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

Supplementary Data

Synthesis, structure and midkine binding of chondroitin sulfate oligosaccharide analogues

Myriam Torres-Rico, Susana Maza, José L. de Paz*, Pedro M. Nieto*

Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), cicCartuja, CSIC and Universidad de Sevilla, Americo Vespucio, 49, 41092 Sevilla, Spain.

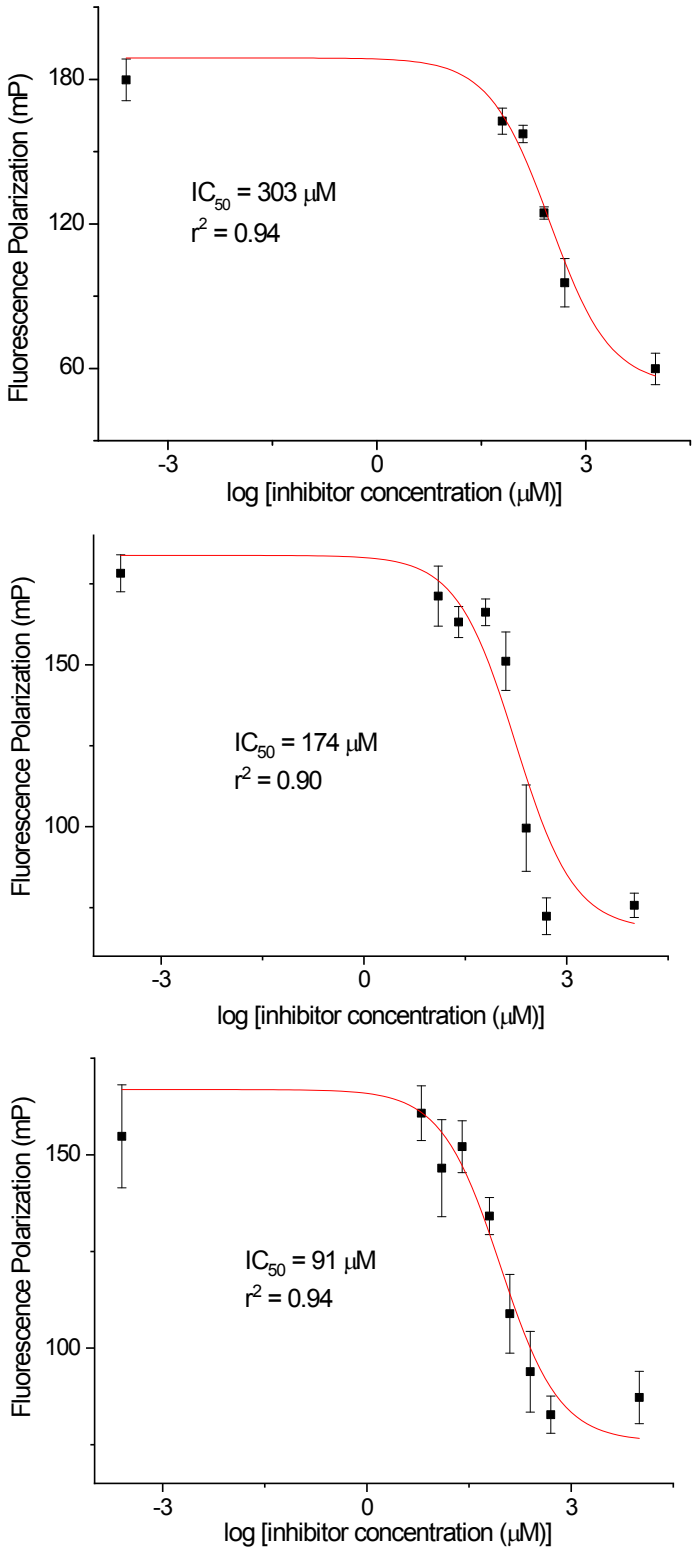


Figure S1. Inhibition curves showing the ability of compound **1** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (189 ± 107 μM) represent the average and the standard deviation from these three independent experiments.

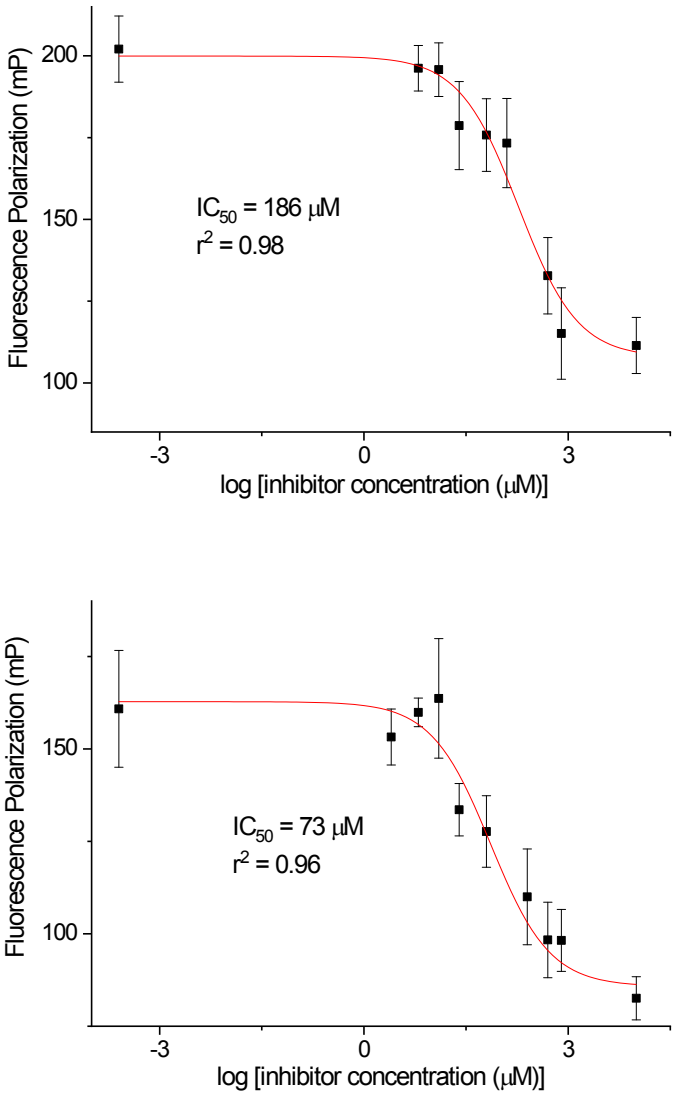


Figure S2. Inhibition curves showing the ability of compound **2** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (130 ± 80 μM) represent the average and the standard deviation from these two independent experiments.

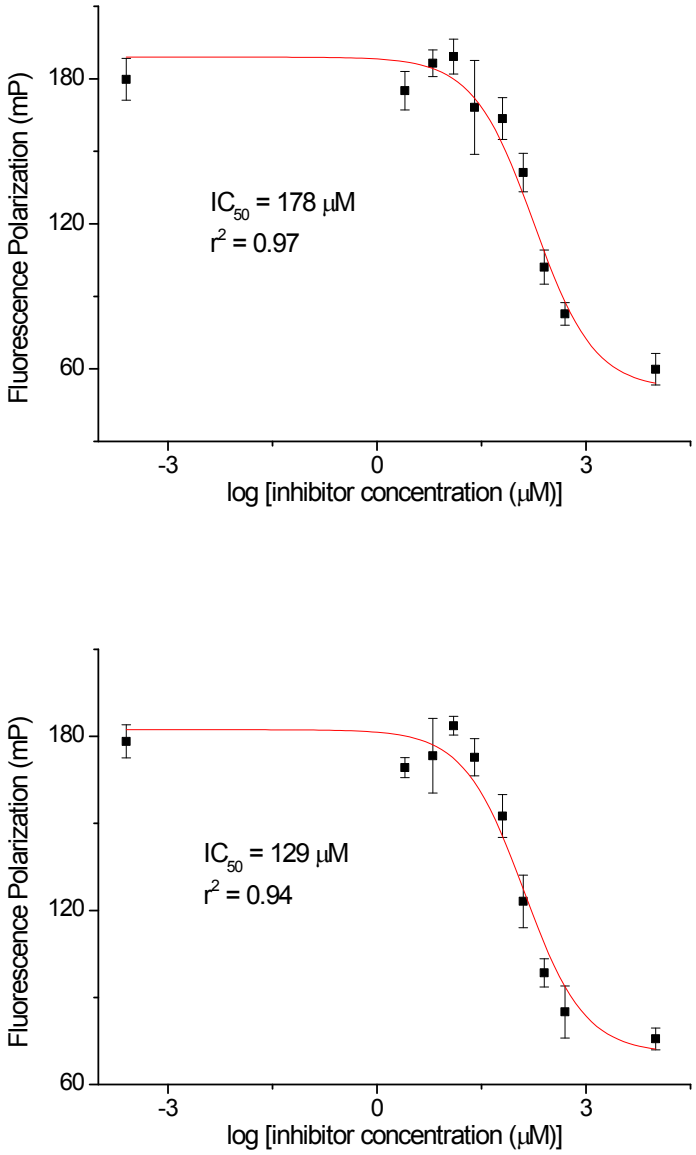


Figure S3. Inhibition curves showing the ability of compound **3** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (154 ± 35 μM) represent the average and the standard deviation from these two independent experiments.

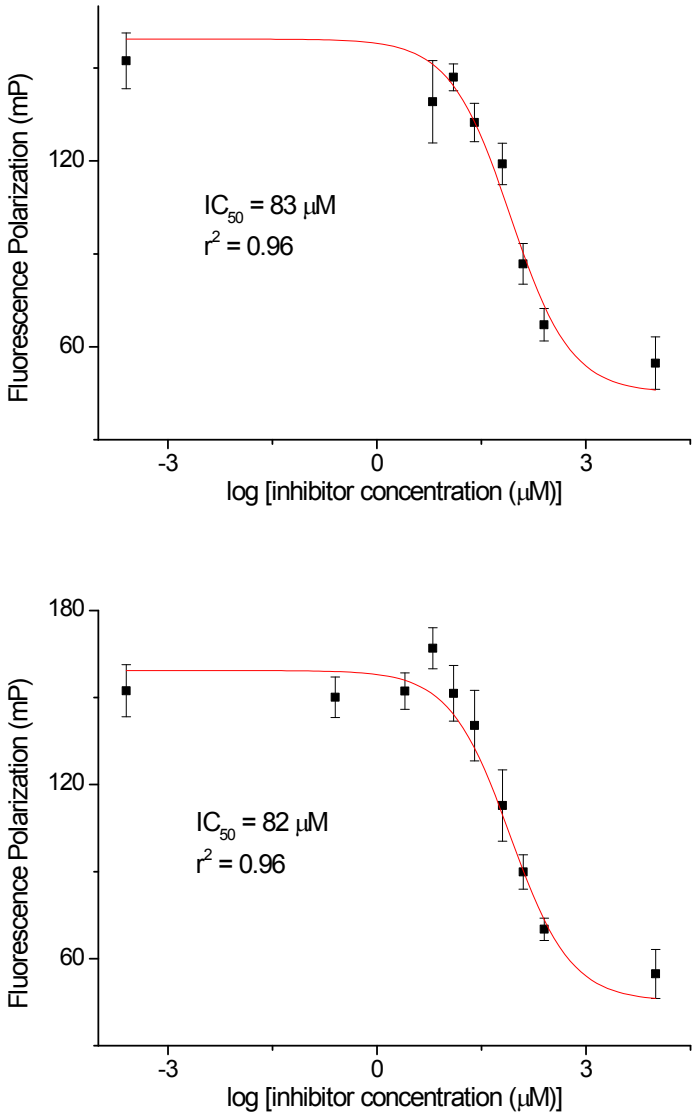


Figure S4. Inhibition curves showing the ability of compound **4** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (83 ± 1 μM) represent the average and the standard deviation from these two independent experiments.

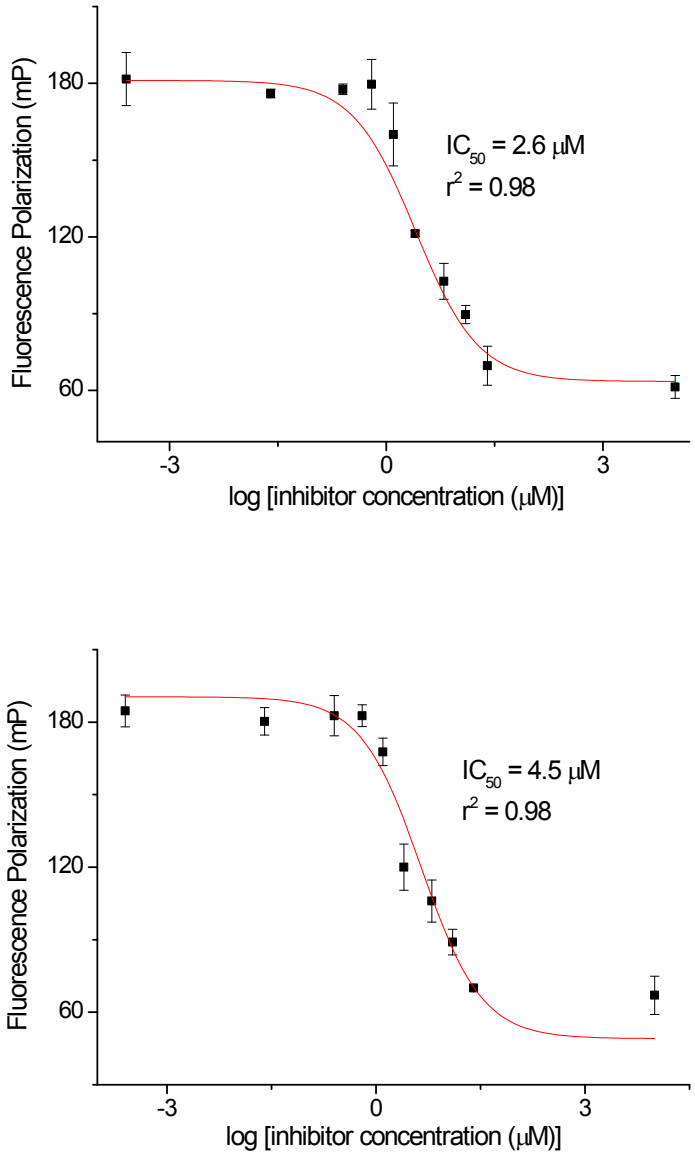


Figure S5. Inhibition curves showing the ability of compound **5** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (3.6 ± 1.3 μM) represent the average and the standard deviation from these two independent experiments.

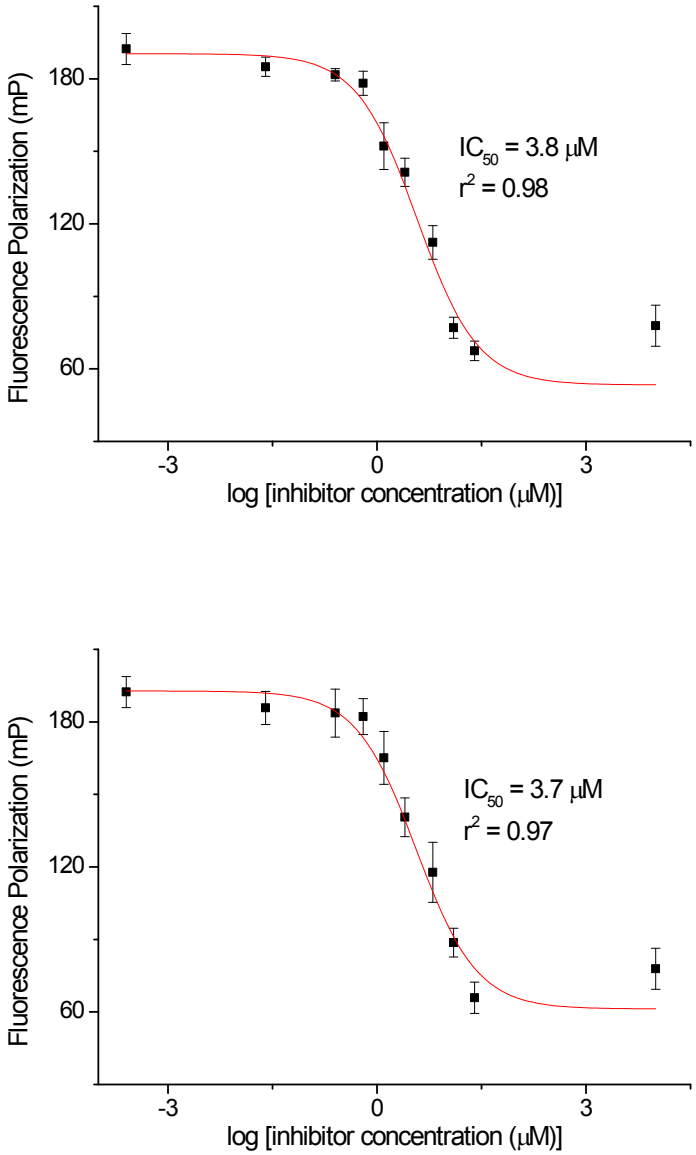
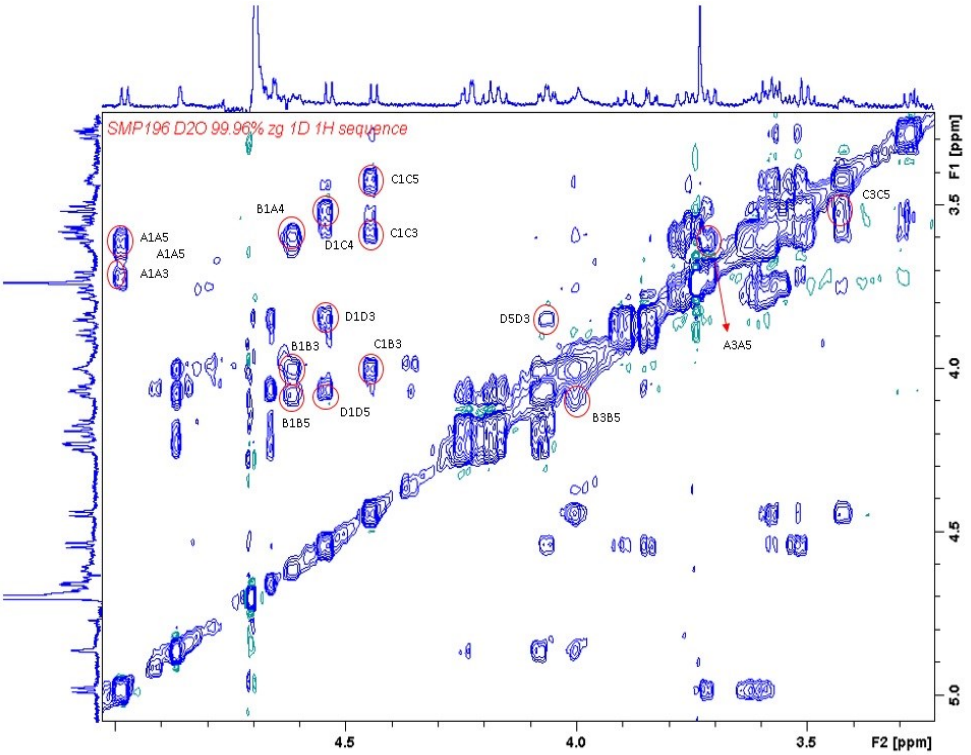
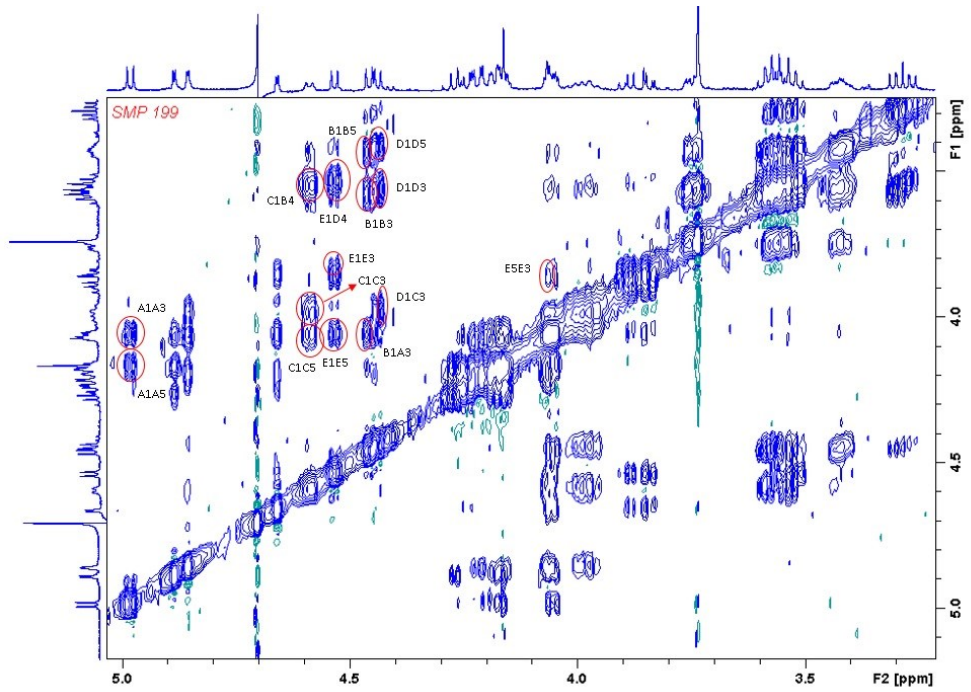


Figure S6. Inhibition curves showing the ability of compound **6** to inhibit the interaction between midkine (63 nM) and fluorescent probe (10 nM). All the FP values are the average of three replicate wells, with error bars showing the standard deviations for these measurements. The reported IC_{50} value and the error (3.8 ± 0.1 μM) represent the average and the standard deviation from these two independent experiments.

Tetra **1**



Penta **2**



Penta **4**

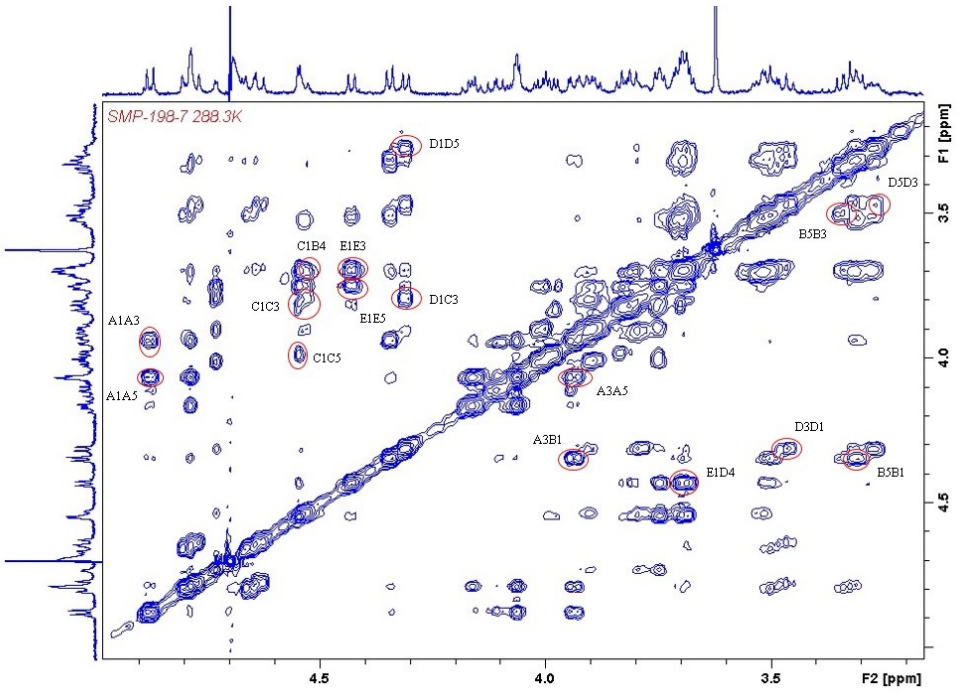


Figure S7. NOESY spectra of oligosaccharides **1**, **2** and **4** at 1.5 mM, 300K and 600MHz in D_2O (mixing time 1000 ms for **1**; 600 ms for **2** and **4**).

$^3J_{HH}$ coupling	1 NMR exp 600 MHz	1 MD 500ns	3 NMR exp 600 MHz	3 MD 250ns	3 MD-tar 8ns	5 MD 250ns	5 MD-tar 8ns
A1A2	8.4	7.3	8.4	7.0	7.5	3.2	7.8
A2A3	10.7	8.6	9.5	9.7	9.8	3.6	9.6
A3A4		8.6		8.9	8.9	3.6	8.9
A4A5		10.2		2.6	8.6	3.6	9.6
B1B2		8.2	8.9	8.4	8.5	8.4	8.2
B2B3	8.9	10.6	11.4	10.3	10.5	10.6	10.4
B3B4	3.1	2.7	2.9	2.8	2.8	2.7	2.9
B4B5	3.3	2.0		1.0	0.9	1.1	1.1
C1C2	8.0	7.3	7.8	2.6	7.6	3.6	7.8
C2C3		8.6	8.0	3.6	9.6	2.8	9.7
C3C4	8.4	8.6	9.2	2.8	8.8	3.3	9.3
C4C5		10.1		5.2	9.5	3.7	9.7
D1D2		8.2	8.2	8.4	8.4	8.5	8.3
D2D3	8.5	10.3		10.4	10.4	10.6	10.5
D3D4		2.7		3.0	2.9	3.1	3.0
D4D5	2.9	1.9		1.2	1.2	1.4	1.0

Table S1. $^3J_{HH}$ coupling constants (Hz) determined by NMR experiments (NMR exp) or calculated from MD or MD-tar simulations for tetrasaccharides **1**, **3** and **5**.

Interprotonic distances	1 NMR exp 600 MHz	1 MD 500ns	3 NMR exp 600 MHz	3 MD 250ns	3 MD-tar 8ns	5 MD 250ns	5 MD-tar 8ns
A1A3	2.8	2.8	3.1	2.7	2.7	3.9	2.8
A1A5	2.4	2.5	2.5	2.5	2.8	3.3	2.5
A3A5		2.7		2.6	2.8	3.9	2.7
B1A4	2.4	2.4	2.4	2.6	2.7	2.4	2.6
B1B3	2.5	2.7	3.1	2.7	2.7	2.5	2.7
B1B5	2.9	2.6		2.5	2.5	2.5	2.6
B3B5	2.6	2.5		2.5	2.5	2.4	2.5
C1B3	2.3	2.4	2.6	2.3	2.5	2.4	2.4
C1C3	2.8	2.8	2.7	4.0	2.7	3.9	2.8
C1C5	2.5	2.5	2.5	2.6	2.5	2.6	2.5
C3C5		2.7	2.6	3.9	2.7	4.0	2.6
D1C4	2.5	2.8	2.2	2.4	2.7	2.3	2.5
D1D3	2.7	2.7	2.7	2.7	2.7	2.5	2.7
D1D5	2.6	2.6	2.5	2.6	2.5	2.5	2.5
D3D5	2.6	2.5		2.5	2.6	2.6	2.5

Table S2. Interprotonic distances (Å) determined by NOESY NMR experiments (NMR exp) or calculated from MD or MD-tar simulations for tetrasaccharides **1**, **3** and **5**.

$^3J_{HH}$ coupling	**2** NMR exp 600 MHz	**2** MD 500ns	**4** NMR exp 600 MHz	**4** MD 500ns	**4** MD-tar 50ns	**6** MD 250ns
A1A2	8.5	8.2	8.9	8.5	8.5	8.5
A2A3	11.7	10.6	11.0	10.4	10.5	10.5
A3A4	3.0	2.5	2.9	2.7	2.7	2.7
A4A5	1.4	1.2		1.0	1.0	1.0
B1B2	7.9	7.3	8.0	7.5	7.5	7.5
B2B3	8.6	8.6	8.9	9.6	9.3	9.3
B3B4	9.2	8.7	9.1	9.1	9.0	9.0
B4B5	9.5	10.2	7.8	9.7	9.8	9.8
C1C2	8.1	8.1	8.2	8.5	8.5	8.5
C2C3	10.5	10.6		10.5	10.4	10.4
C3C4	2.5	2.5	2.7	2.8	2.8	2.8
C4C5		1.2		1.1	1.1	1.1
D1D2	7.7	7.6	7.8	1.5	7.6	7.6
D2D3	9.1	9.4		3.4	9.4	9.4
D3D4	9.5	9.2	9.2	1.7	9.2	9.2
D4D5	9.3	9.6		6.1	9.7	9.7
E1E2	8.4	8.2	8.3	8.5	8.5	8.5
E2E3	10.8	10.4		10.4	10.5	10.5
E3E4	3.0	2.8	3.0	2.9	3.2	3.2
E4E5		1.3		1.2	1.4	1.4

Table S3. $^3J_{HH}$ coupling constants (Hz) determined by NMR experiments (NMR exp) or calculated from MD or MD-tar simulations for pentasaccharides **2**, **4** and **6**.

Interprotonic distances	2 NMR exp 600 MHz	2 MD 500ns	4 NMR exp 600 MHz	4 MD 500ns	4 MD-tar 50ns	6 MD 250ns
A1A3	2.4	2.6	2.6	2.7	2.7	2.6
A1A5	2.8	2.6	2.7	2.4	2.5	2.5
A3A5		2.5	2.6	2.5	2.5	2.5
B1A3	2.2	2.4	2.4	2.5	2.4	2.4
B1B3	2.8	2.8	2.7	2.8	2.7	2.7
B1B5	2.5	2.5	2.4	2.4	2.4	2.5
B3B5		2.7	2.5	2.6	2.6	2.5
C1B4	2.6	2.3	2.6	2.4	2.5	2.4
C1C3	2.9	2.7	2.8	2.7	2.7	2.6
C1C5	2.5	2.6	2.6	2.4	2.5	2.6
C3C5		2.5		2.6	2.6	2.5
D1C3	2.6	2.5	2.3	2.2	2.4	2.4
D1D3		2.8	2.8	**4.2**	2.7	2.7
D1D5		2.5	2.4	2.6	2.5	2.5
D3D5		2.7	2.9	**3.9**	2.6	2.6
E1D4		2.3	2.3	2.3	2.4	2.4
E1E3		2.7	2.5	2.6	2.7	2.5
E1E5		2.6	2.7	2.5	2.6	2.5
E3E5		2.5		2.5	2.6	2.5

Table S4. Interprotonic distances (Å) determined by NOESY NMR experiments (NMR exp) or calculated from MD or MD-tar simulations for pentasaccharides **2**, **4** and **6**.

