Supporting information for:

Synthesis and NMR spectroscopic analysis of acylated pentasaccharide fragments of mycobacterial arabinogalactan

Chunjuan Liu, Michele R. Richards, Todd L. Lowary*

Alberta Ingenuity Centre for Carbohydrate Science and Department of Chemistry,

The University of Alberta, Gunning-Lemieux Chemistry Centre,

Edmonton, Alberta, T6G 2G2 Canada

Email:tlowary@ualberta.ca

Table of Contents

General methods.

Reactions were carried out 1n oven-dried glassware. Reactions solvents were distilled from appropriate drying agent before use. Unless stated otherwise, all reactions were carried out at room temperature under a positive pressure of argon and were monitored by TLC on Silica Gel 60 F_{254} (0.25 mm, E. Merck). Spots were detected under UV light or by charring with acidified *p*-anisaldehyde solution in EtOH. All column chromatography experiments were performed on silica gel (40–60 μM) or Iatrobeads, which refers to a beaded Silica Gel 6RS-8060 manufactured by Iatron Laboratories (Tokyo). The ratio between silica gel and crude product ranged from 100 to 50:1 (w:w). Optical rotations were measured at 22 ± 2 °C. ^{1}H NMR spectra were recorded at 400, 500, or 600 MHz, and chemical shifts are referenced to either TMS (0.0 ppm, $CDCl_3$) or CD_3OD (3.30 ppm, CD_3OD). ^{1}H data are reported as though they were first order. ^{13}C NMR APT spectra were recorded at 125 MHz, and ^{13}C chemical shifts are referenced to internal $CDCl_3$ (77.23 ppm, $CDCl_3$) or CD_3OD (48.9 ppm, CD_3OD). Assignments of resonance in NMR spectra were made on the basis of 2D NMR (^{1}H–^{1}H COSY, TOCSY, HMQC, HSQC, and HMBC) experiments. In the processing of reaction mixtures, solutions of organic solvent were washed with equal amounts of aqueous solutions. Organic solutions of crude products were dried over anhydrous $MgSO_4$. Solvents were concentrated under vacuum at <40 °C. Electrospray mass spectra were recorded on samples suspended in mixtures of THF with CH_3OH and added NaCl.

General methods for NMR spectra used in conformational analysis.

Individual 1D-ge-CSSF-TOCSY[1,2] spectra of **4–8** were recorded at 300 K on a 600 MHz spectrometer, following the pulse sequence described by Duncan, et al.[2] The solvent was either CD_3OD or 1:1 CD_3OD–$CDCl_3$, depending on the solubility of the compound. For the 1D ^{1}H

NMR spectra, presaturation was used to eliminate any residual H_2O signals; presaturation was not used for the 1D-ge-CSSF-TOCSY spectra. Figure 4 in the main text and Figures S1–S5 below show the TOCSY spectra for **4–8**. The captions indicate the following parameters used to acquire each spectrum: the mixing times (mix) and the frequency differences (Δ) between the irradiated peak and its nearest neighbor.

During the data analysis, no window functions or line broadening were applied. Chemical shifts and coupling constants were determined by a combination of first-order spectral analysis and simulation in WinDNMR[3] until a good correlation (i.e. Pearson's $r \geq 0.95$) was obtained between the simulated and real spectra. The resultant simulated TOCSY spectra were then summed and compared with the 1D ^{1}H NMR spectrum for each compound. See Figure S6 for a representative simulation of a 1D-ge-CSSF-TOCSY spectrum, and Figure S7 for a representative simulation of a 1D ^{1}H NMR spectrum. The chemical shifts for the compounds **4–8** can be found in Table S1.

Individual 1D-ge-CSSF-DPFGSE-NOESY[1,2] spectra of **4** and **5** were recorded at 300 K on a 600 MHz spectrometer, following the pulse sequence described by Duncan, et al.[2] The solvent was either CD_3OD or 1:1 CD_3OD–$CDCl_3$, depending on the solubility of the compound. A cosine-squared window function of 1.5 (one-half the acquisition time) was applied to each NOESY spectrum when the data were Fourier transformed, but no line broadening was applied. Figure S8 shows the NOESY spectra for **4** and Figure S10 shows the NOESY spectra for **5**. The captions for both figures indicate the following parameters used to acquire each spectrum: the mixing times (mix) and the frequency differences (Δ) between the irradiated peak and its nearest neighbor.

