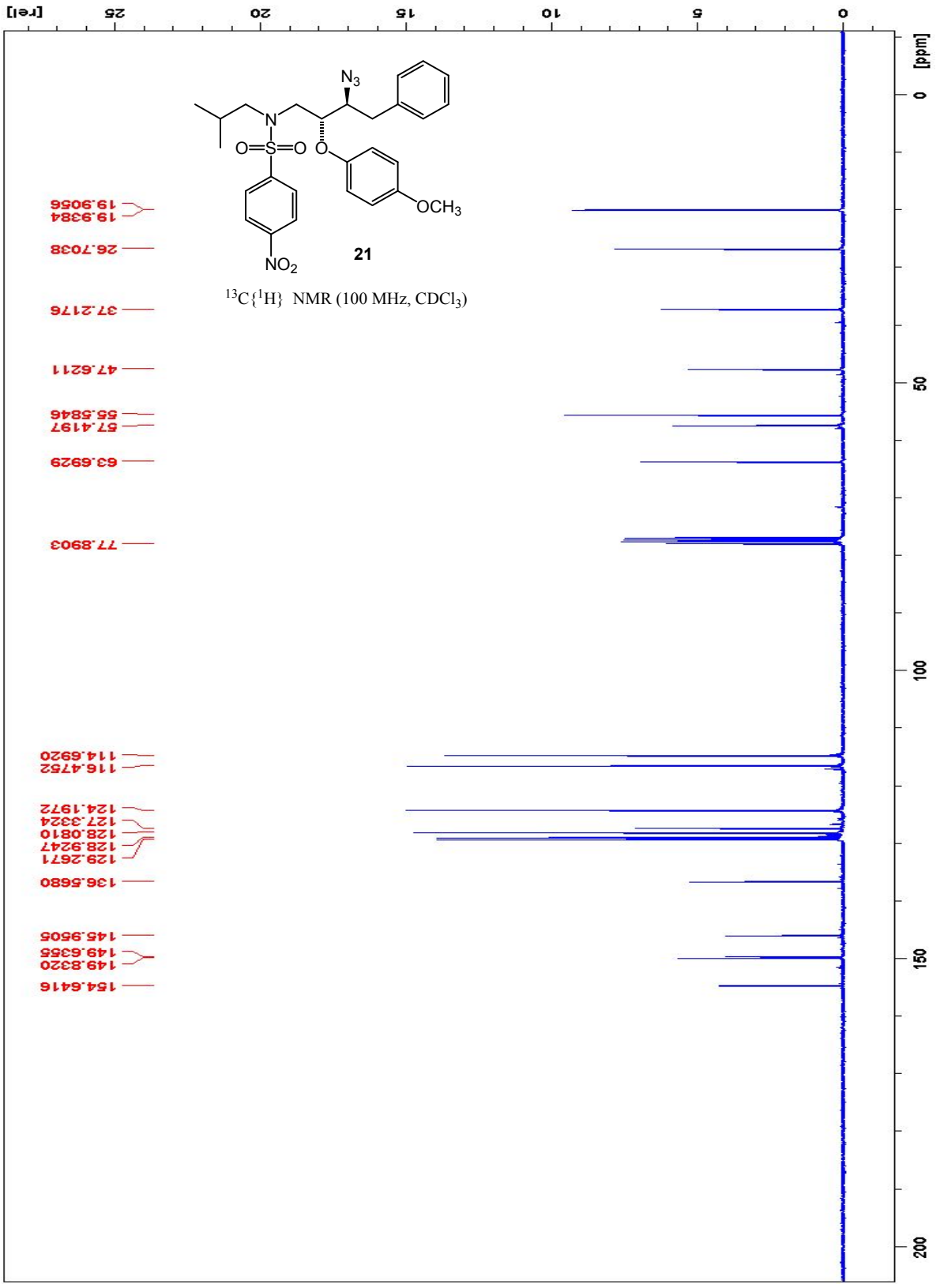


^{1}H NMR (500 MHz, $CDCl_3$) of **20**

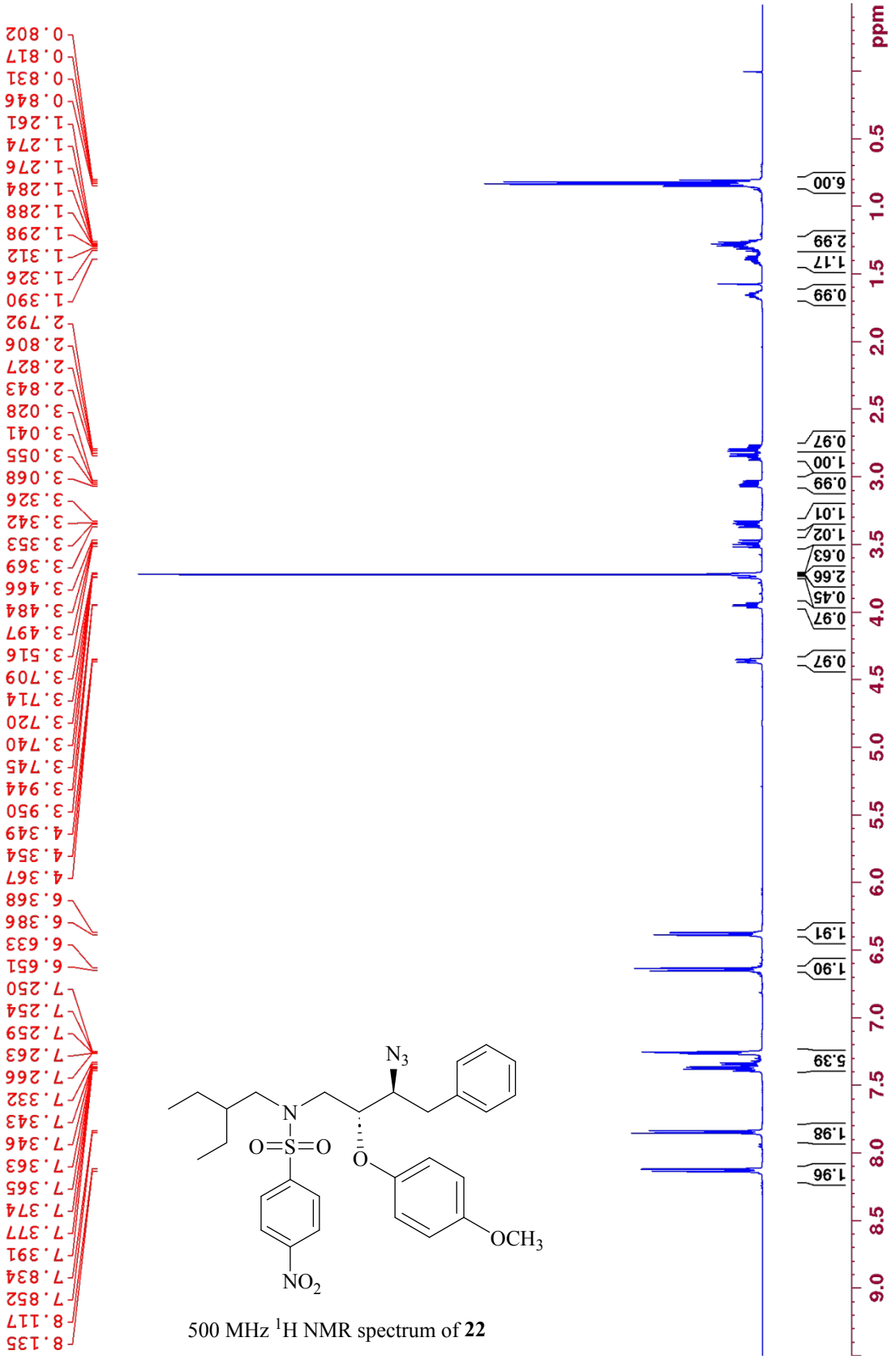


$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of **20**

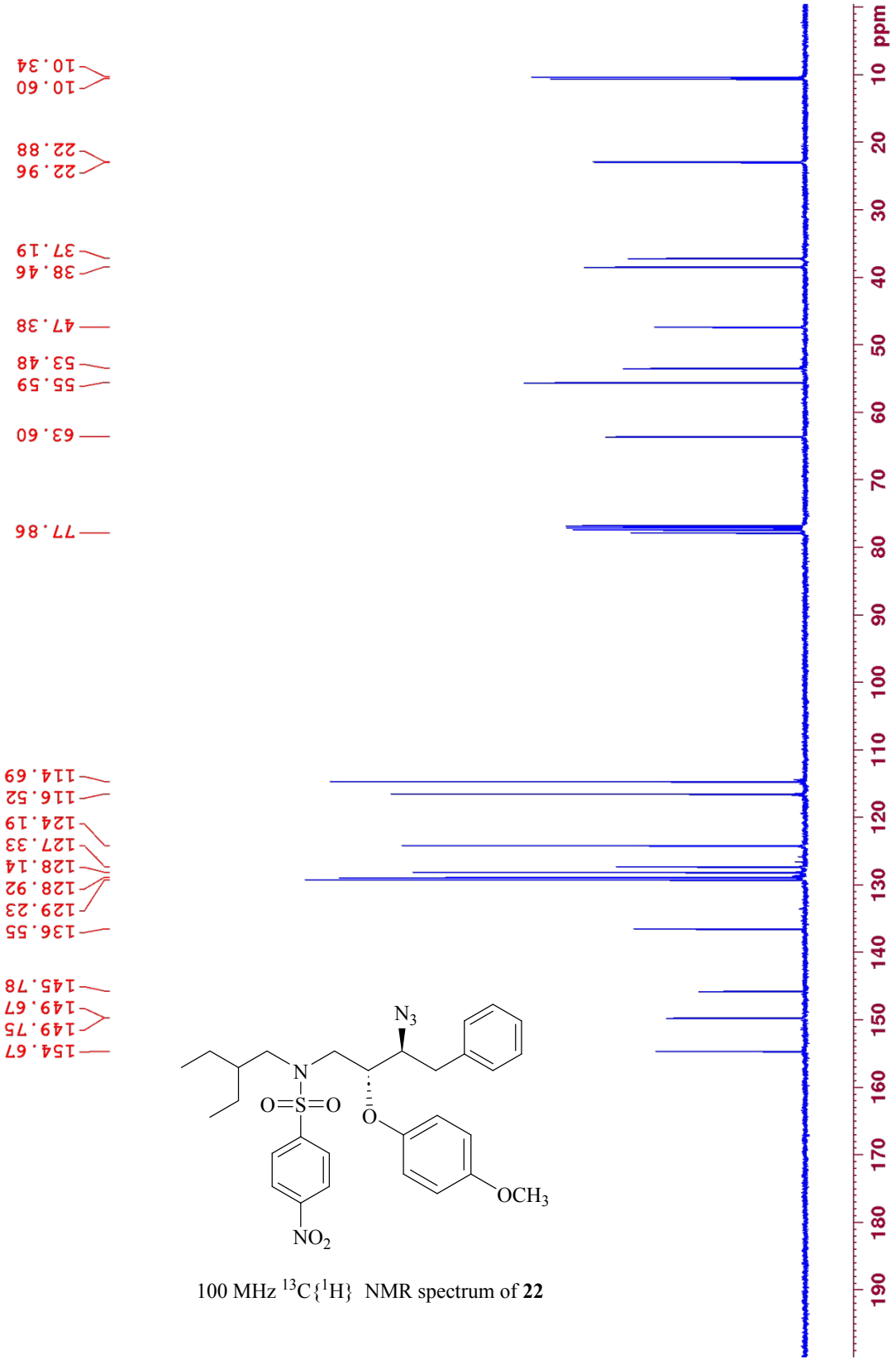

N3
N
O=S=O
O
OCH3
NO2
21
1H NMR (500 MHz, CDCl3)
[rel]
[ppm]



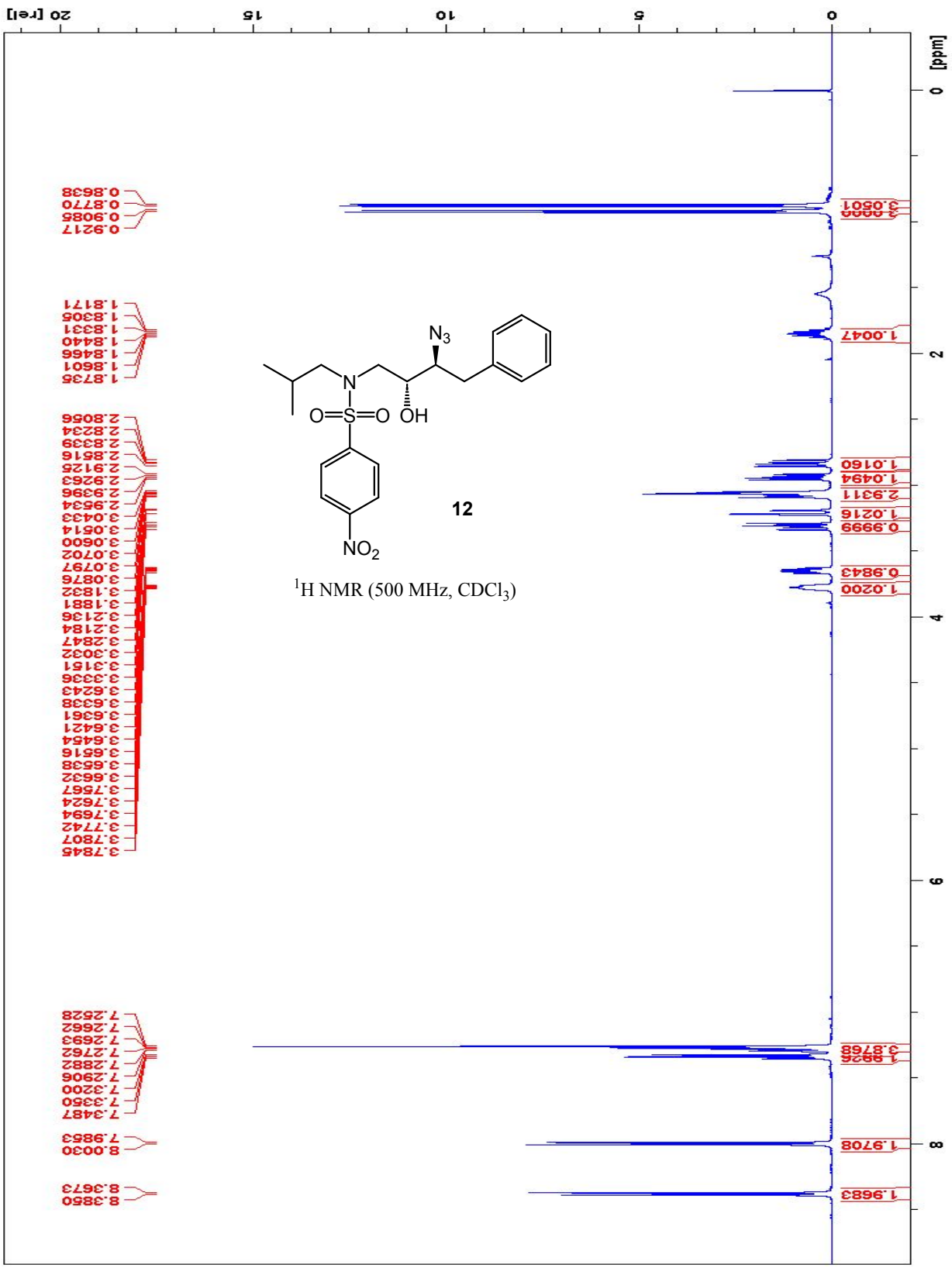
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$)



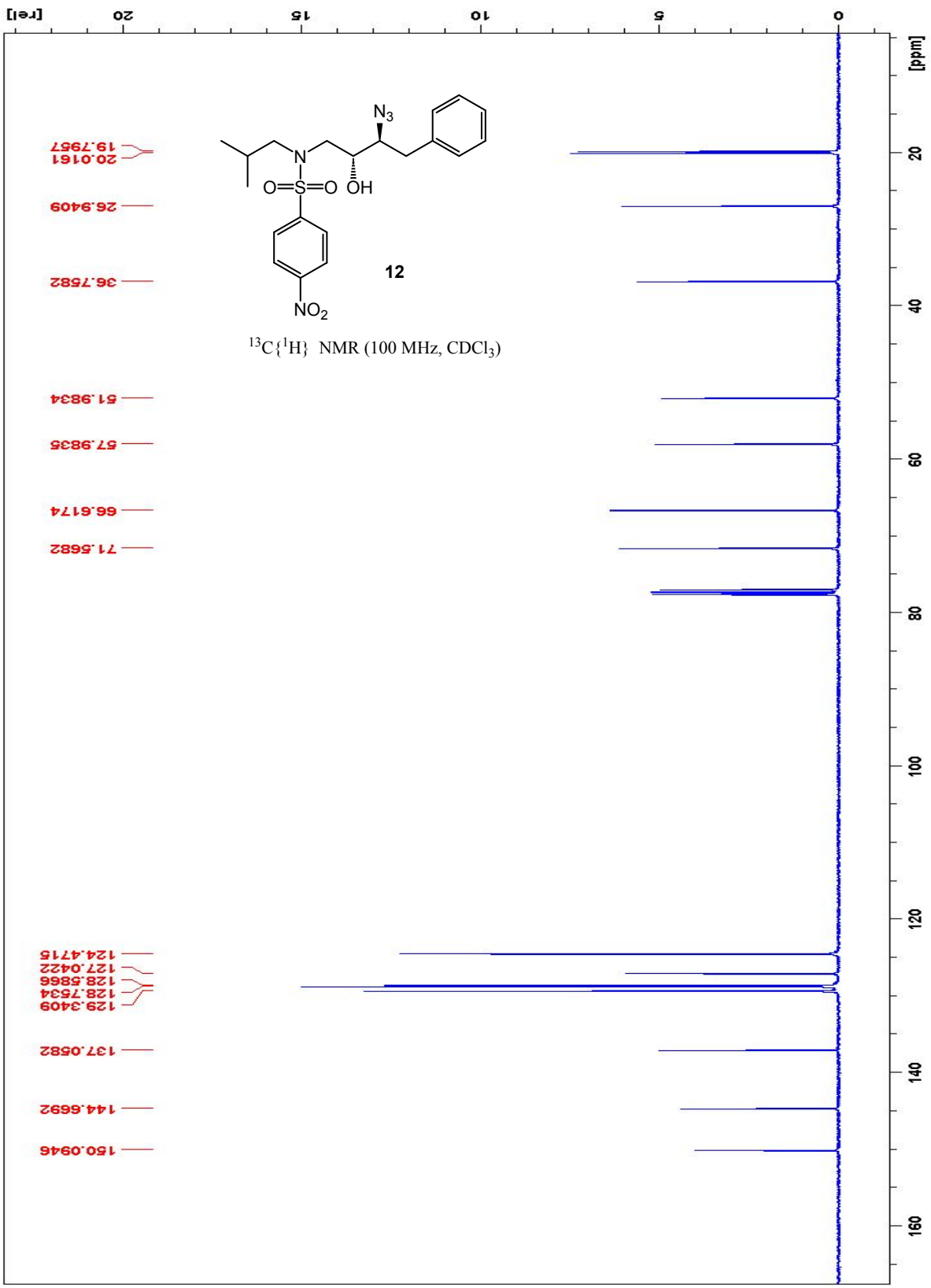
500 MHz ^{1}H NMR spectrum of **22**



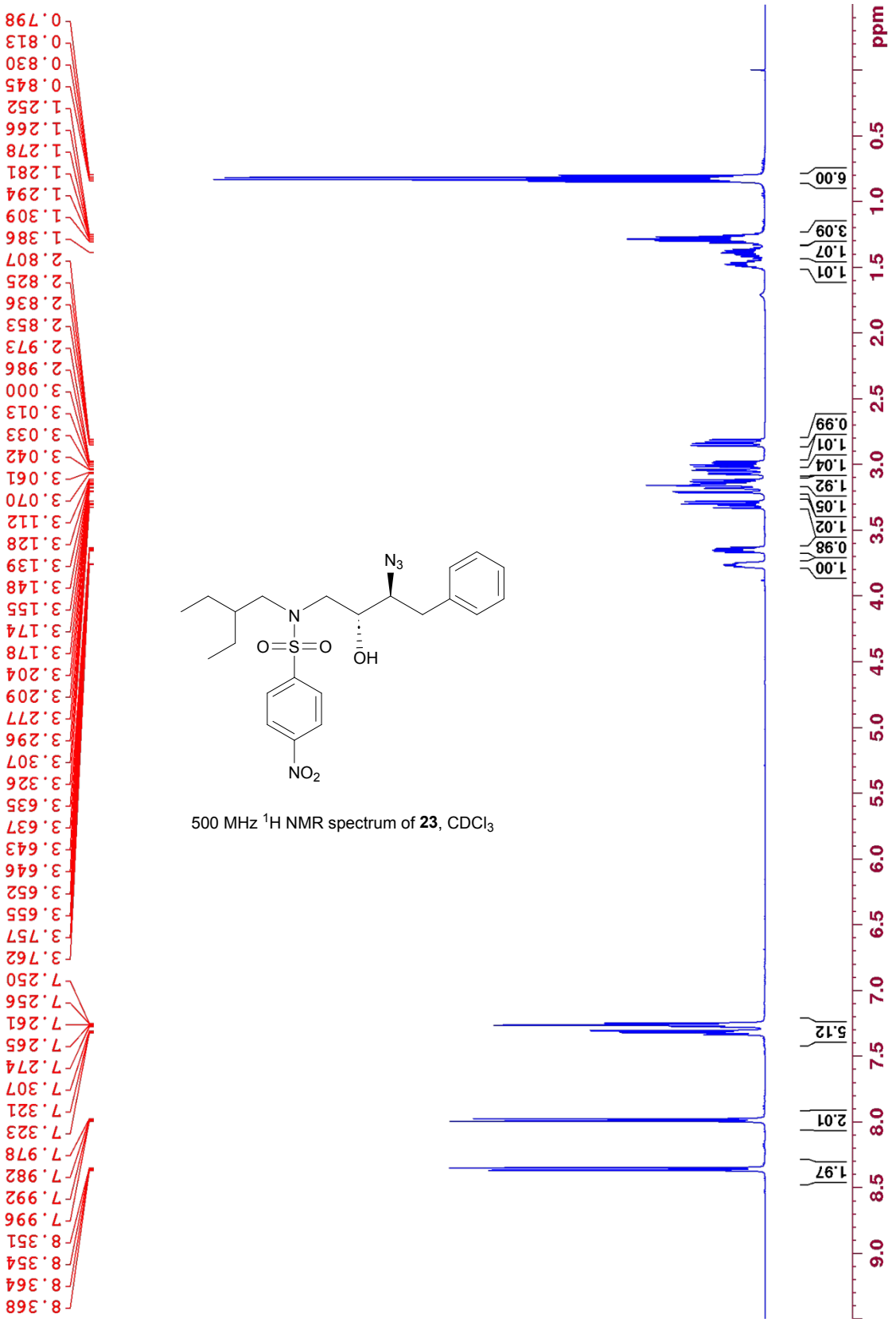
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **22**