Molecular dynamics (MD)

Input preparation: The input files, as topology and coordinates, for molecular dynamics were generated with t*LEAP* module AMBER12[1] package, where each molecule was immersed in the TIP3P water box and equilibrated with sodium ions. For this, it was necessary to have previously all the prep and frcmod files for each of the sugar residues. We have reused most of the parameters from the protocol of previous works.[2] However, it was necessary to calculate the partial charges for the CF_3 group. We used the RESP method, available in the ante-R.E.D 2.0 and R.E.D IV, with a procedure employed in the development of GLYCAM06 force fields and HF/6-31G* as levels of theory for the structure optimization and the molecular electrostatic potential (MEP) calculations[3]. Also, we used the parm99 package for the parametrization beside *Antechamber* and *Parmchk* modules.

Molecular simulation procedure: The molecular dynamics were run on the Carbohidratos cluster, hosted by the *Centro de Investigaciones Científicas Isla de la Cartuja de Sevilla* (CSIC, CicCartuja), Spain. All simulations were carried out with AMBER12 with a protocol similar to previous group work[2]. The procedure consisted in several steps: first, we performed an initial minimization of the oligosaccharide in the water box followed by a minimization of the entire system, including sodium ions. Then, the system was heated from 0K to 300K at constant volume (4000 steps) and equilibrated at a constant pressure (1bar). Finally, the production MD simulations were performed at a constant temperature, pressure and volume and the overall length was 250-500 ns depending on the complexity of the mimetic. When structures presented any distortion, Time-Average Distance Restrained MD simulations (MD-tar) were carried out. These are short simulations (8-50ns) that use NOE-derived distances between H_1-H_3, H_1-H_5, H_3-H_5 as restrictions in the molecule. All dynamics were carried out using the PME (Particle Mesh Ewald Molecular Dynamics)[4] as simulation program with *sander.MPI* or *pmemd.MPI* modules and the results were analyzed by *cpptraj* module of AMBER12.

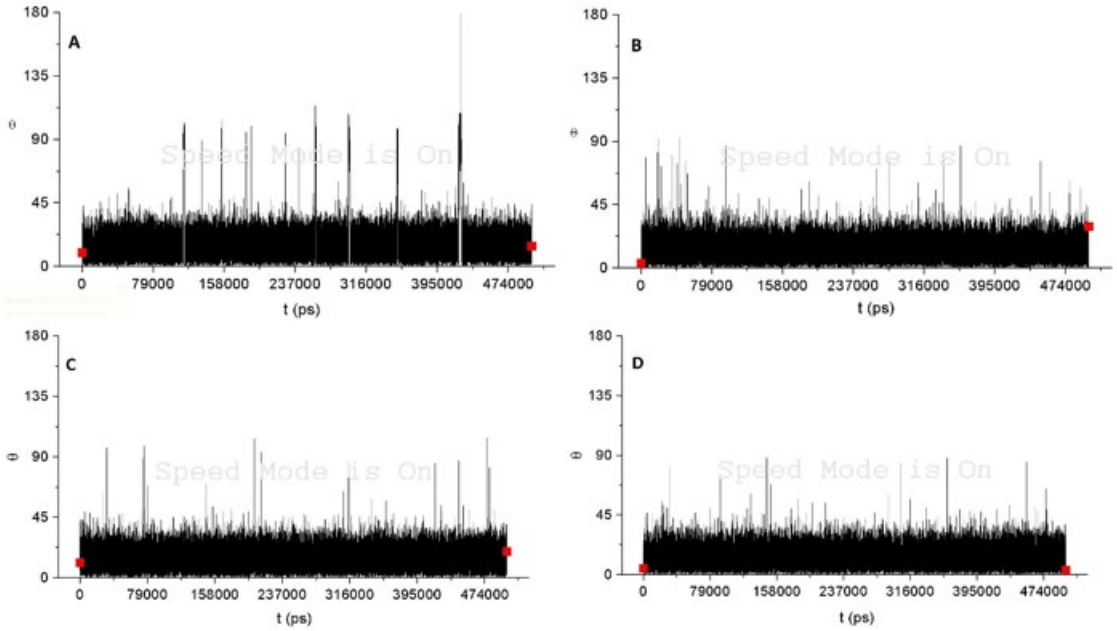


Figure S8. Puckering coordinates for the tetrasaccharide **1** (500 ns MD).

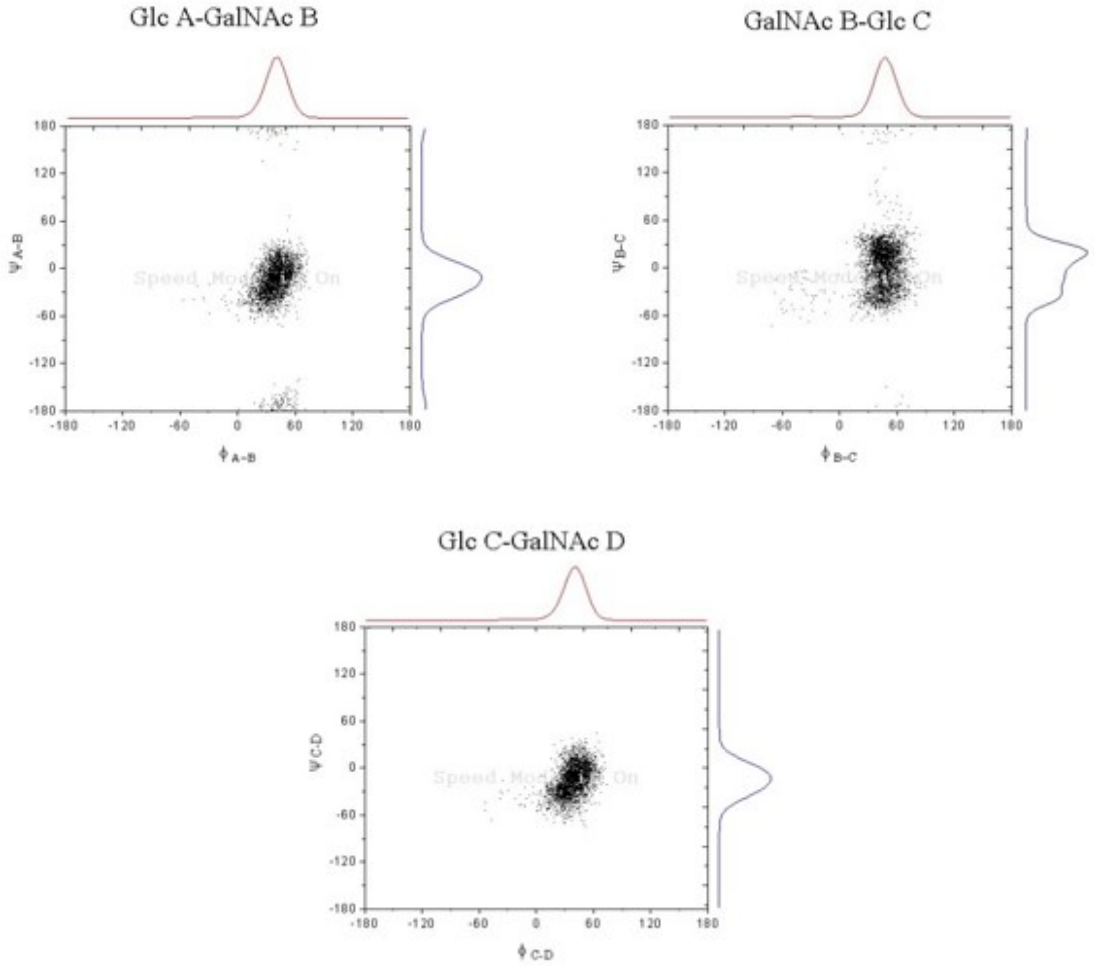


Figure S9. Trajectory (ϕ,ψ) for compound **1** (500 ns MD).

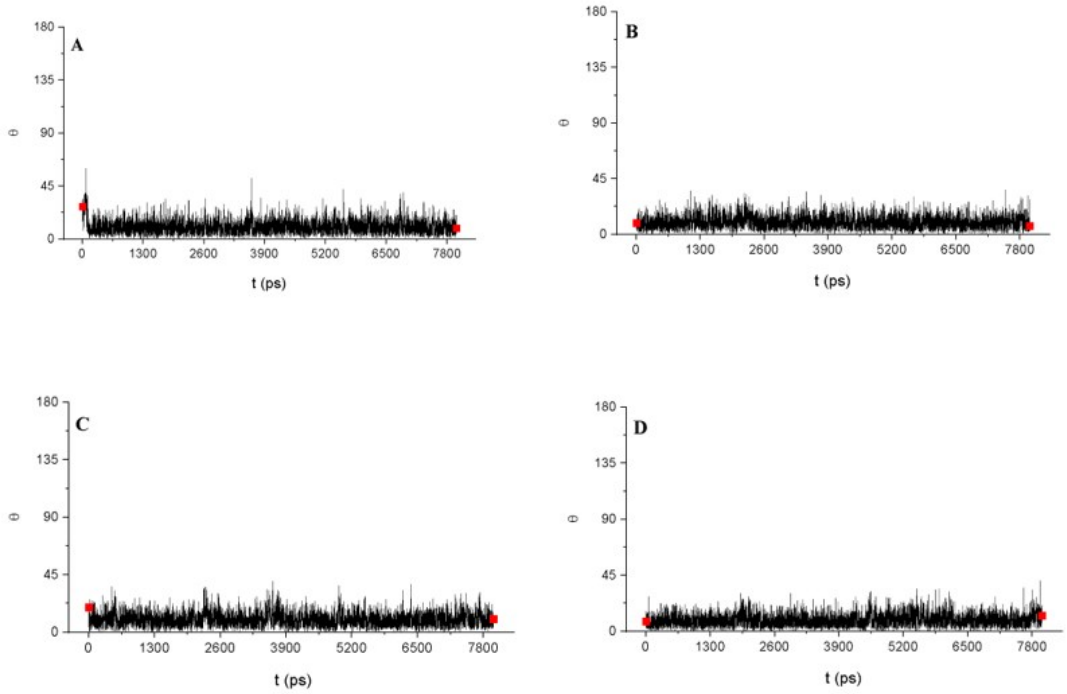
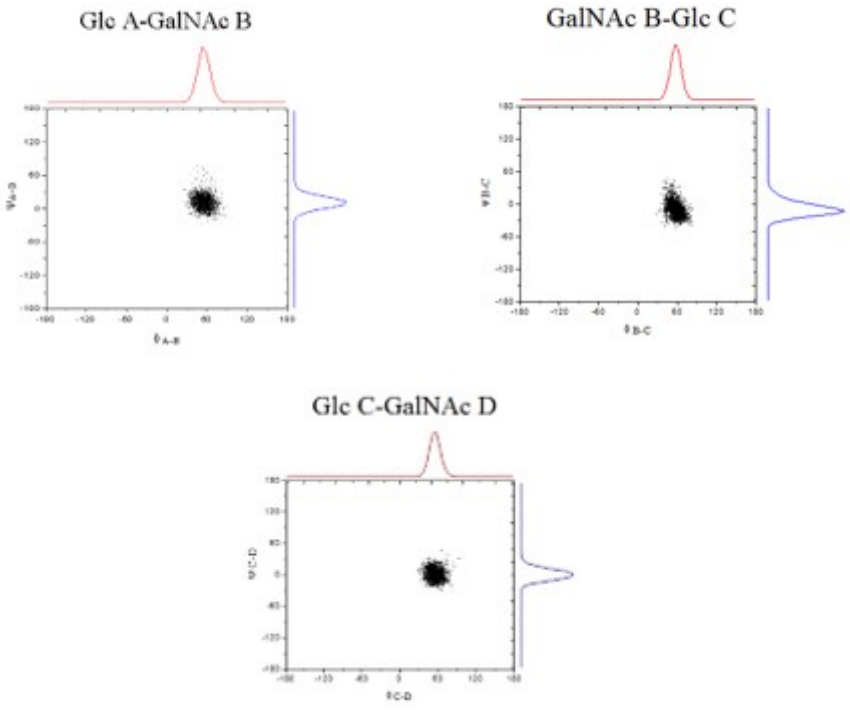


Figure S10. Puckering coordinates for tetrasaccharide **5** along a MD-tar of 8ns.



.

Figure S11. Trajectory (ϕ,ψ) of the MD-tar simulation of 8ns for compound **5**.

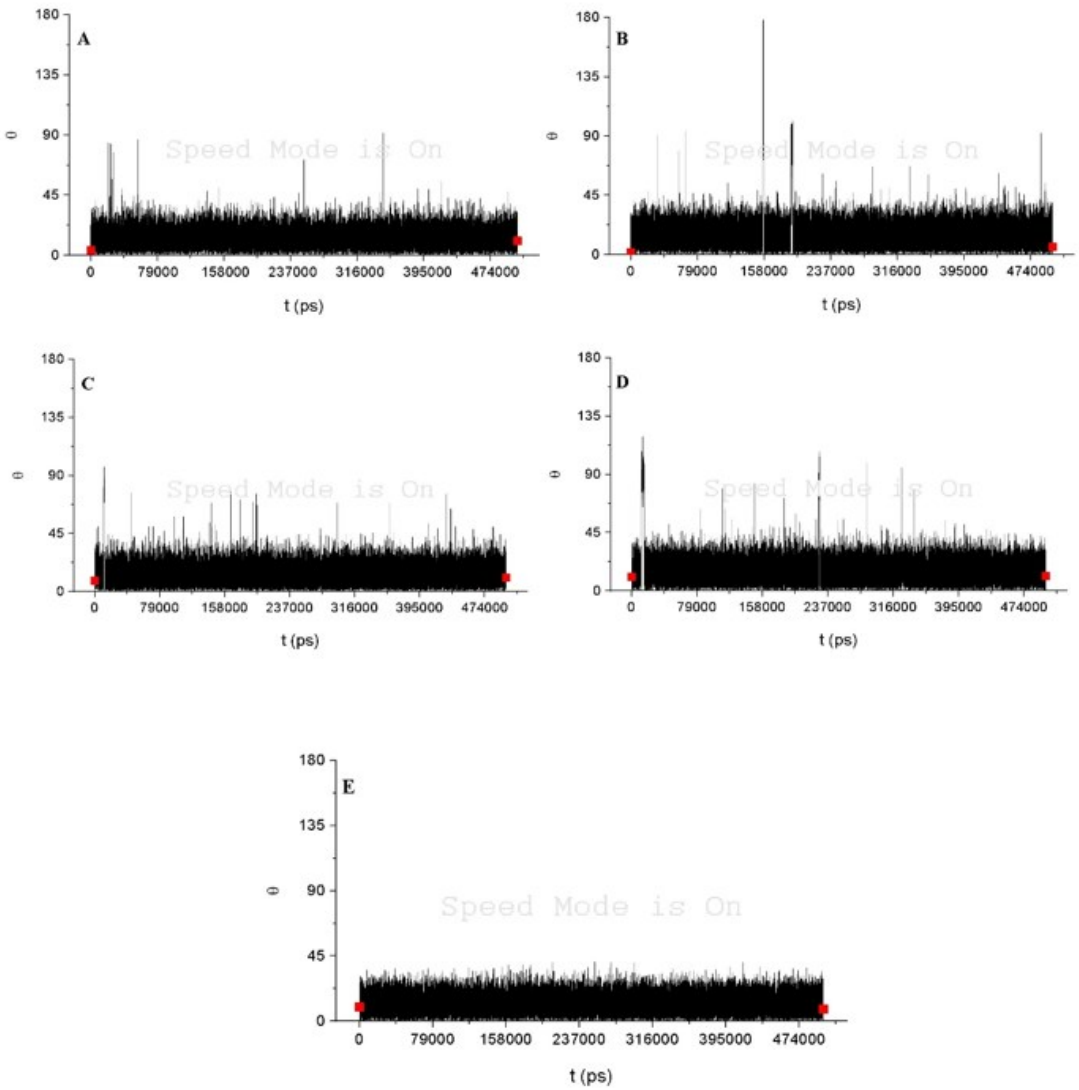


Figure S12. Puckering representation for pentasaccharide **2** (500 ns MD).

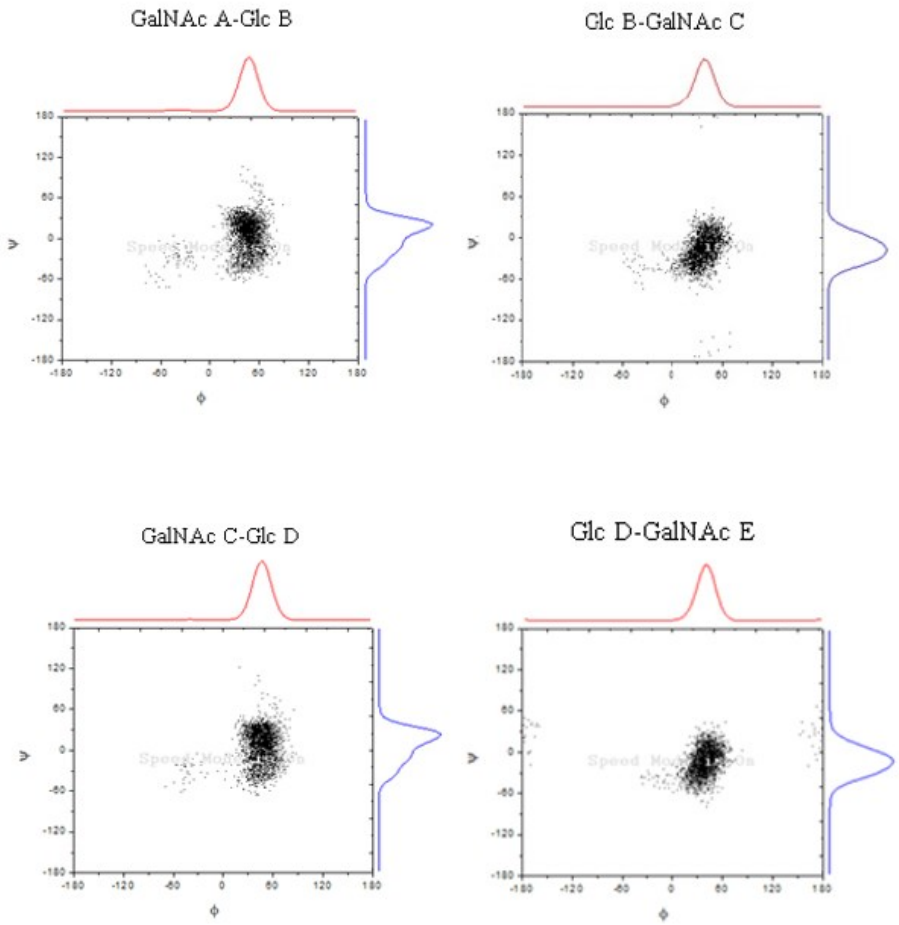


Figure S13. Trajectory (ϕ,ψ) of the unrestrained MD for compound **2**.

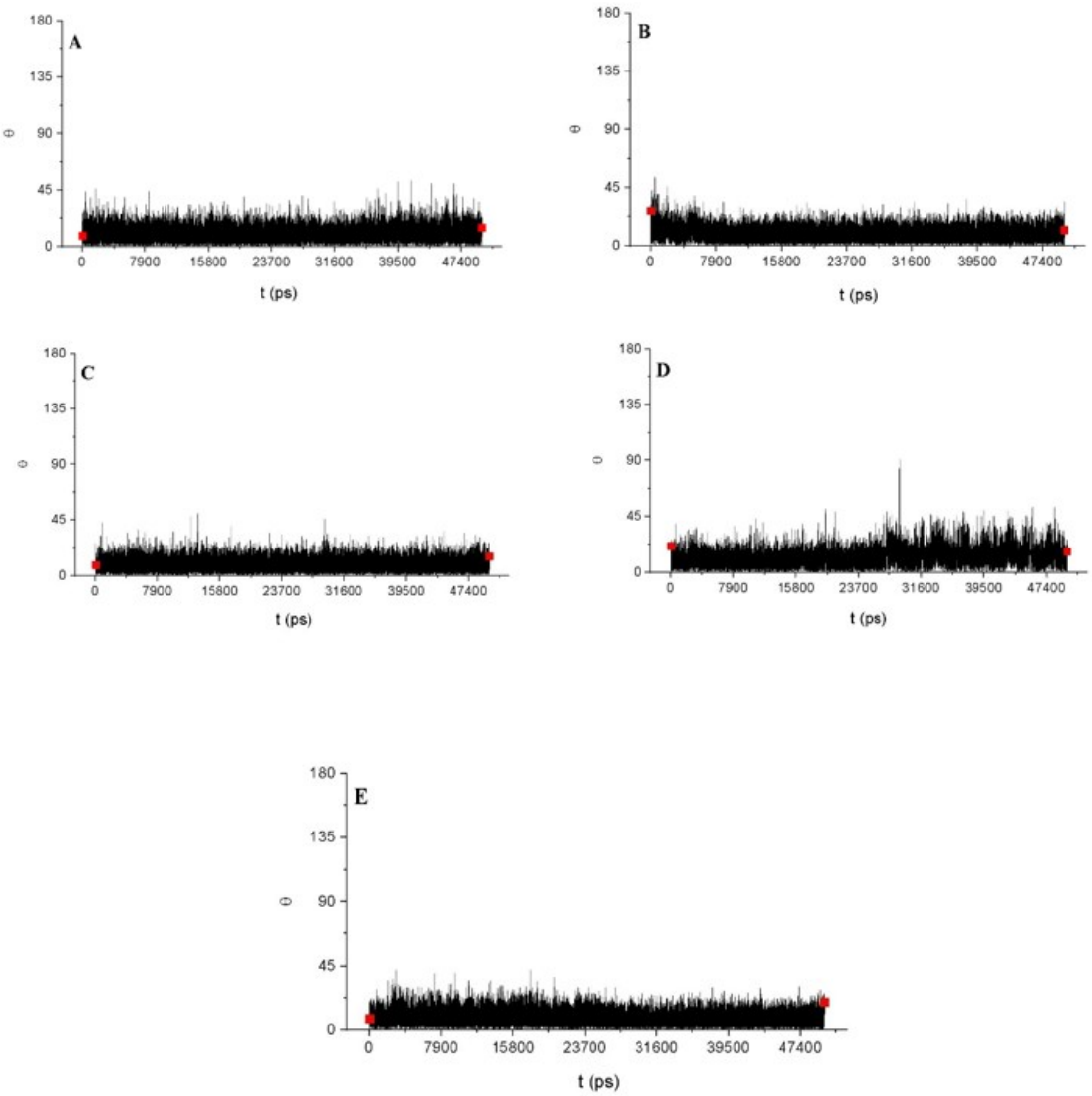


Figure S14. Puckering coordinates for the pentasaccharide **4** (50 ns MD-tar).

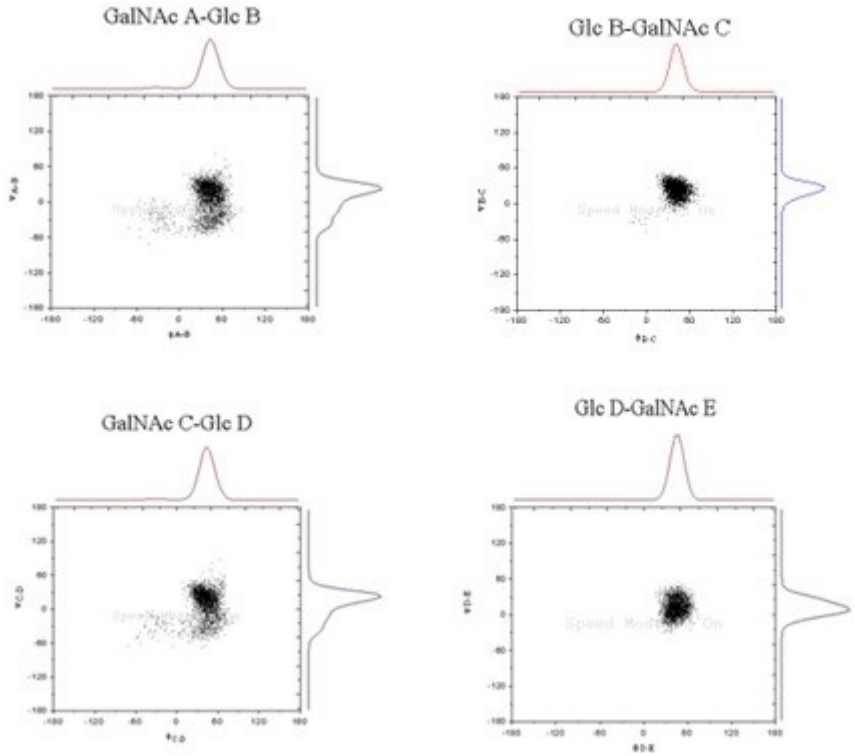


Figure S15. Trajectory (ϕ,ψ) for the compound **4** (50 ns MD-tar).

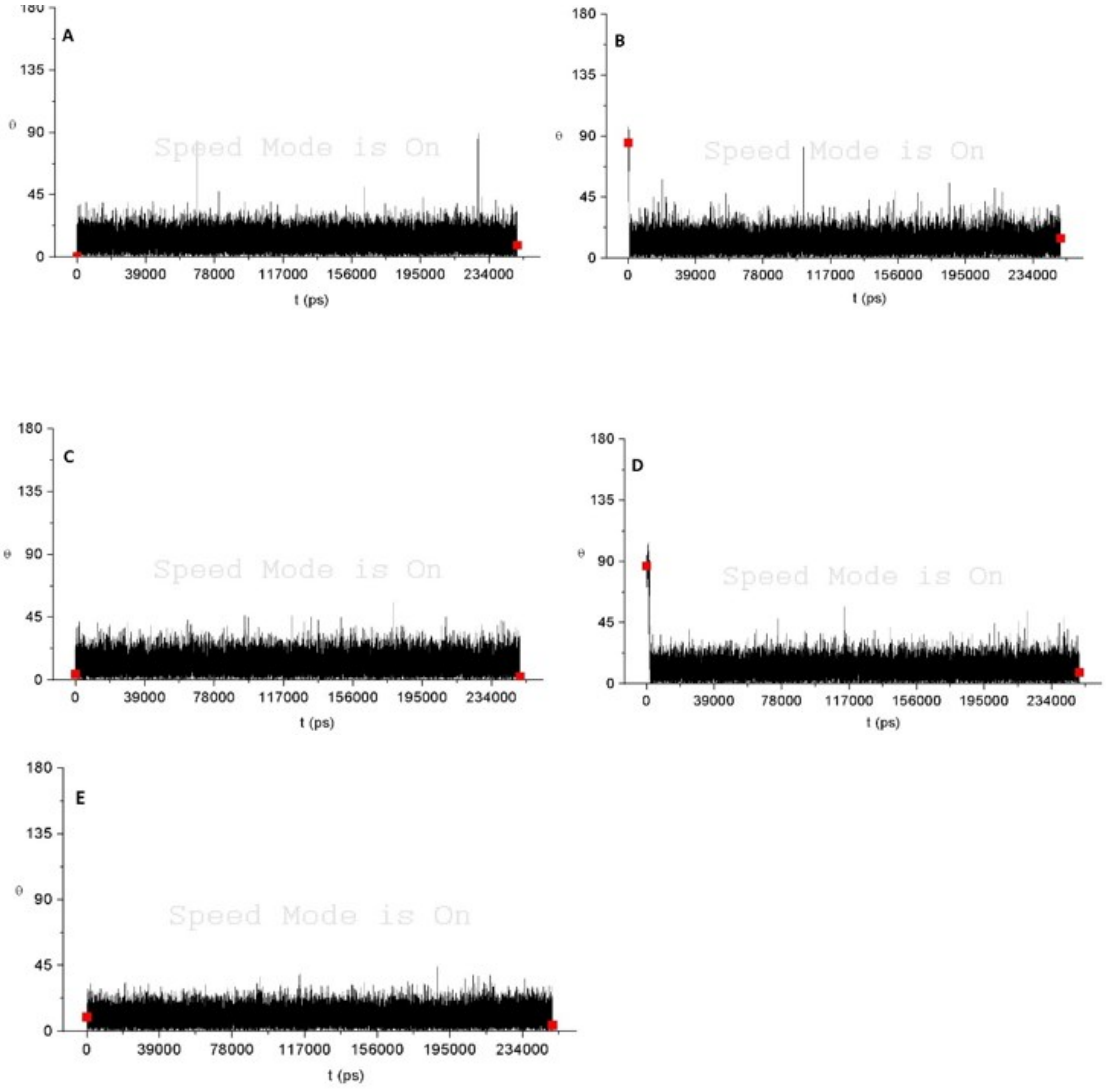


Figure S16. Puckering coordinates for the pentasaccharide **6** (250 ns MD).

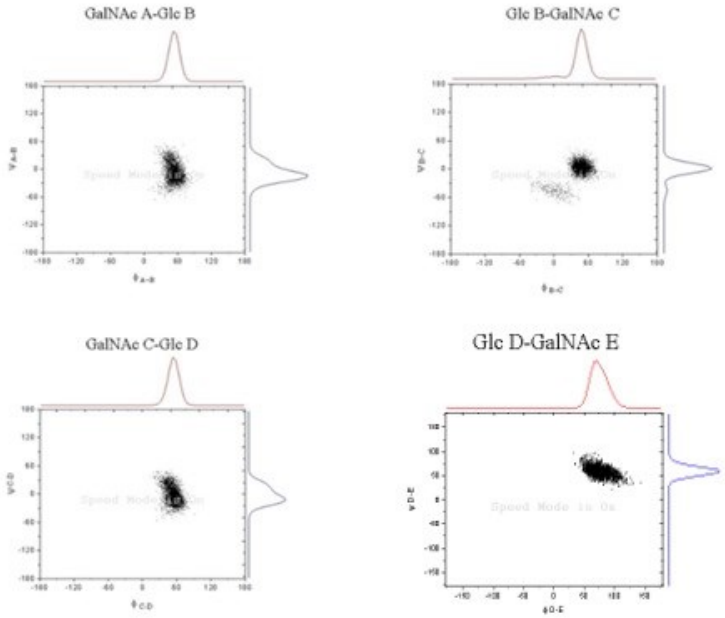


Figure S17. Trajectory (ϕ,ψ) of the unrestrained MD for the compound **6**.