Determining the rotamer populations about the C-4–C-5 bond.

Using methods described previously,[4-6] rotamer populations about the C-4–C-5 bond were calculated for the terminal rings in compounds **4**–**8**. Equations 1 and 2, Karplus relationships[7] derived specifically for arabinofuranose,[8] were used for the parent saccharide **4**.

$$^{3}J_{H-4,H-5R} = 5.23 + 0.02\cos(\phi + 15.1^{\circ}) + 4.67\cos(2\phi + 30.2^{\circ}) \qquad (1)$$

$$^{3}J_{H-4,H-5S} = 4.95 - 0.42\cos(\phi) + 4.03\cos(2\phi) \qquad (2)$$

The angle ϕ is the dihedral angle between the coupled protons. These equations were solved to give the limiting coupling constants for the three staggered conformations about the C-4–C-5 bond, where ϕ = 60º, 180º, and –60º. The coupling constant measured in solution represents a weighted average of these staggered conformations; the population of each conformation was determined for each ring using equations 3–5 below, which were solved for compound **4**, and the results can be found in Table 2 in the paper.

$$^{3}J_{H-4,H-5R} = 1.18X_{gg} + 9.25X_{gt} + 5.26X_{tg} \qquad (3)$$

$$^{3}J_{H-4,H-5S} = 2.73X_{gg} + 2.73X_{gt} + 9.40X_{tg} \qquad (4)$$

$$1 = X_{gg} + X_{gt} + X_{tg} \qquad (5)$$

For the acylated saccharides **4**–**8**, we developed new Karplus-type relationships for $^{3}J_{\text{H-4,H-5}R}$ and $^{3}J_{\text{H-4,H-5}S}$ using DFT calculations.[9] The newly derived equations 6 and 7 describe the $^{3}J_{\text{H-4,H-5}R}$ and $^{3}J_{\text{H-4,H-5}S}$ relationships for both anomers of methyl 5-*O*-acetyl-D-arabinofuranoside.

$$^{3}J_{H-4,H-5R} = 5.27 - 0.51\cos(\phi) + 4.83\cos(2\phi) + 0.02\sin(\phi) - 0.04\sin(2\phi) \qquad (6)$$

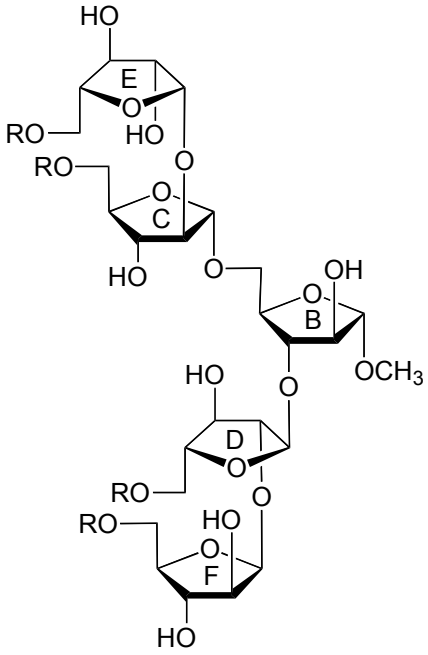
$$^{3}J_{H-4,H-5S} = 5.34 - 0.62\cos(\phi) + 4.50\cos(2\phi) + 0.07\sin(\phi) + 2.06\sin(2\phi) \qquad (7)$$

Again, the angle ϕ is the dihedral angle between the coupled protons. These equations were solved to give the coupling constants for the staggered conformations about the C-4–C-5 bond, as described above, to give equations 8–10 below. These equations were solved for compounds **5**–**8**, and the results can also be found in Table 2 in the paper.

$$^{3}J_{H-4,H-5R} = 2.58X_{gg} + 10.62X_{gt} + 2.63X_{tg} \qquad (8)$$

$$^{3}J_{H-4,H-5S} = 4.50X_{gg} + 1.05X_{gt} + 10.46X_{tg} \qquad (9)$$

$$1 = X_{gg} + X_{gt} + X_{tg} \qquad (10)$$



4 R = H
5 R = $C(O)C_3H_7$
6 R = $C(O)C_9H_{19}$
7 R = $C(O)C_{15}H_{31}$
8 R = $C(O)C_{21}H_{43}$

Chart S1. Target pentasaccharides **4–8** with the rings labeled for clarity.