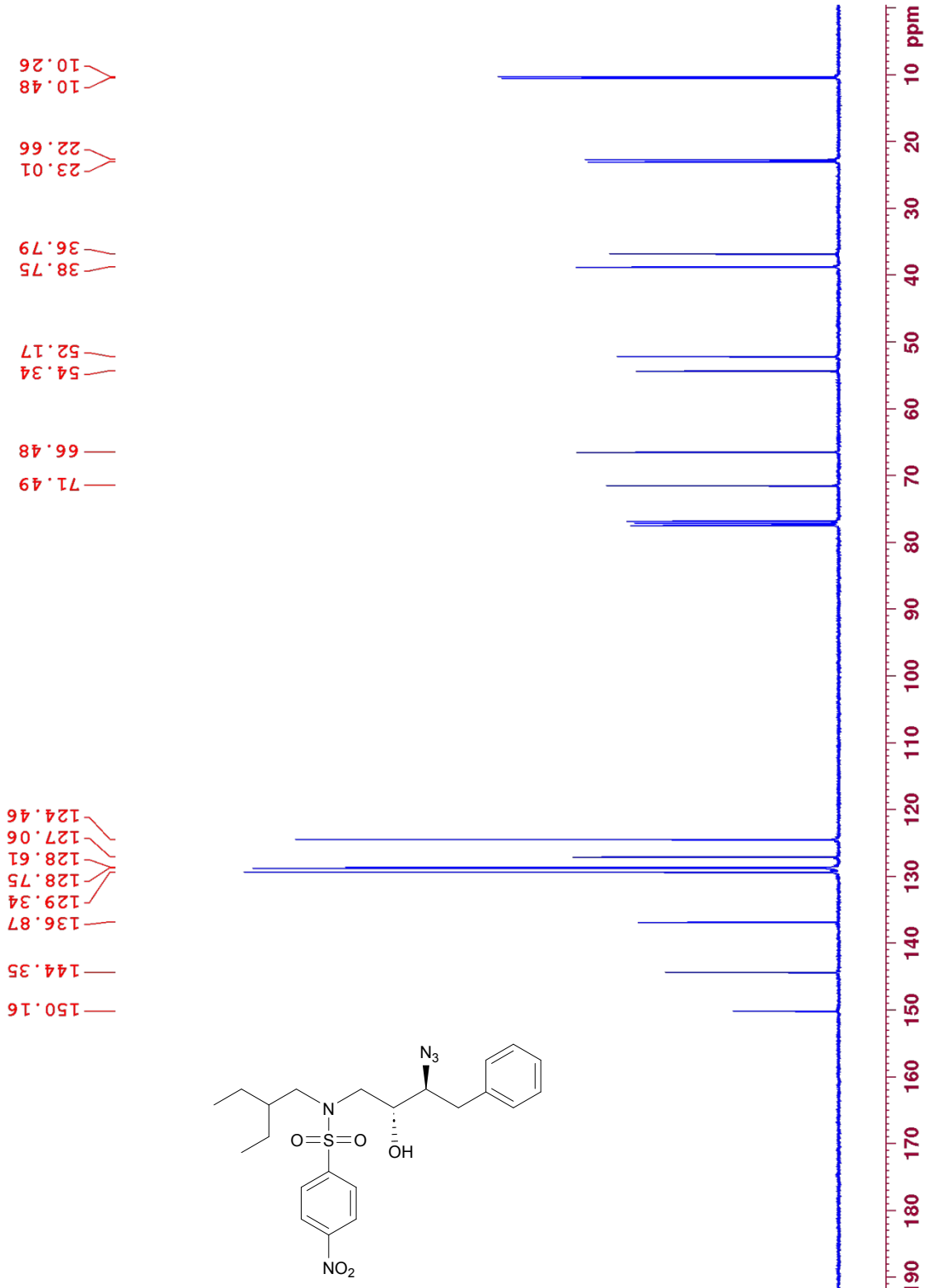
^{1}H NMR (500 MHz, $CDCl_3$)



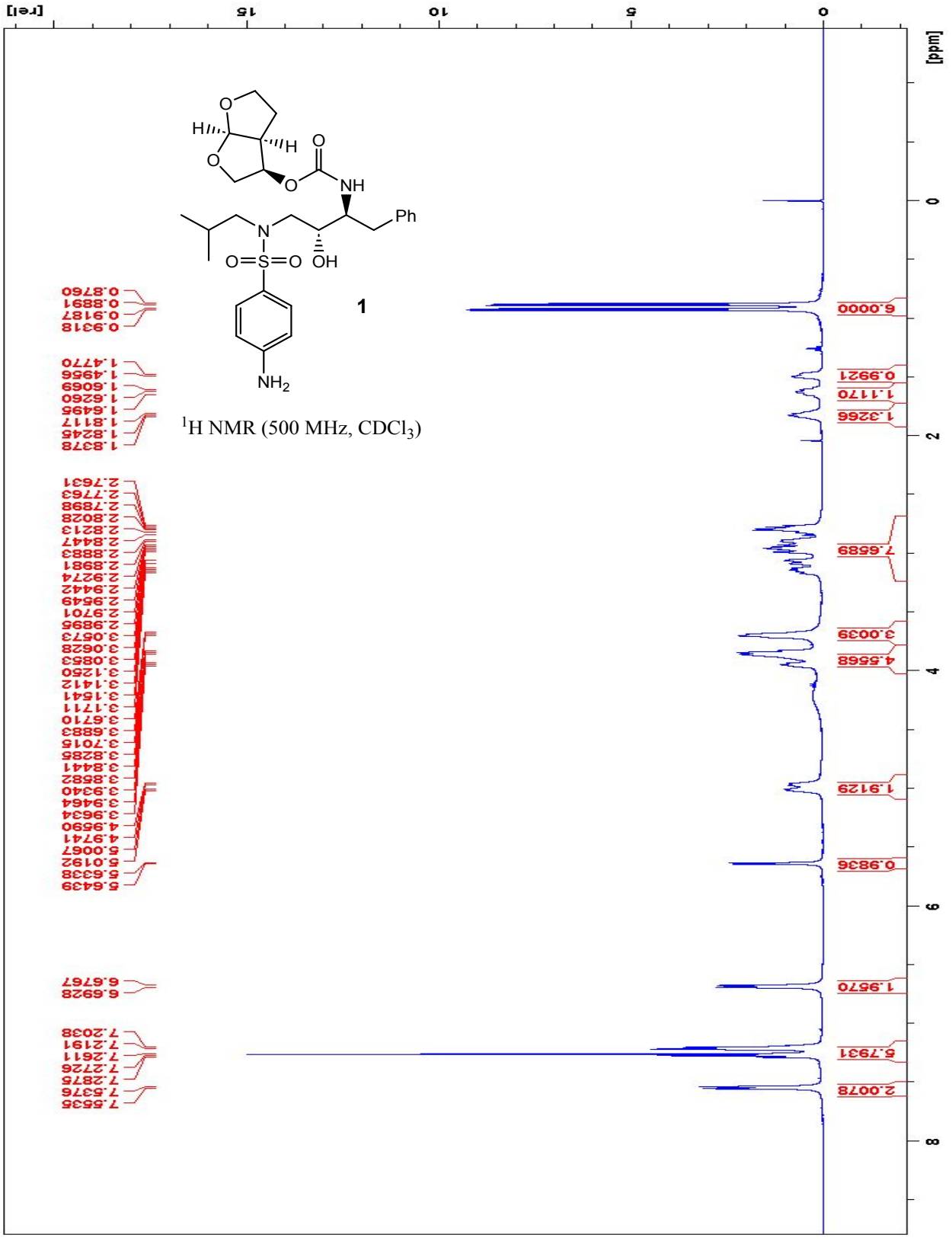
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$)

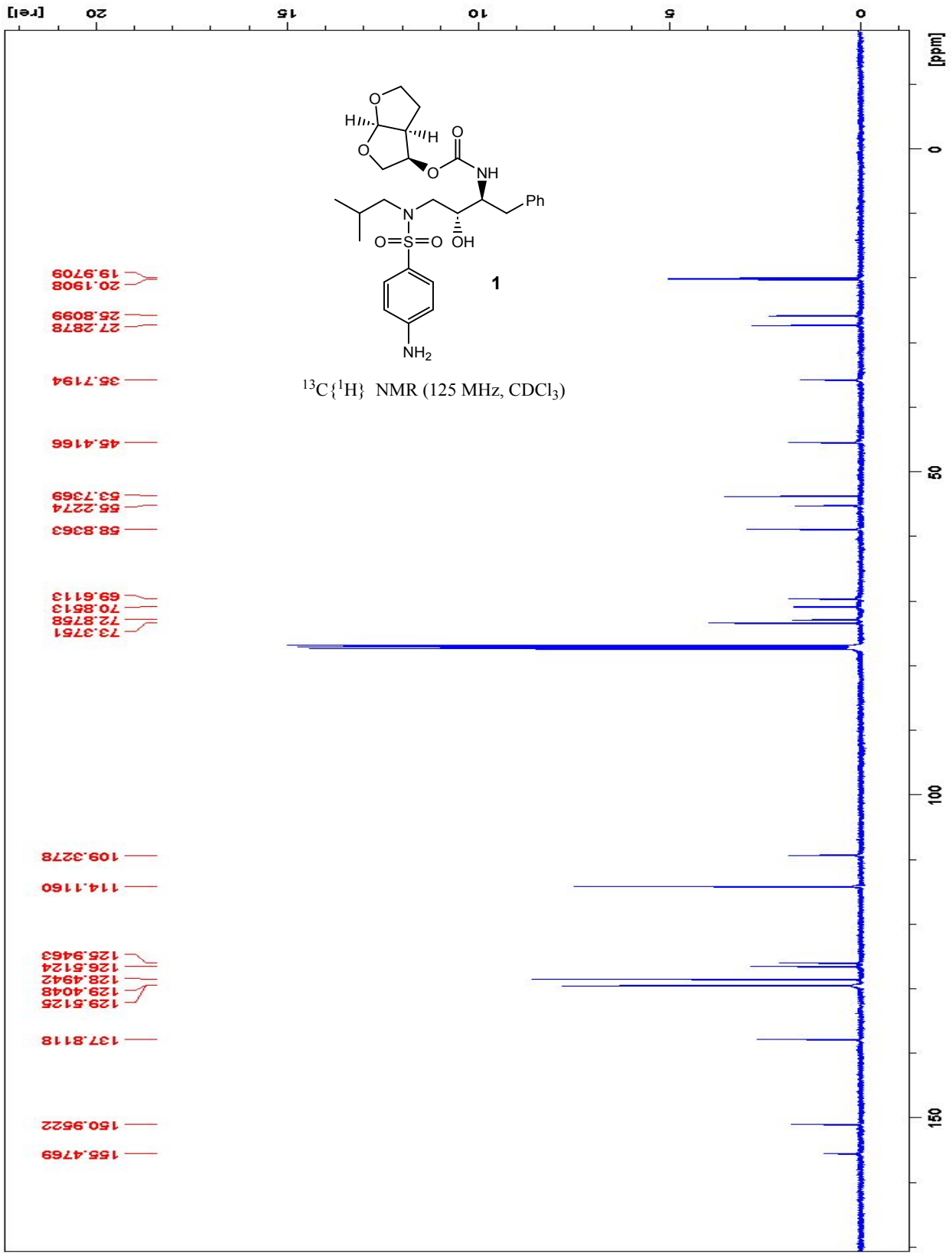


500 MHz ^{1}H NMR spectrum of **23**, $CDCl_3$

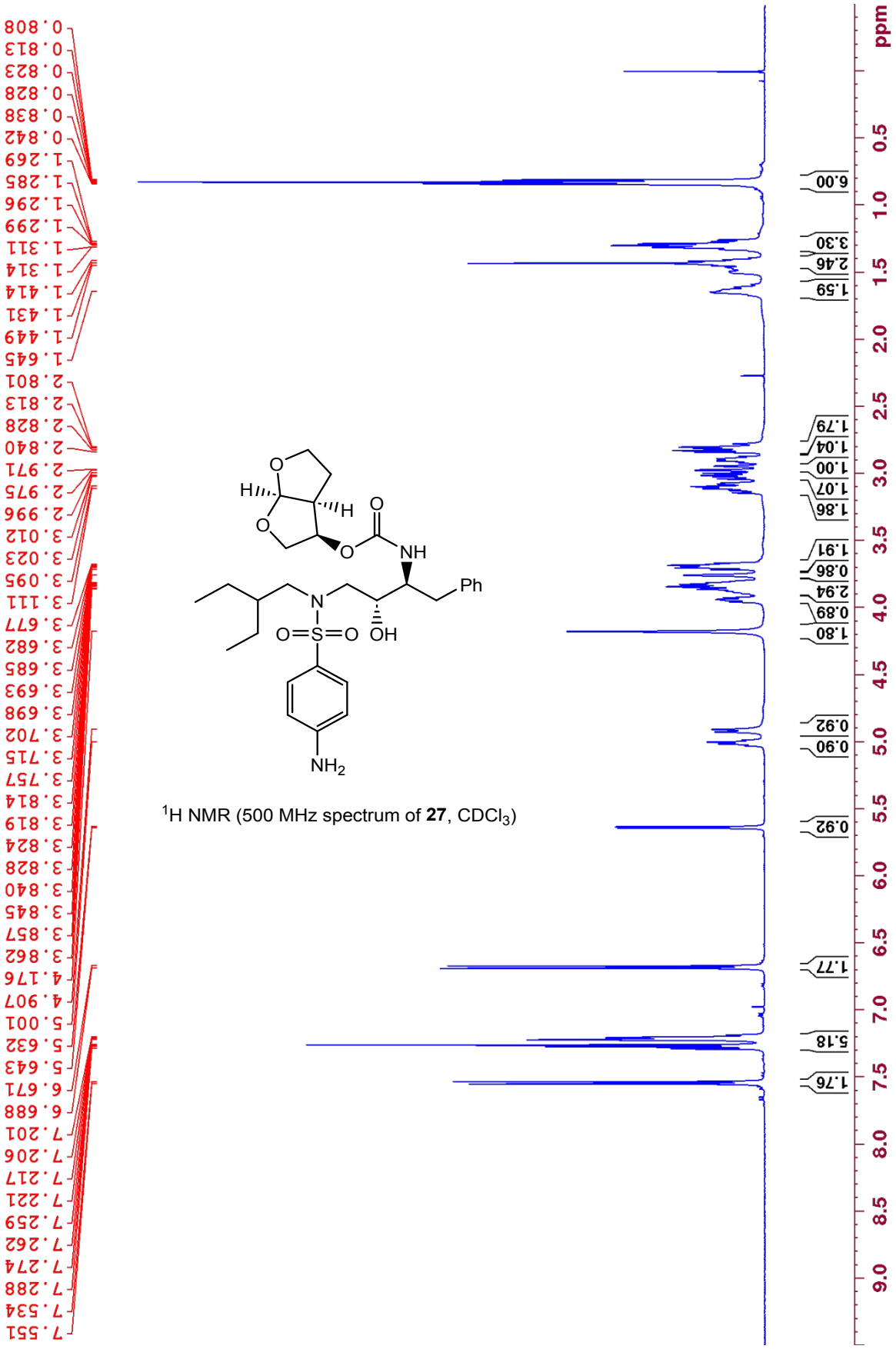


100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **23**, $CDCl_3$

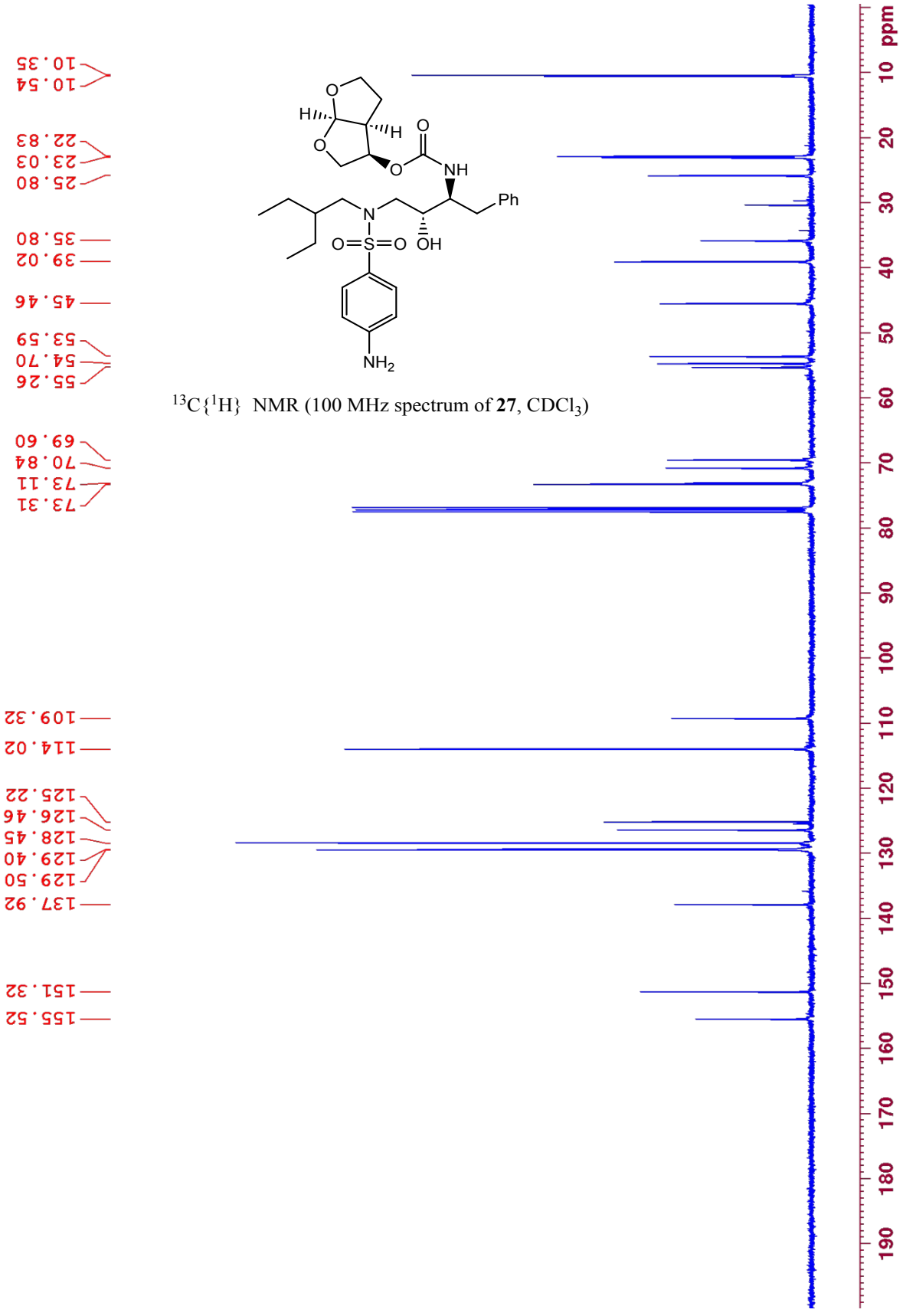

1
^{1}H NMR (500 MHz, $CDCl_3$)
[ppm]