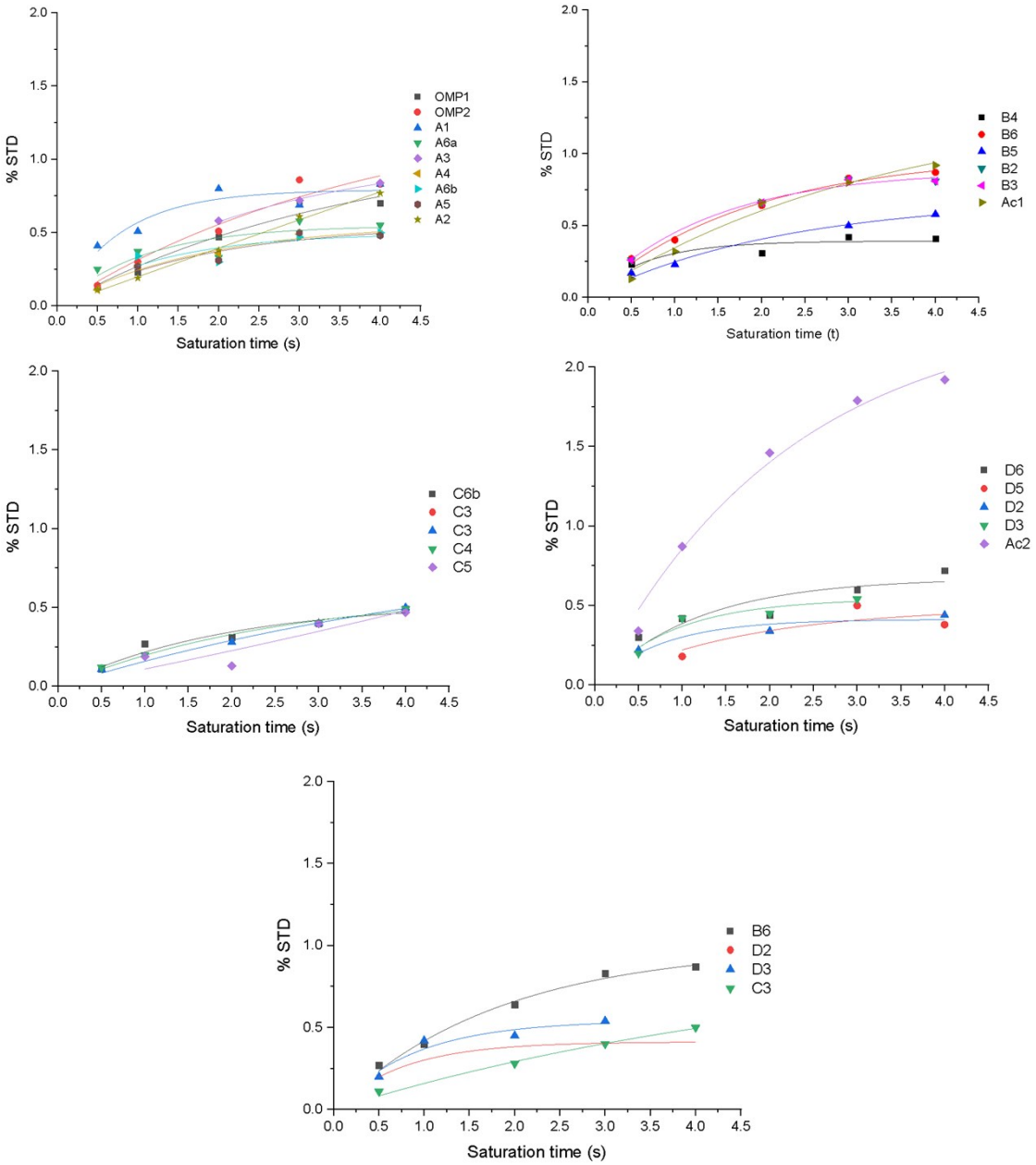


Figure S18. STD build-up curves of the interaction between the ligand **1** and midkine (600 MHz). The curves are divided by residue.

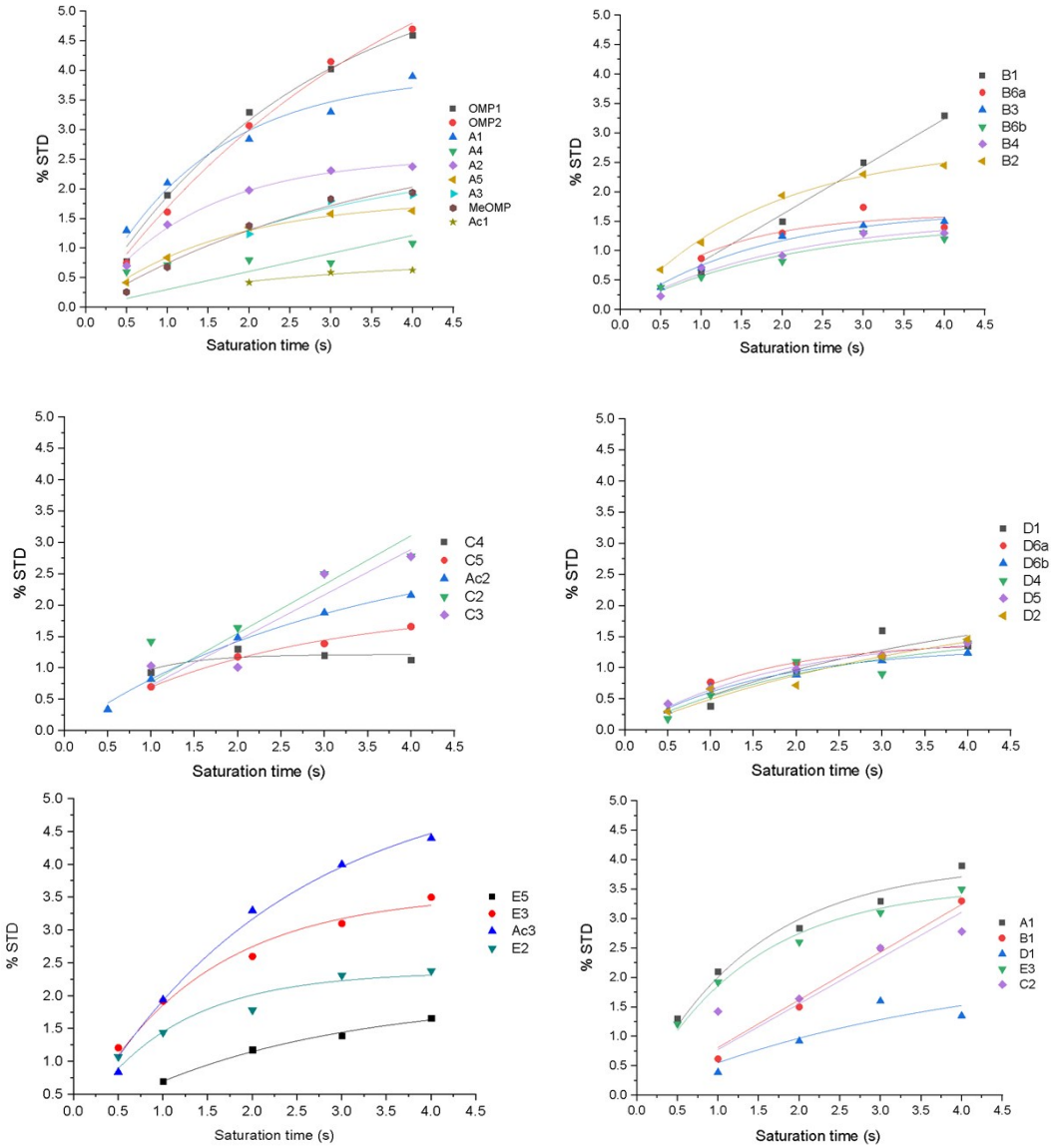


Figure S19. STD build-up curves of the interaction between the ligand **2** and midkine (600 MHz). The curves are divided by residue.

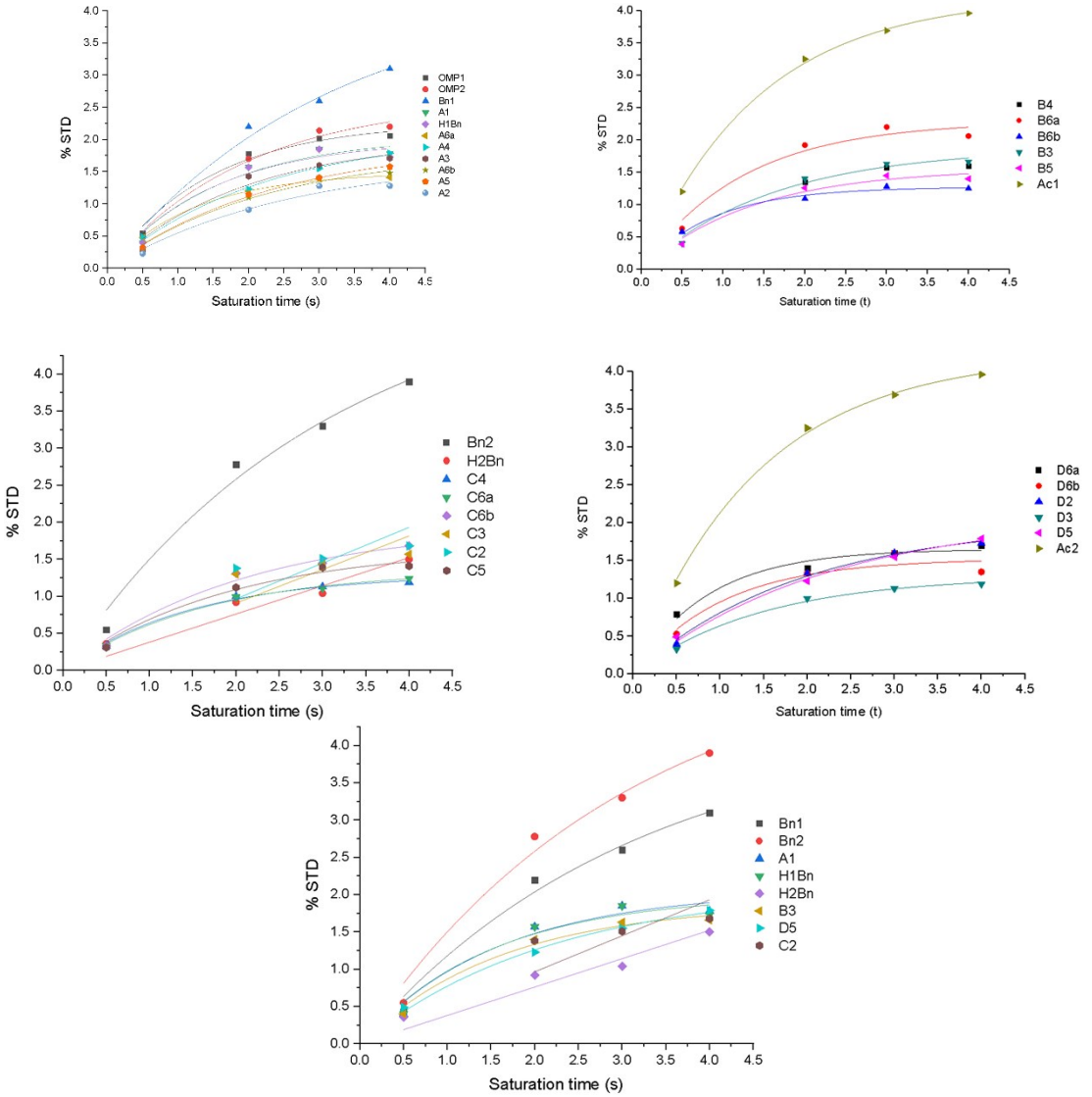


Figure S20. STD build-up curves of the interaction between the ligand **3** and midkine (600 MHz). The curves are divided by residue.

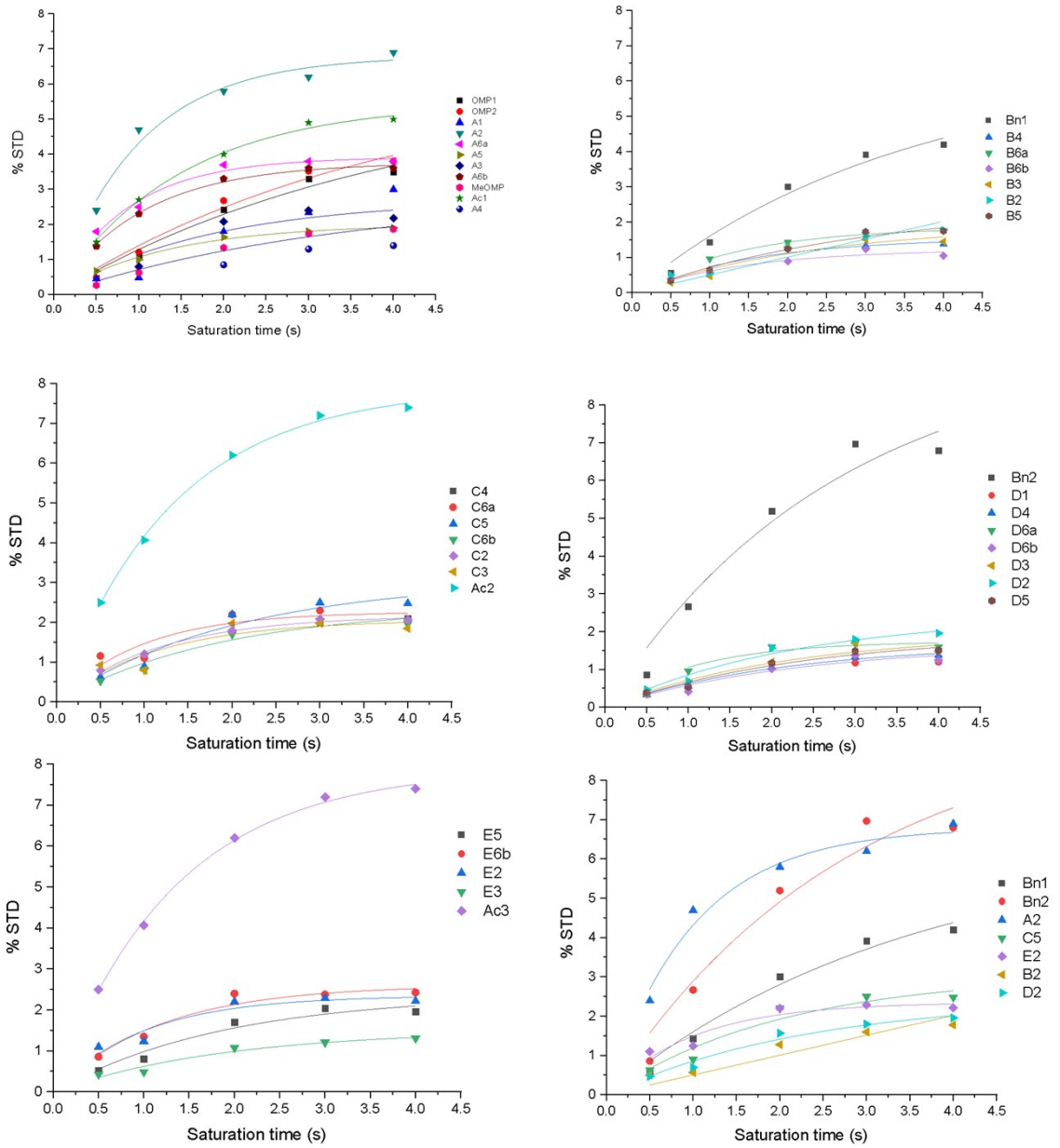
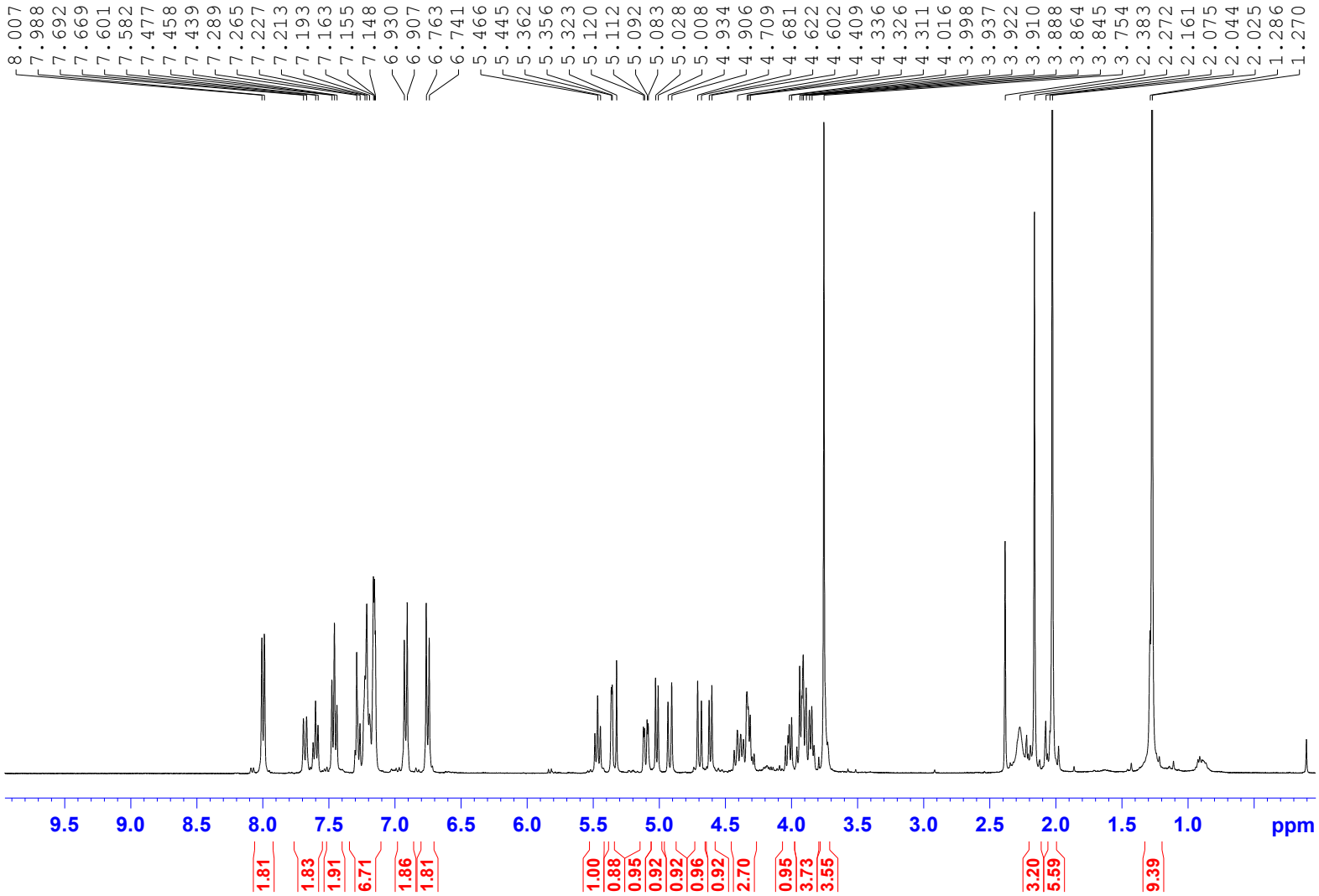


Figure S21. STD build-up curves of the interaction between the ligand **4** and midkine (600 MHz). The curves are divided by residue.

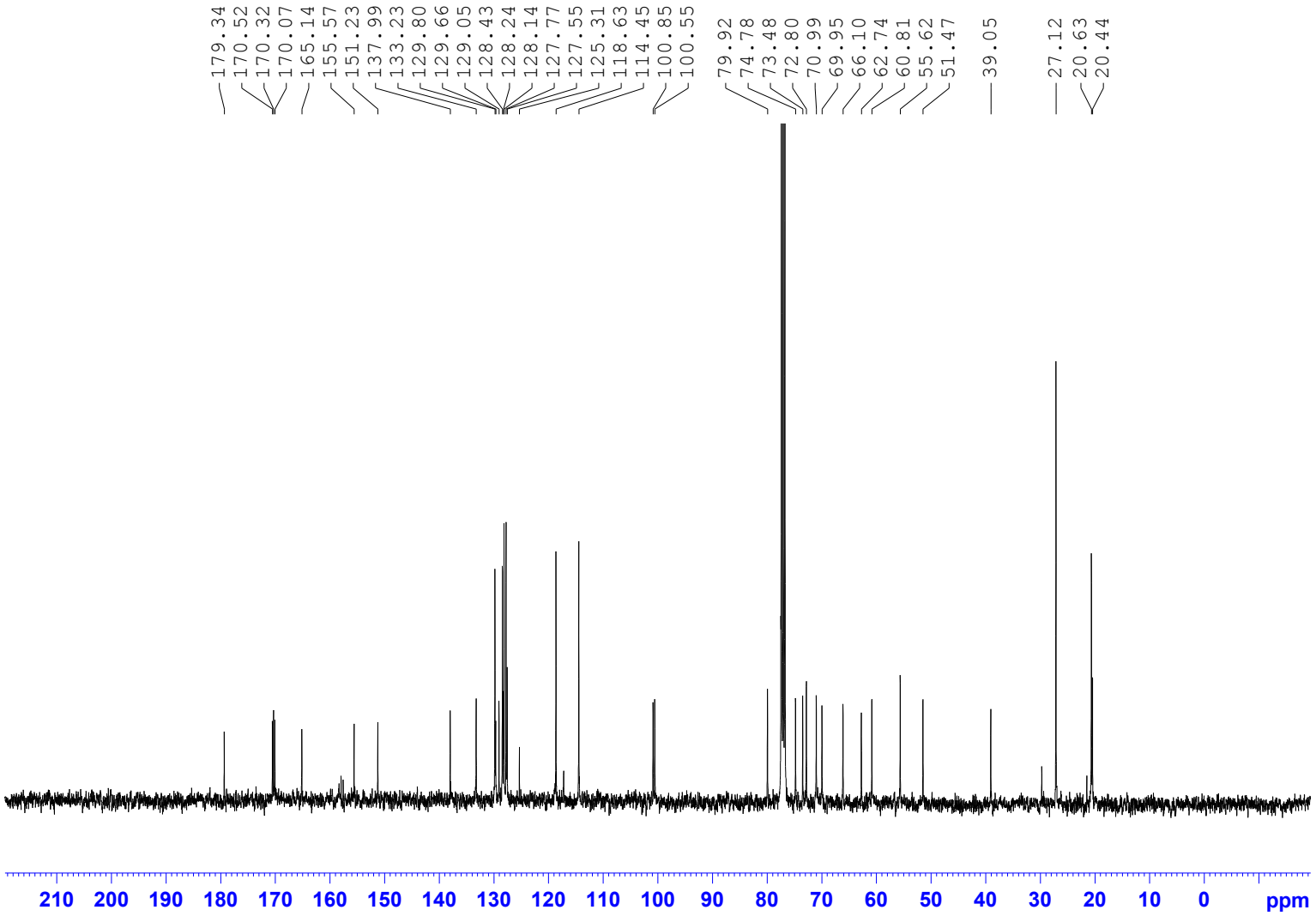
References

1. Case, D. a. *et al.* Amber 12 Reference Manual. *Univ. California, San Fr.* 350 (2012) doi:citeulike-article-id:10779586.
2. Muñoz-García, J. C. *et al.* Langerin-heparin interaction: Two binding sites for small and large ligands as revealed by a combination of NMR spectroscopy and cross-linking mapping experiments. *J. Am. Chem. Soc.* **137**, 4100–4110 (2015).
3. Kirschner, K., Yongye, A. & Tschampel, S. Glycam06. *J Comput Chem2* **29**, 622–655 (8AD).
4. Petersen, H. G. Accuracy and efficiency of the particle mesh Ewald method. *Phys. Fluids* **10**, 3668–3679 (1967).

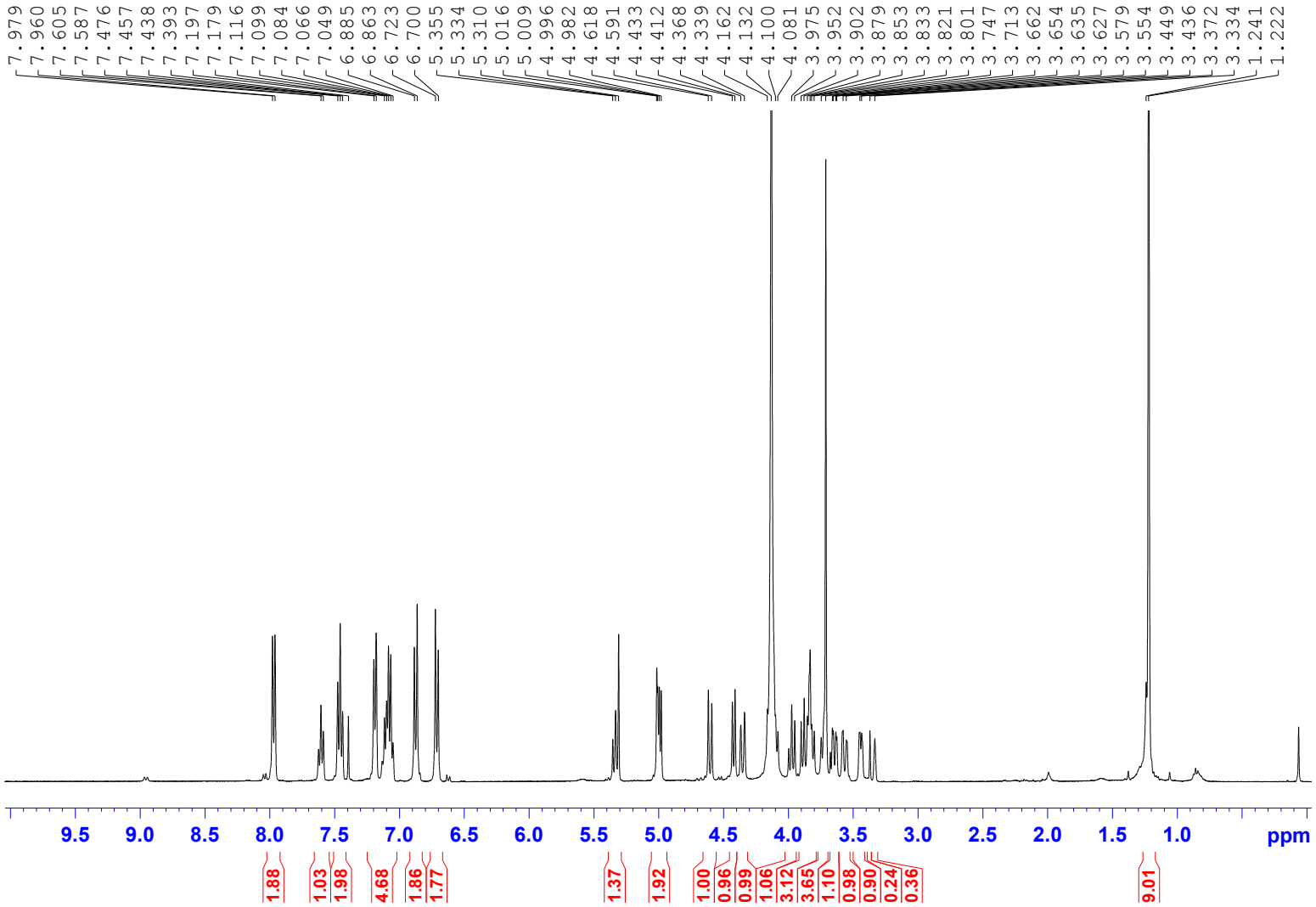
Compound 9, ^{1}H-NMR, 400 MHz, $CDCl_3$



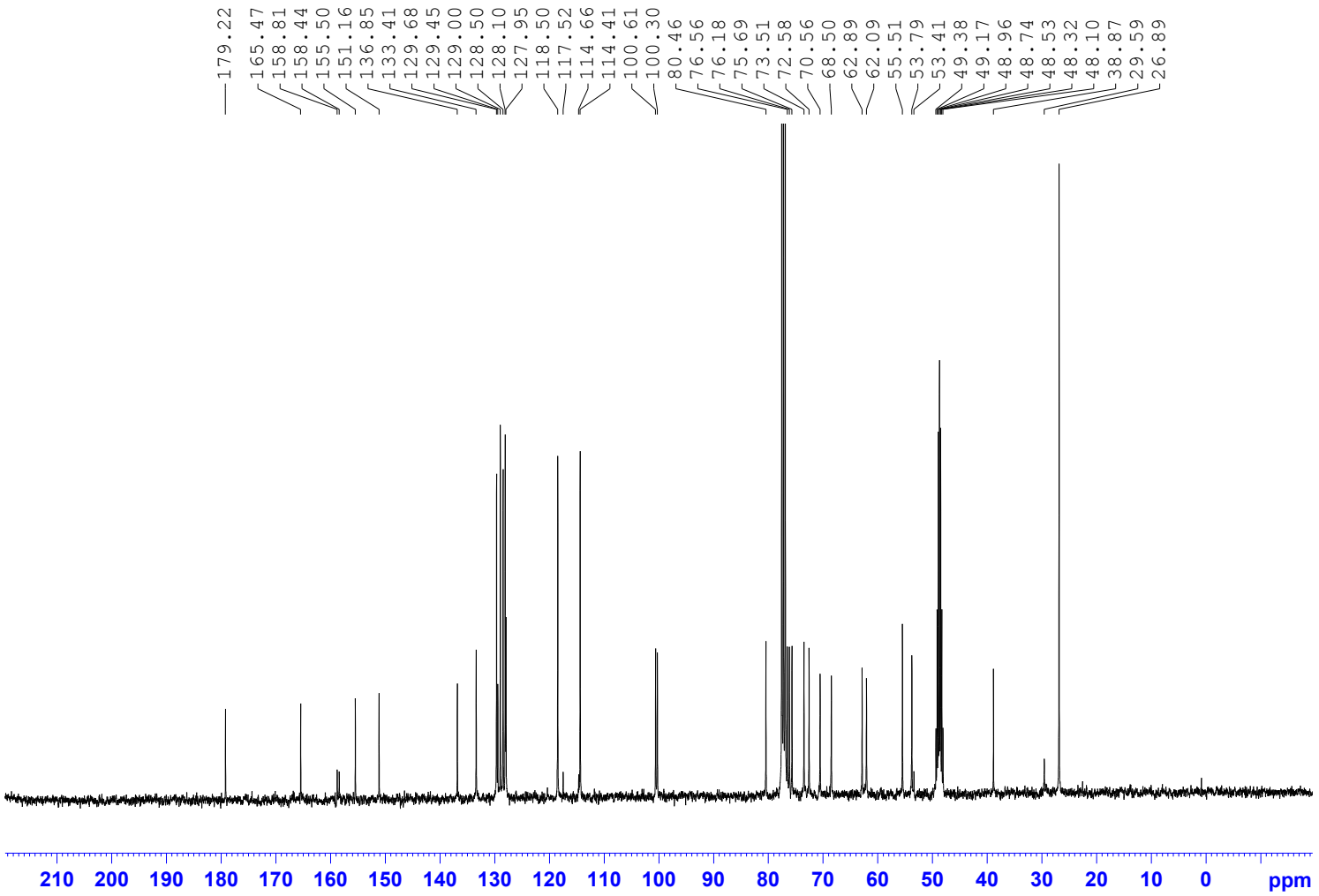
Compound 9, ^{13}C-NMR, 100 MHz, $CDCl_3$