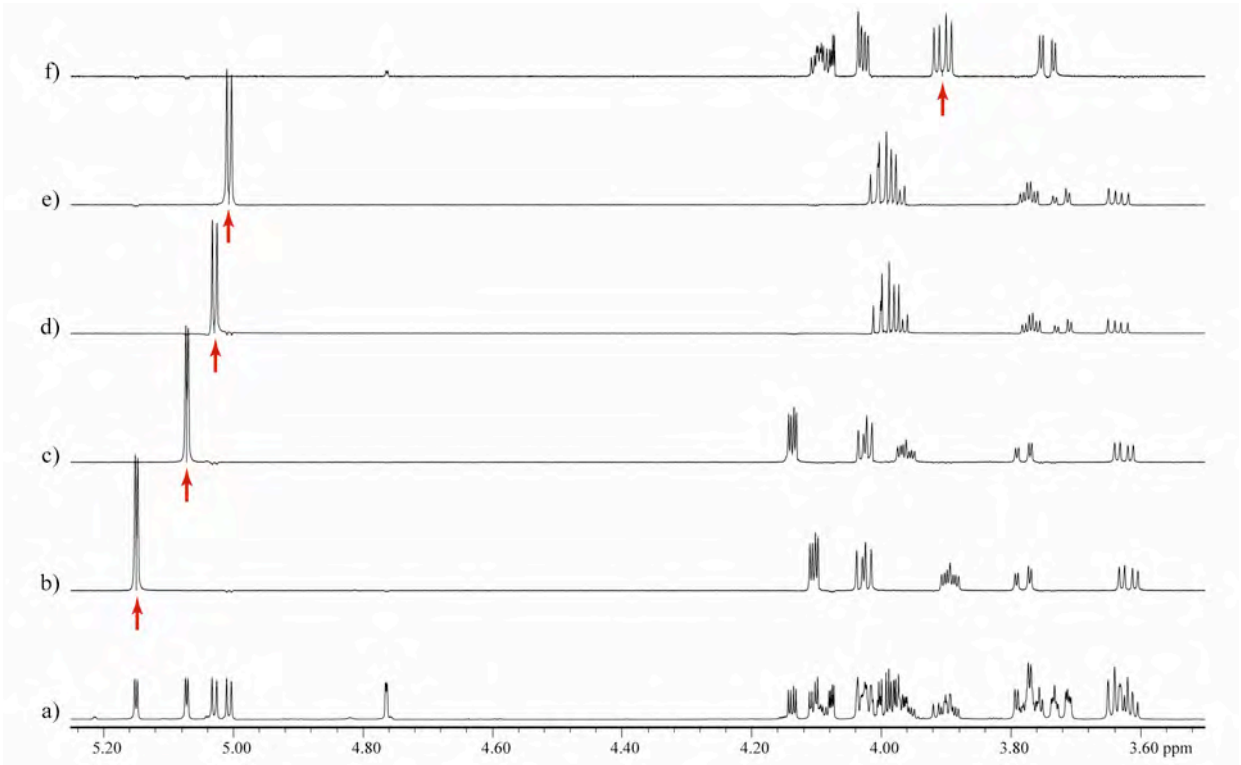


Figure S1. Chemical shift selective filtering TOCSY spectra for **4** in CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *f* show 1D-ge-CSSF-TOCSY of each ring with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.15 ppm, mix = 0.2 s, Δ = 23 Hz; c) ring C, selectively irradiated at 5.07 ppm, mix = 0.2 s, Δ = 17 Hz; d) ring E, selectively irradiated at 5.03 ppm, mix = 0.12 s, Δ = 3.9 Hz; e) ring F, selectively irradiated at 5.00 ppm, mix = 0.12 s, Δ = 12 Hz; f) ring B, selectively irradiated at 3.91 ppm, mix = 0.12 s, Δ = 1.0 Hz.