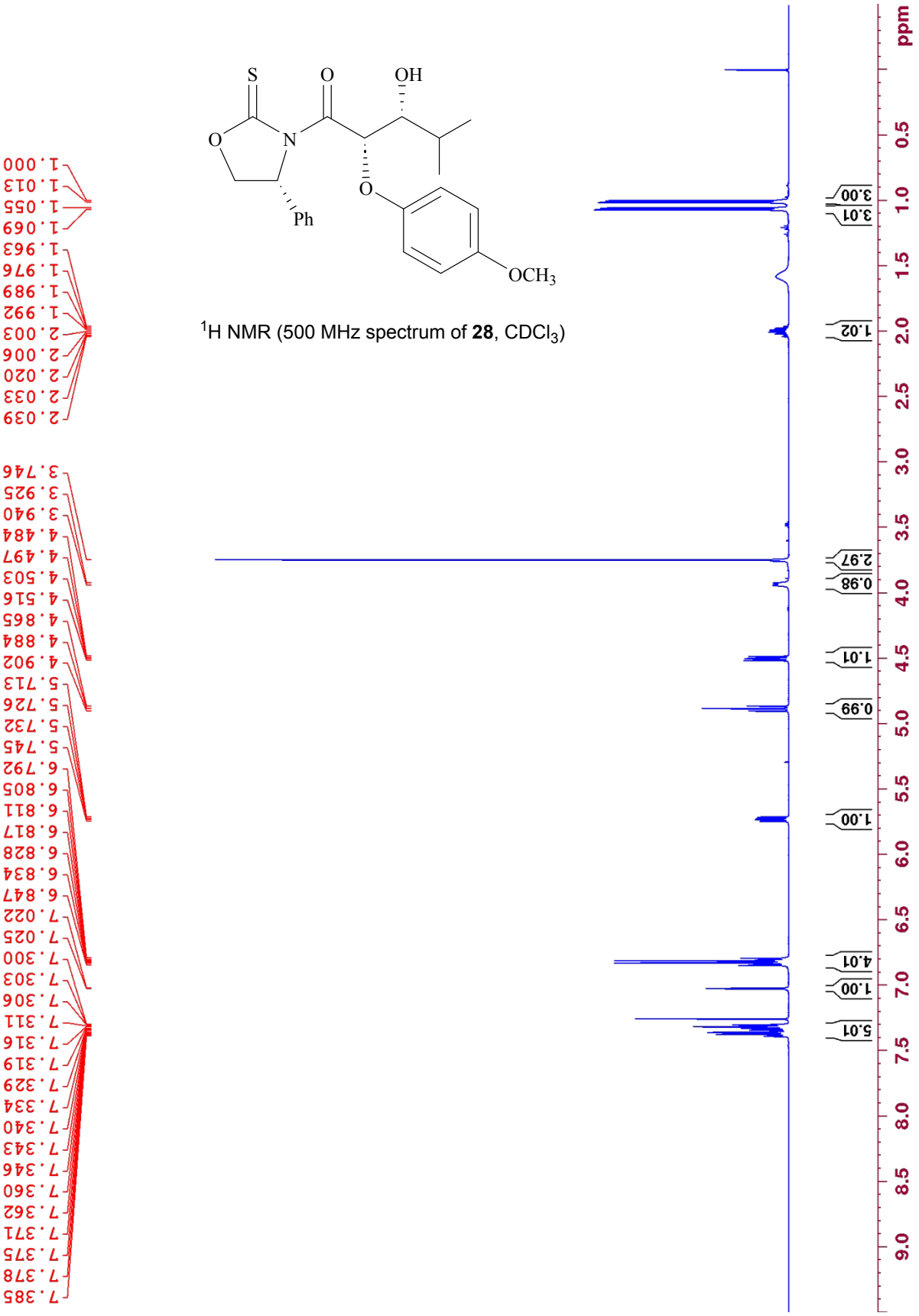
$^{13}C\{^{1}H\}$ NMR (125 MHz, $CDCl_3$)



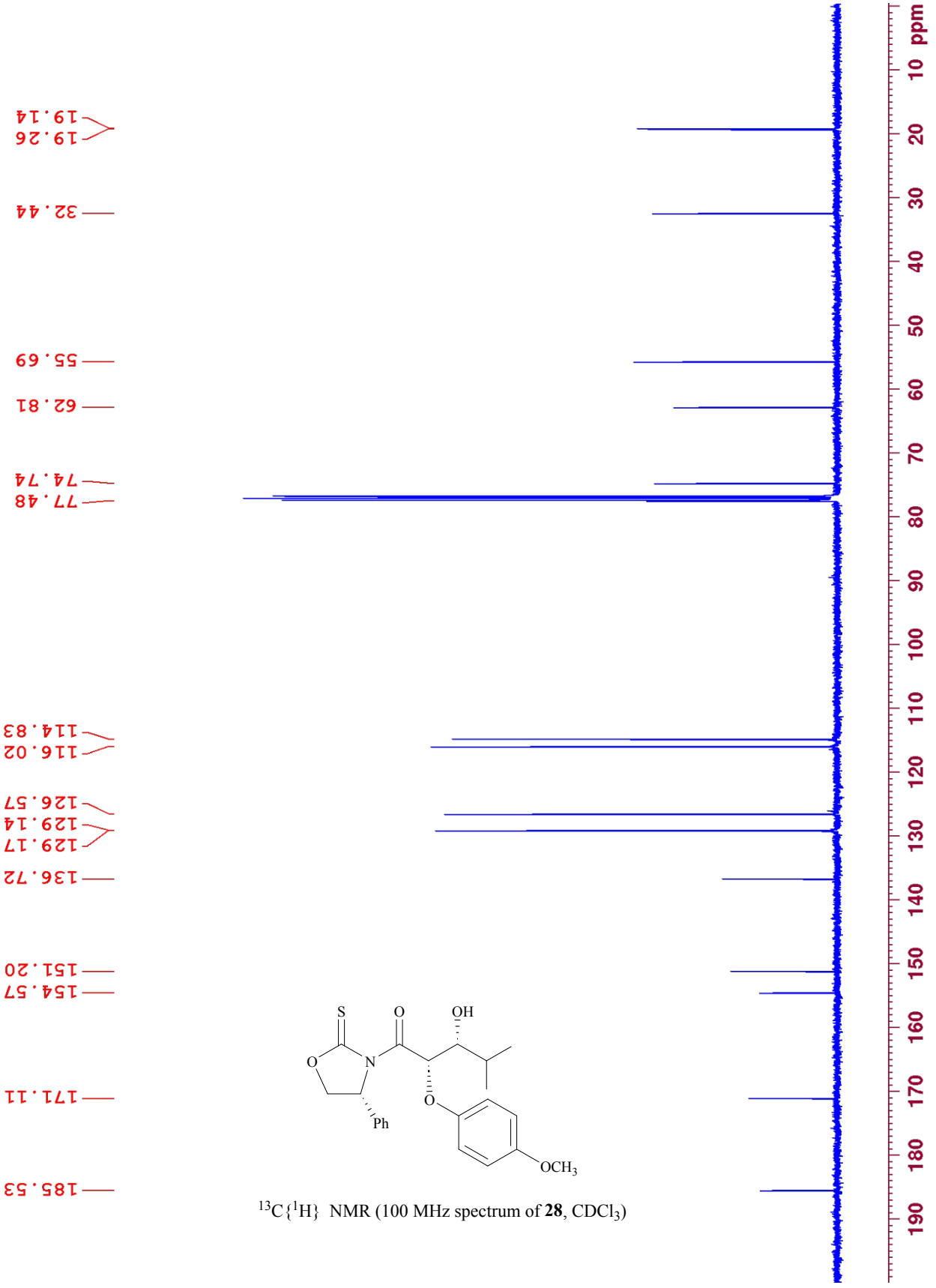
^{1}H NMR (500 MHz spectrum of **27**, $CDCl_3$)



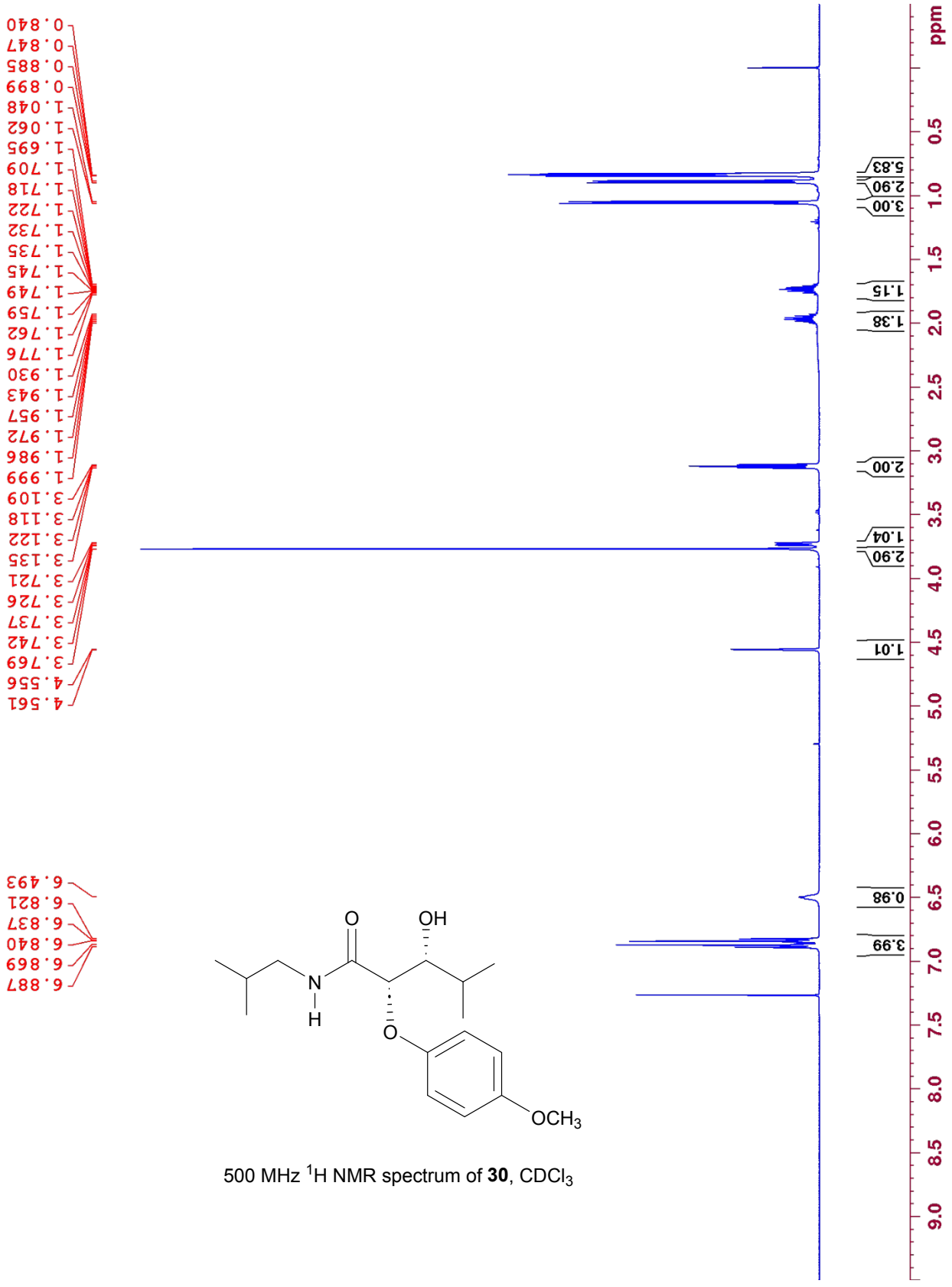
$^{13}C\{^{1}H\}$ NMR (100 MHz spectrum of **27**, $CDCl_3$)



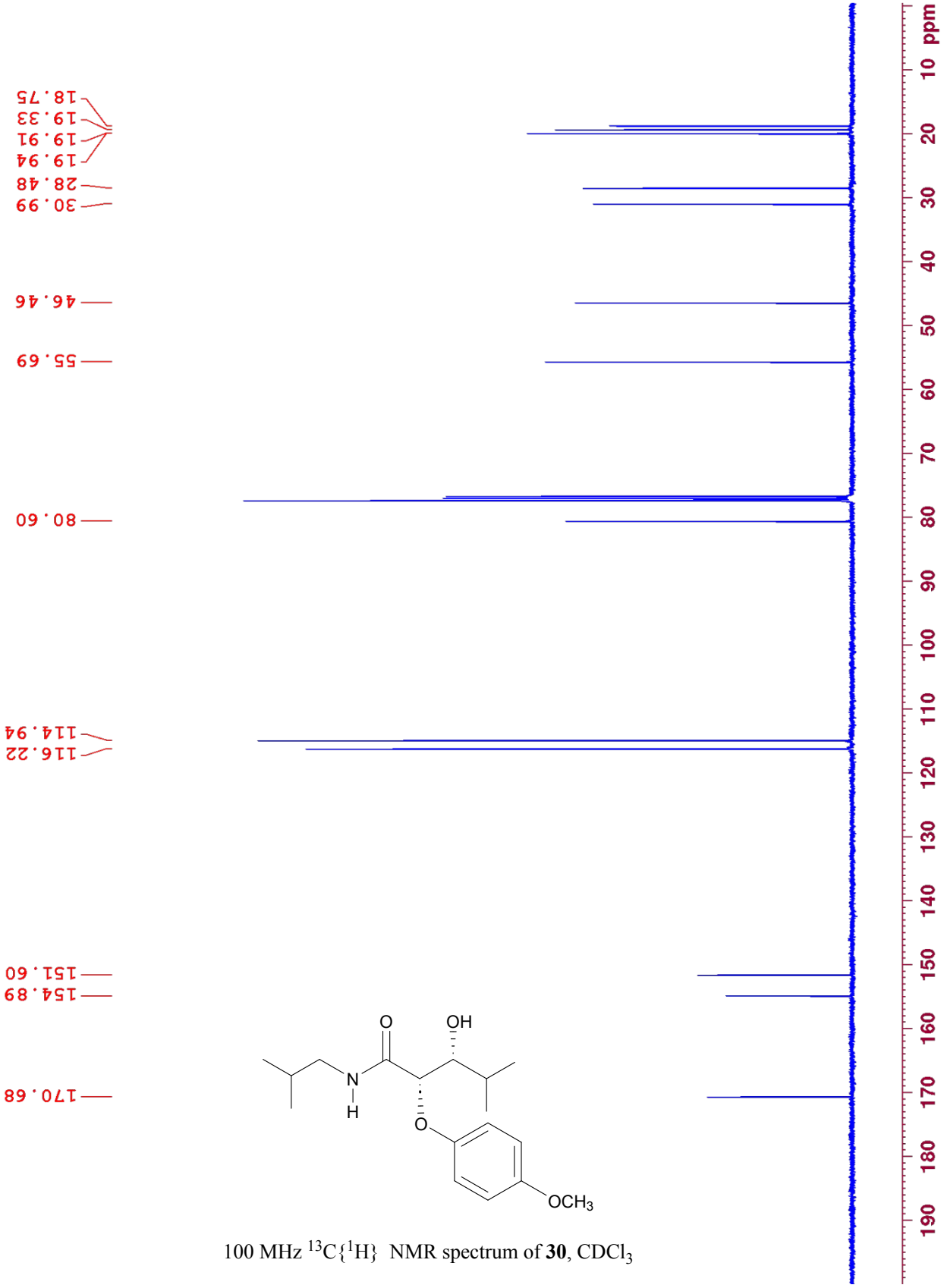
^{1}H NMR (500 MHz spectrum of **28**, $CDCl_3$)



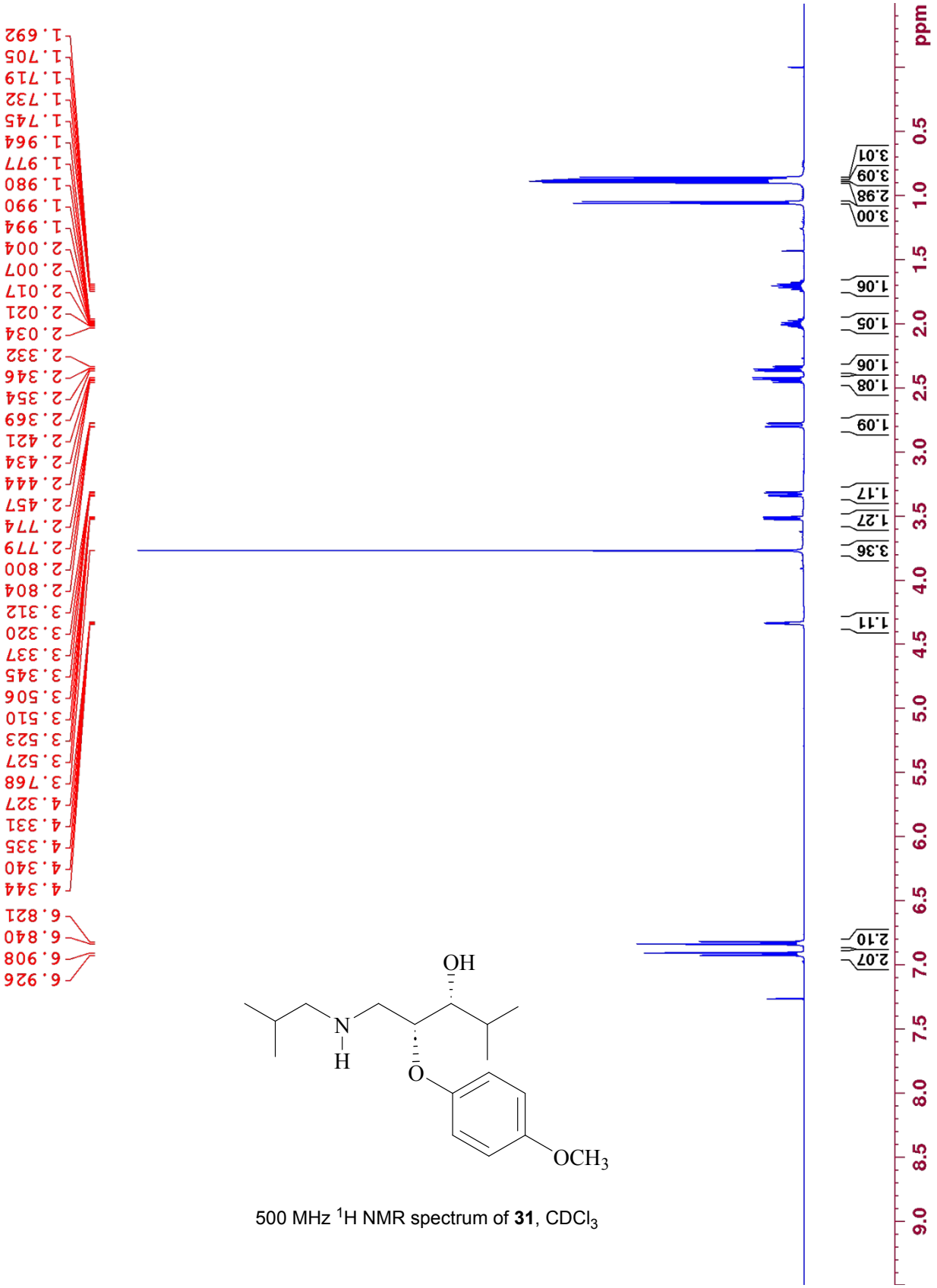
$^{13}C\{^{1}H\}$ NMR (100 MHz spectrum of **28**, $CDCl_3$)