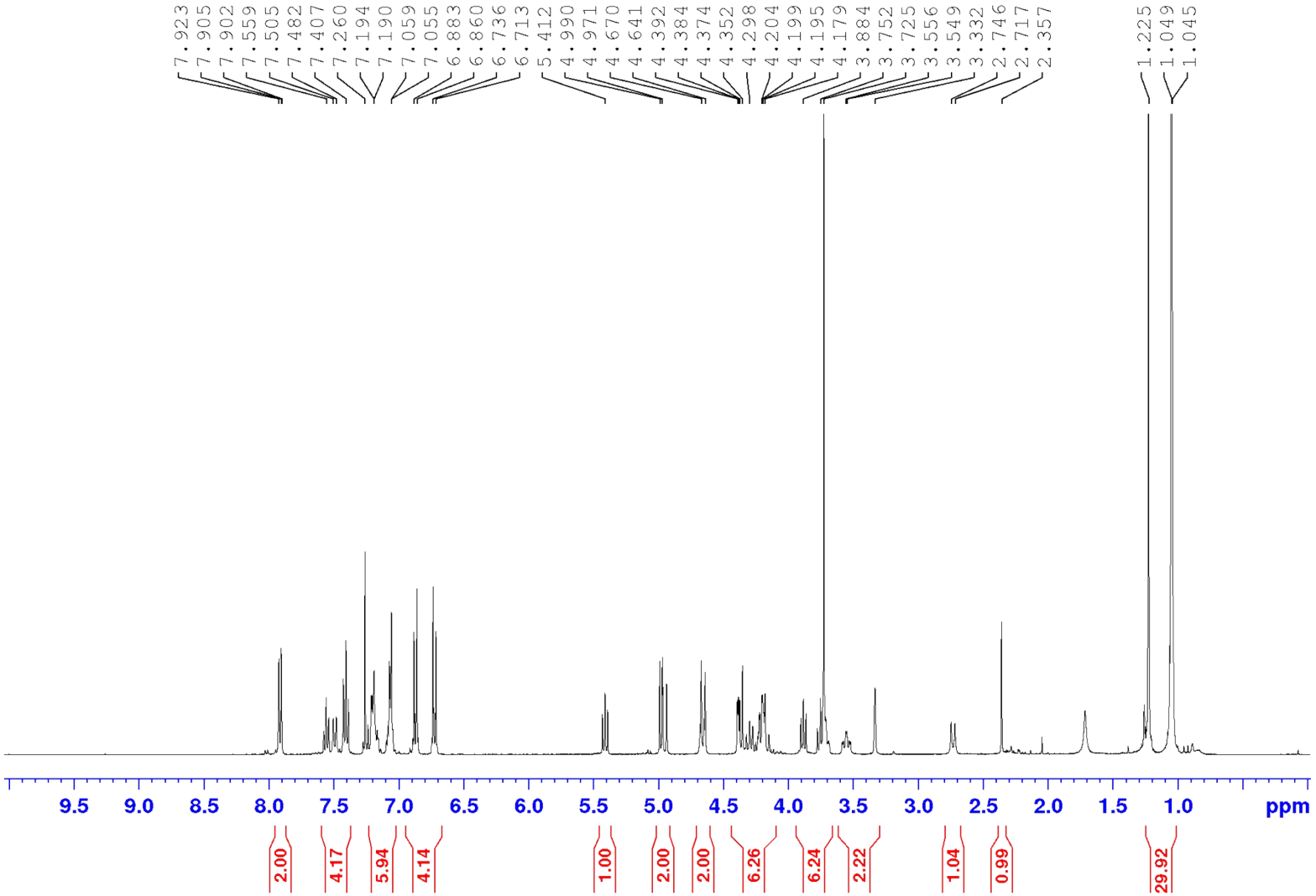
Compound 10, ^{1}H-NMR, 400 MHz, $CDCl_3/CD_3OD$ 9:1



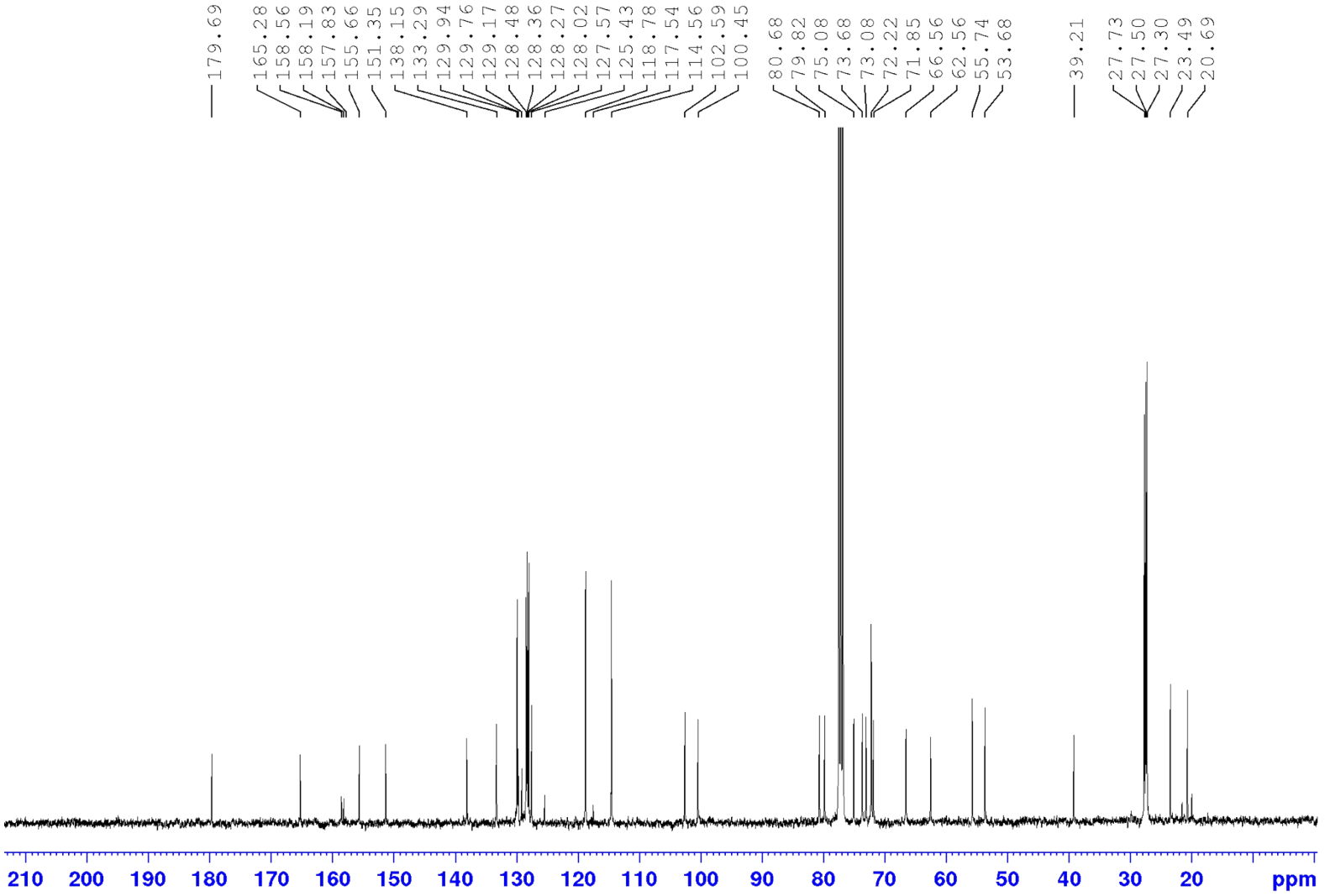
Compound 10, ^{13}C-NMR, 100 MHz, $CDCl_3/CD_3OD$ 9:1



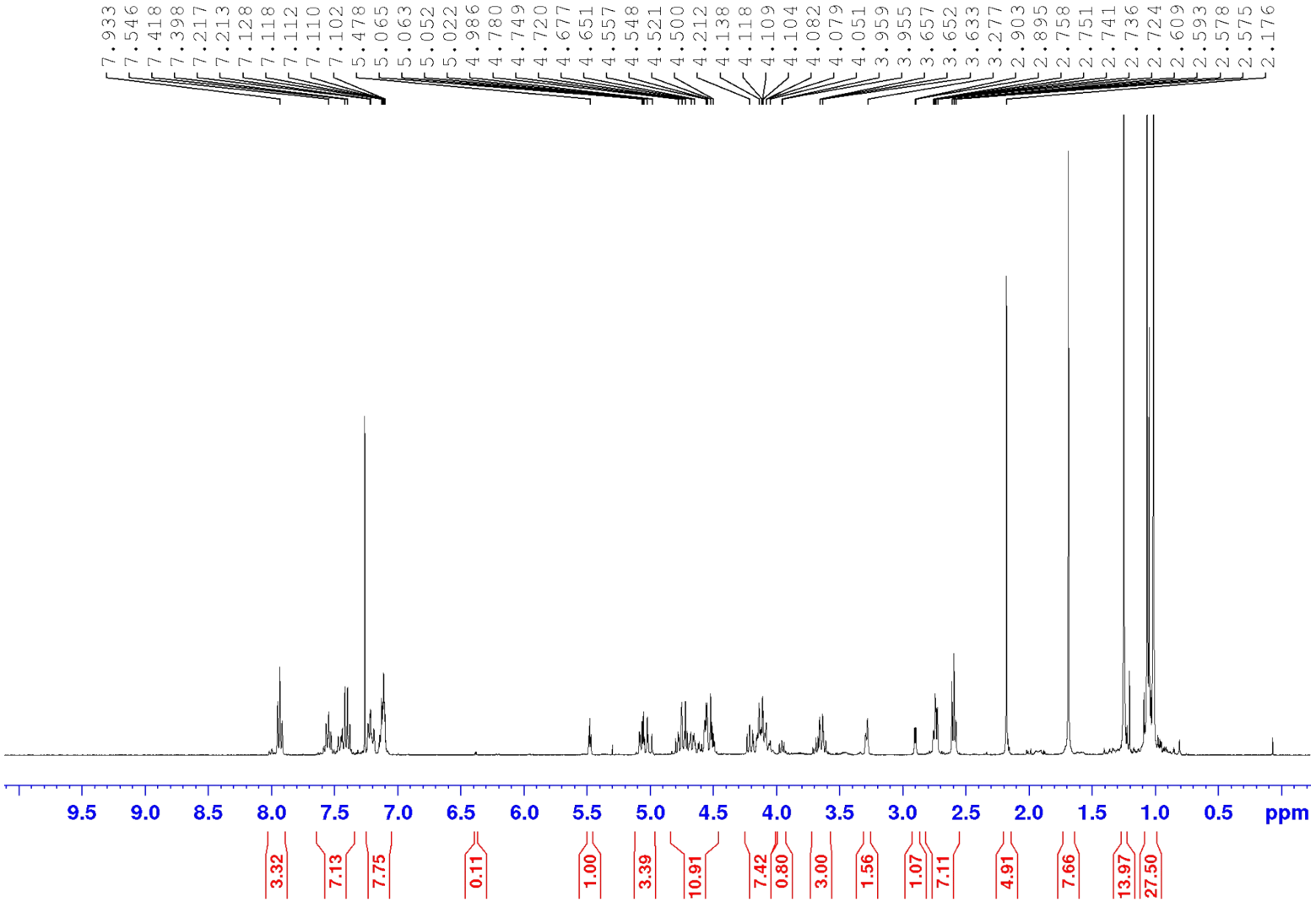
Compound 11, ^{1}H-NMR, 400 MHz, $CDCl_3$



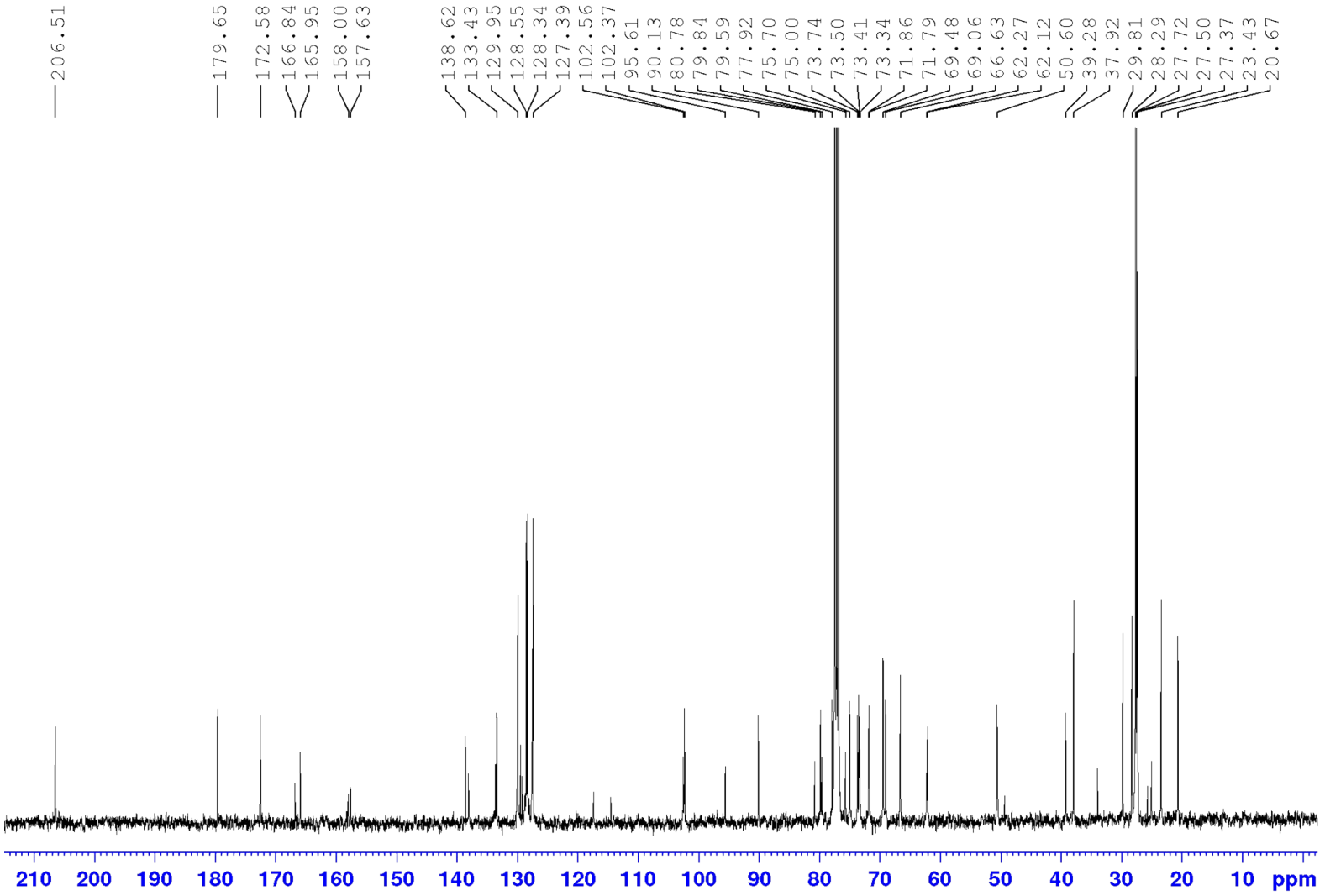
Compound 11, ^{13}C-NMR, 100 MHz, $CDCl_3$



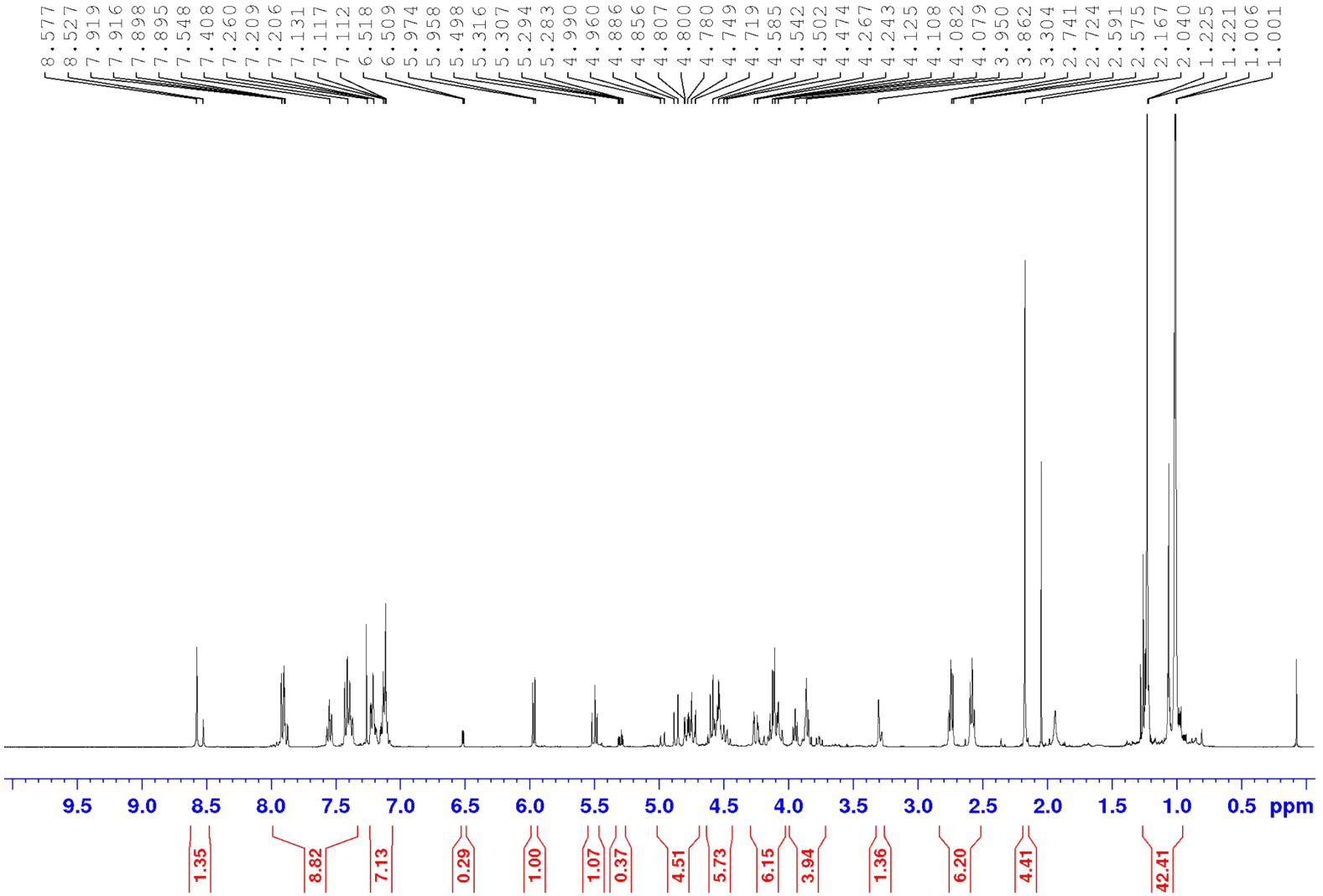
Compound 12, ^{1}H-NMR, 400 MHz, $CDCl_3$



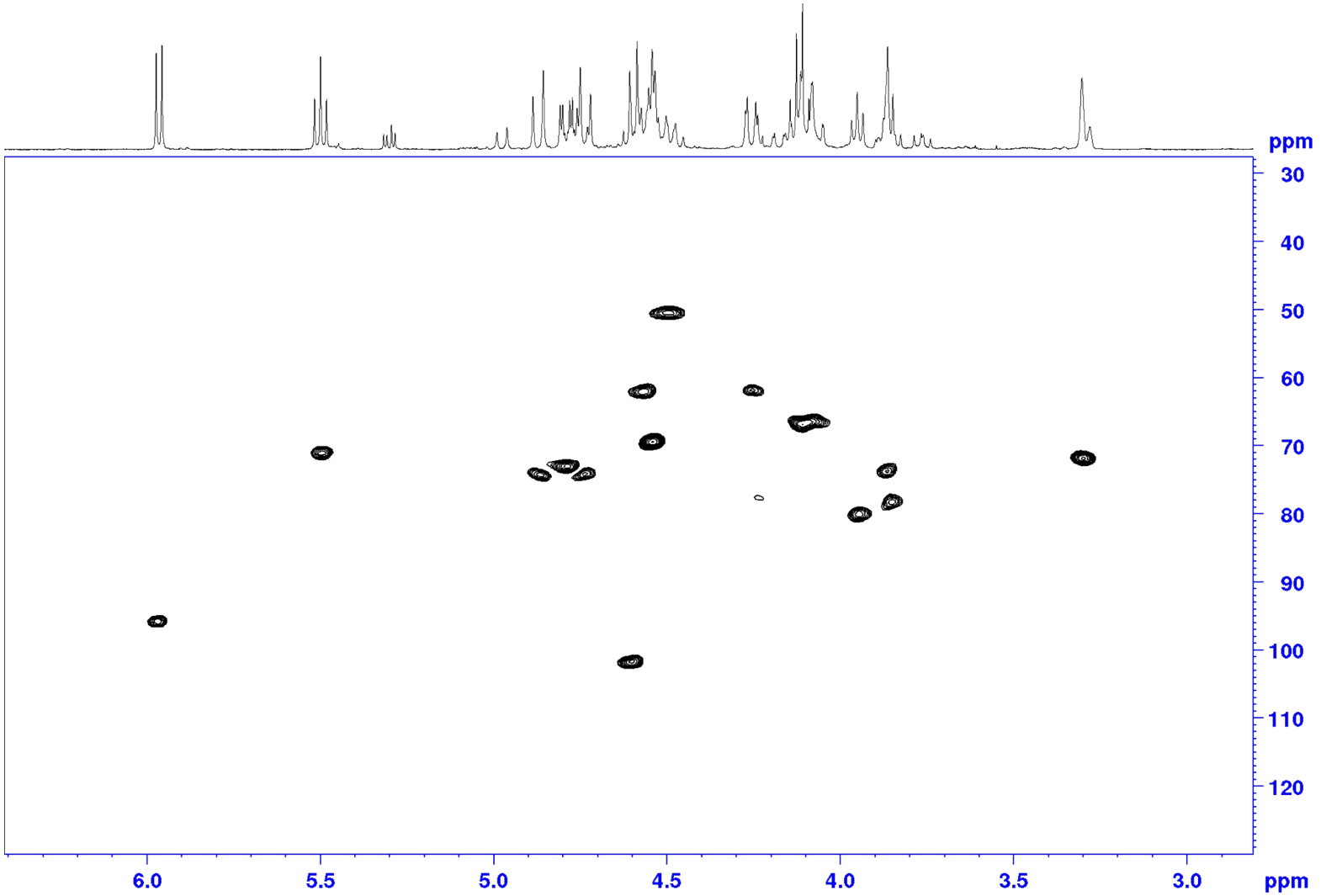
Compound 12, ^{13}C-NMR, 100 MHz, $CDCl_3$



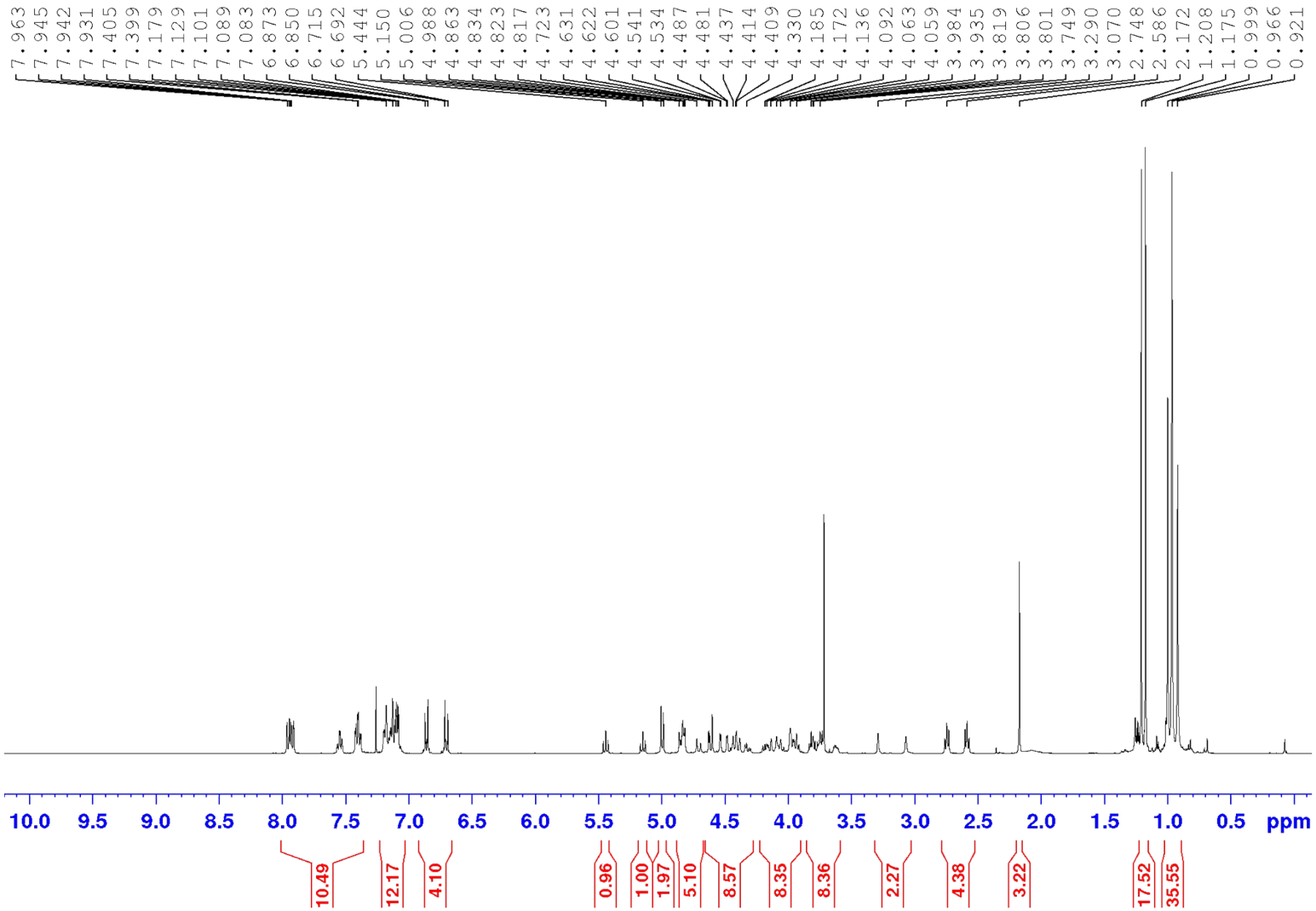
Compound 13, ^{1}H-NMR, 400 MHz, $CDCl_3$