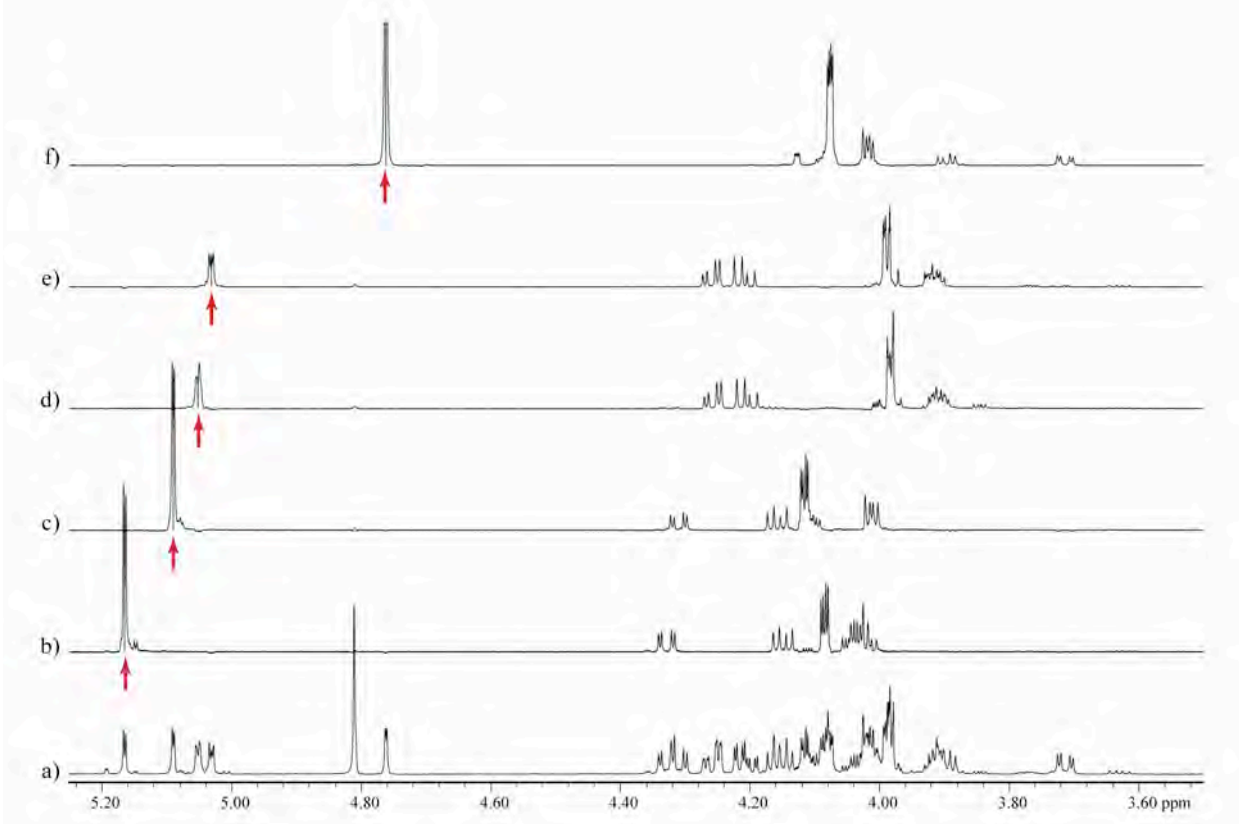


Figure S2. Chemical shift selective filtering TOCSY spectra for **5** in CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *f* show 1D-ge-CSSF-TOCSY of each ring with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.17 ppm, mix = 0.2 s, Δ = 45 Hz; c) ring C, selectively irradiated at 5.09 ppm, mix = 0.2 s, Δ = 22 Hz; d) ring E, selectively irradiated at 5.05 ppm, mix = 0.16 s, Δ = 1.2 Hz; e) ring F, selectively irradiated at 5.03 ppm, mix = 0.16 s, Δ = 1.2 Hz; f) ring B, selectively irradiated at 4.76 ppm, mix = 0.2 s, Δ = 29 Hz.