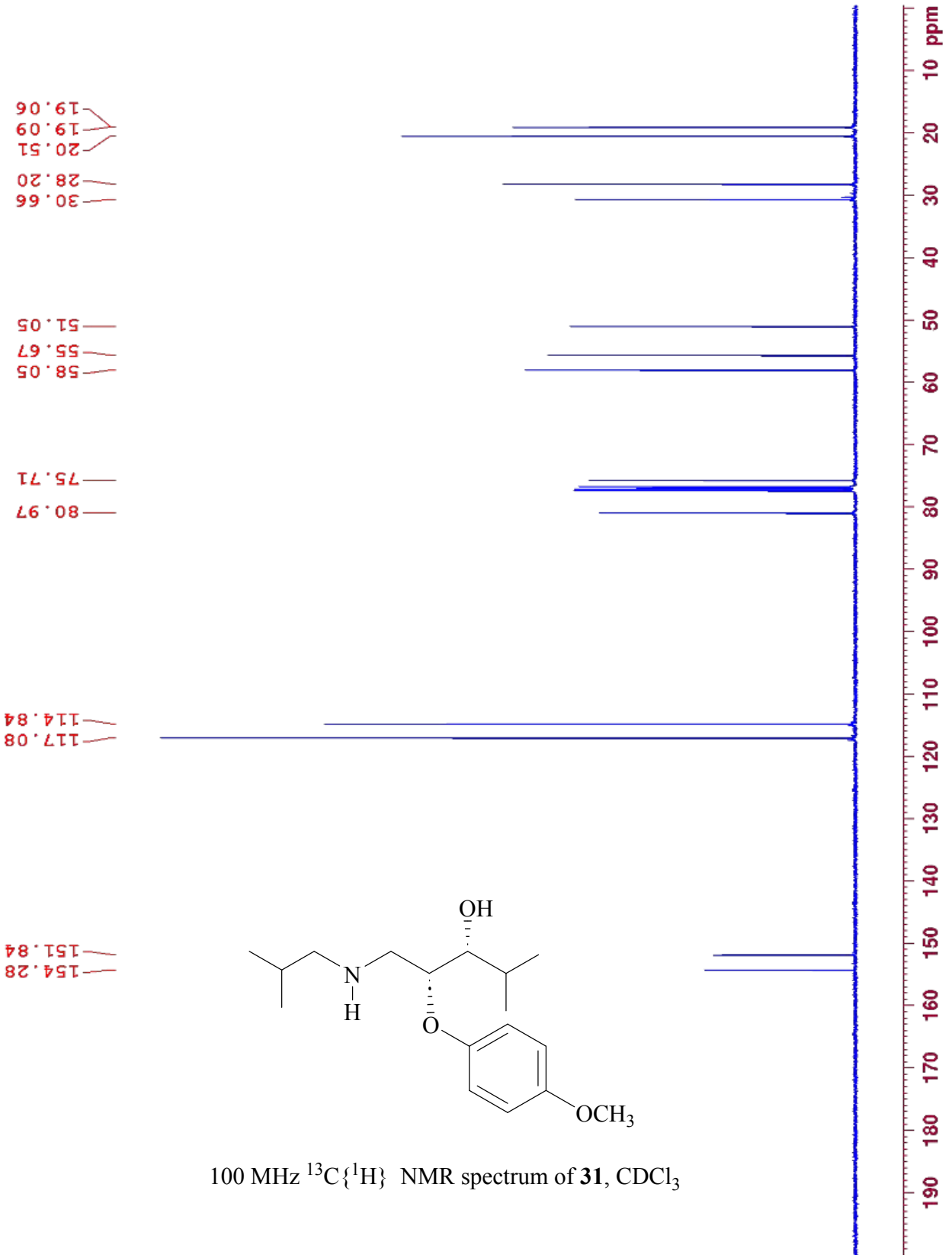
500 MHz ^{1}H NMR spectrum of **30**, $CDCl_3$



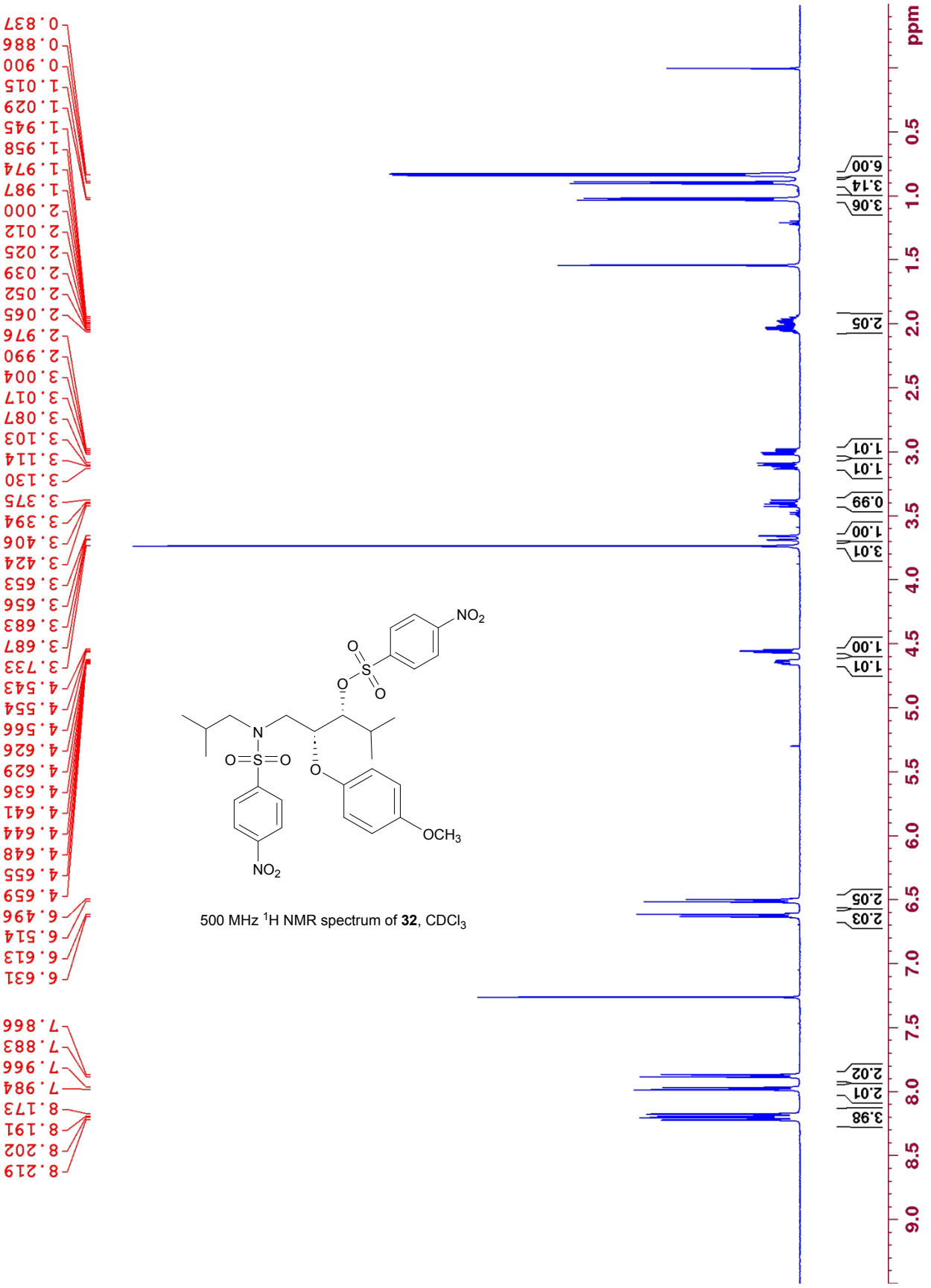
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **30**, $CDCl_3$



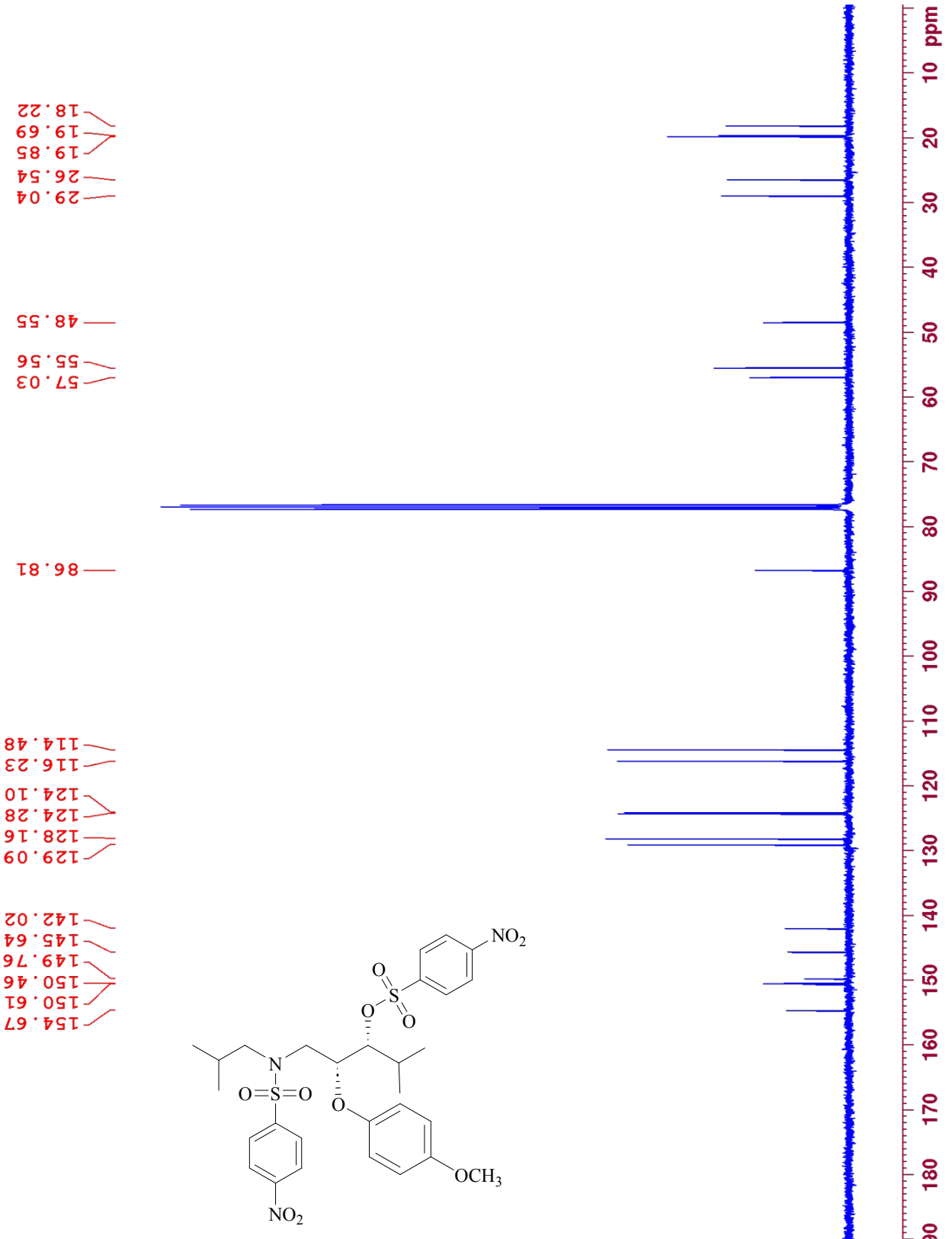
500 MHz ^{1}H NMR spectrum of **31**, $CDCl_3$



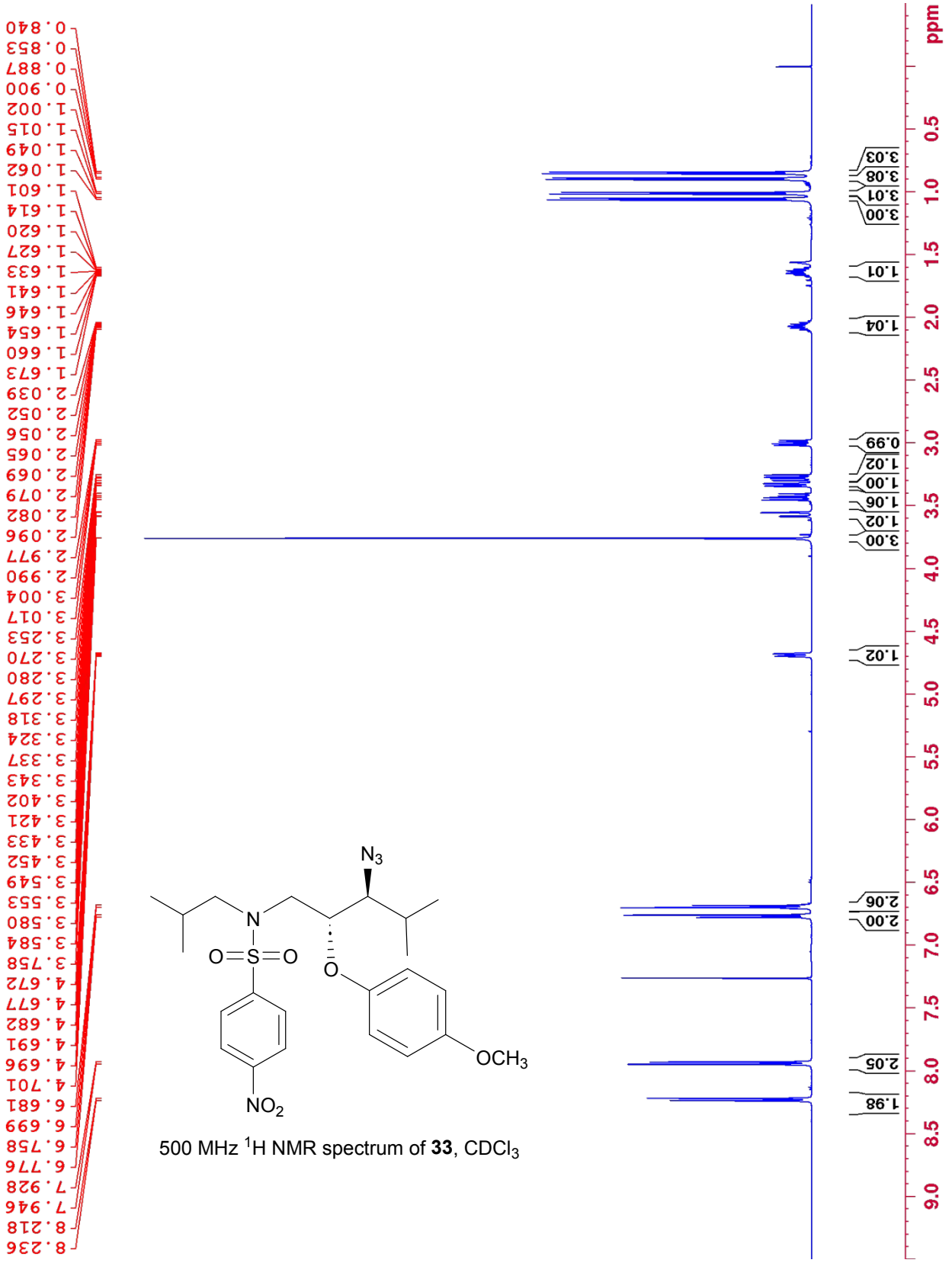
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **31**, $CDCl_3$



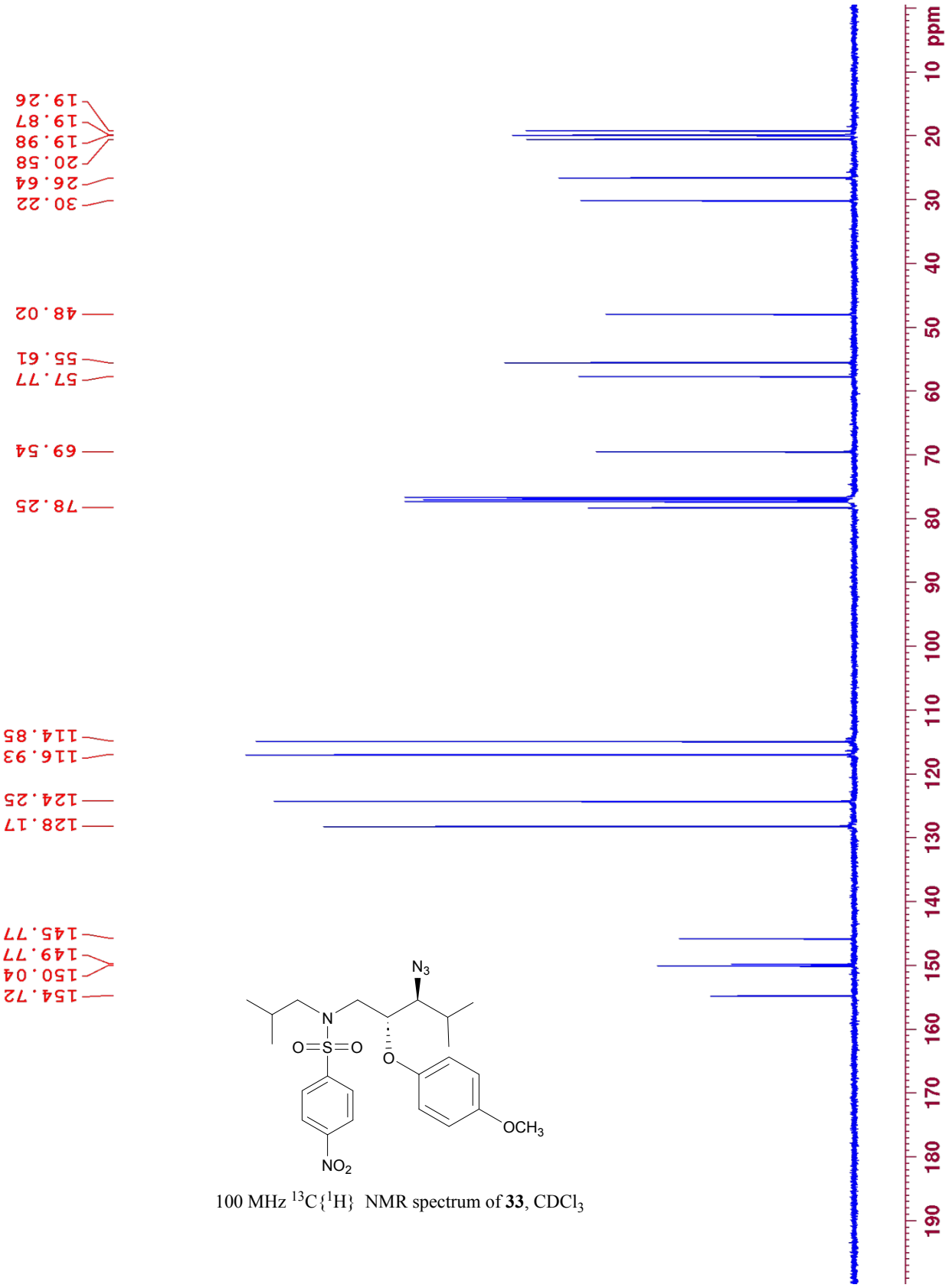
500 MHz ^{1}H NMR spectrum of **32**, $CDCl_3$