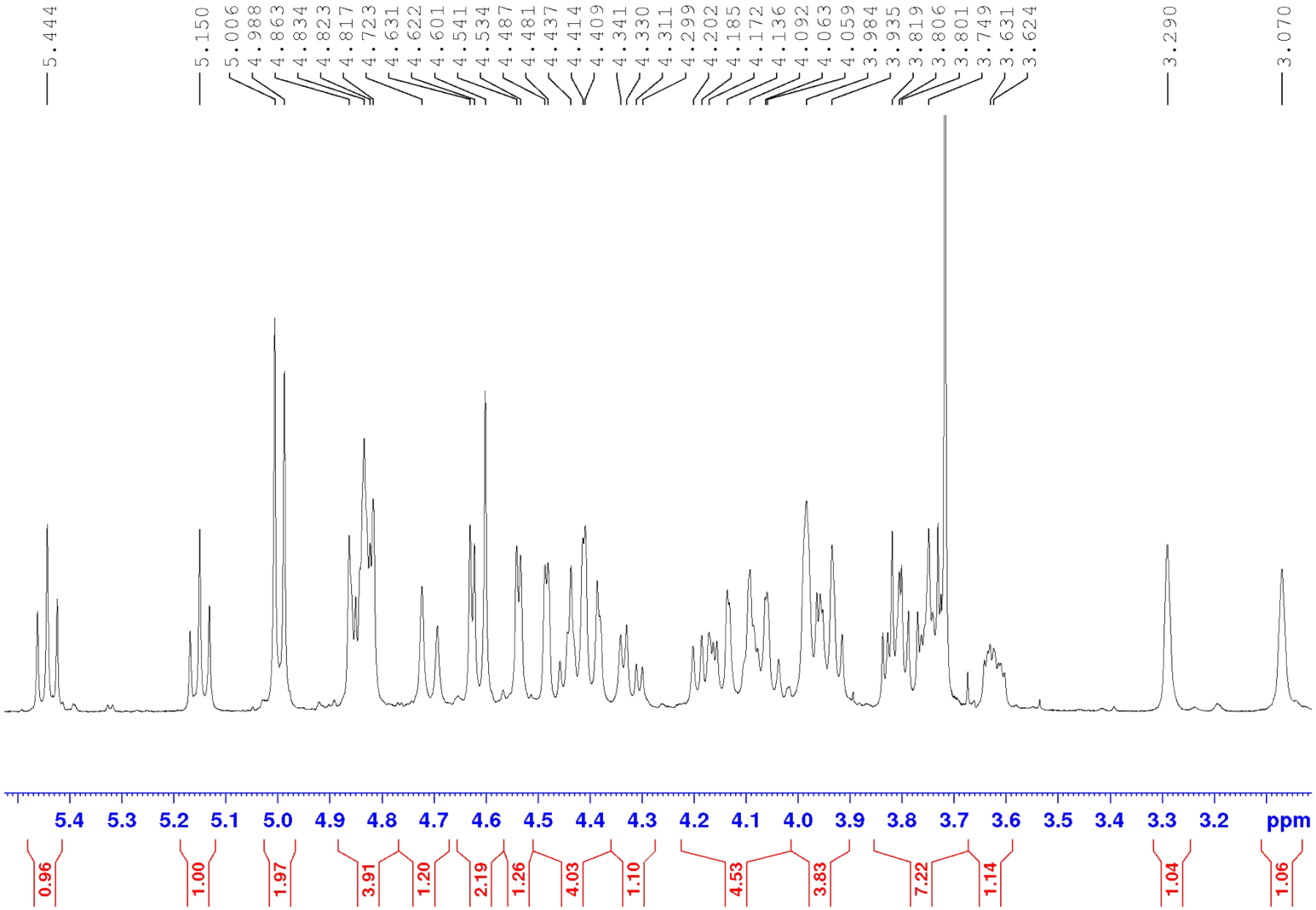
Compound 13, HSQC, 400 MHz, $CDCl_3$



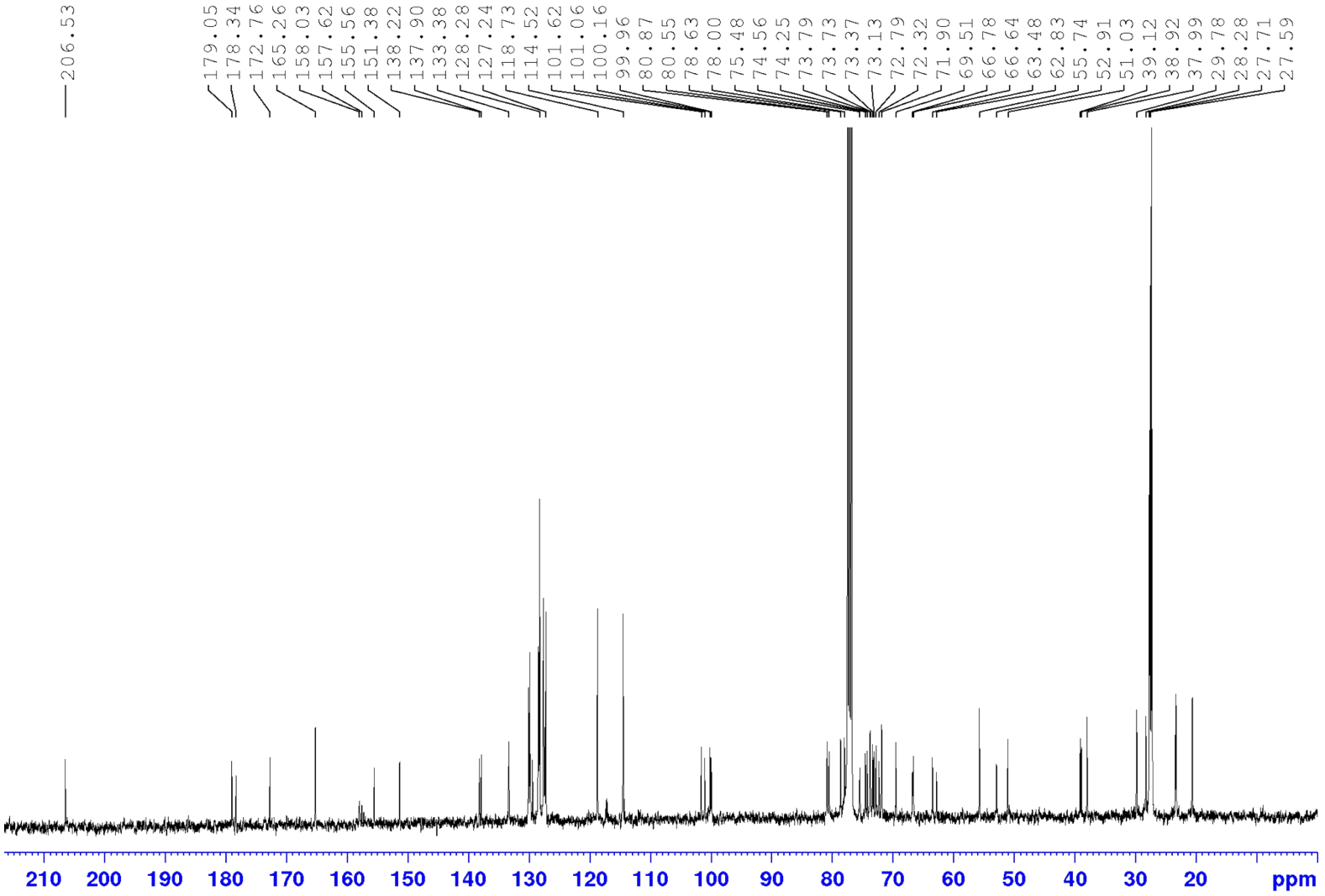
Compound 14, ^{1}H-NMR, 400 MHz, $CDCl_3$



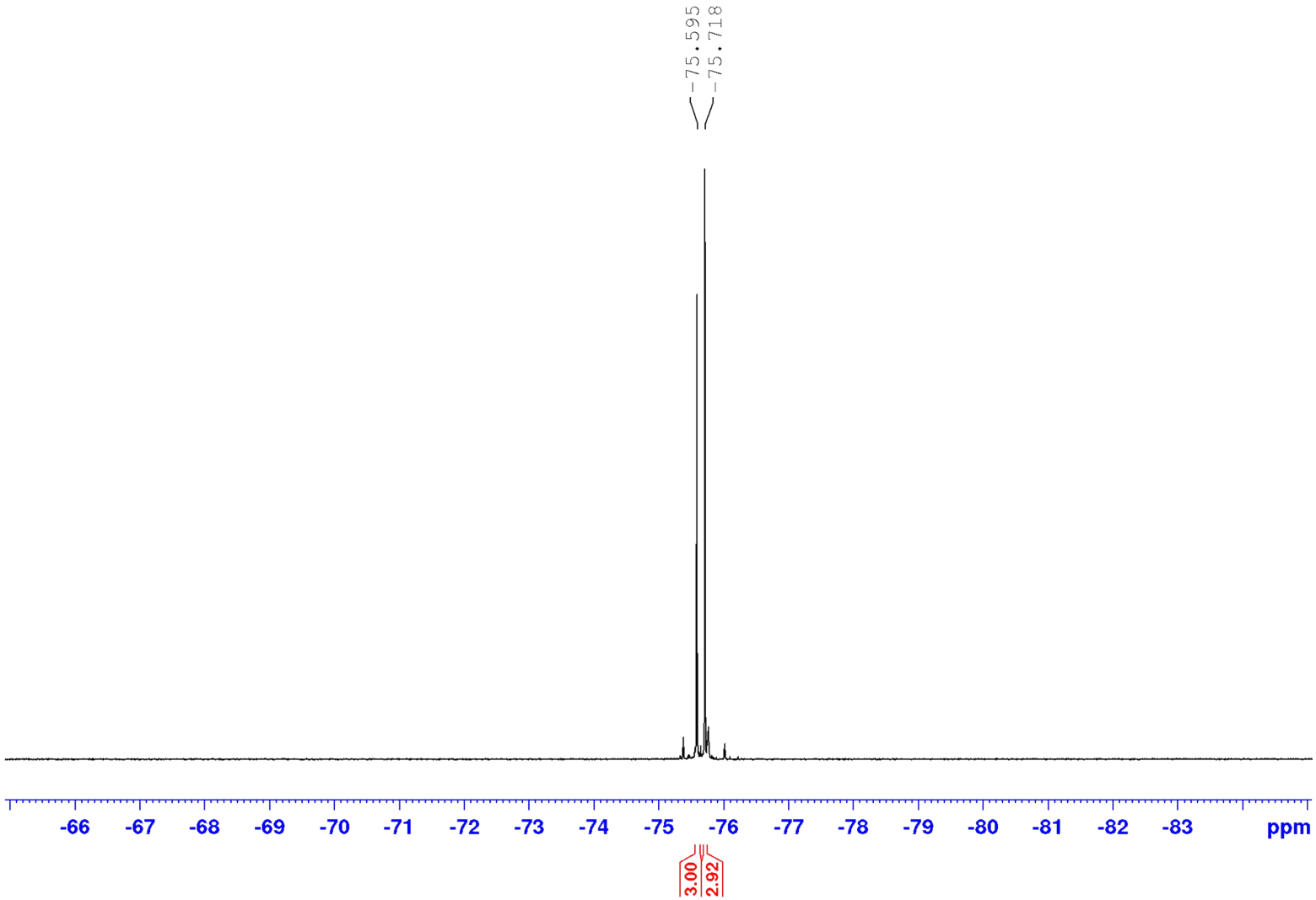
Compound 14, ^{1}H-NMR, 400 MHz, $CDCl_3$



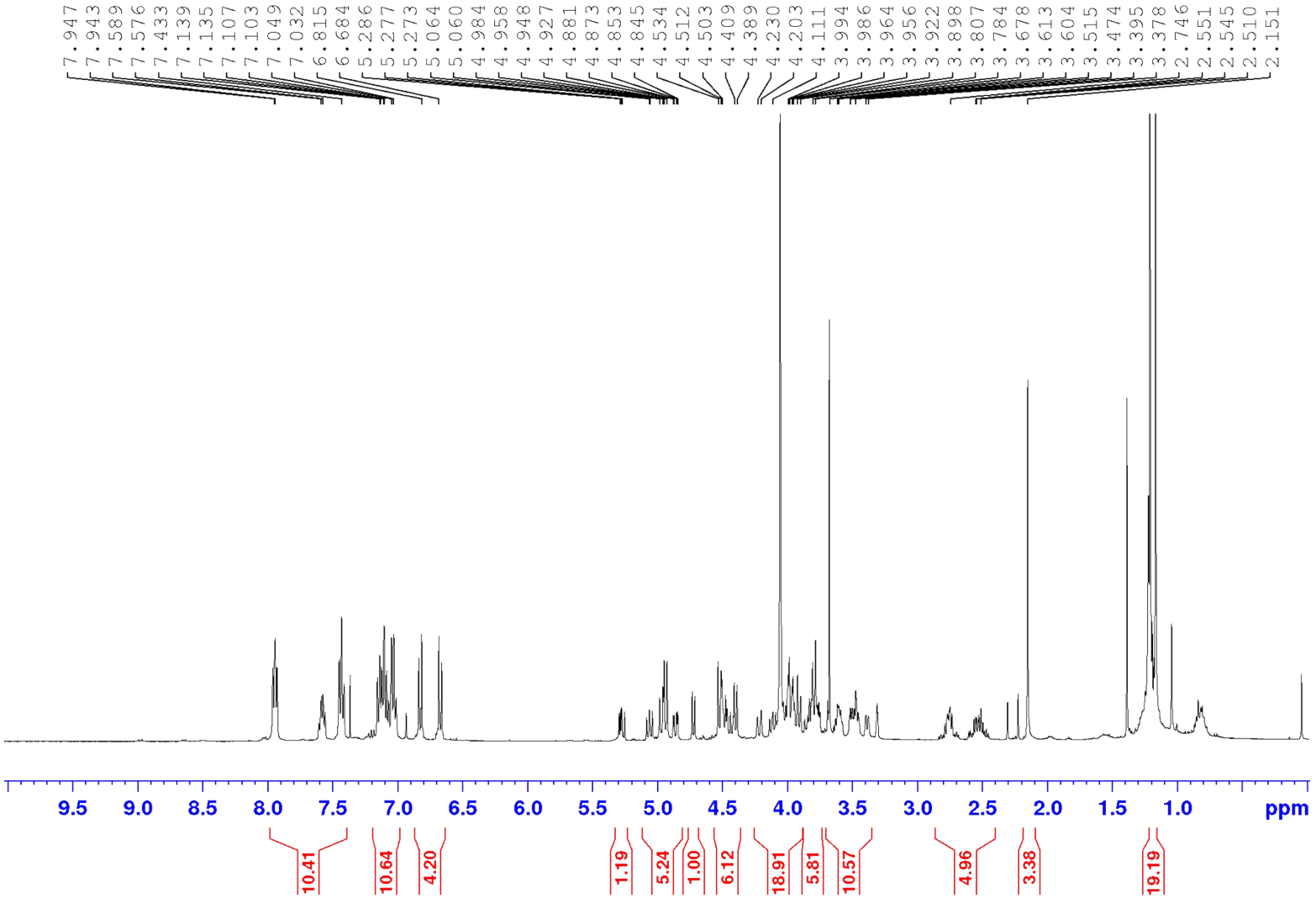
Compound 14, ^{13}C-NMR, 100 MHz, $CDCl_3$



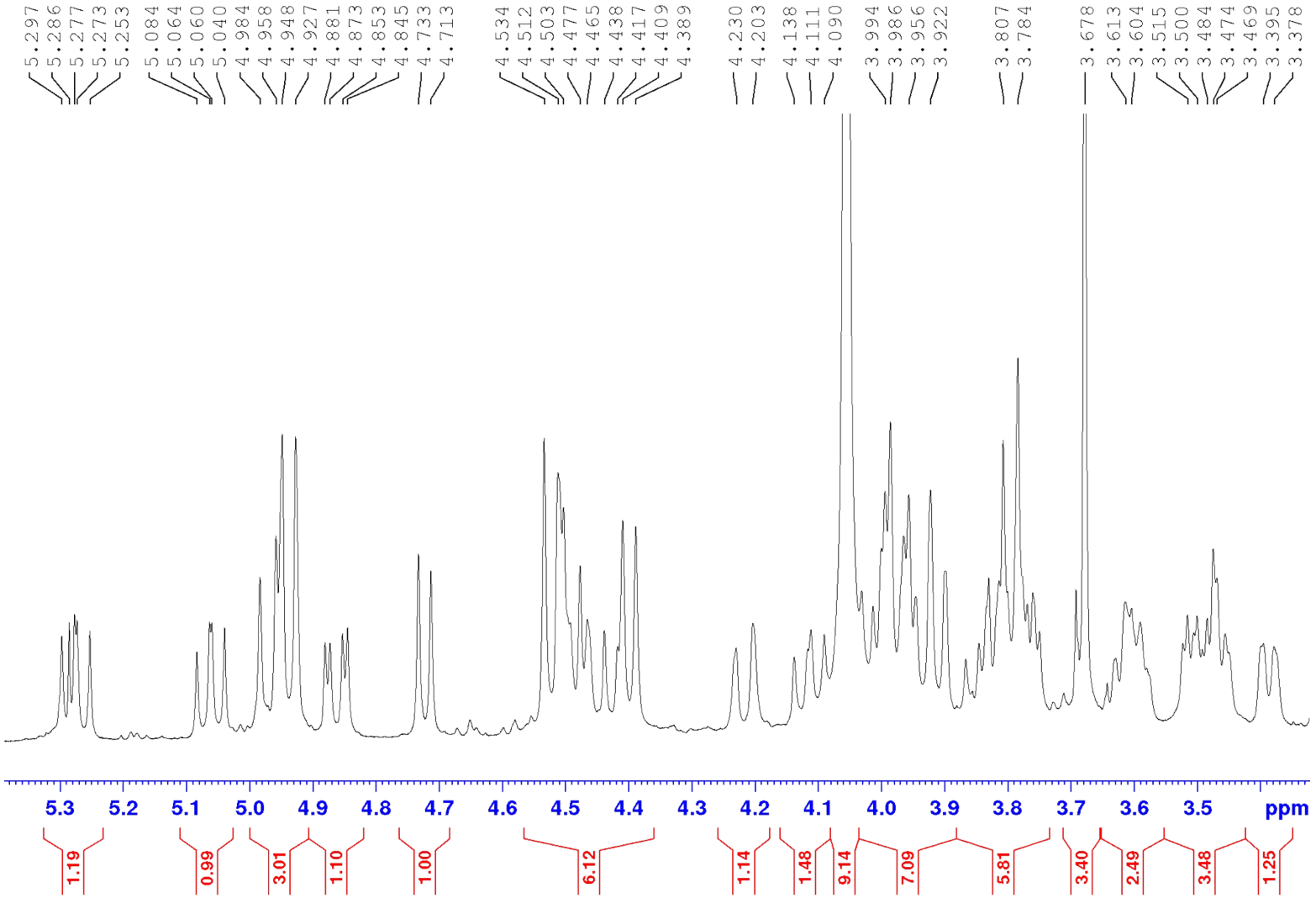
Compound 14, ^{19}F-NMR, 376 MHz, $CDCl_3$



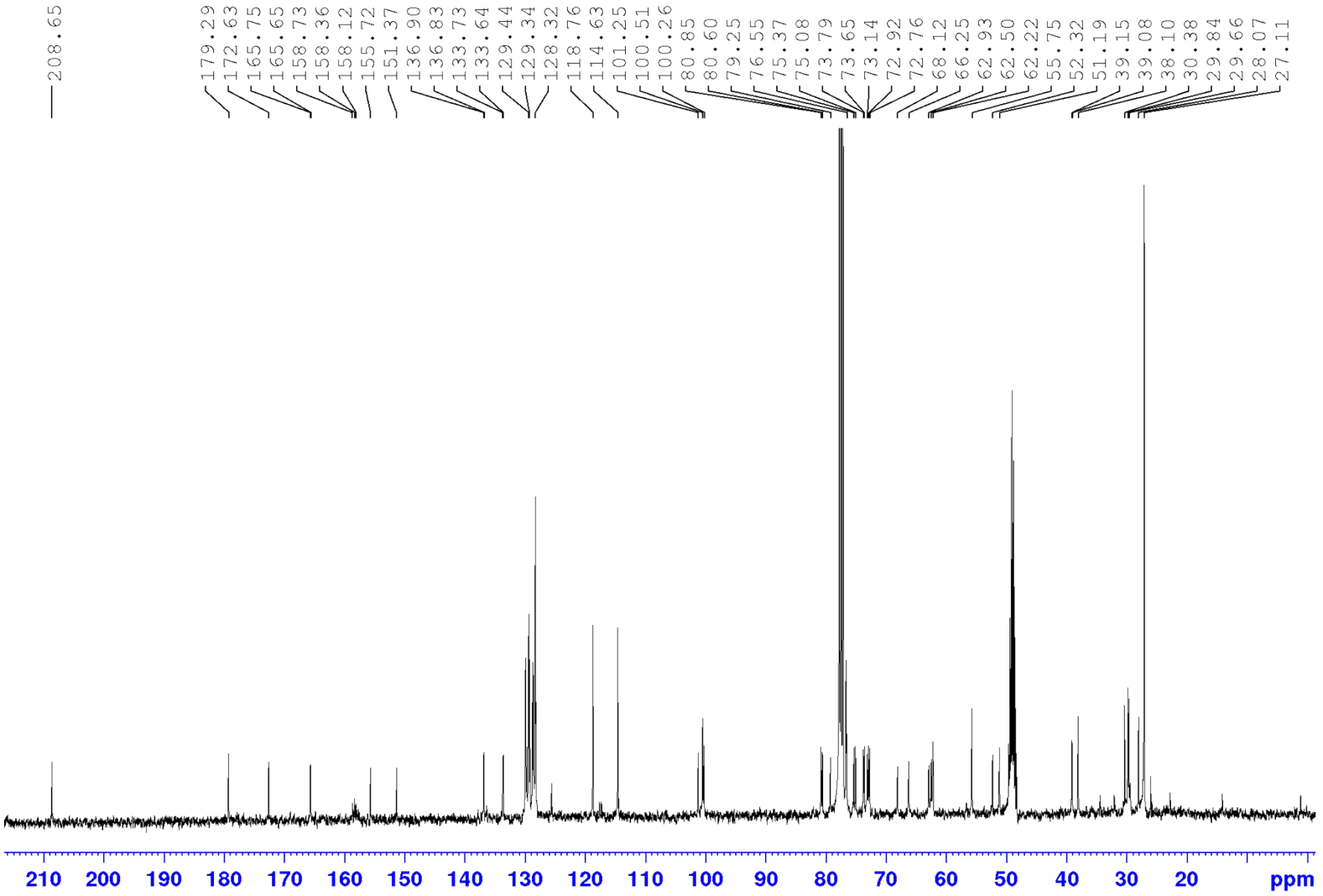
Compound 15, ^{1}H-NMR, 400 MHz, $CDCl_3/CD_3OD$ 4:1



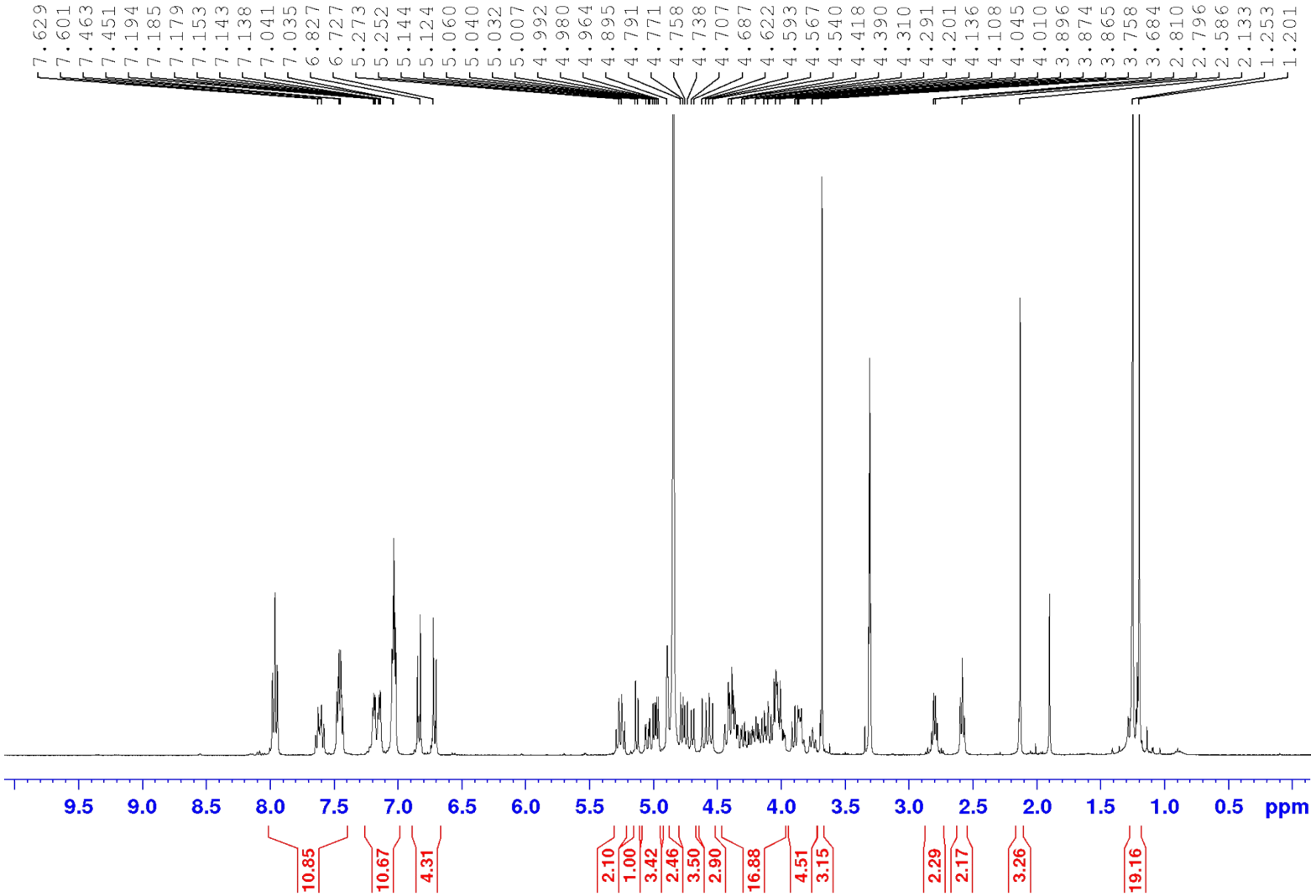
Compound 15, ^{1}H-NMR, 400 MHz, $CDCl_3/CD_3OD$ 4:1



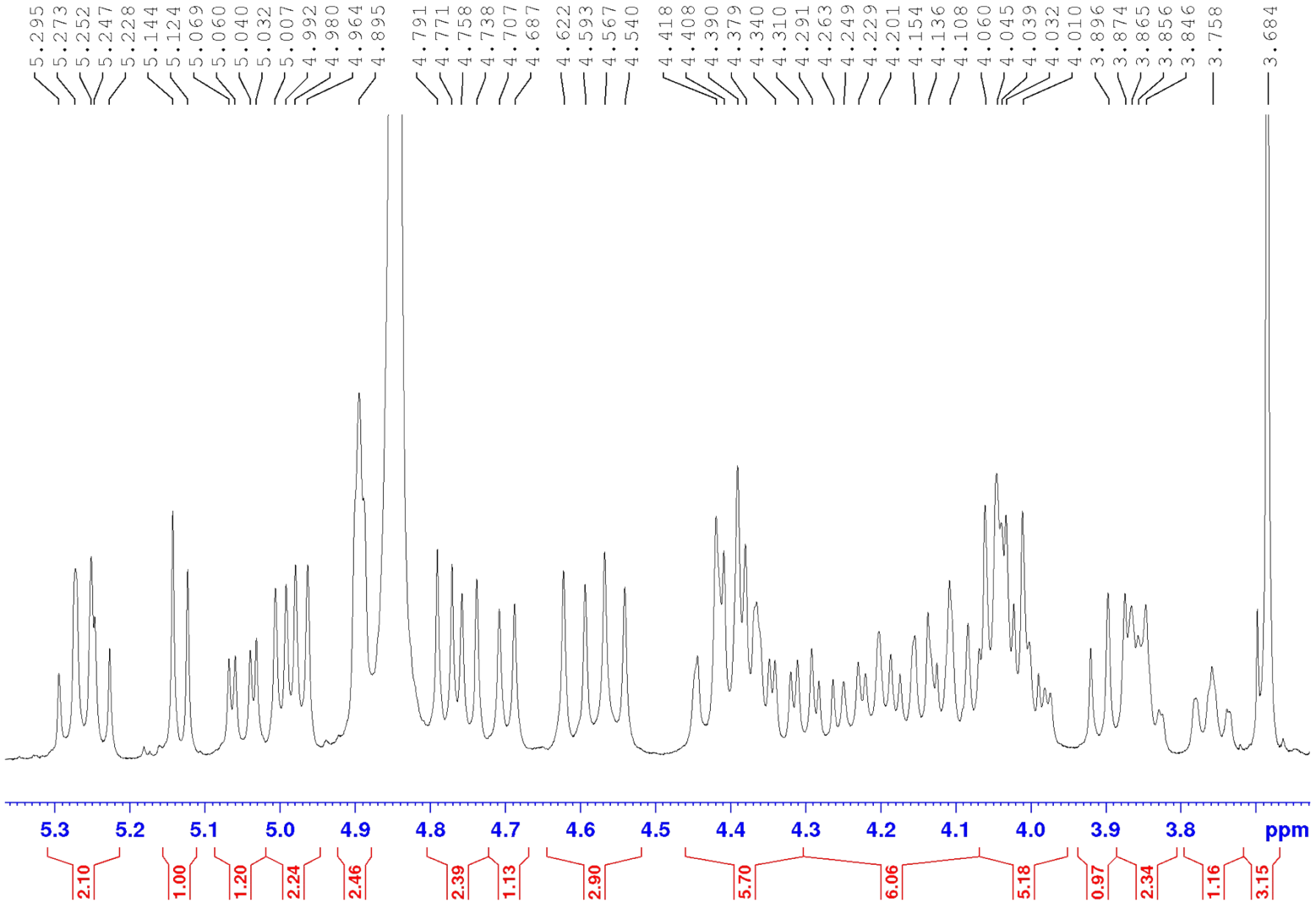
Compound 15, ^{13}C-NMR, 100 MHz, $CDCl_3/CD_3OD$ 4:1