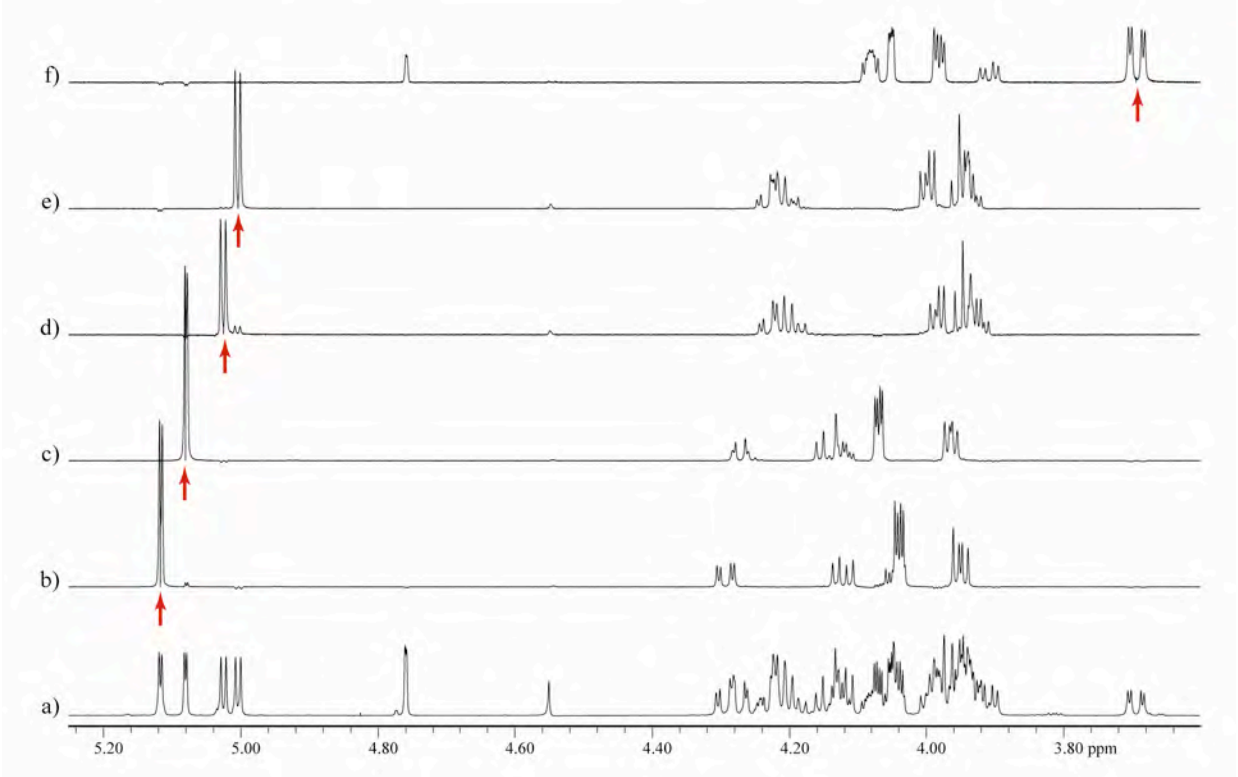


Figure S3. Chemical shift selective filtering TOCSY spectra for **6** in 1:1 $CDCl_3$–CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *f* show 1D-ge-CSSF-TOCSY of each ring with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.12 ppm, mix = 0.2 s, Δ = 1.8 Hz; c) ring C, selectively irradiated at 5.08 ppm, mix = 0.2 s, Δ = 17 Hz; d) ring E, selectively irradiated at 5.03 ppm, mix = 0.16 s, Δ = 1.8 Hz; e) ring F, selectively irradiated at 5.00 ppm, mix = 0.16 s, Δ = 1.8 Hz; f) ring B, selectively irradiated at 3.69 ppm, mix = 0.2 s, Δ = 1.8 Hz.