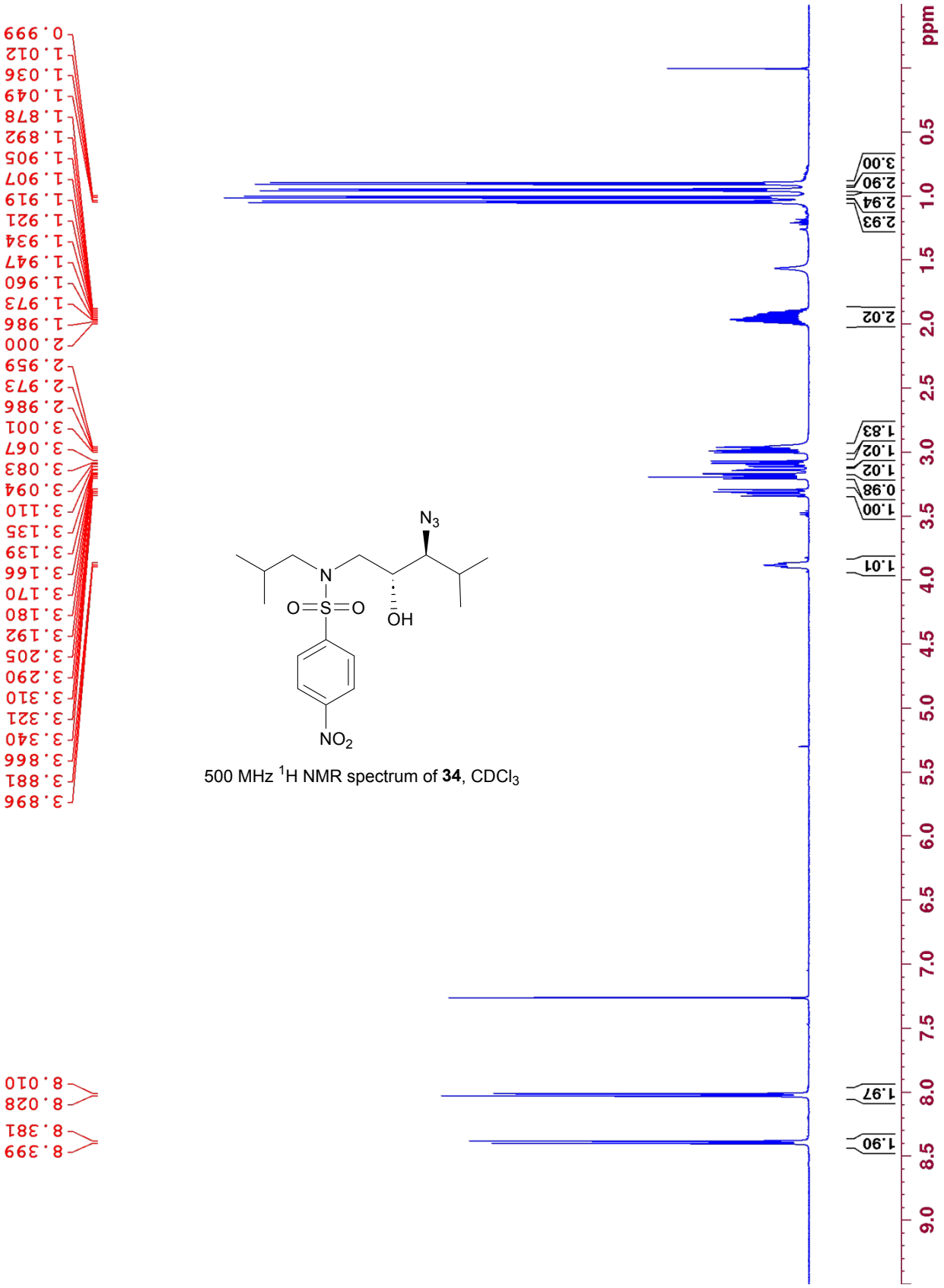
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **32**, $CDCl_3$



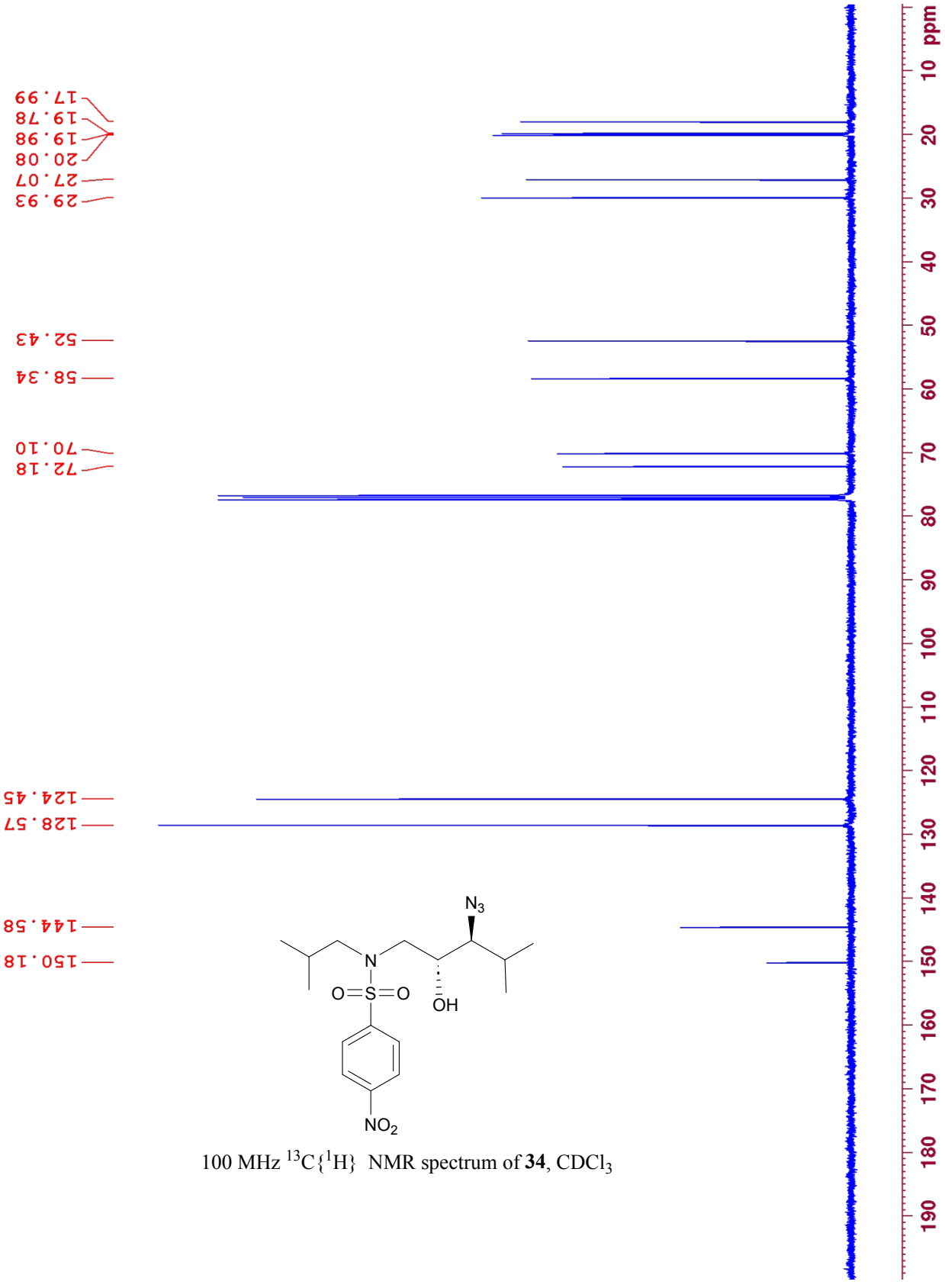
500 MHz ^{1}H NMR spectrum of **33**, $CDCl_3$



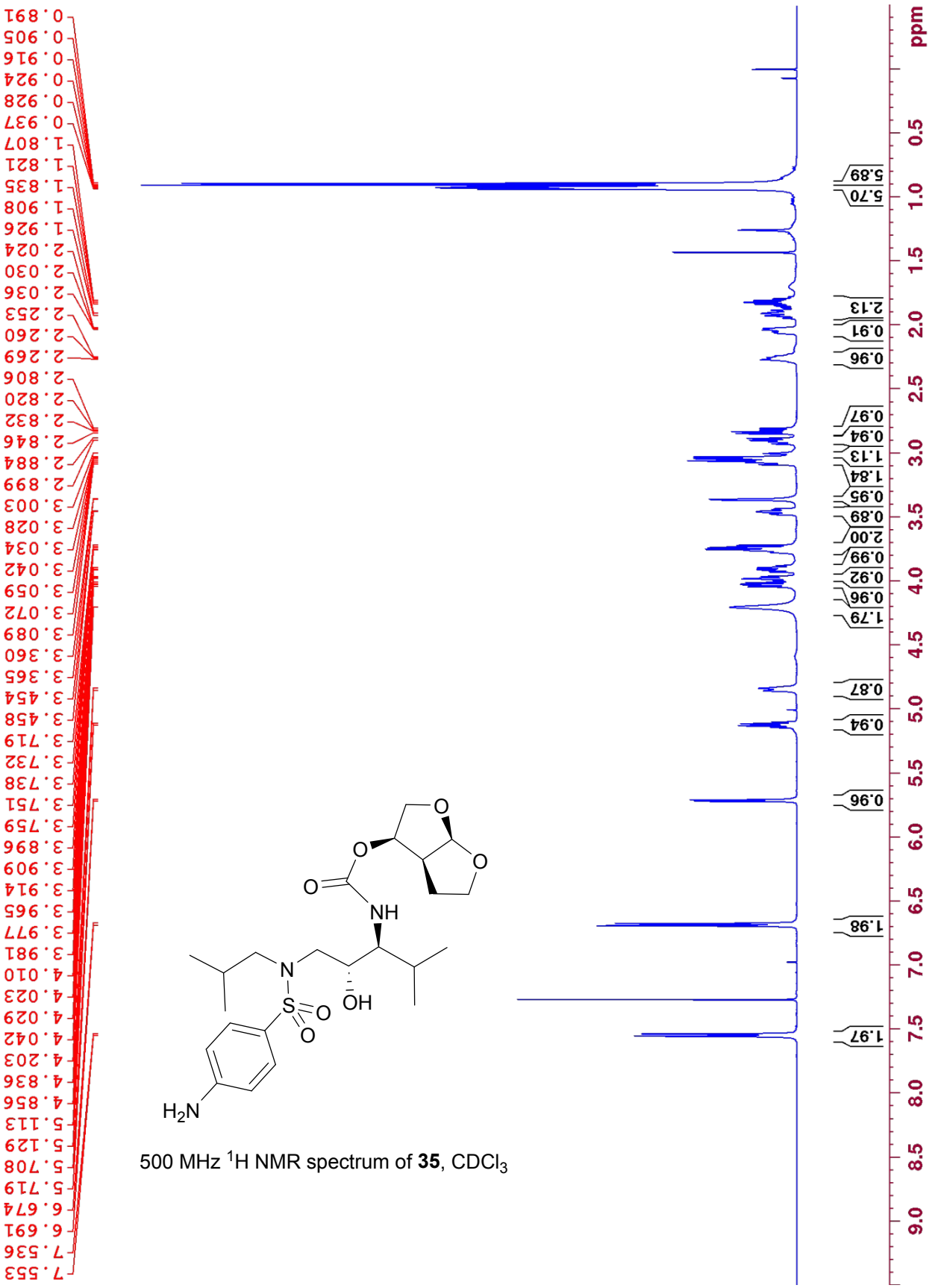
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **33**, $CDCl_3$



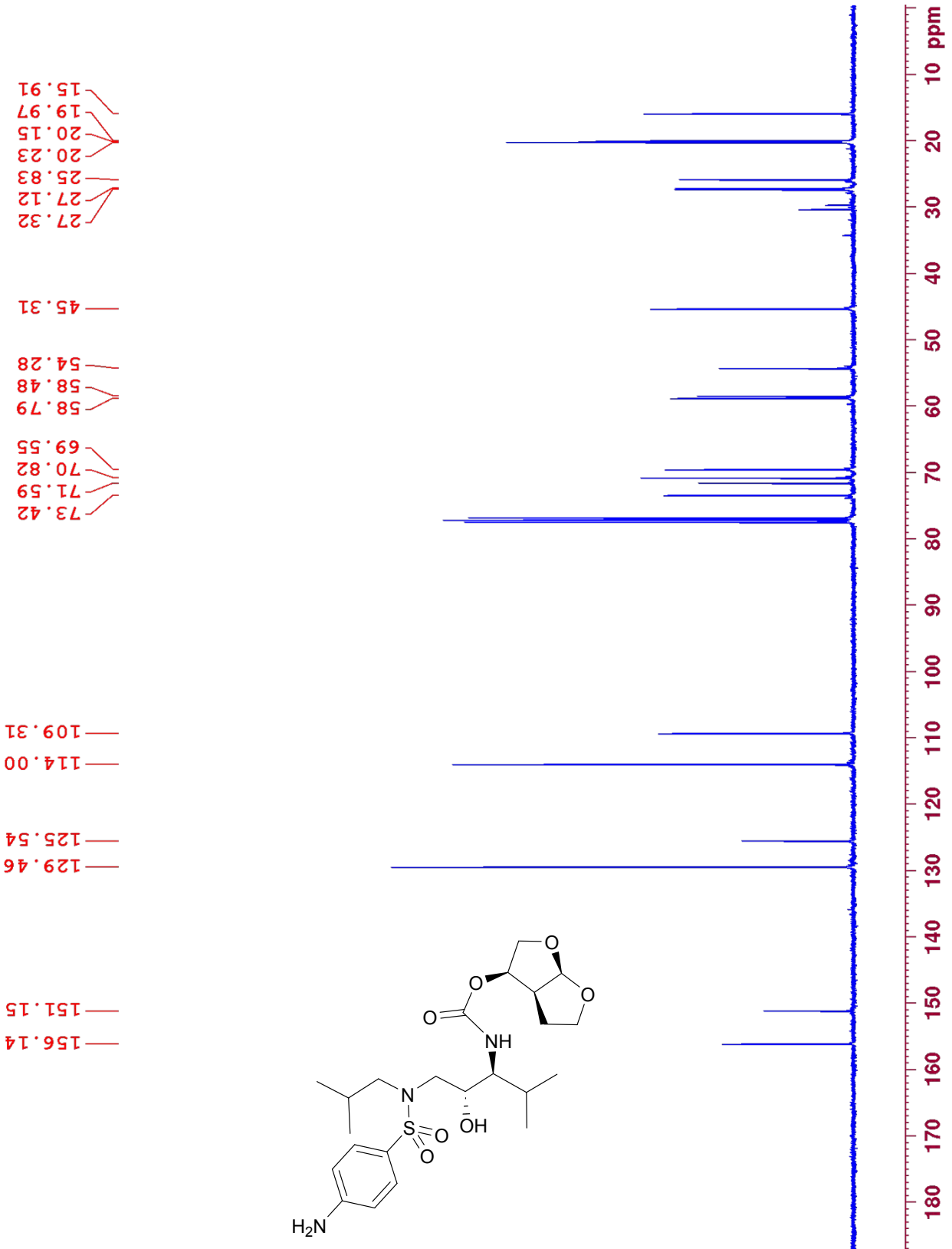
500 MHz ^{1}H NMR spectrum of **34**, $CDCl_3$



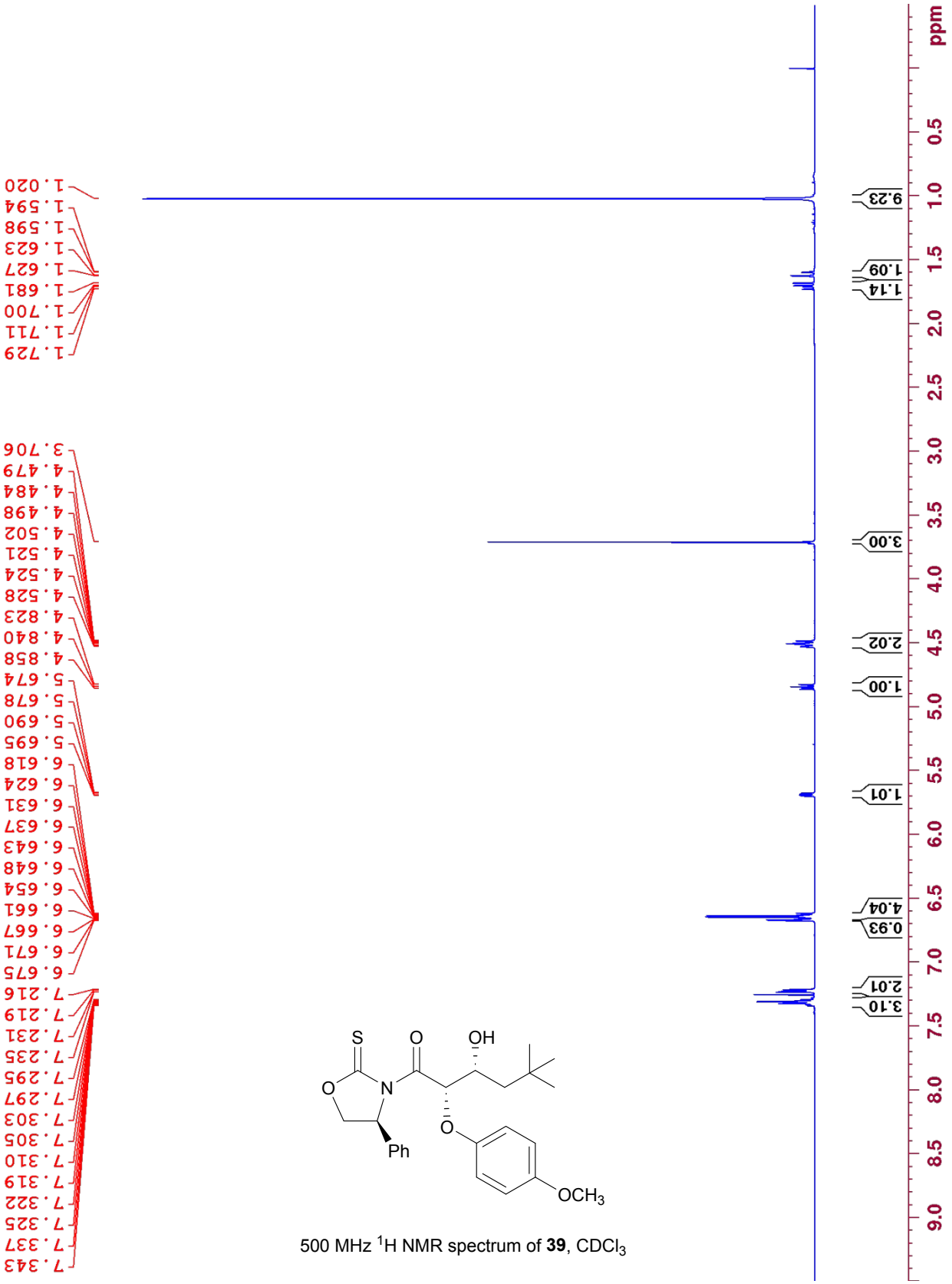
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **34**, $CDCl_3$



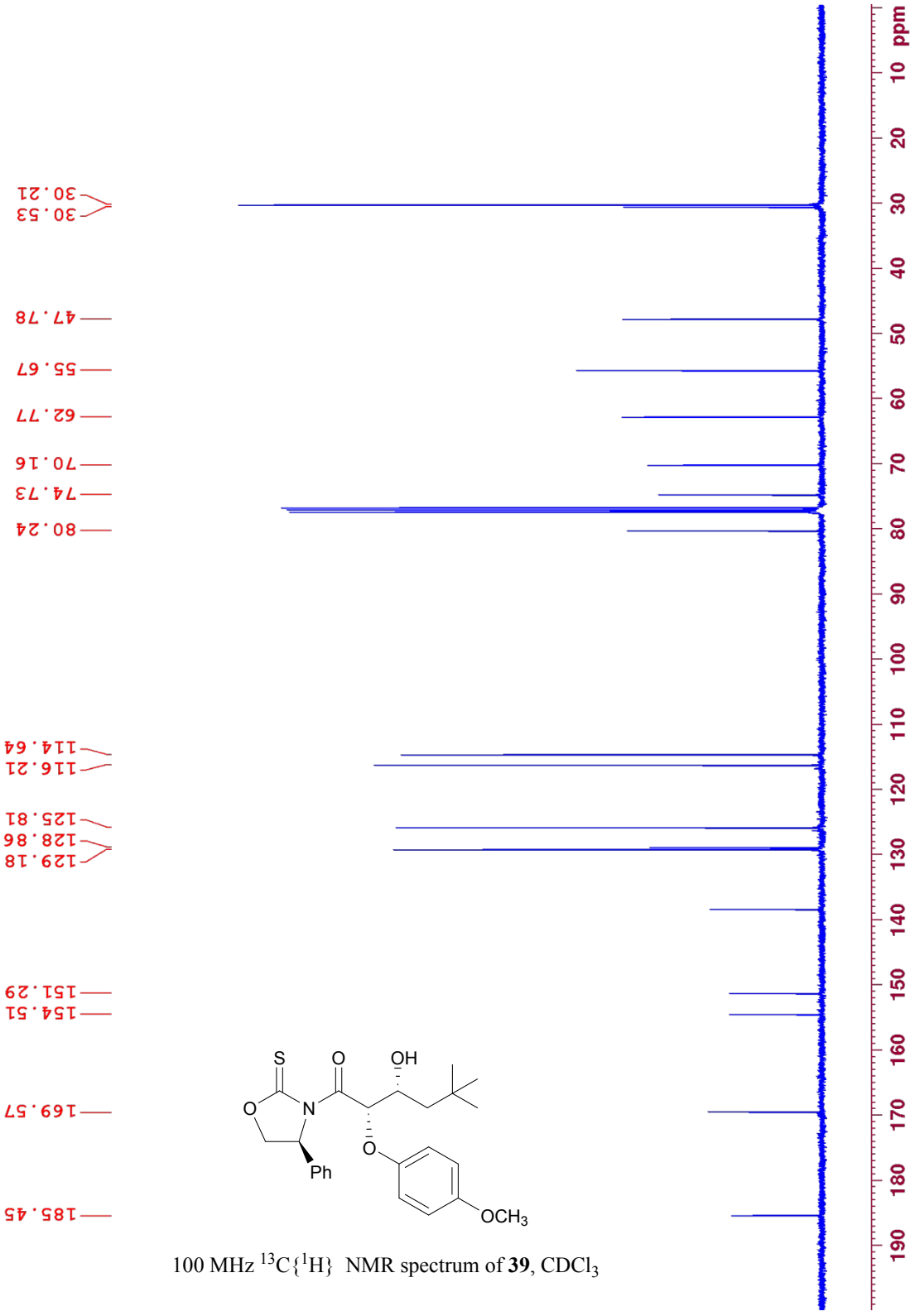
500 MHz ^{1}H NMR spectrum of **35**, $CDCl_3$