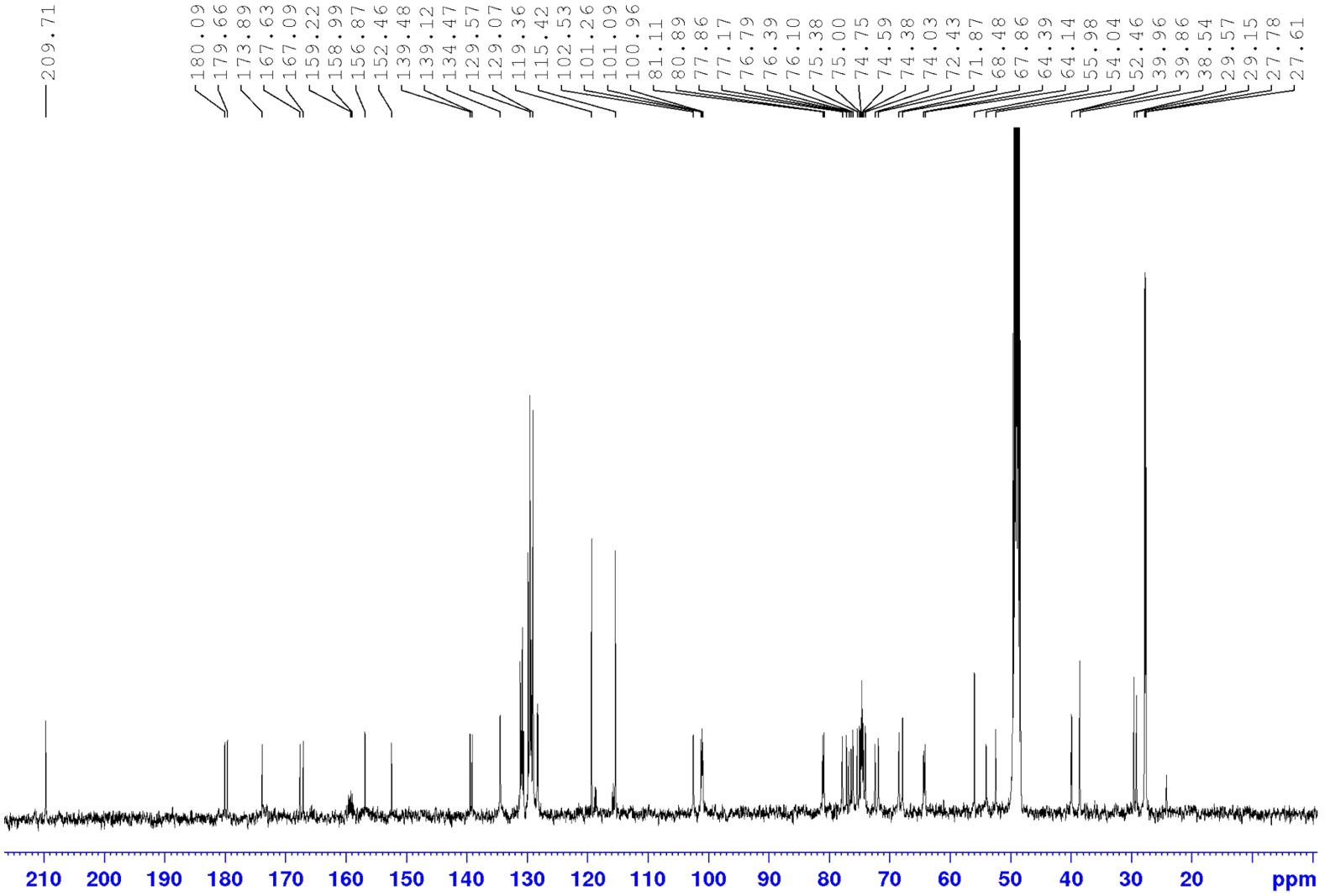
Compound 5, ^{1}H-NMR, 400 MHz, CD_3OD



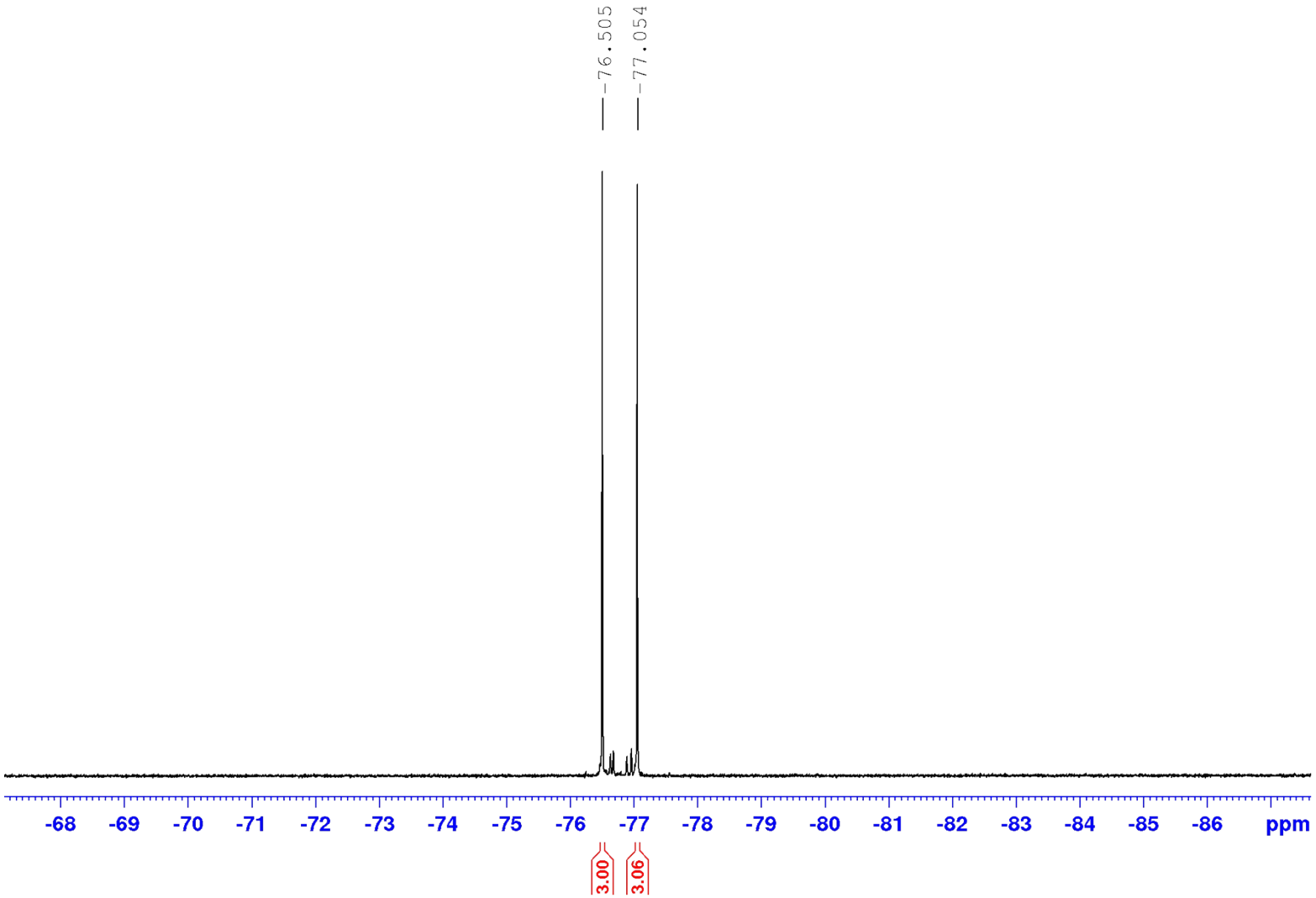
Compound 5, ^{1}H-NMR, 400 MHz, CD_3OD



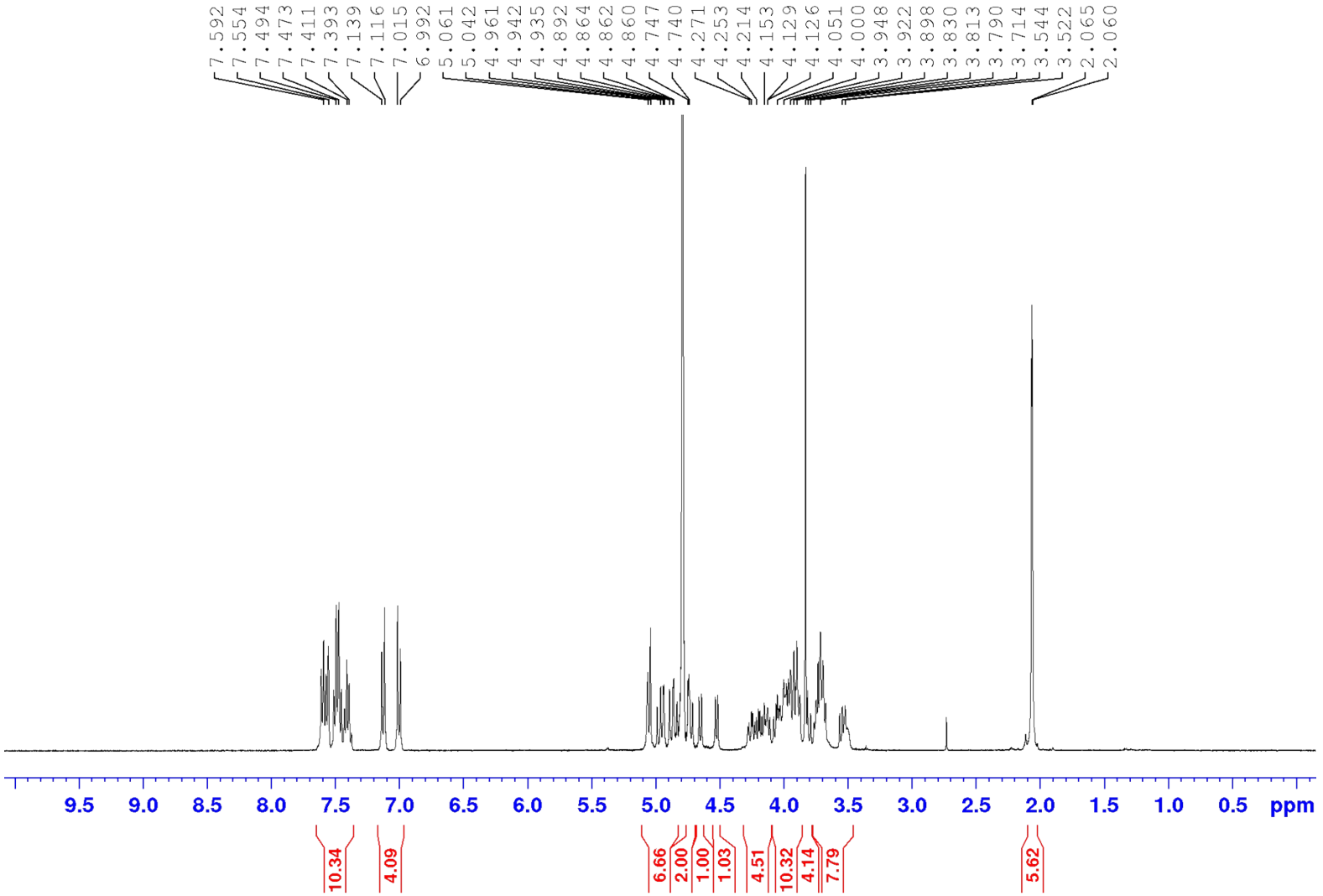
Compound 5, ^{13}C-NMR, 100 MHz, CD_3OD



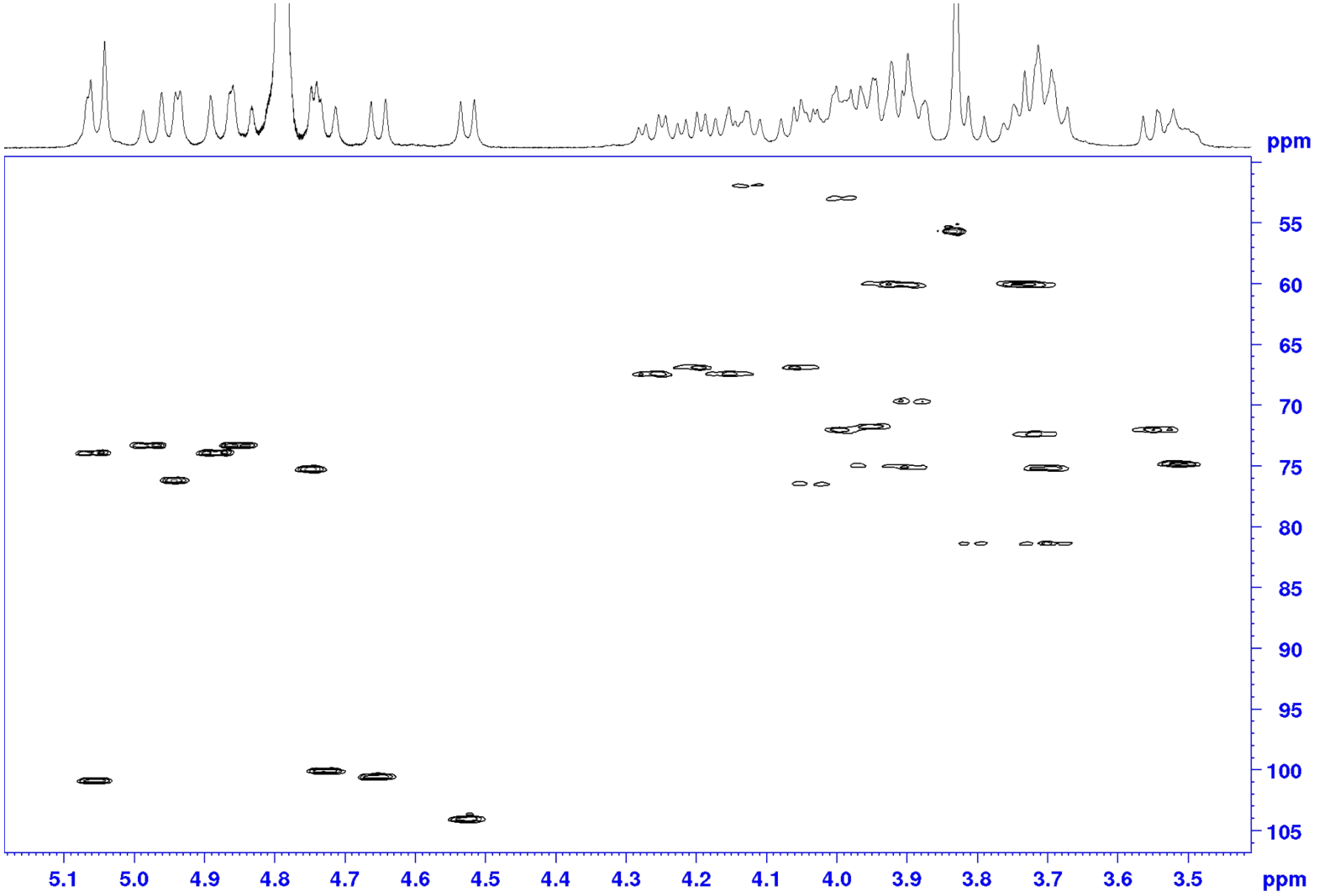
Compound 5, ^{19}F-NMR, 376 MHz, CD_3OD



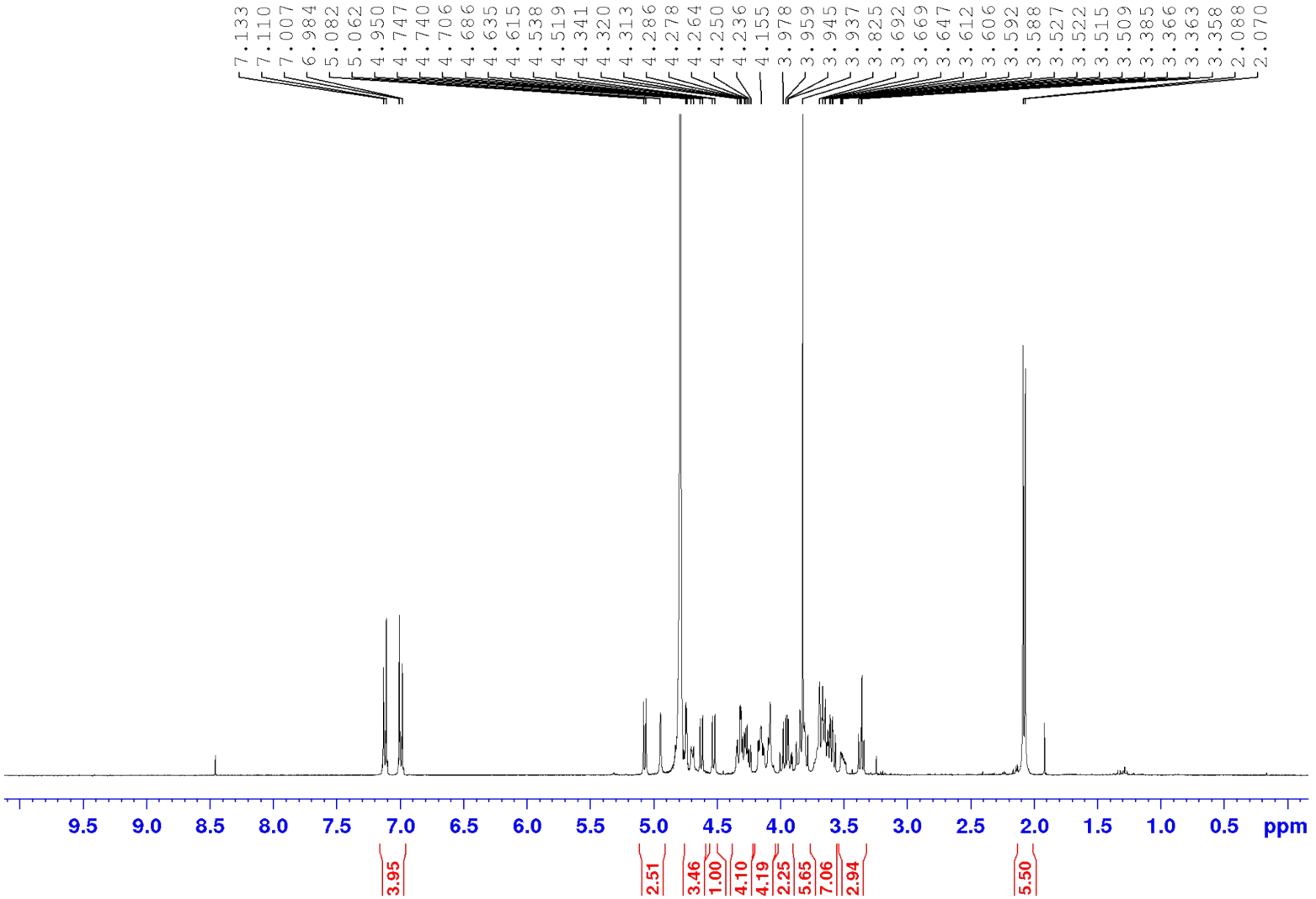
Compound 3, ^{1}H-NMR, 400 MHz, D_2O



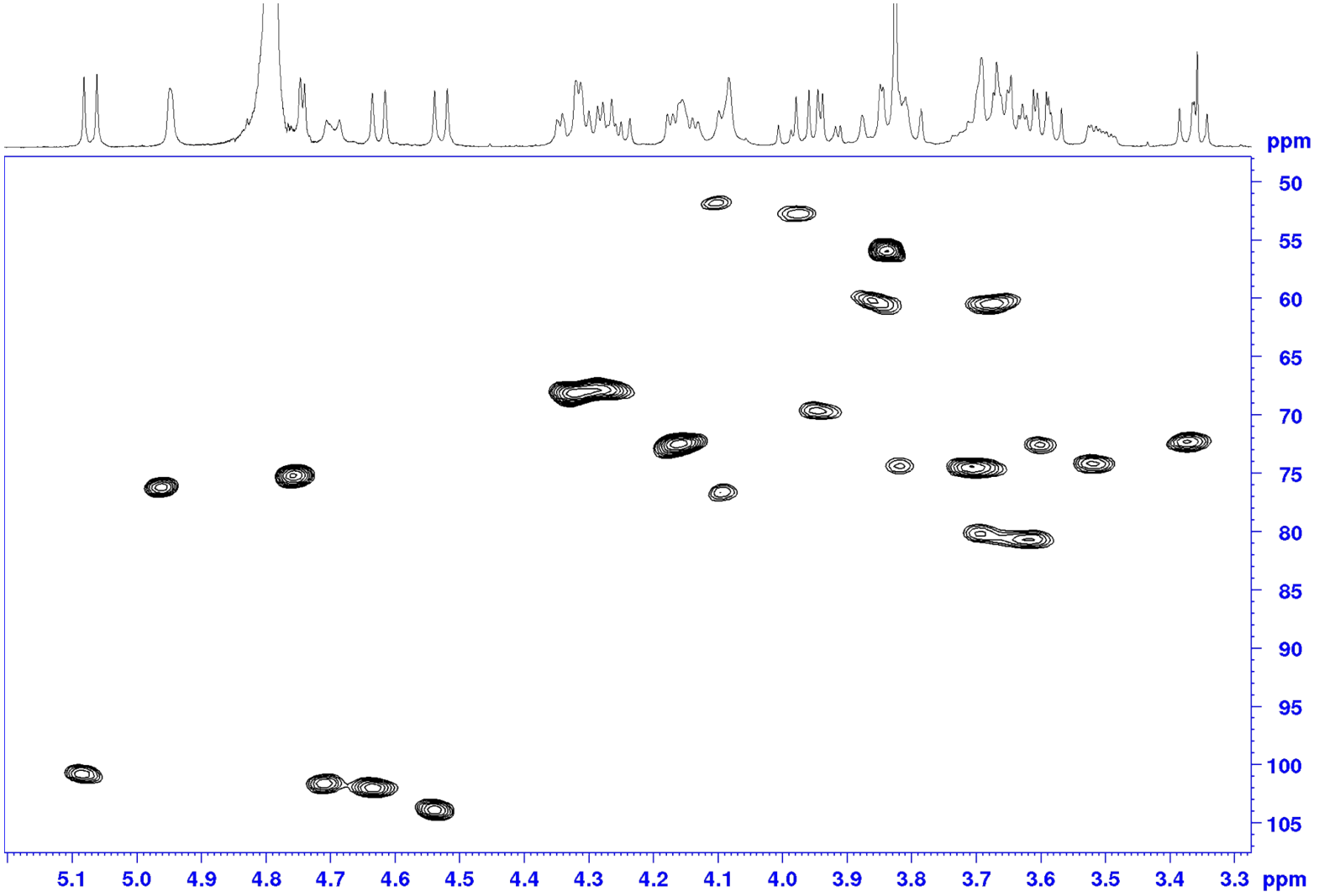
Compound 3, HSQC, 400 MHz, D_2O



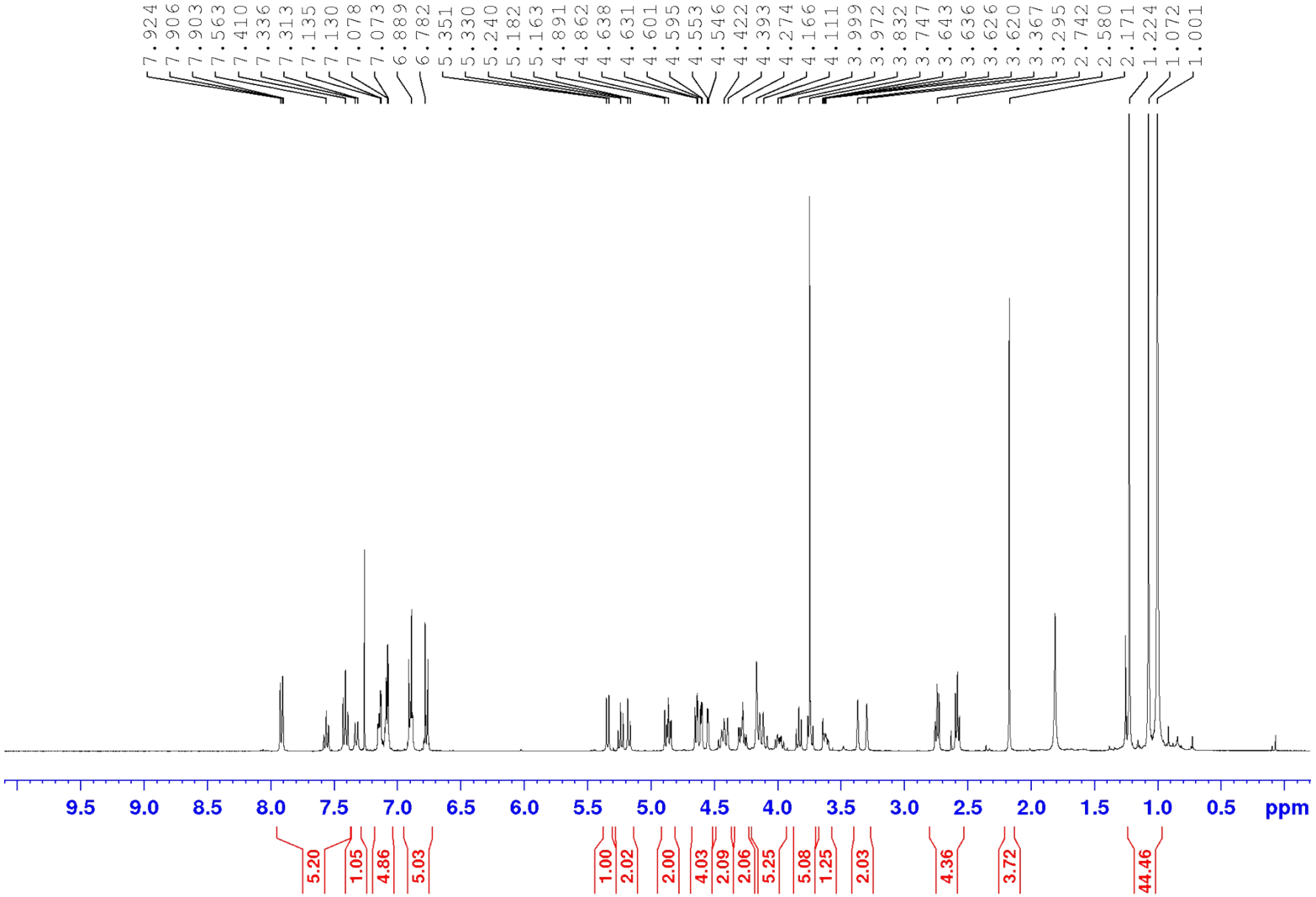
Compound 1, ^{1}H-NMR, 400 MHz, D_2O



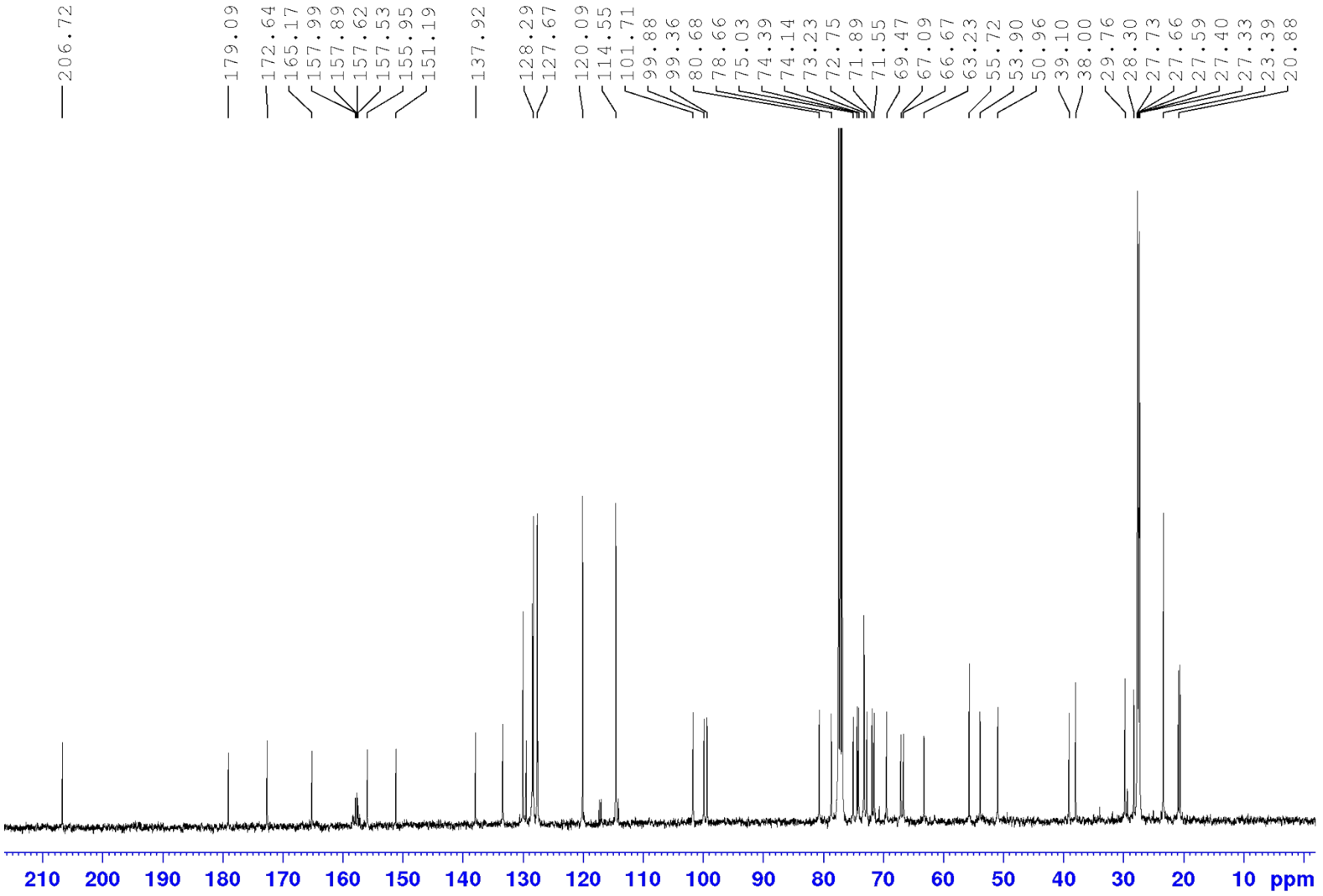
Compound 1, HSQC, 400 MHz, D_2O



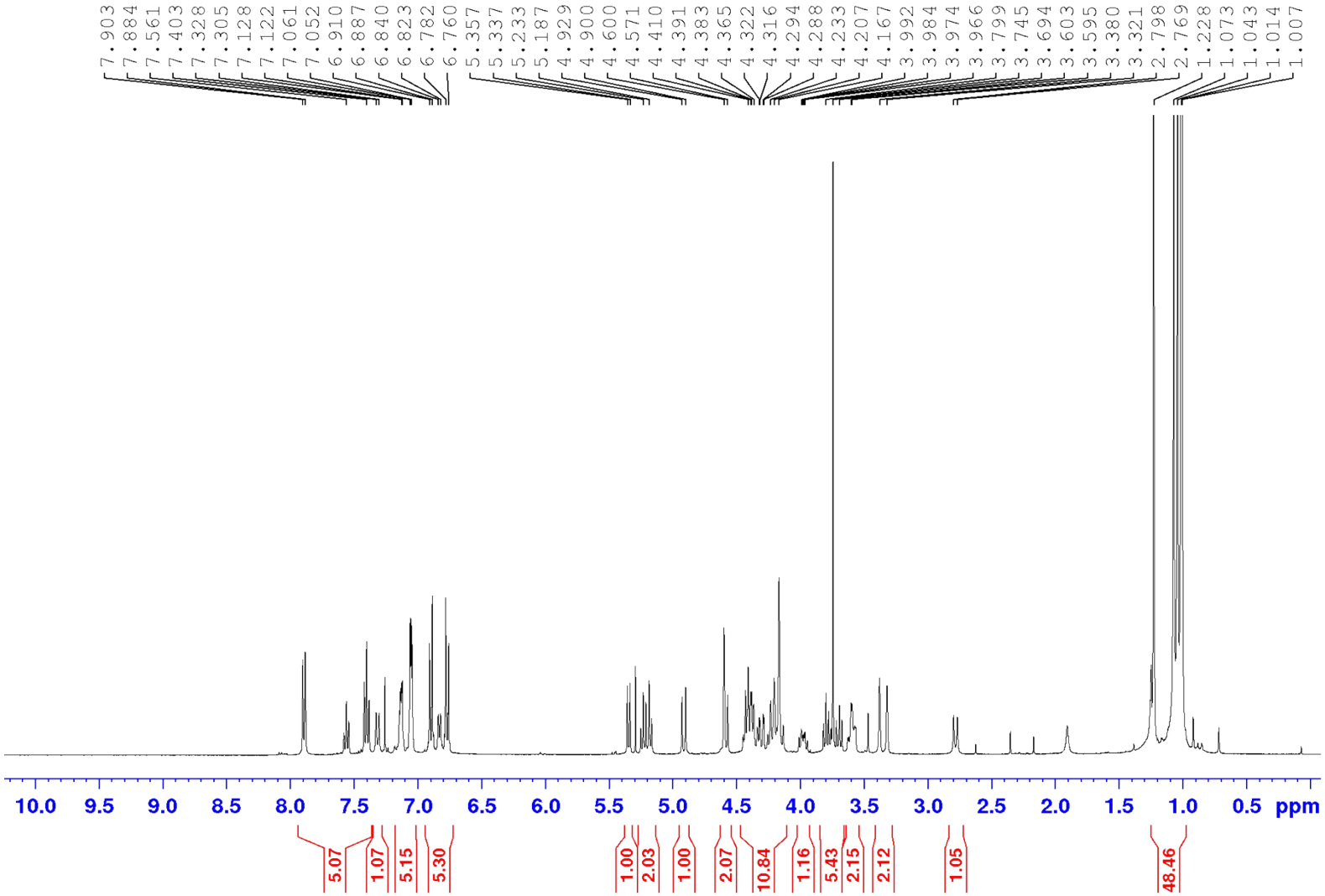
Compound 17, ^{1}H-NMR, 400 MHz, $CDCl_3$