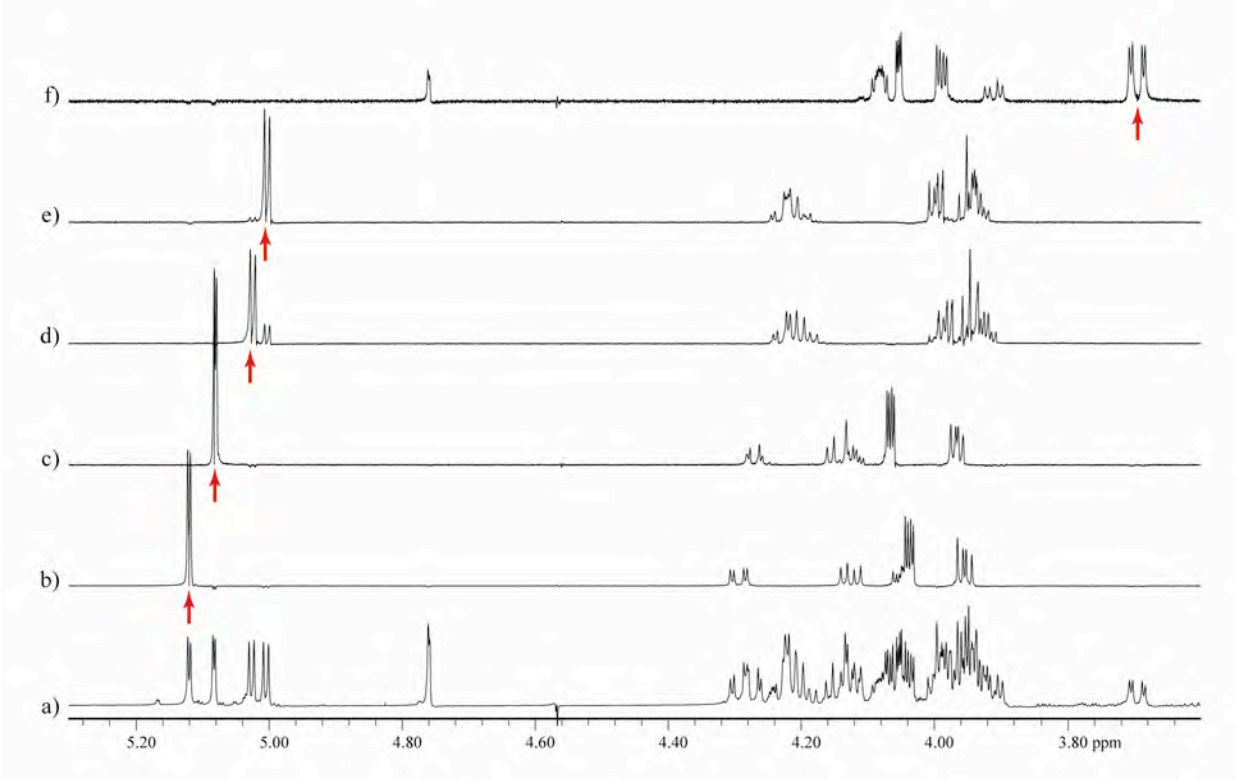


Figure S4. Chemical shift selective filtering TOCSY spectra for **7** in 1:1 $CDCl_3$–CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *f* show 1D-ge-CSSF-TOCSY of each ring with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.12 ppm, mix = 0.2 s, Δ = 2.4 Hz; c) ring C, selectively irradiated at 5.08 ppm, mix = 0.2 s, Δ = 5.0 Hz; d) ring E, selectively irradiated at 5.03 ppm, mix = 0.16 s, Δ = 5.0 Hz; e) ring F, selectively irradiated at 5.00 ppm, mix = 0.16 s, Δ = 5.0 Hz; f) ring B, selectively irradiated at 3.69 ppm, mix = 0.2 s, Δ = 4.0 Hz.