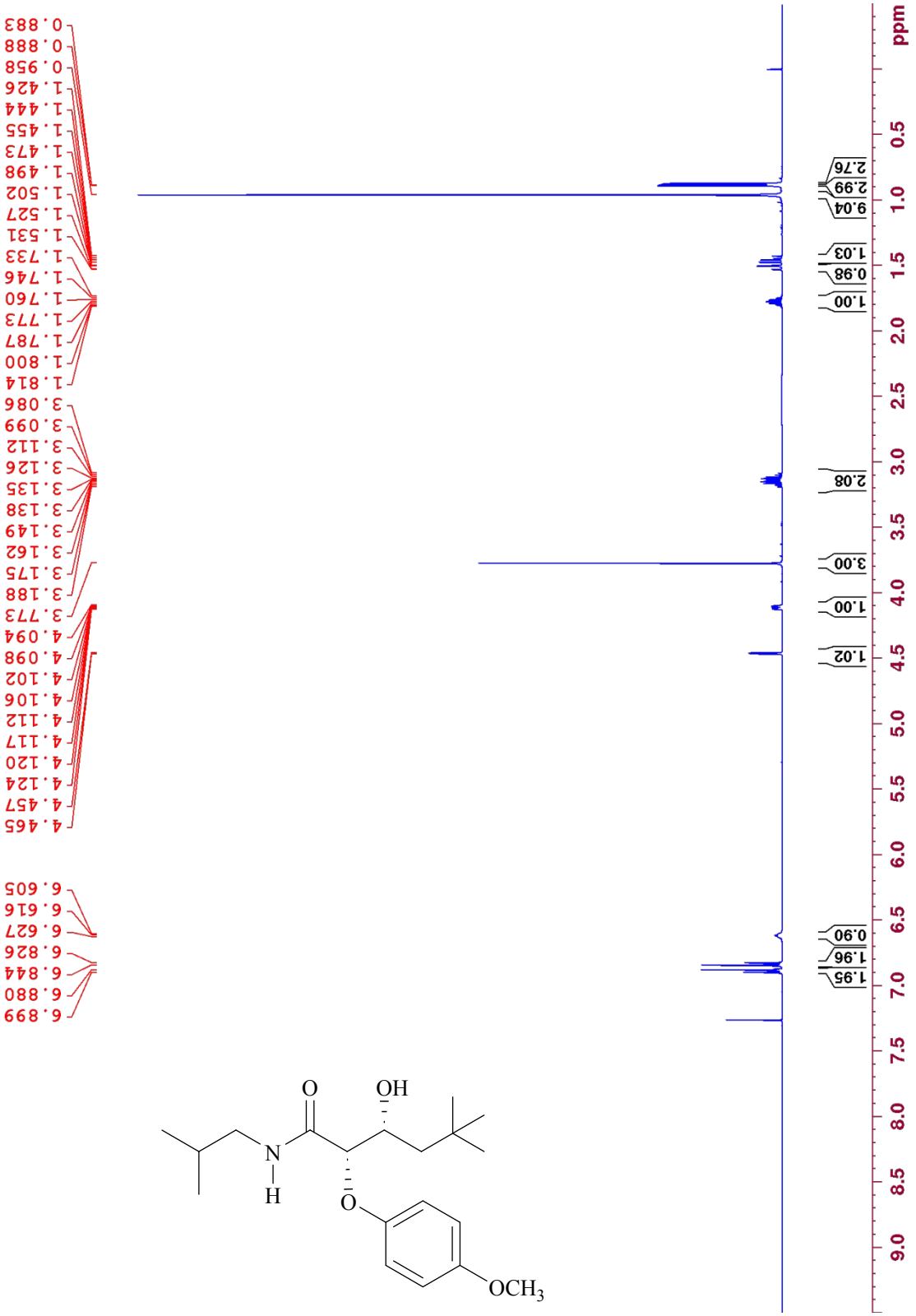
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **35**, $CDCl_3$



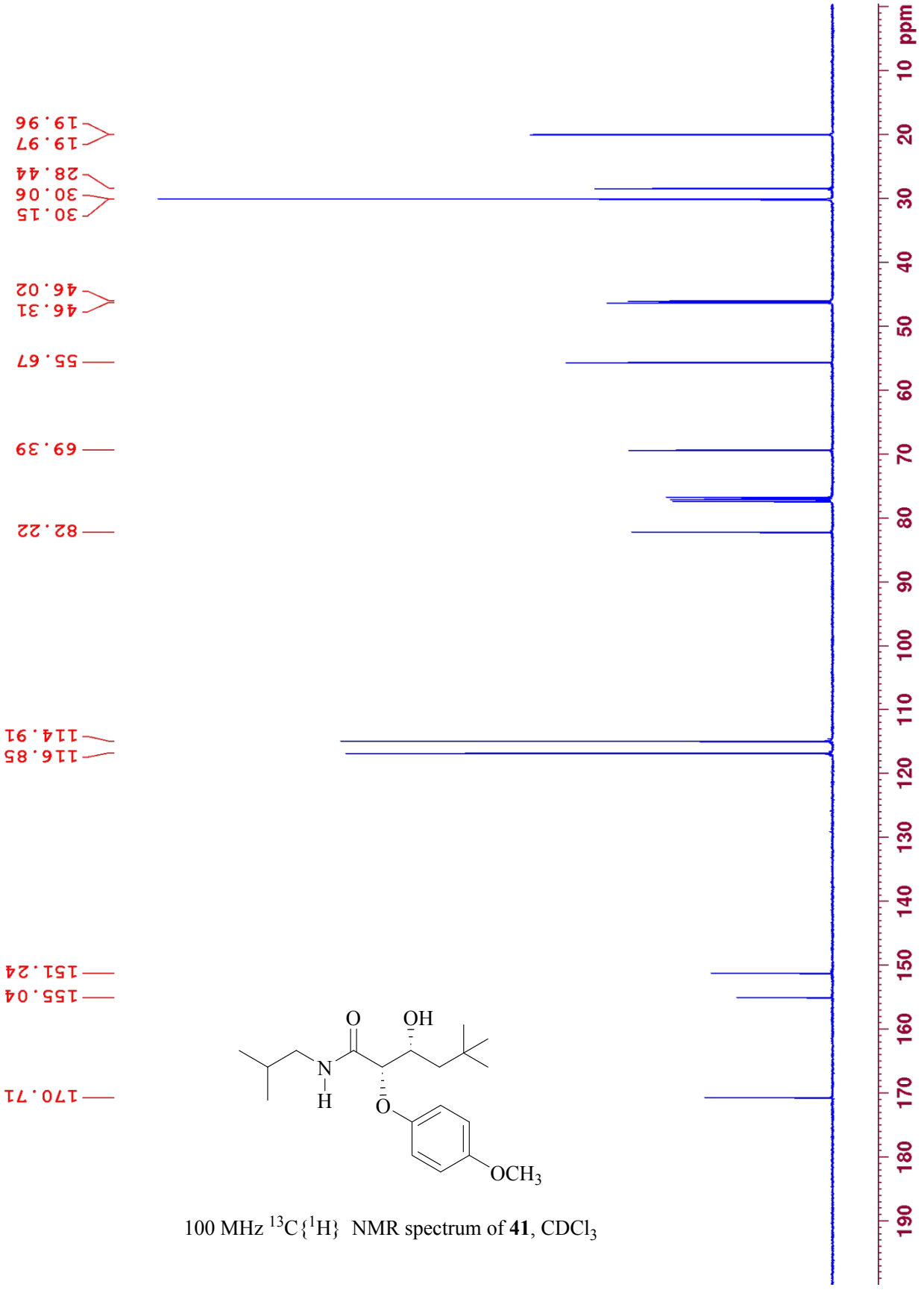
500 MHz ^{1}H NMR spectrum of **39**, $CDCl_3$



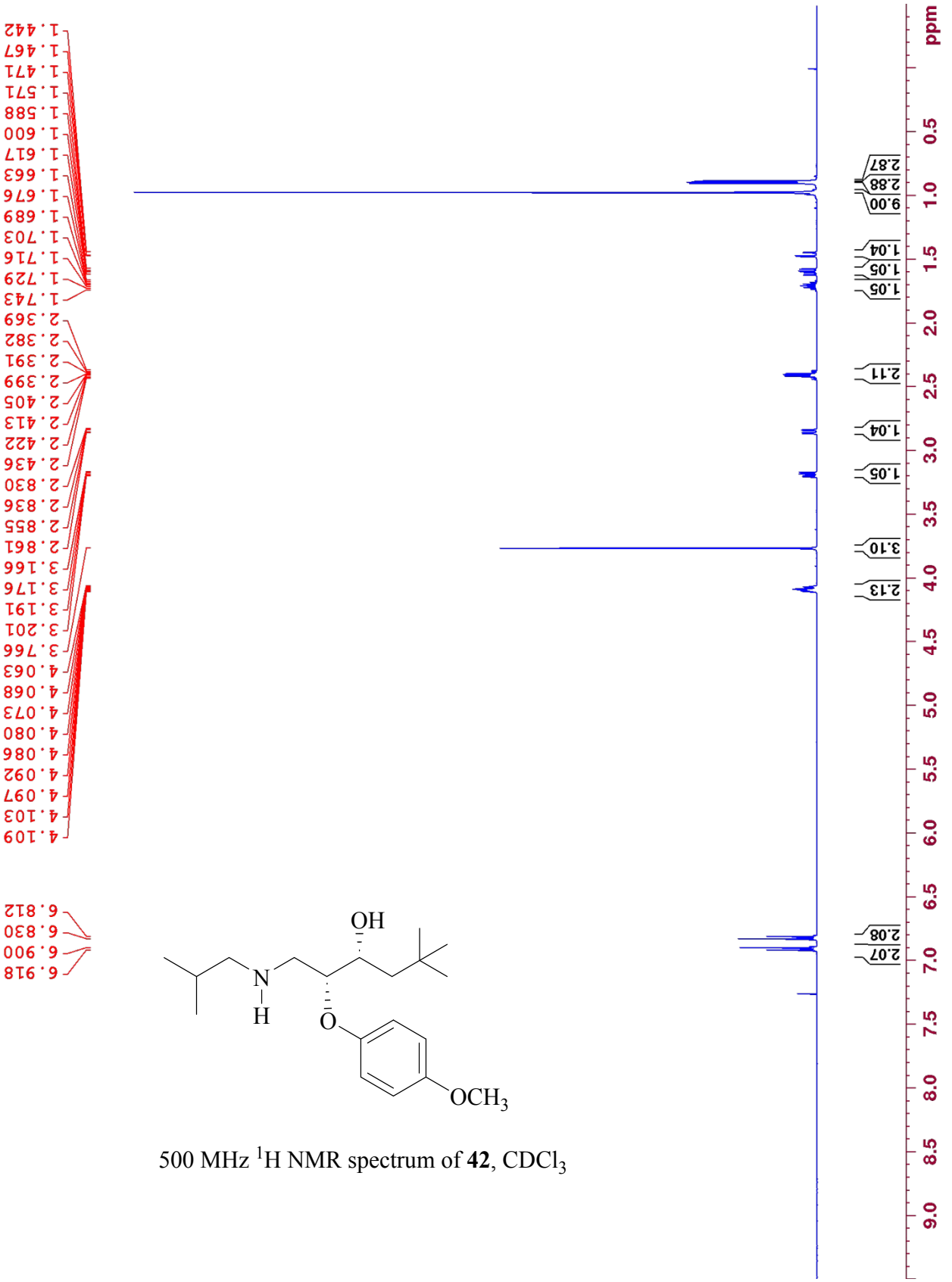
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **39**, $CDCl_3$



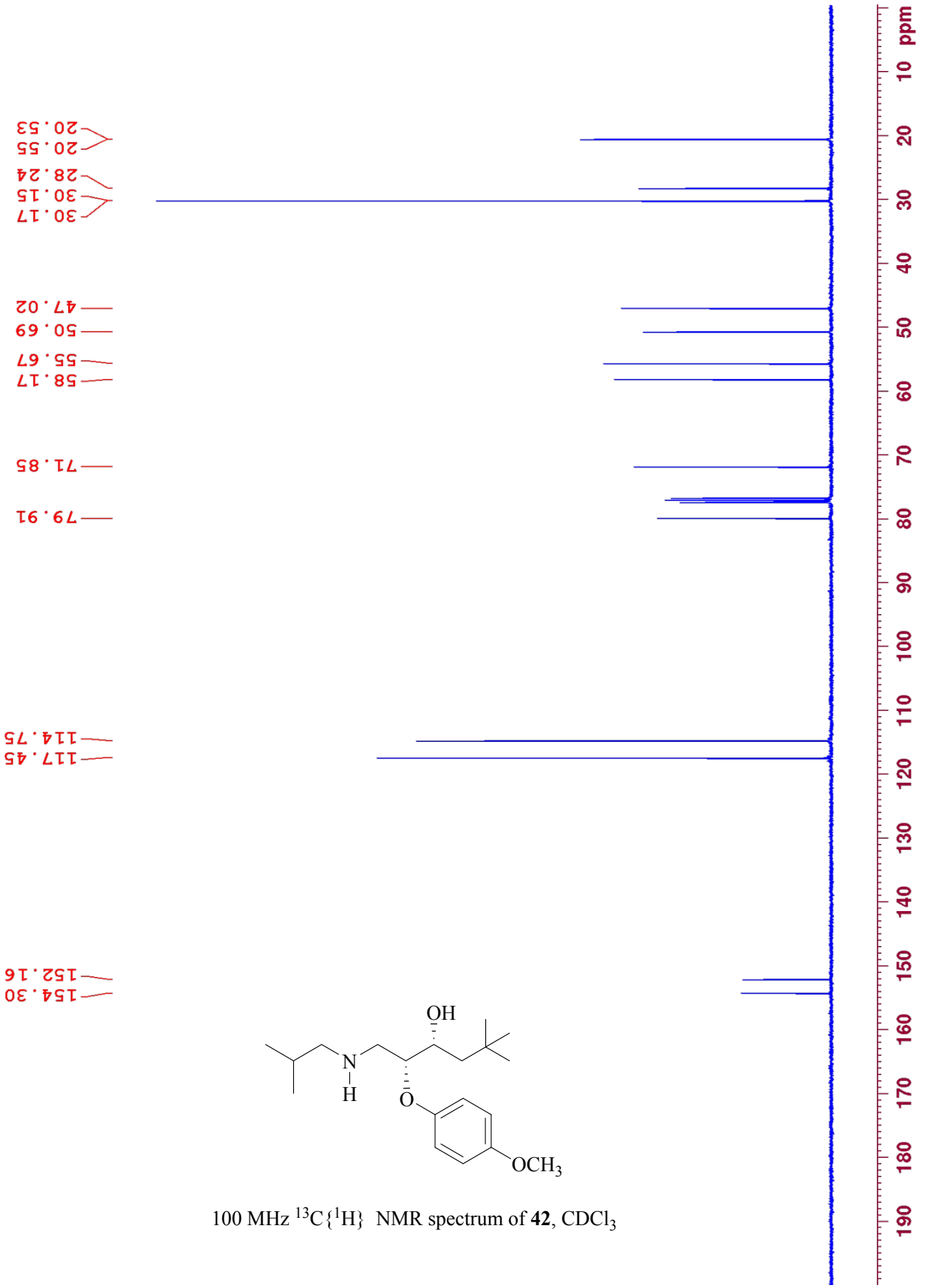
500 MHz ^{1}H NMR spectrum of **41**, $CDCl_3$



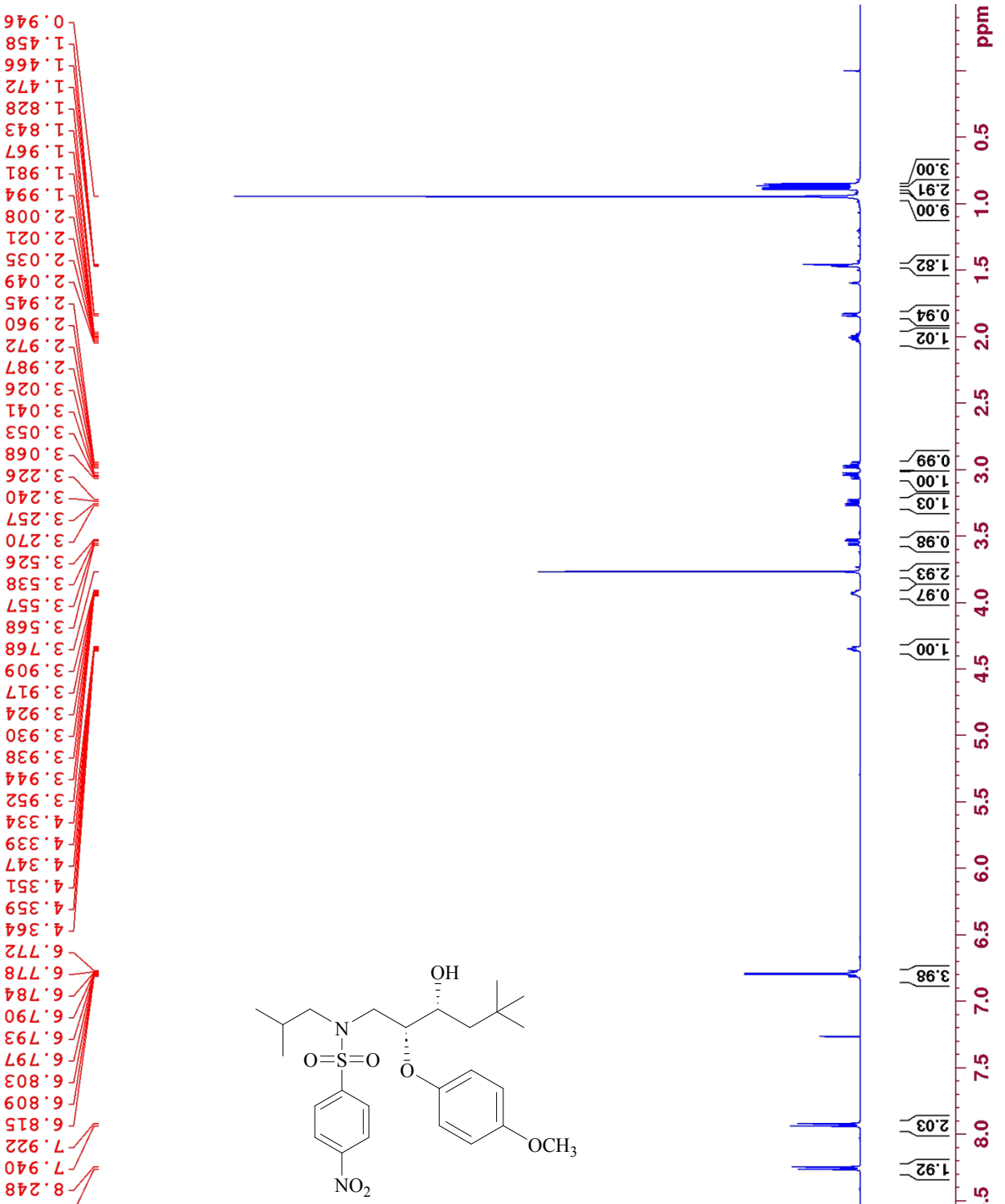
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **41**, $CDCl_3$