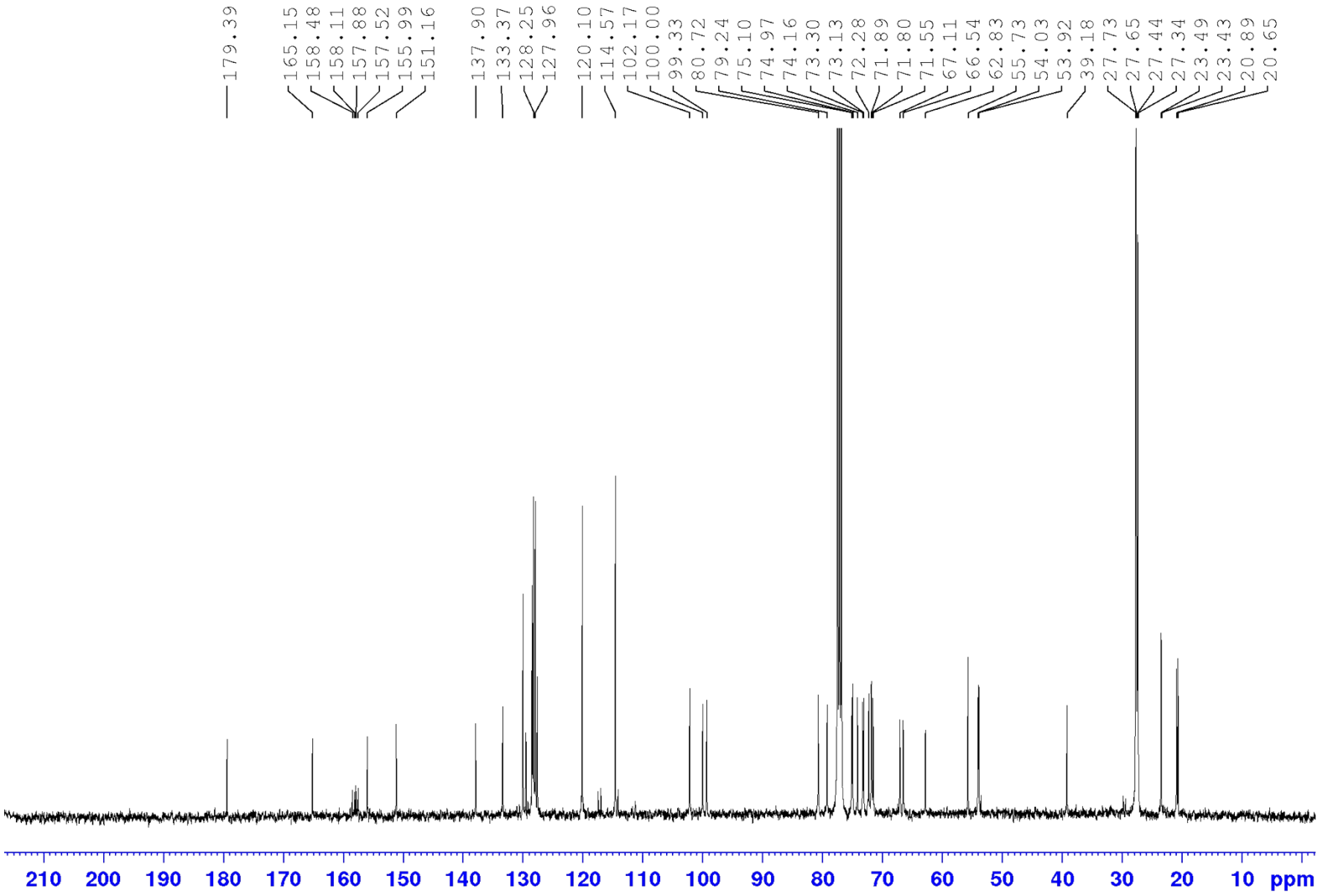
Compound 17, ^{13}C-NMR, 100 MHz, $CDCl_3$



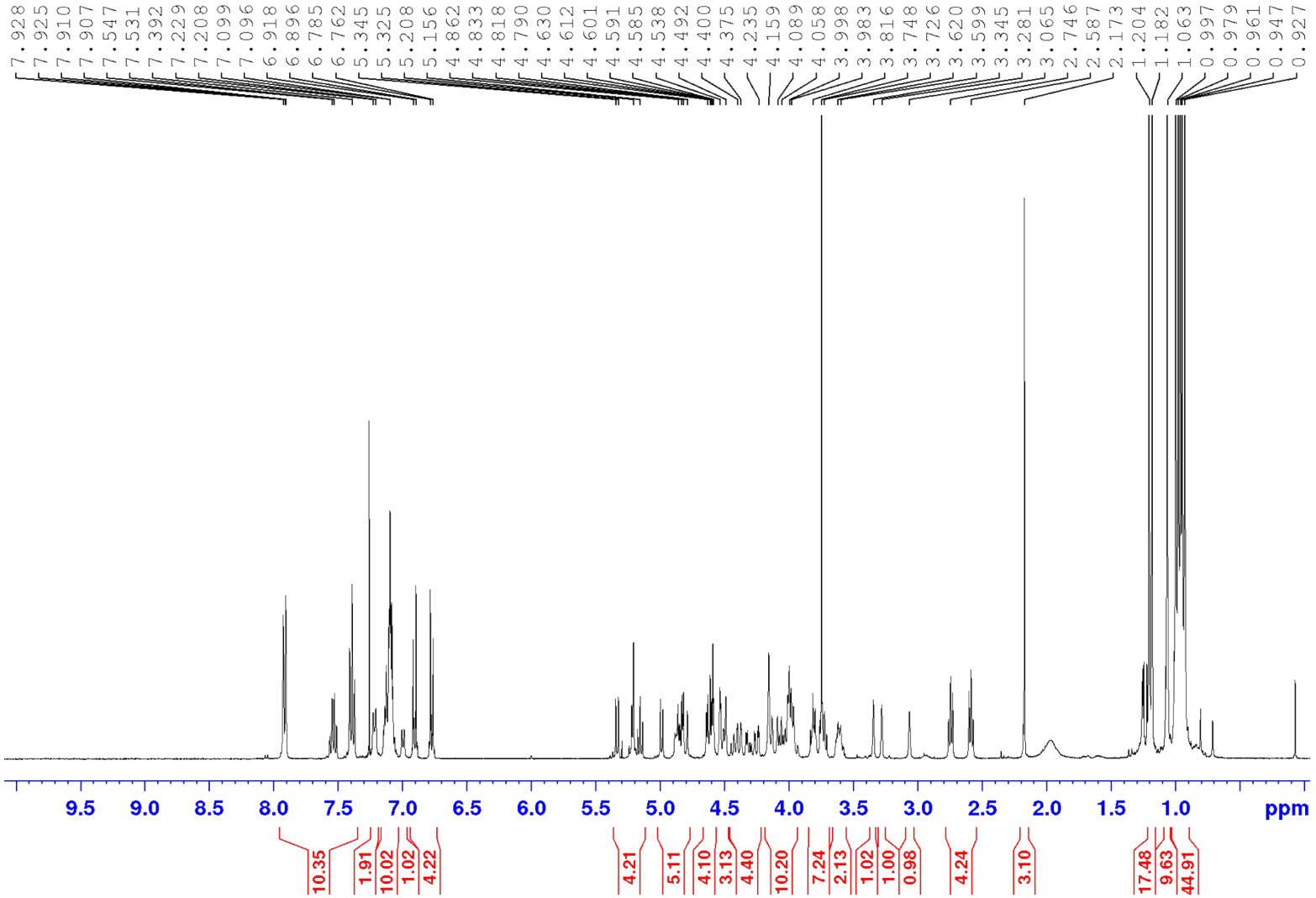
Compound 18, ^{1}H-NMR, 400 MHz, $CDCl_3$



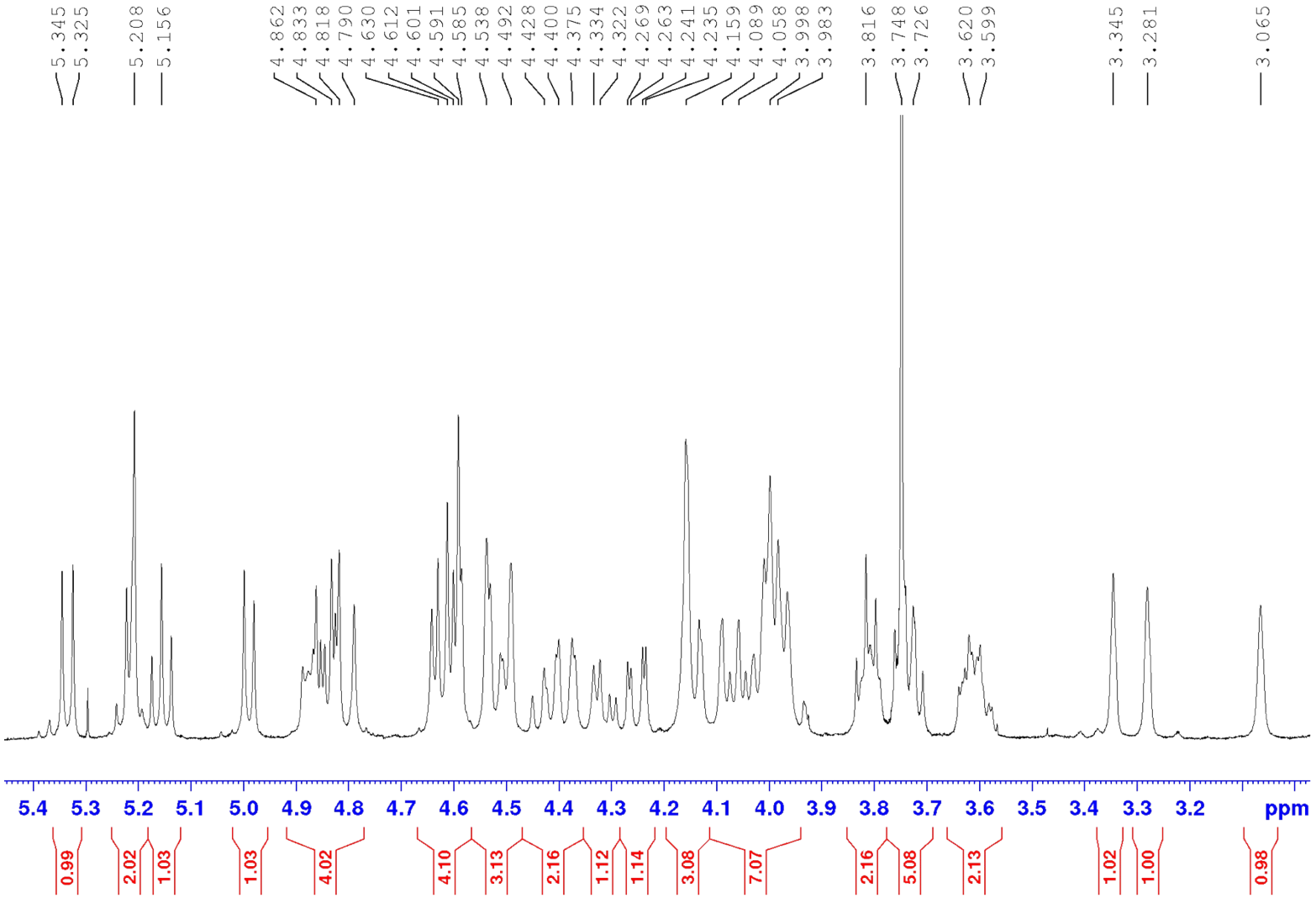
Compound 18, ^{13}C-NMR, 100 MHz, $CDCl_3$



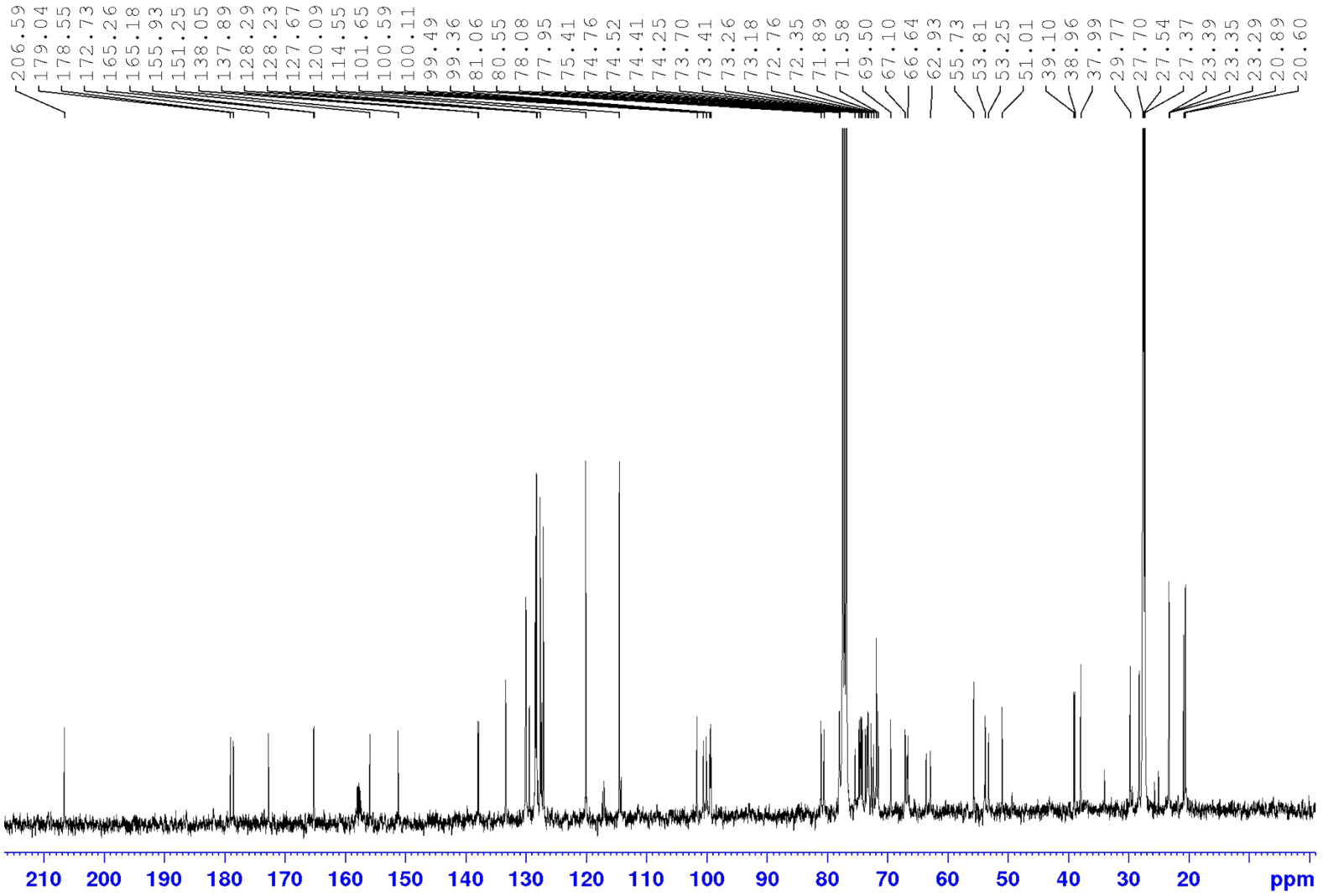
Compound 19, ^{1}H-NMR, 400 MHz, $CDCl_3$



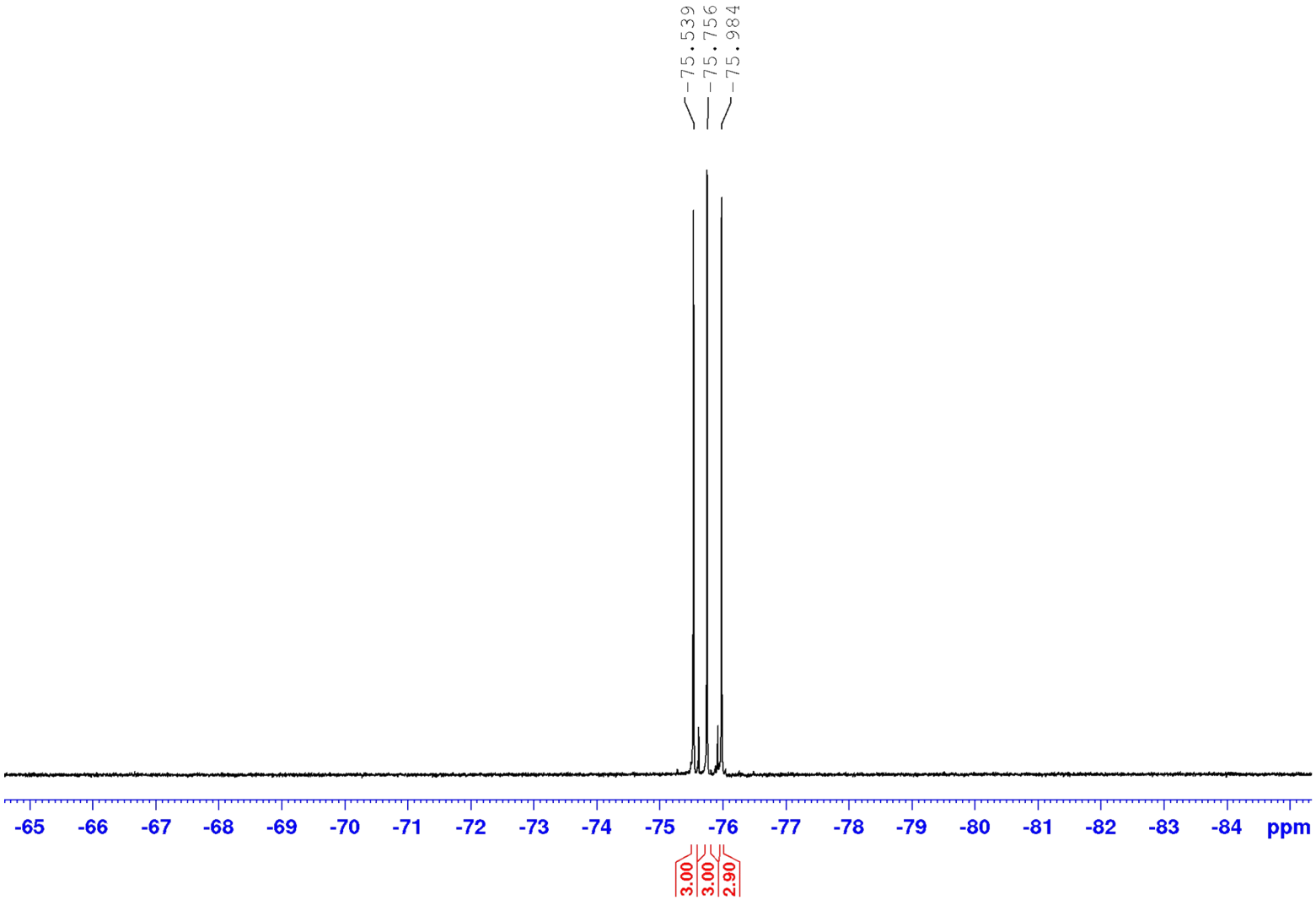
Compound 19, ^{1}H-NMR, 400 MHz, $CDCl_3$



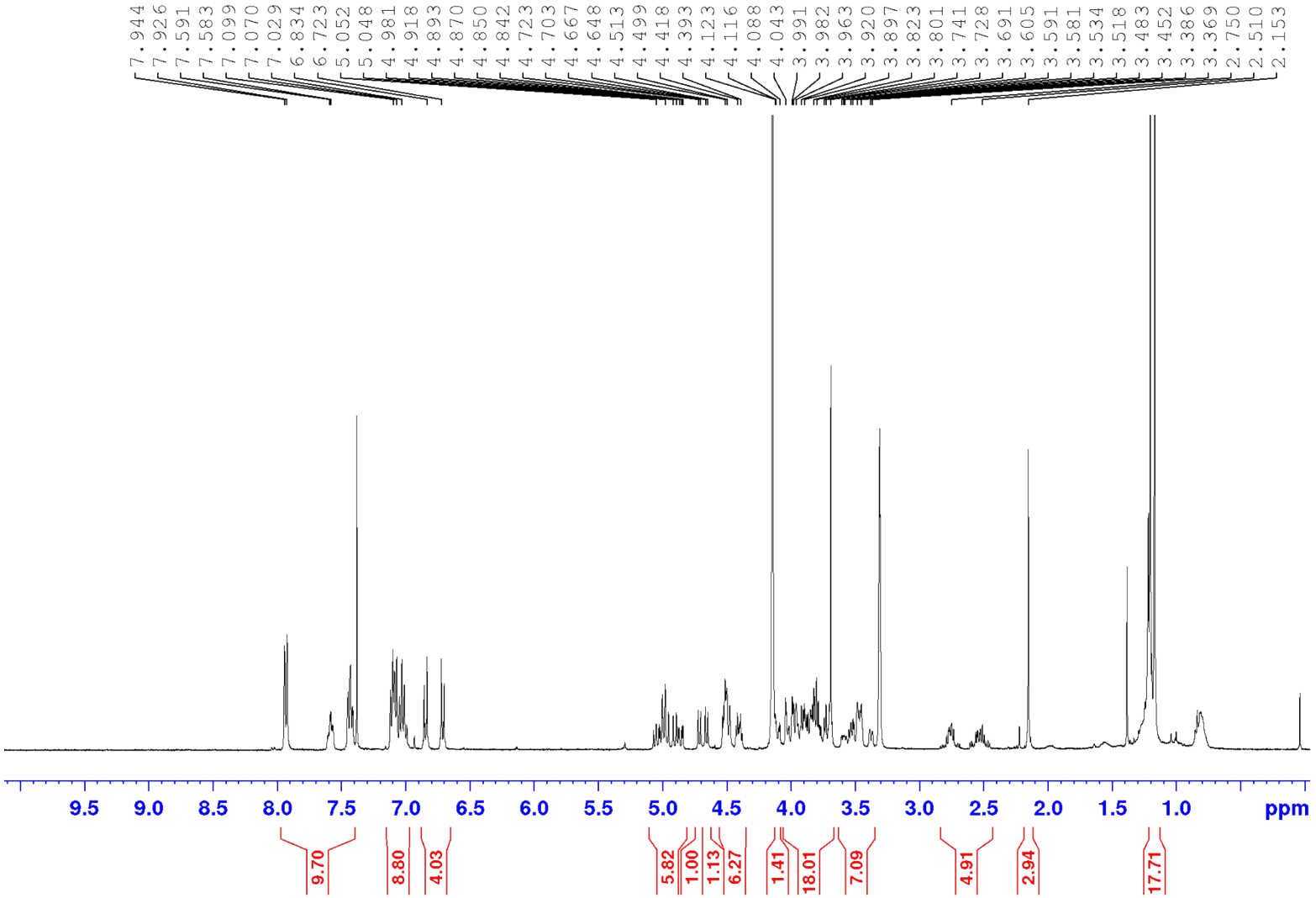
Compound 19, ^{13}C-NMR, 100 MHz, $CDCl_3$



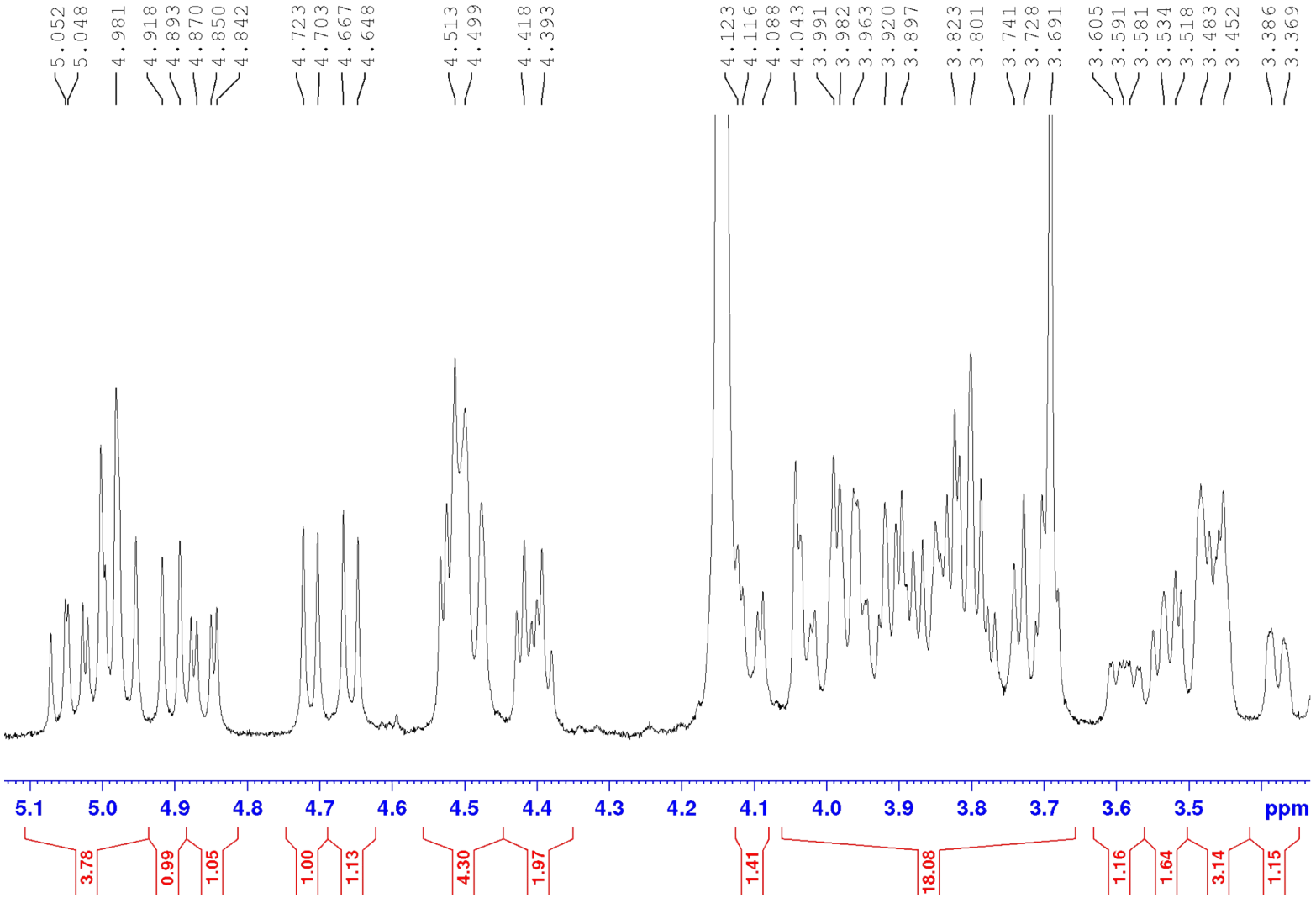
Compound 19, ^{19}F-NMR, 376 MHz, $CDCl_3$