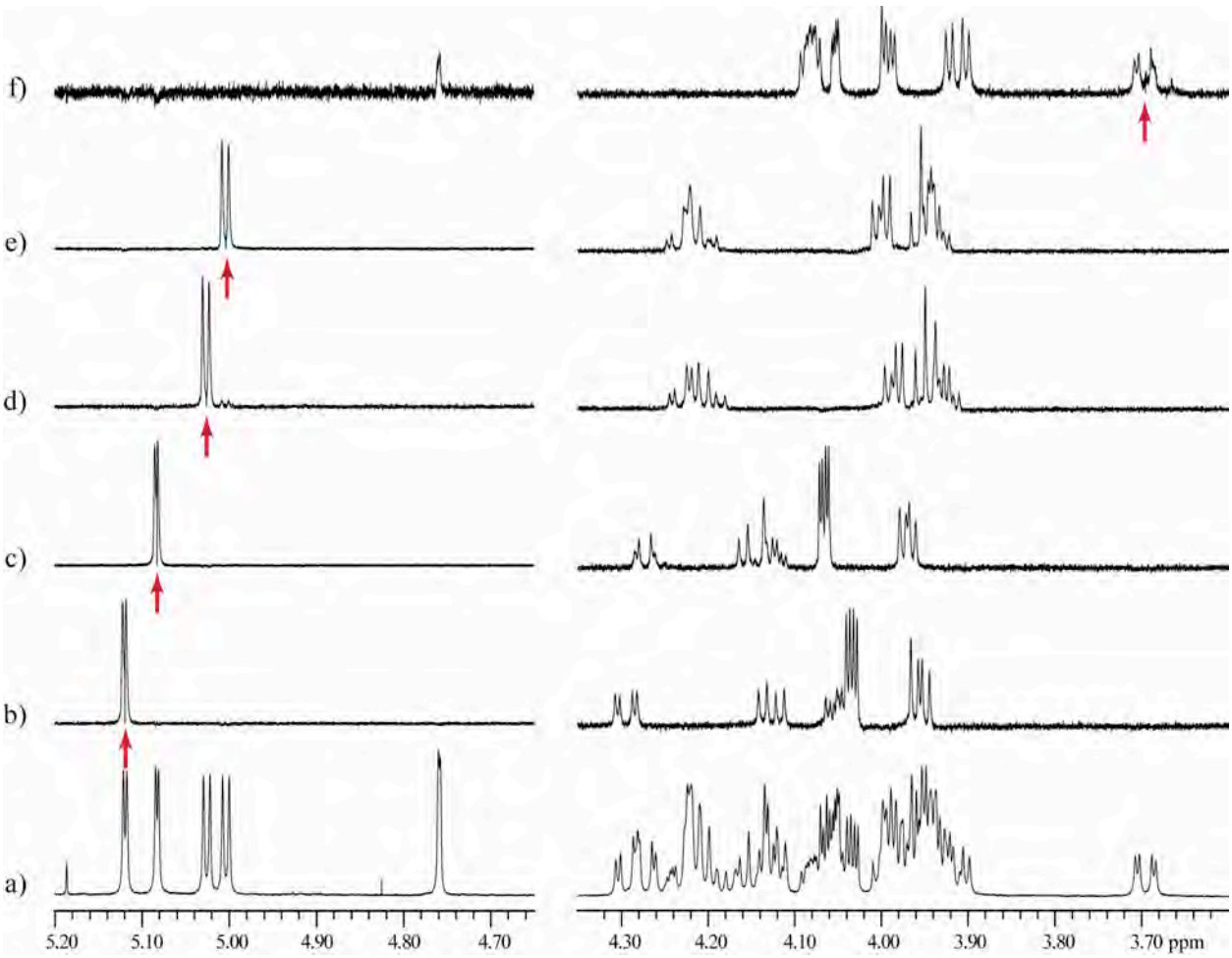


Figure S5. Chemical shift selective filtering TOCSY spectra for **8** in 1:1 $CDCl_3$–CD_3OD. The portions of the spectra from 4.65–4.35 ppm contains the residual OH/H_2O signal and was removed for clarity. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *f* show 1D-ge-CSSF-TOCSY of each ring with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.12 ppm, mix = 0.2 s, Δ = 4.0 Hz; c) ring C, selectively irradiated at 5.08 ppm, mix = 0.2 s, Δ = 4.0 Hz; d) ring E, selectively irradiated at 5.03 ppm, mix = 0.16 s, Δ = 2.0 Hz; e) ring F, selectively irradiated at 5.00 ppm, mix = 0.16 s, Δ = 2.0 Hz; f) ring B, selectively irradiated at 3.69 ppm, mix = 0.16 s, Δ = 60 Hz.

Table S1. ^{1}H chemical shifts (ppm) of the pentasaccharides, **4–8**.