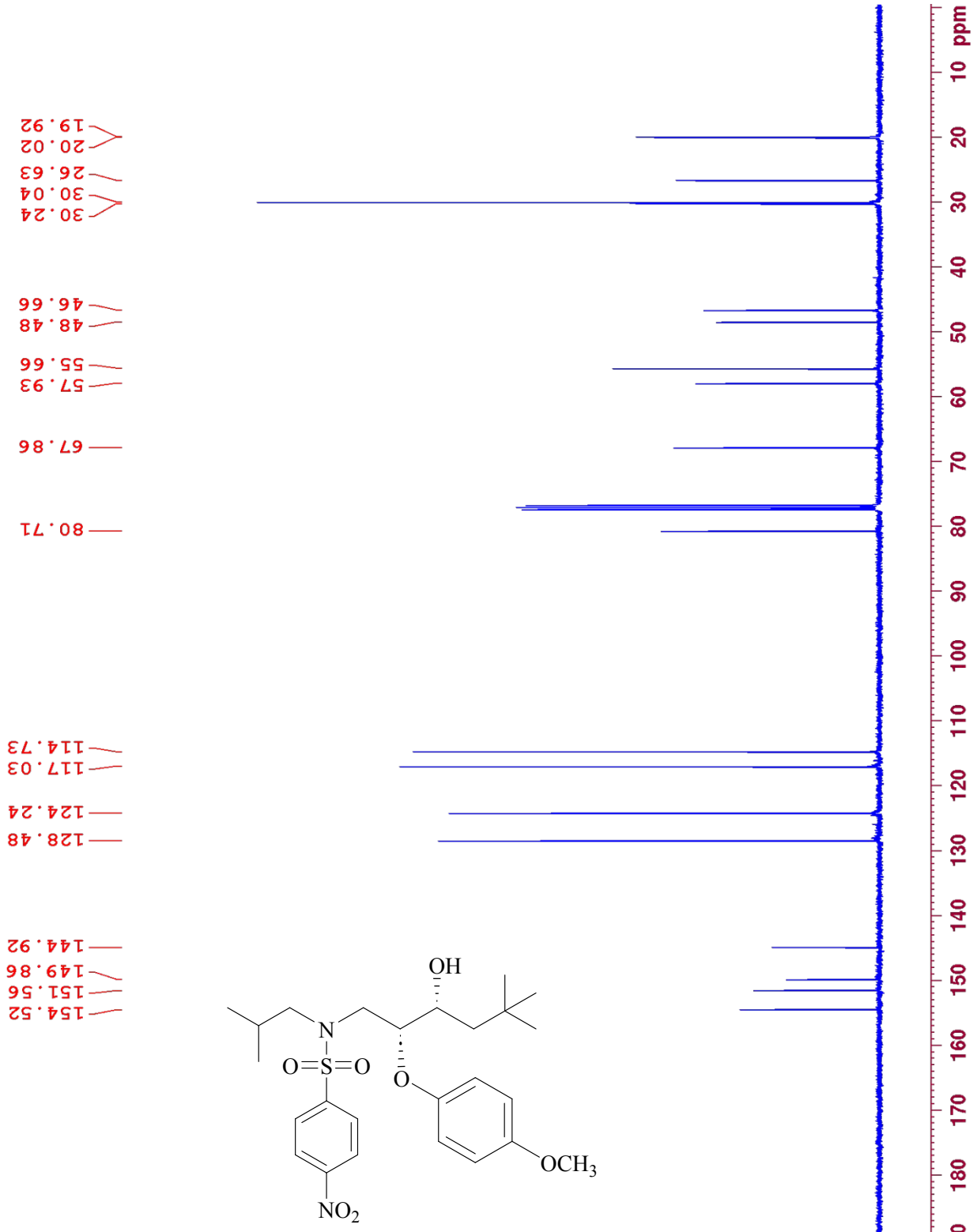
500 MHz ^{1}H NMR spectrum of **42**, $CDCl_3$



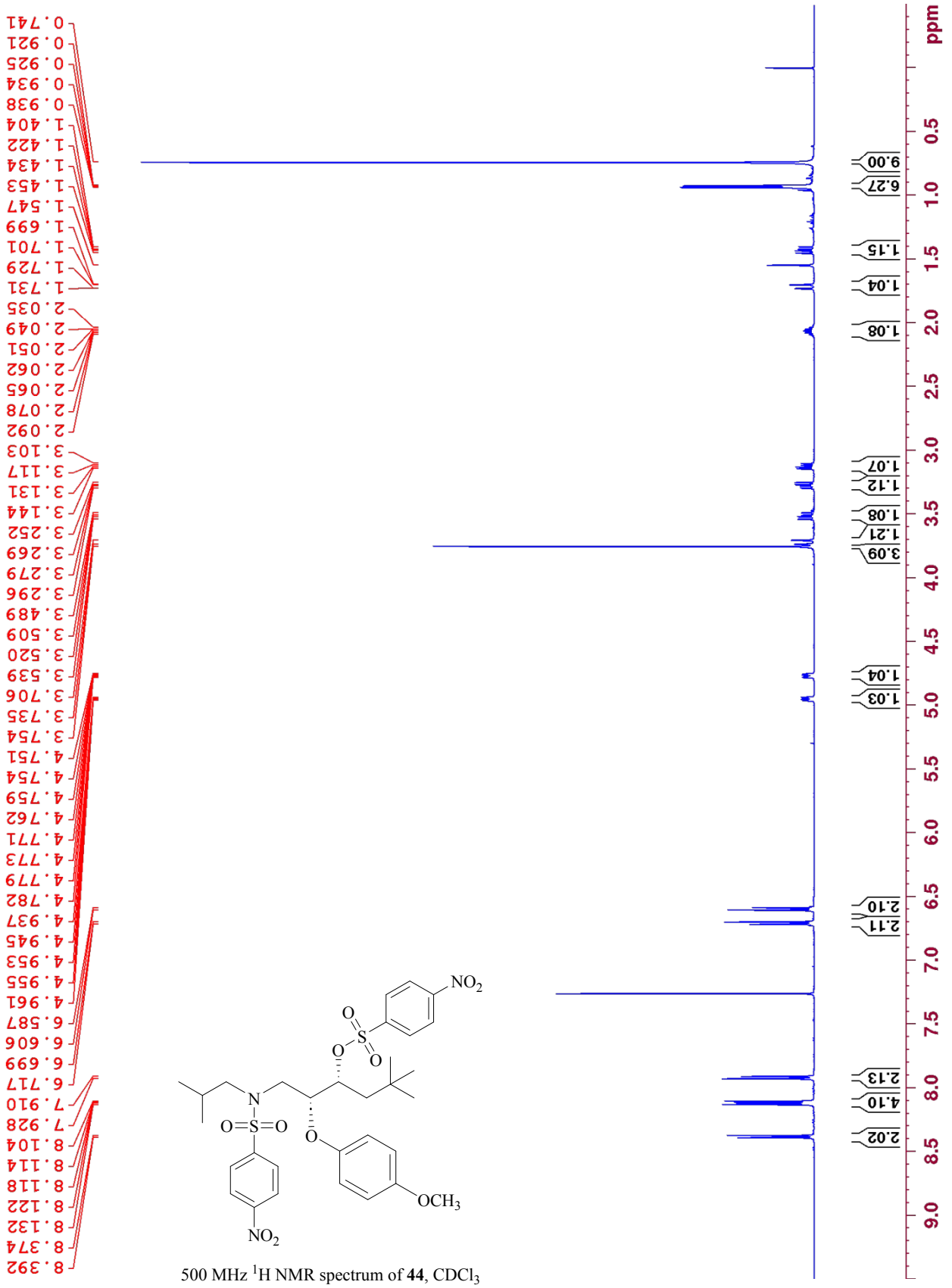
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **42**, $CDCl_3$



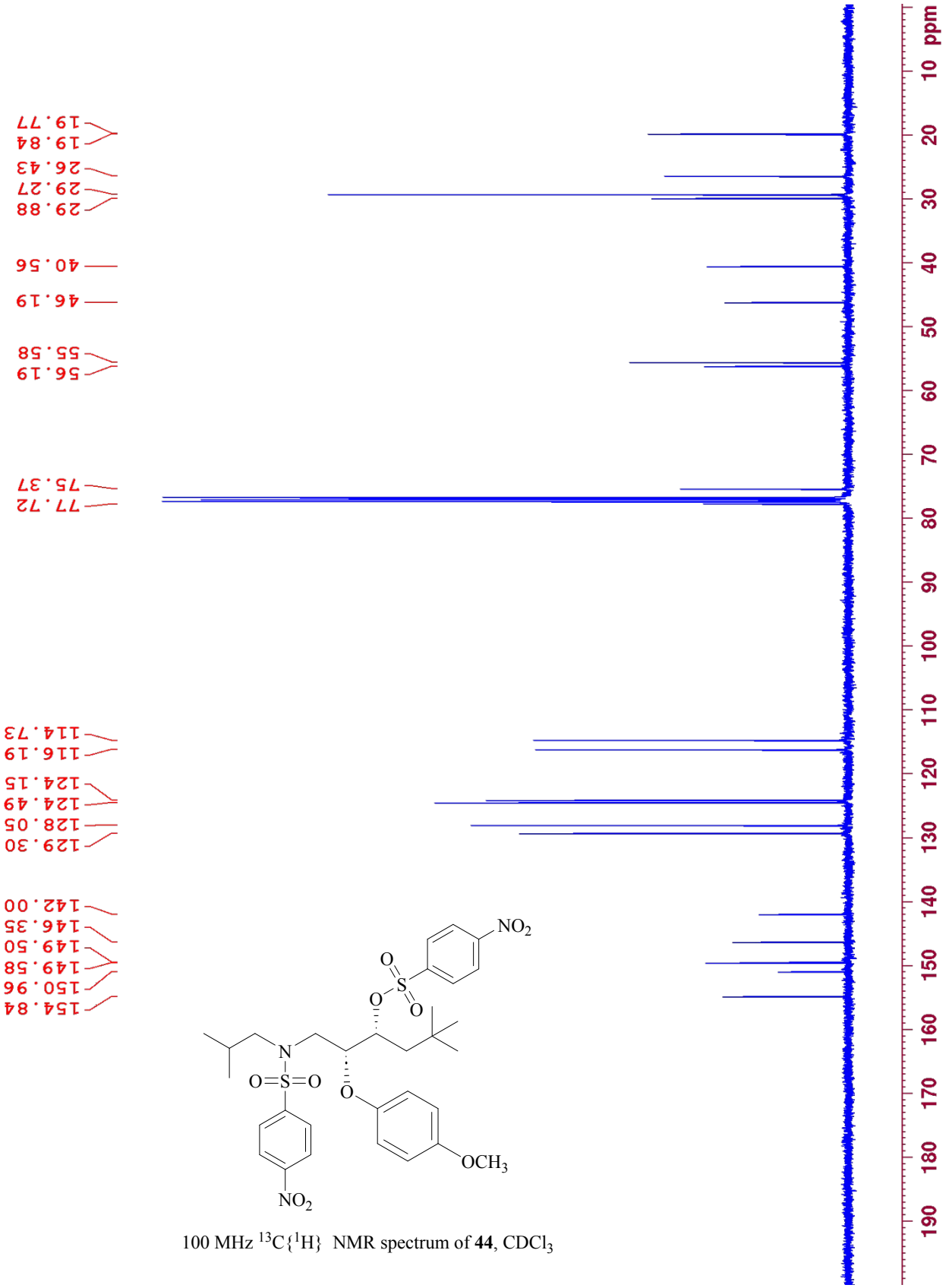
500 MHz ^{1}H NMR spectrum of **43**, $CDCl_3$



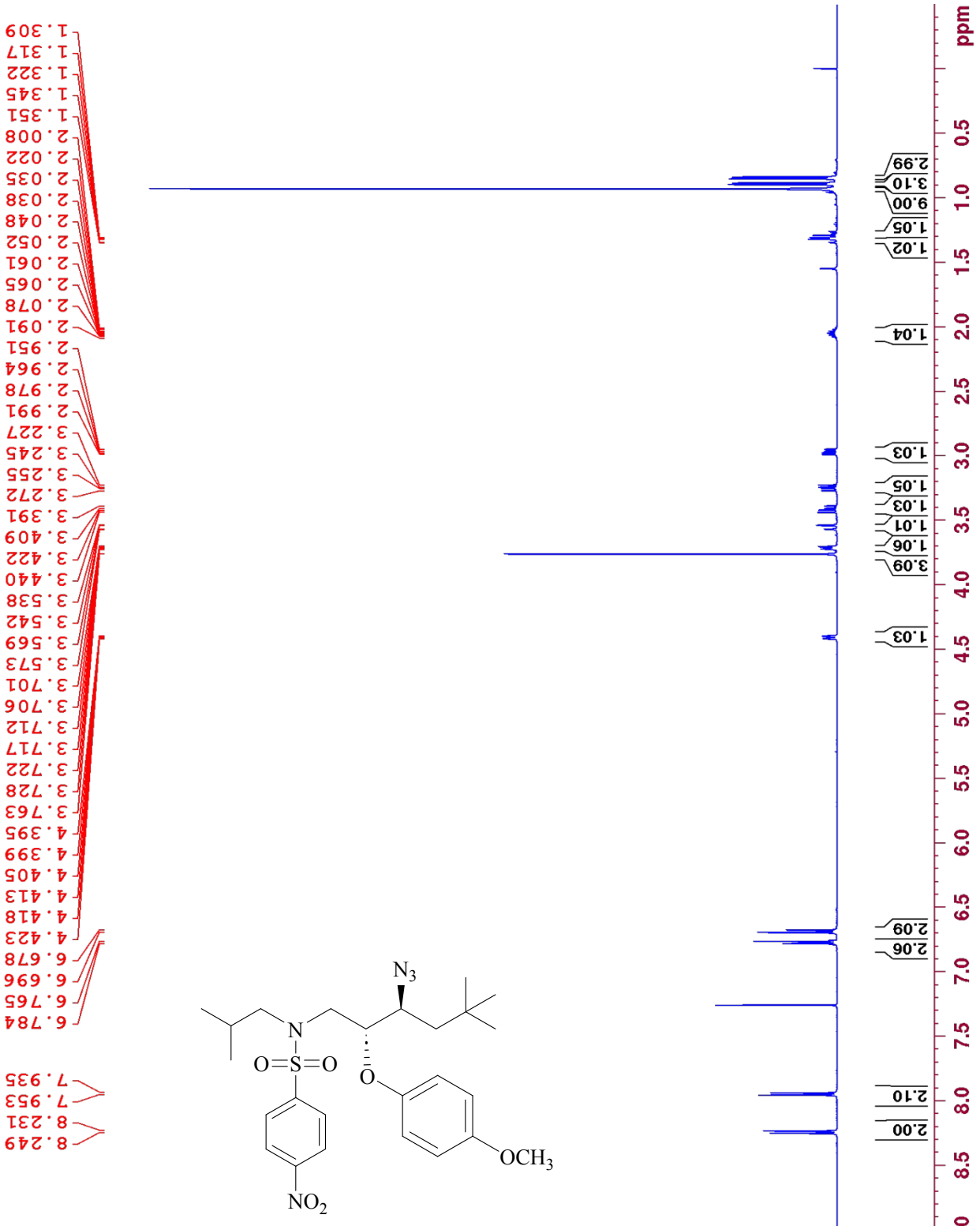
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **43**, $CDCl_3$



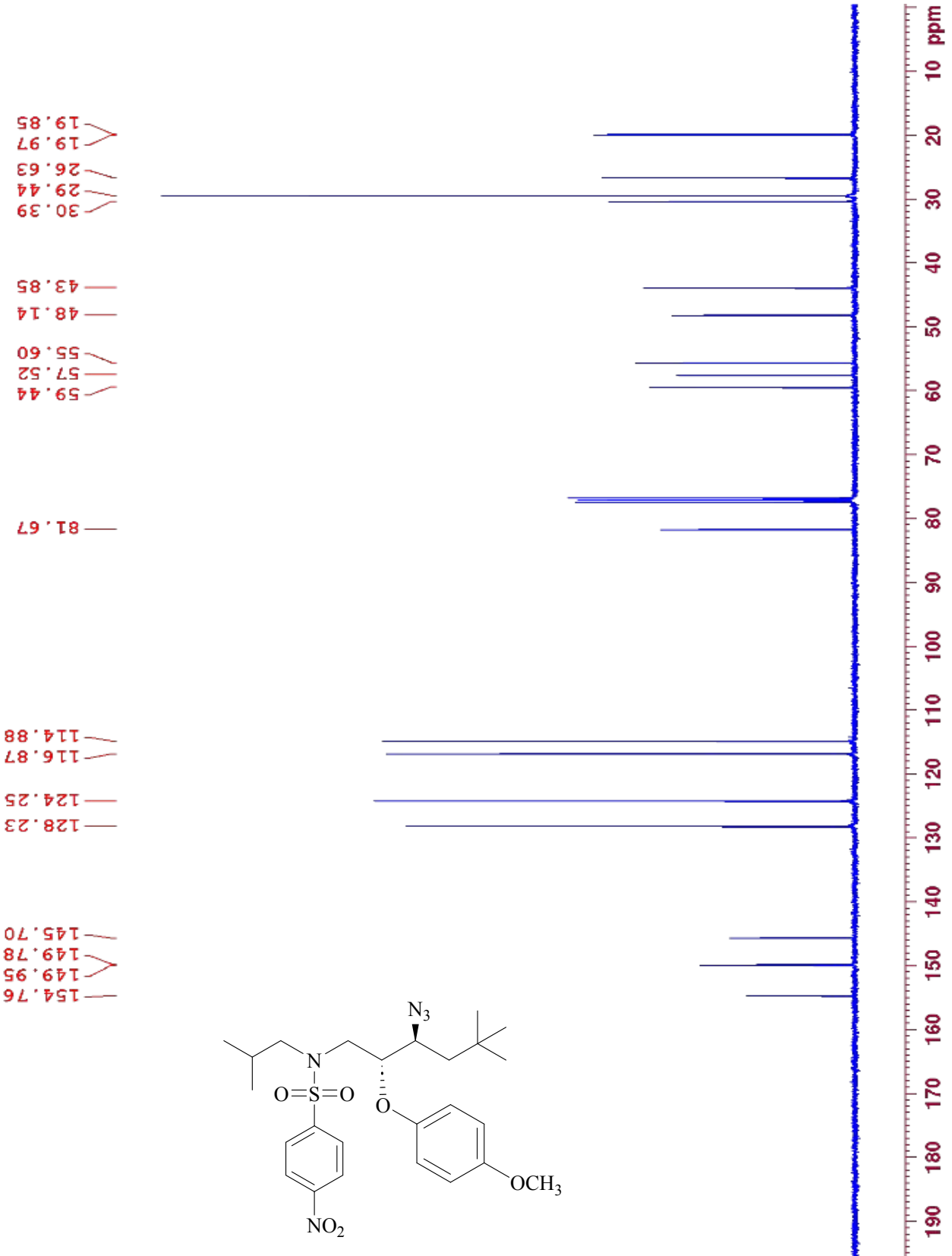
500 MHz ^{1}H NMR spectrum of **44**, $CDCl_3$



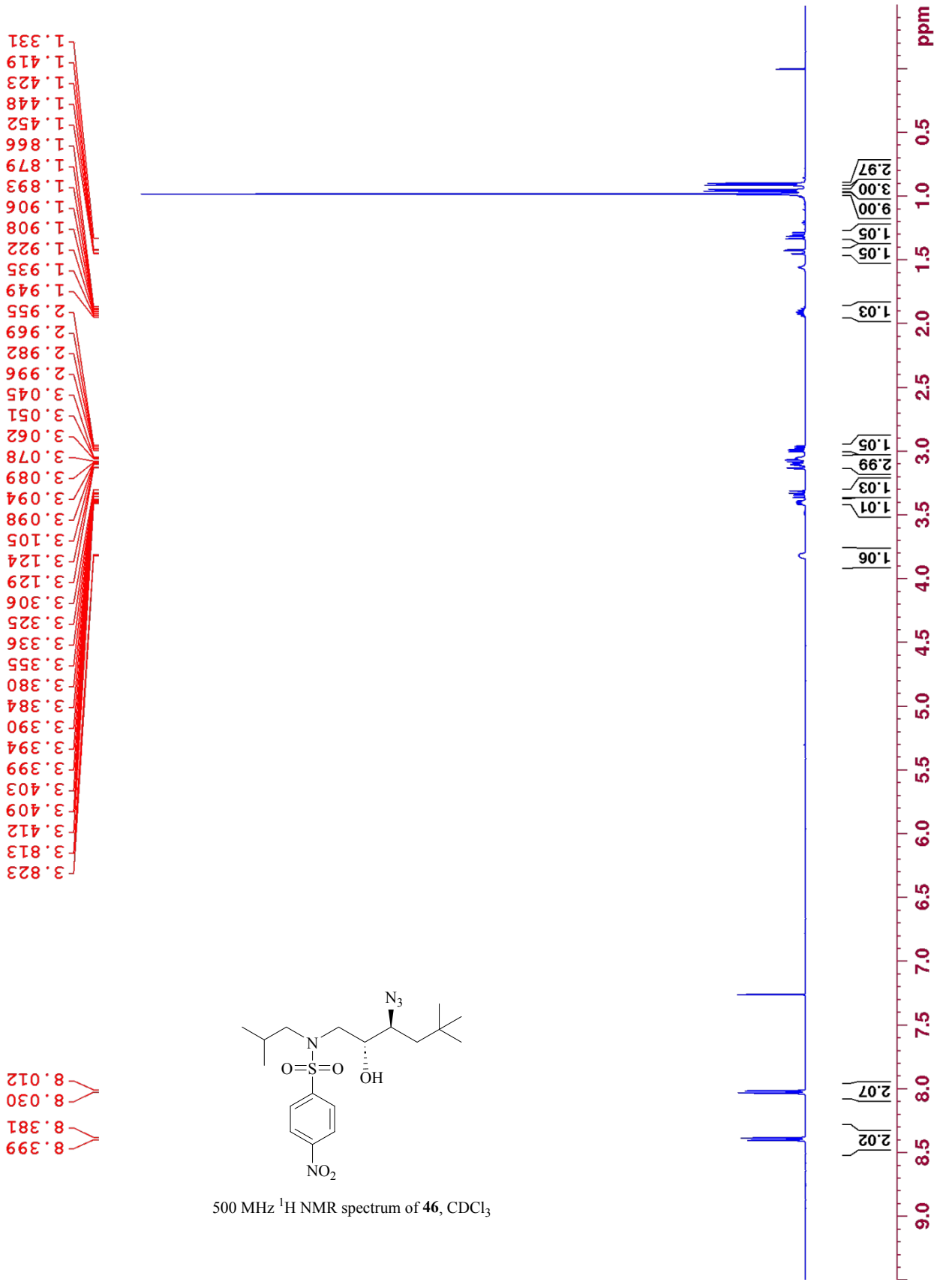
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **44**, $CDCl_3$