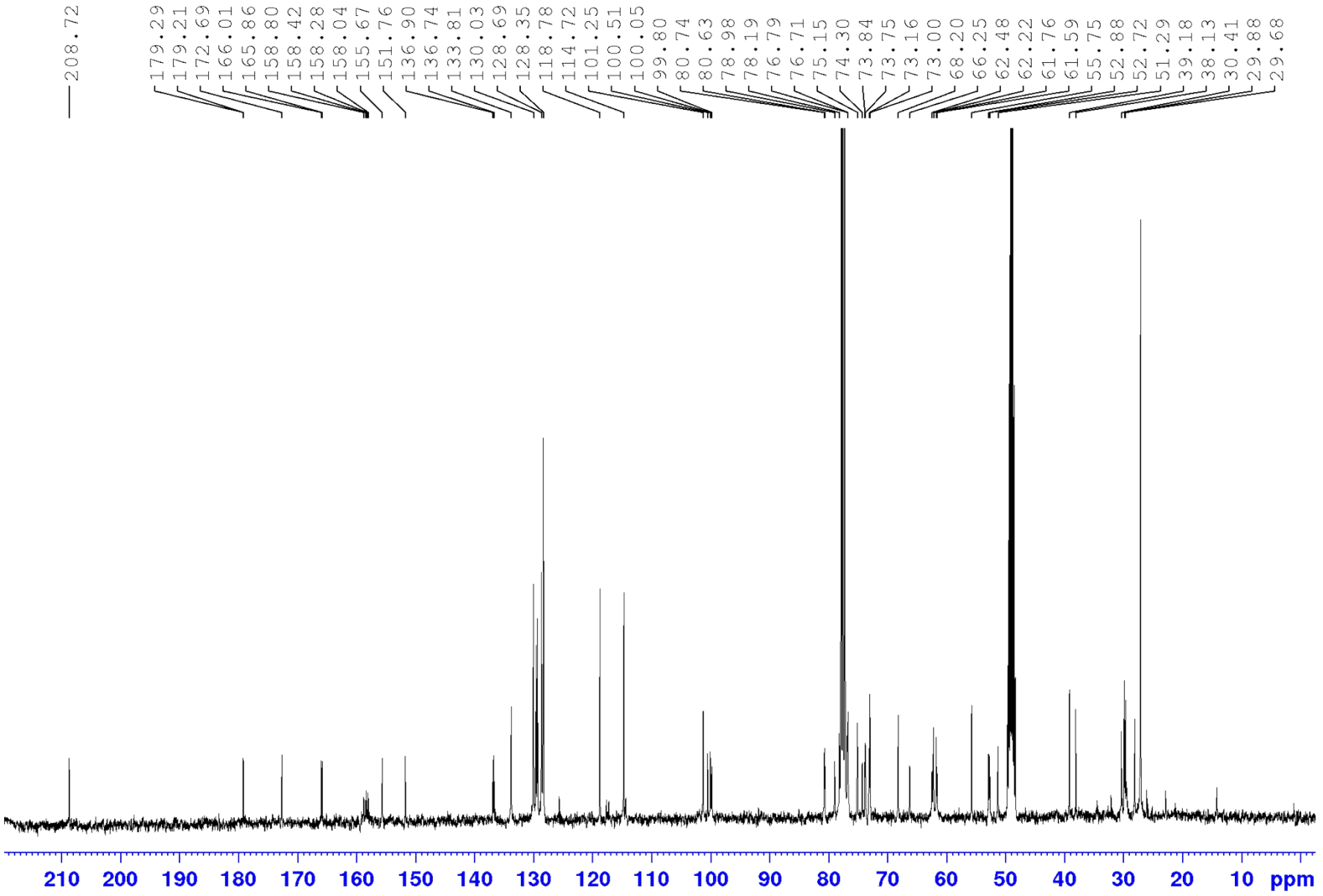
Compound 20, ^{1}H-NMR, 400 MHz, $CDCl_3/CD_3OD$ 4:1



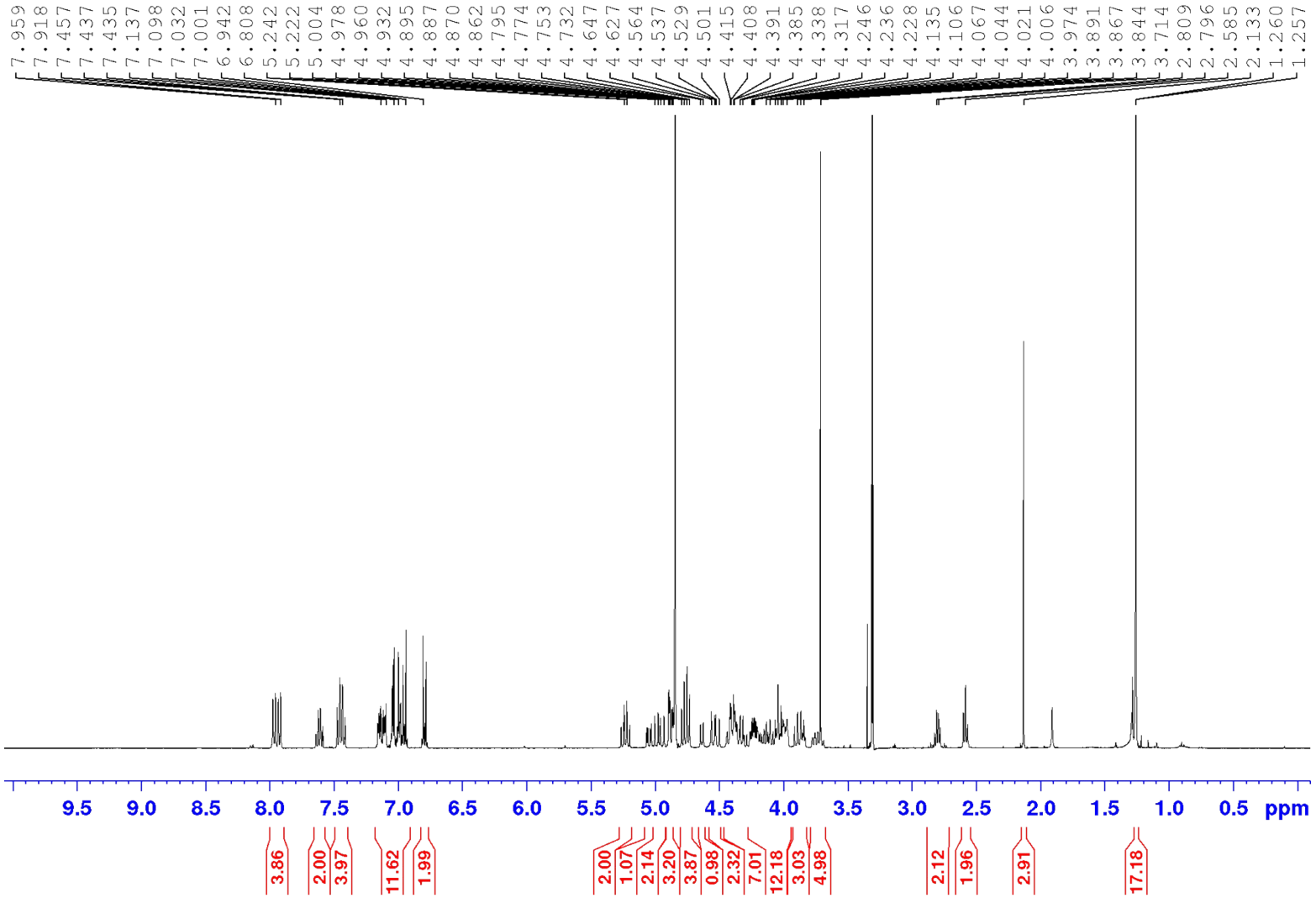
Compound 20, ^{1}H-NMR, 400 MHz, $CDCl_3/CD_3OD$ 4:1



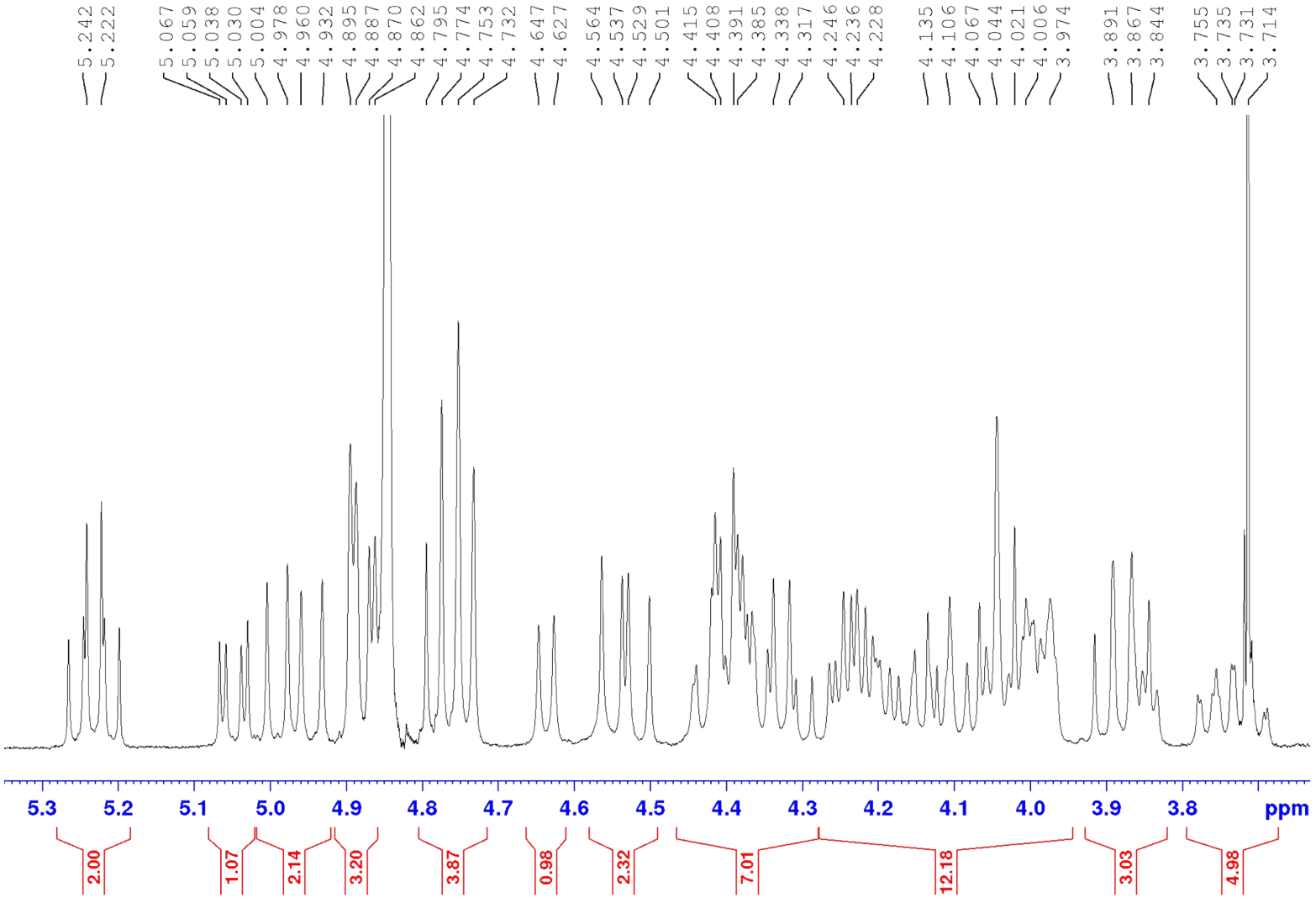
Compound 20, ^{13}C-NMR, 100 MHz, $CDCl_3/CD_3OD$ 4:1



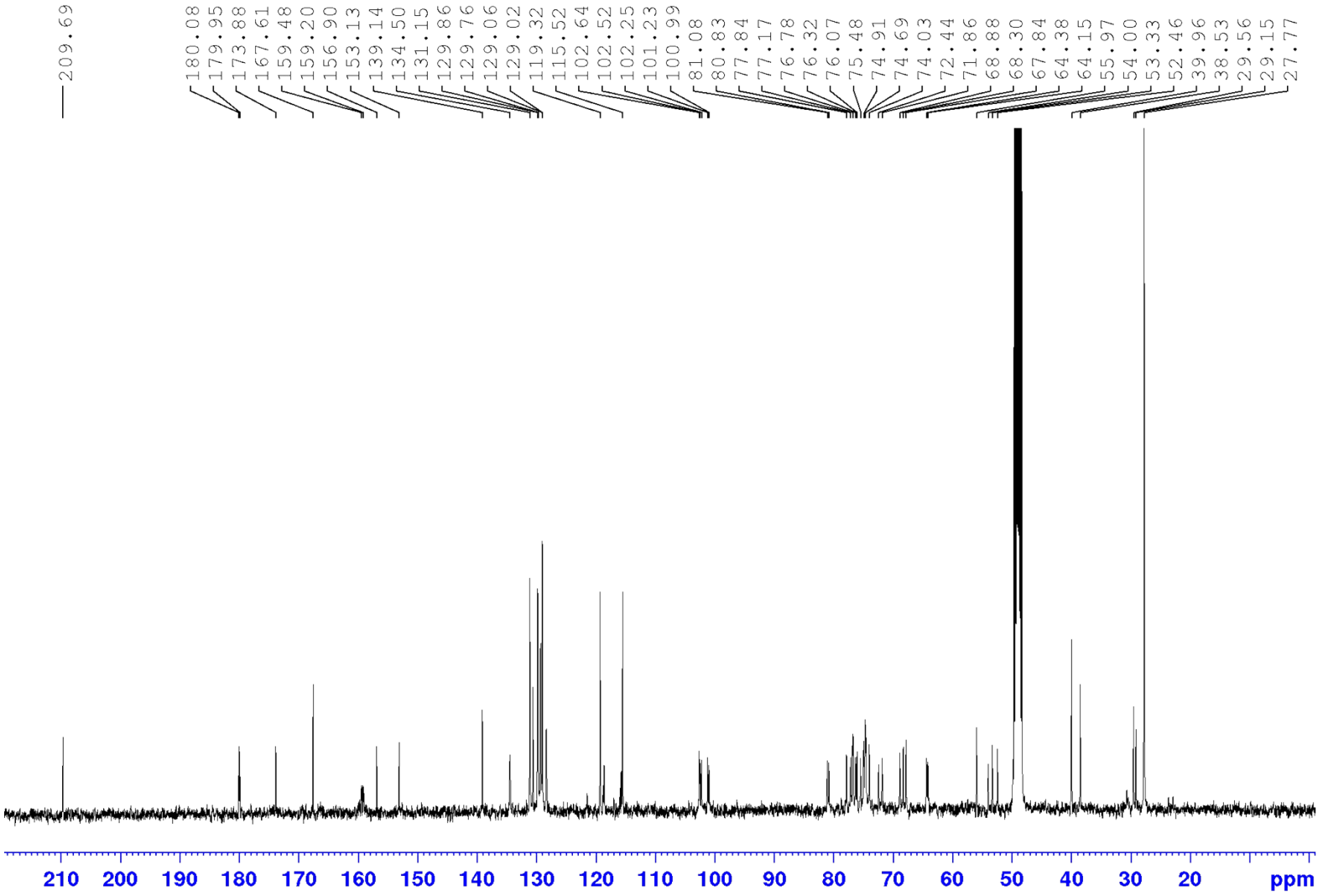
Compound 6, ^{1}H-NMR, 400 MHz, CD_3OD



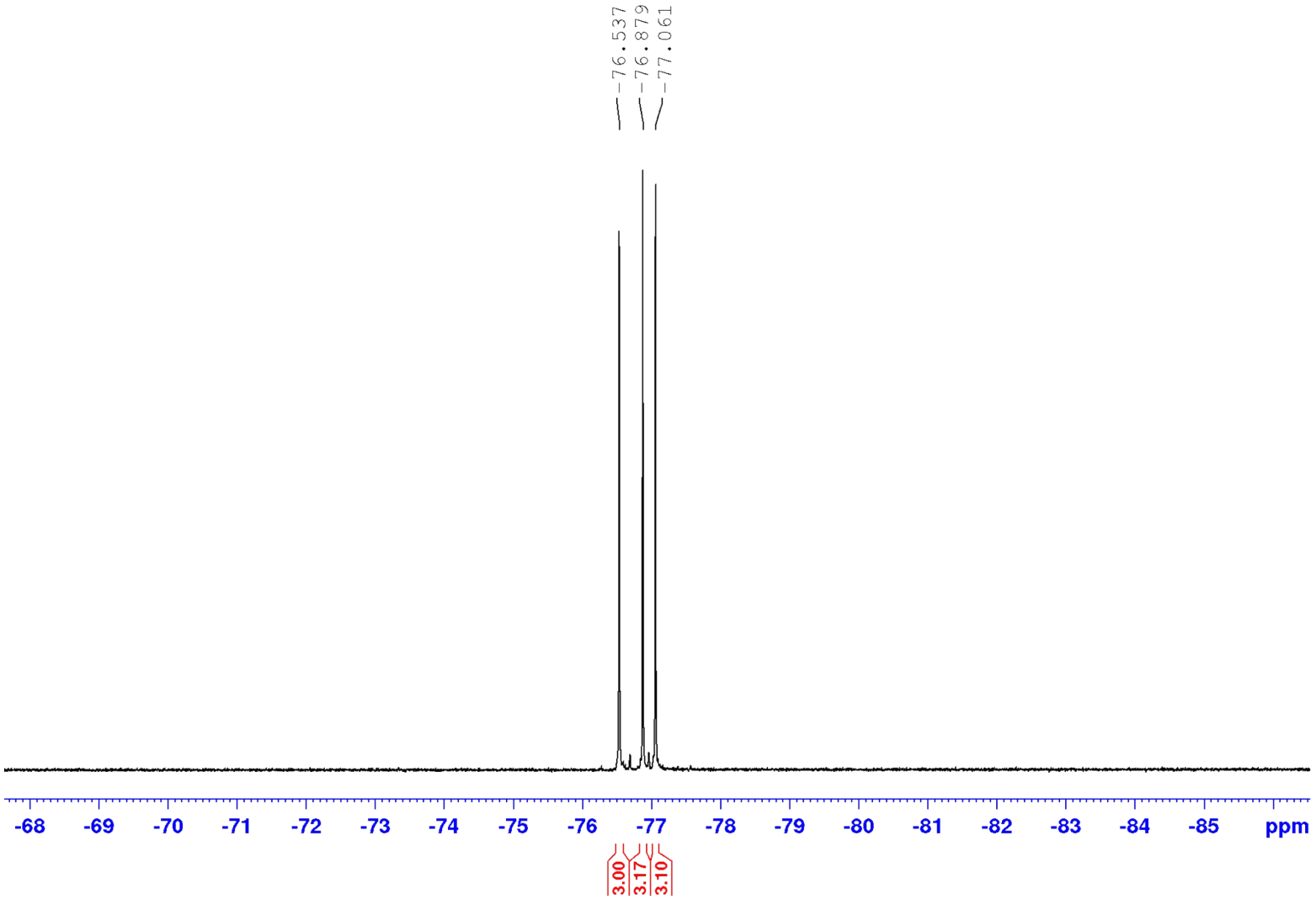
Compound 6, ^{1}H-NMR, 400 MHz, CD_3OD



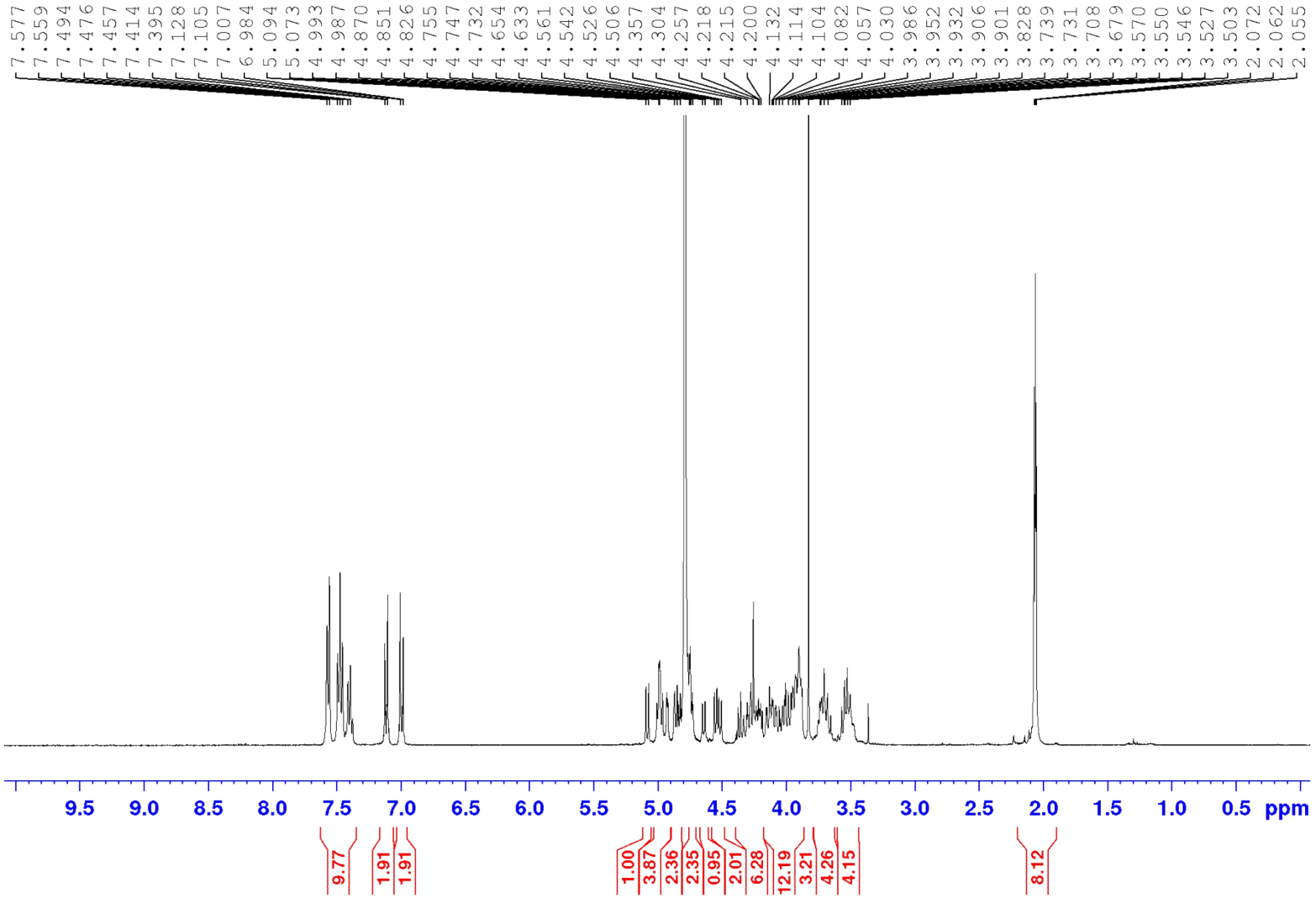
Compound 6, ^{13}C-NMR, 100 MHz, CD_3OD



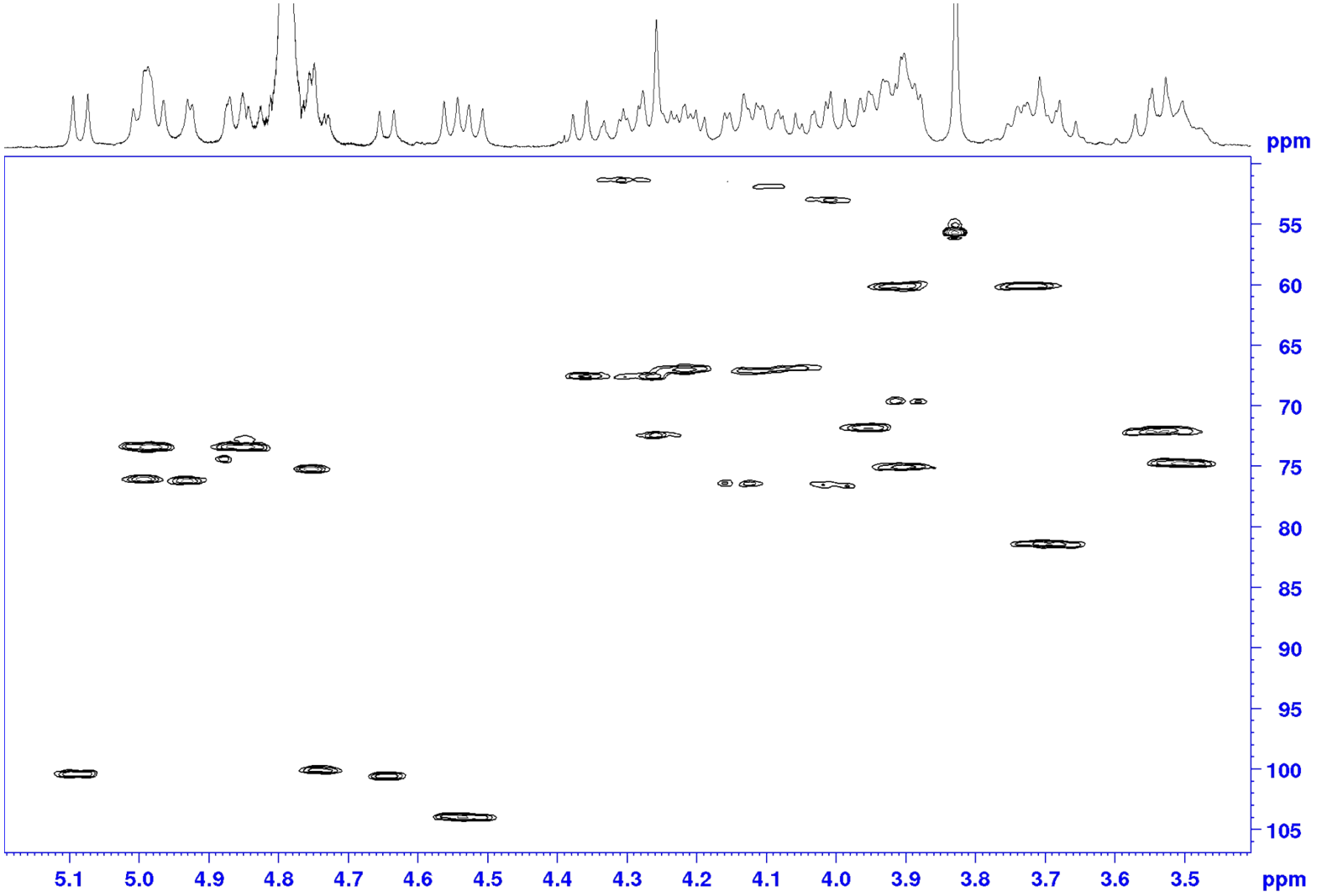
Compound 6, ^{19}F-NMR, 376 MHz, CD_3OD



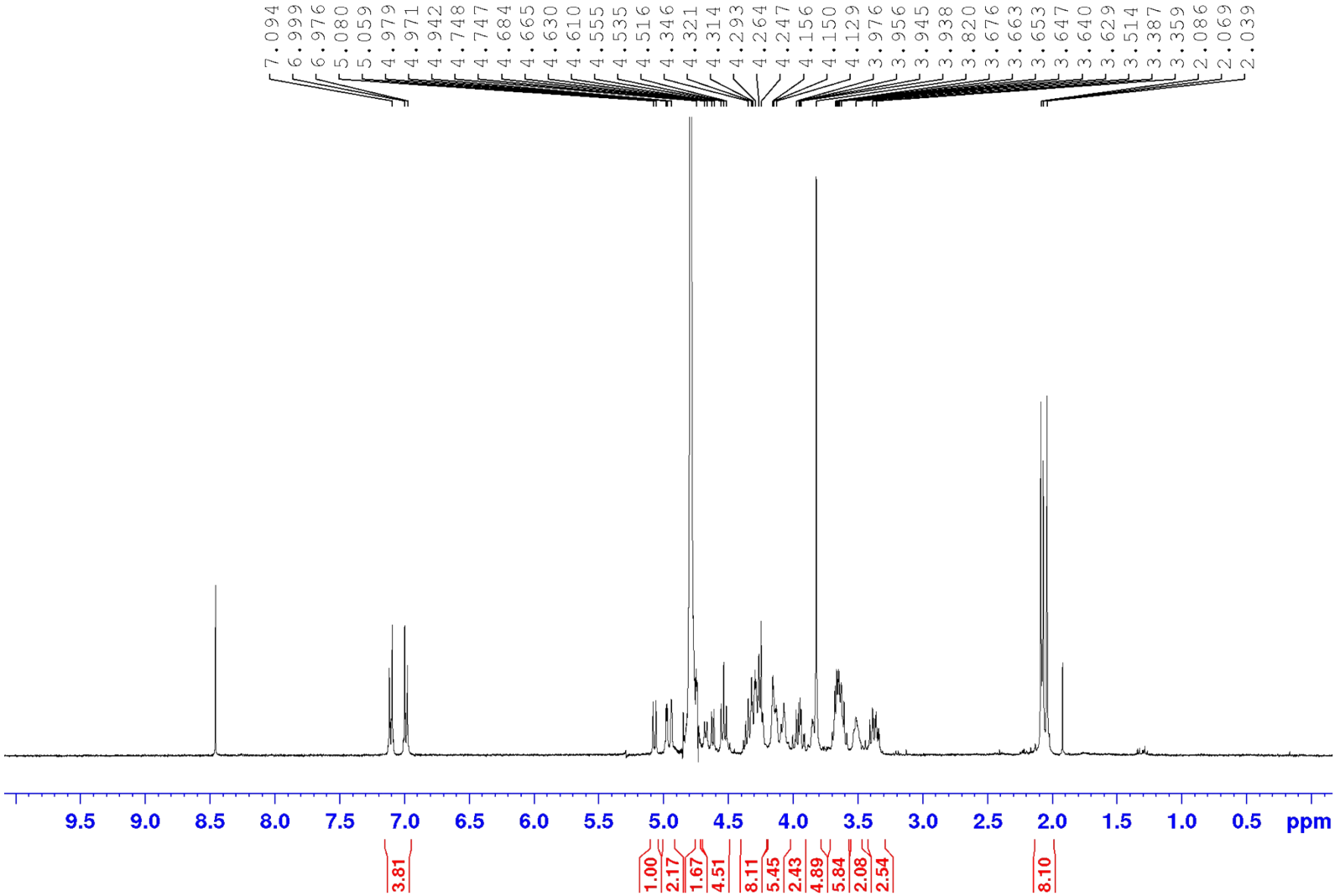
Compound 4, ^{1}H-NMR, 400 MHz, D_2O



Compound 4, HSQC, 400 MHz, D_2O



Compound 2, ^{1}H-NMR, 400 MHz, D_2O



Compound 2, HSQC, 400 MHz, D_2O