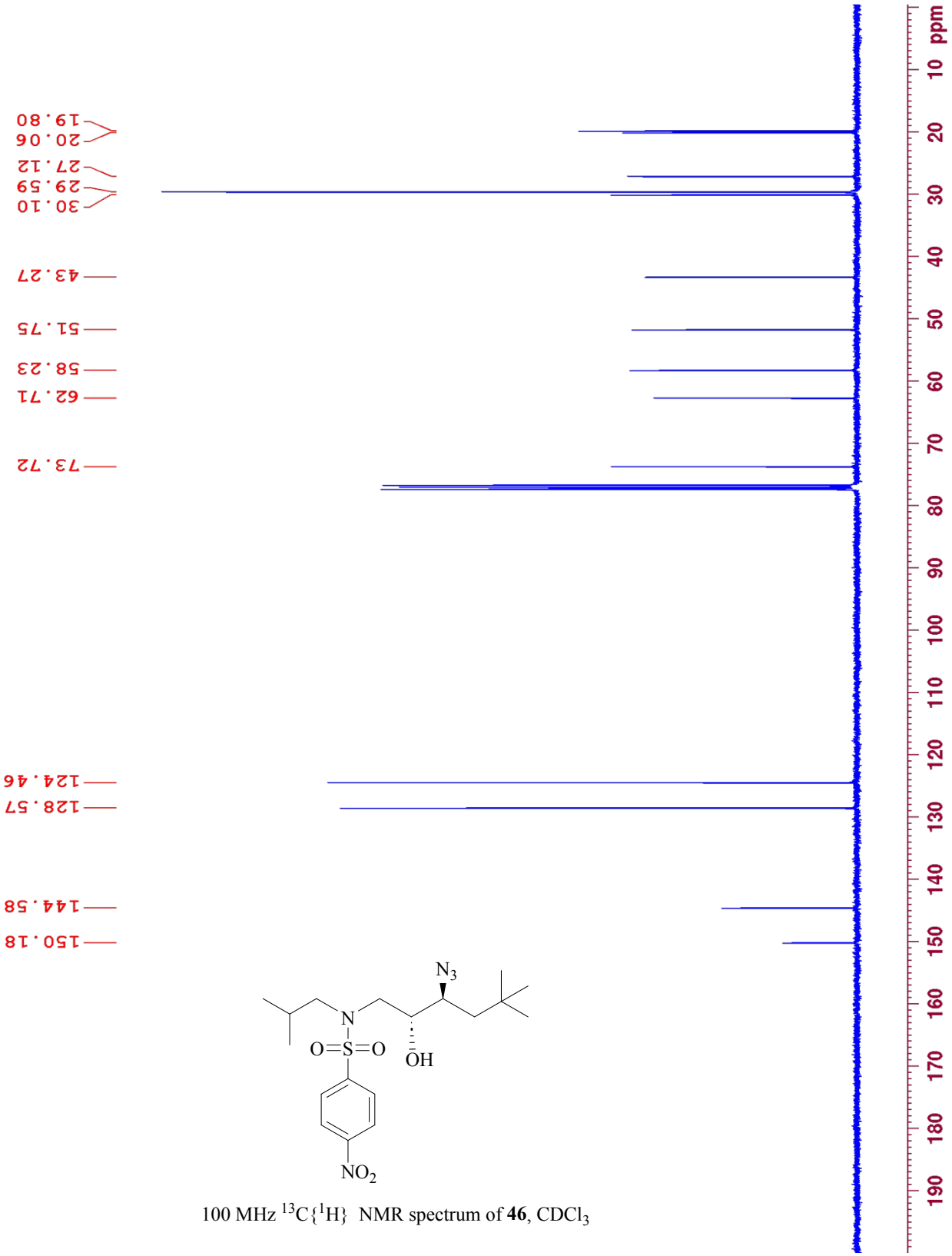
500 MHz ^{1}H NMR spectrum of **45**, $CDCl_3$



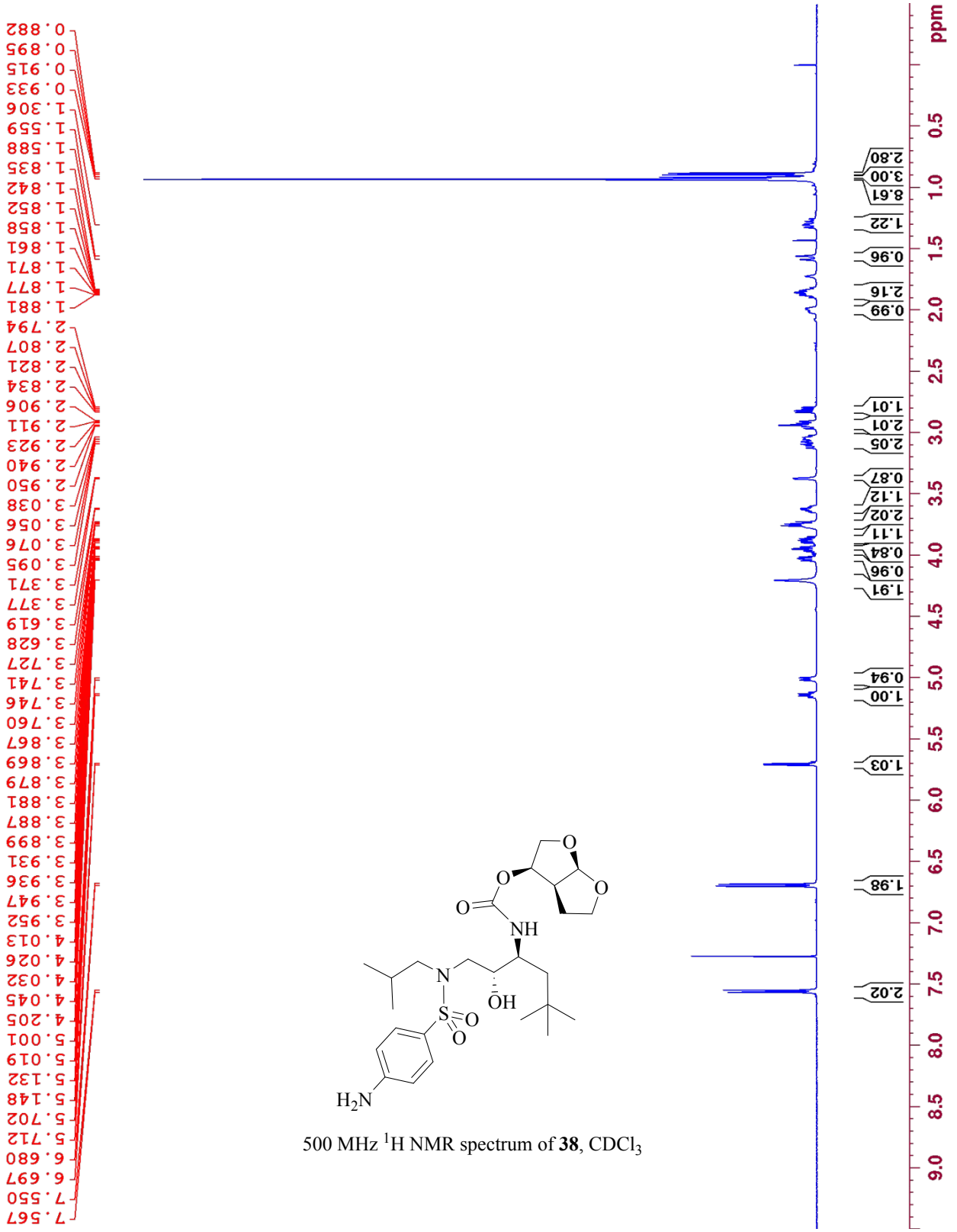
100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **45**, $CDCl_3$



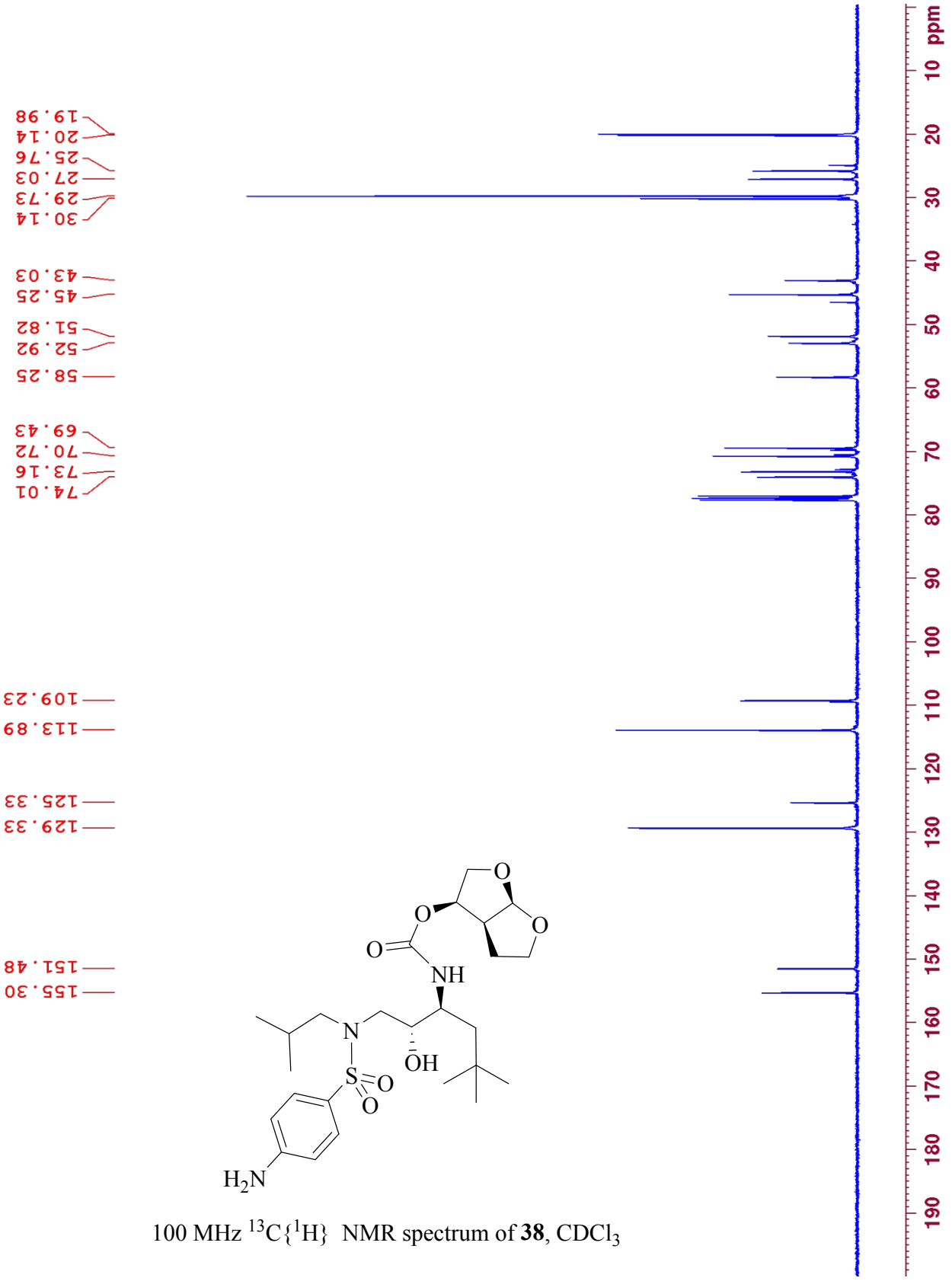
500 MHz ^{1}H NMR spectrum of **46**, $CDCl_3$



100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **46**, $CDCl_3$



500 MHz ^{1}H NMR spectrum of **38**, $CDCl_3$



100 MHz $^{13}C\{^{1}H\}$ NMR spectrum of **38**, $CDCl_3$

General Crystallographic Experimental Details of (*R*,*S*,*R*)-10

Samples of X-ray quality crystals were suspended in mineral oil at ambient temperature and a suitable crystal was selected, mounted on a MiTeGen Micromount and transferred to a Bruker AXS SMART APEXII CCD X-ray diffractometer. The X-ray diffraction data were collected at 100(2) K using Mo-Kα (λ = 0.71073 Å) radiation. The frames were integrated with the Bruker SAINT software package using a narrow-frame algorithm.[1] Data were corrected for absorption effects using the Multi-Scan method (SADABS).[1] Structure were solved and refined using the Bruker SHELXTL Software Package,[2] Molecular diagrams were generated using Mercury.[3]

Specifically, slow vapor diffusion of pentane into a solution of **10** dissolved in ethyl acetate led to formation of SCXRD quality crystals within 24 hours. A mineral oil coated clear colourless plate-like specimen of **10** ($C_{26}H_{25}NO_5S$), with approximate dimensions 0.122 mm x 0.273 mm x 0.312 mm, was harvested and placed in the 100(2) K cryogenic stream on the goniometer. The X-ray intensity data were measured. A total of 3672 frames were collected. The total exposure time was 4.08 hours. The integration of the frame data using a triclinic unit cell yielded a total of 45216 reflections to a maximum θ angle of 33.10° (0.65 Å resolution), of which 7908 were independent (average redundancy 5.718, completeness = 95.8%, R_{int} = 5.12%, R_{sig} = 4.41%) and 6792 (85.89%) were greater than $2\sigma(F^2)$. The final cell constants of a = 10.5787(3) Å, b = 10.2358(3) Å, c = 10.8871(3) Å, β = 109.878(2)°, volume = 1108.63(6) Å^3, are based upon the refinement of the XYZ-centroids of 7908 reflections above 20 $\sigma(I)$ with 4.638° < 2θ < 59.13°. Data were corrected for absorption effects. The ratio of minimum to maximum apparent transmission was 0.938. The calculated minimum and maximum transmission coefficients (based

on crystal size) are 0.9440 and 0.97800. The structure was solved and refined using the space group $P2_1$, with $Z = 2$ for the formula unit, $C_{26}H_{25}NO_5S$. All non-H atoms were refined anisotropically. All H atoms were identifiable in the difference Fourier, but with the exception of the H atom attached to O, were included in the final refinement using the riding-model approximation. d(C--H) = 0.93 Å, $U_{iso} = 1.2U_{eq}$(C) for aromatic, 0.97 Å , $U_{iso} = 1.2\ U_{eq}$(C) for methylene, 0.98 Å, $U_{iso} = 1.2\ U_{eq}$(C) for methine, and 0.96 Å, $U_{iso} = 1.5\ U_{eq}$(C) for methyl H atoms. Coordinates of the H atom attached to the O atom were freely refined with $U_{iso} = 1.5U_{eq}$(O). The final anisotropic full-matrix least-squares refinement on F^2 with 302 variables converged at $R_1 = 4.01\%$, for the observed data and $wR_2 = 9.35\%$ for all data. The goodness-of-fit was 1.026. The largest peak in the final difference electron density synthesis was 0.545 $e^-/Å^3$ and the largest hole was a – 0.393 $e^-/Å^3$ with an RMS deviation of 0.053 $e^-/Å^3$. On the basis of the final model, the calculated density was 1.389 g/cm^3 and F(000), 488 e^-. The absolute configuration was confirmed using 2811 quotients in determination of a Flack parameter of 0.10(2). All residual electron density was within accepted norms and was deemed of no chemical significance.

CCDC-2285488 contains the supplementary crystallographic data for this structure. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre via https://summary.ccdc.cam.ac.uk/structure-summary?ccdc=2285488.

1. Bruker, *APEX4 v2021.10-0*, Bruker AXS Inc., Madison, Wisconsin, USA, 2021.

2. Sheldrick, G. M. Crystal Structure Refinement with SHELXL. *Acta Cryst.* **2015**, *C71*, 3-8.

3. Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek,P. T. A.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M.; Wood, P. A. Mercury 4.0: from Visualization to Analysis, Design and Prediction. *J. Appl. Cryst.* **2020**, *53*, 226-235.

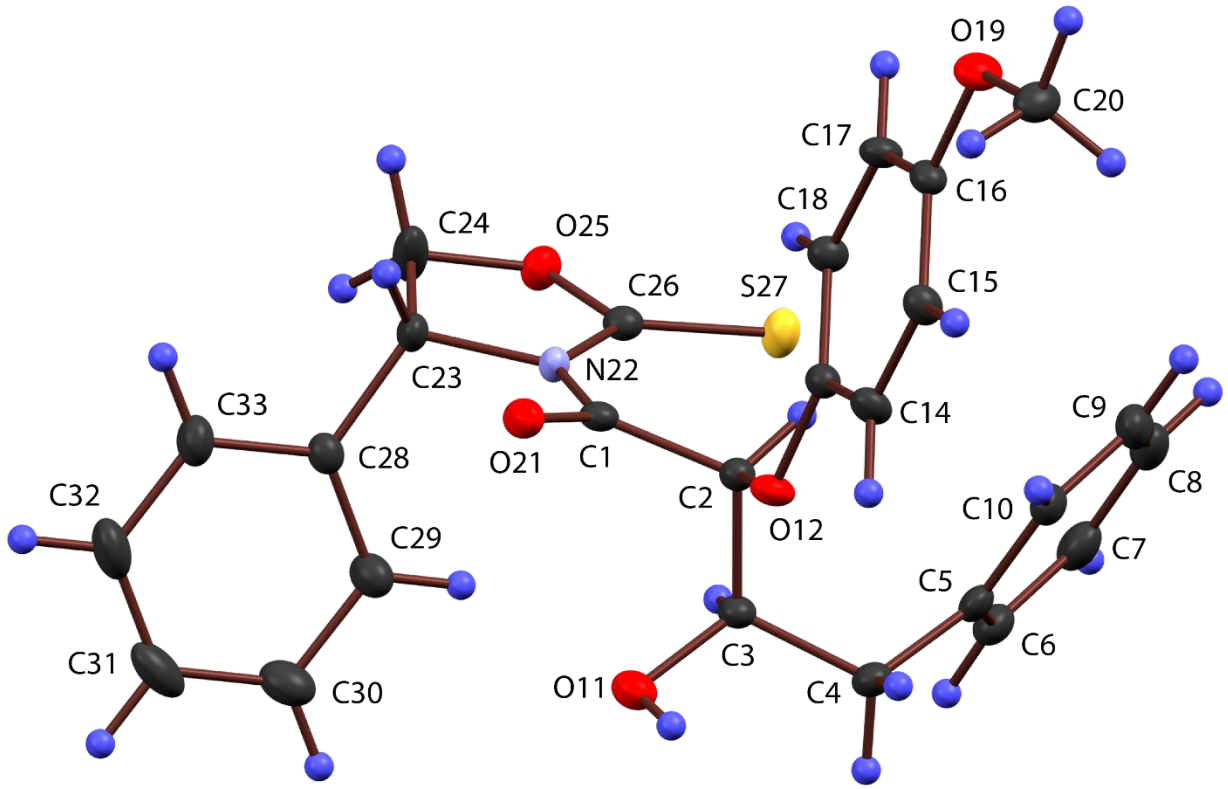


Figure S-1: X-ray crystal structure ORTEP rendering of **10**. The ellipsoids of all non-hydrogen atoms are shown at the 50% probability level; hydrogen atoms are drawn arbitrarily small.