Ring	Proton	**4**[a]	**5**[a]	**5**[b]	**6**[b]	**7**[b]	**8**[b]
B	**H1**	4.76	4.76	4.76	4.76	4.76	4.76
	H2	4.08	4.08	4.05	4.05	4.05	4.05
	H3	4.03	4.02	3.98	3.98	3.99	3.99
	H4	4.10	4.08	4.08	4.08	4.08	4.08
	H5*R*	3.91	3.90	3.90	3.91	3.91	3.91
	H5*S*	3.74	3.71	3.69	3.69	3.70	3.70
C	**H1**	5.07	5.09	5.08	5.08	5.08	5.08
	H2	4.14	4.11	4.08	4.07	4.07	4.06
	H3	4.02	4.01	3.97	3.96	3.97	3.97
	H4	3.96	4.11	4.12	4.12	4.12	4.12
	H5*R*	3.78	4.31	4.28	4.27	4.27	4.27
	H5*S*	3.63	4.16	4.15	4.15	4.15	4.15
D	**H1**	5.15	5.16	5.12	5.12	5.12	5.12
	H2	4.10	4.08	4.05	4.04	4.04	4.03
	H3	4.03	4.02	3.96	3.95	3.95	3.95
	H4	3.89	4.04	4.04	4.05	4.05	4.05
	H5*R*	3.78	4.33	4.30	4.29	4.29	4.29
	H5*S*	3.62	4.15	4.12	4.12	4.13	4.13
E	**H1**	5.03	5.05	5.03	5.03	5.02	5.03
	H2	3.97	3.98	3.98	3.98	3.98	3.98
	H3	4.00	3.98	3.95	3.95	3.95	3.95
	H4	3.77	3.91	3.92	3.92	3.93	3.93
	H5*R*	3.72	4.25	4.23	4.23	4.23	4.23
	H5*S*	3.64	4.21	4.20	4.19	4.20	4.20
F	**H1**	5.01	5.03	5.01	5.00	5.00	5.00
	H2	3.98	3.99	4.00	4.00	4.00	4.00
	H3	4.00	3.99	3.96	3.95	3.95	3.95
	H4	3.77	3.91	3.93	3.94	3.94	3.94
	H5*R*	3.72	4.26	4.23	4.23	4.23	4.23
	H5*S*	3.63	4.21	4.21	4.20	4.21	4.21

[a]CD_3OD as solvent. [b]1:1 CD_3OD–$CDCl_3$ as solvent.

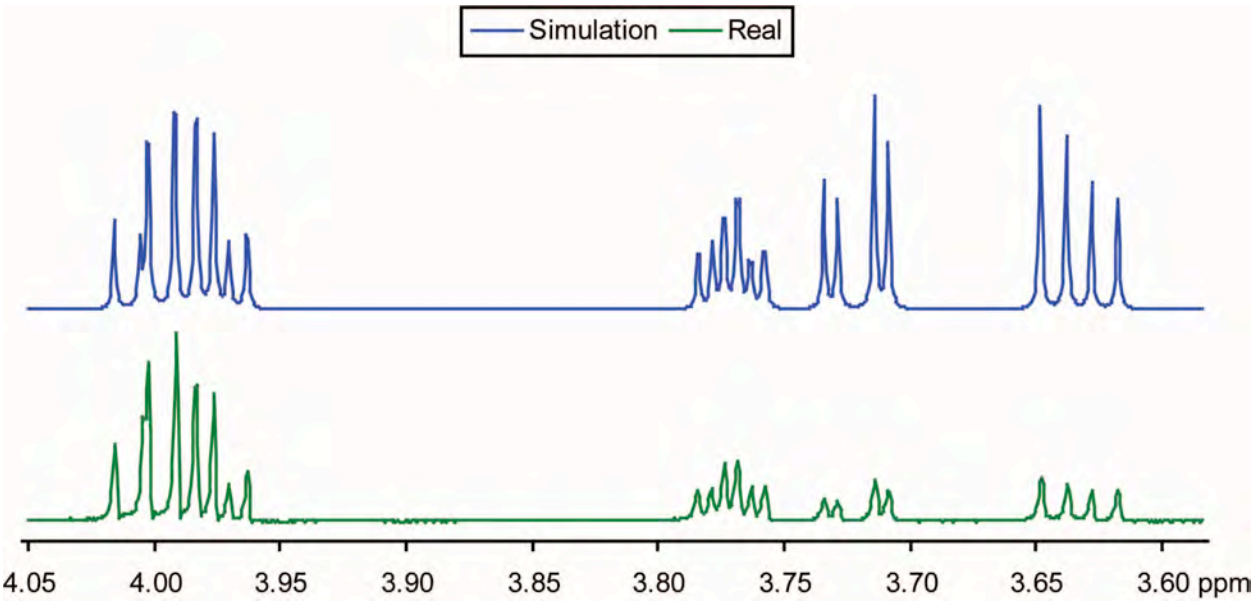


Figure S6. A comparison of the acquired 1D-ge-CSSF-TOCSY spectrum (bottom, green) to the simulation in WinDNMR[3] (top, blue) for ring F of **4**. (See also Figure S1.) The correlation coefficient (i.e. Pearson correlation) between the real and simulated data is r = 0.824 due to the intensities falling off as the protons are further from the irradiated signal (i.e. H2 is of a higher intensity than H5*R* or H5*S*).

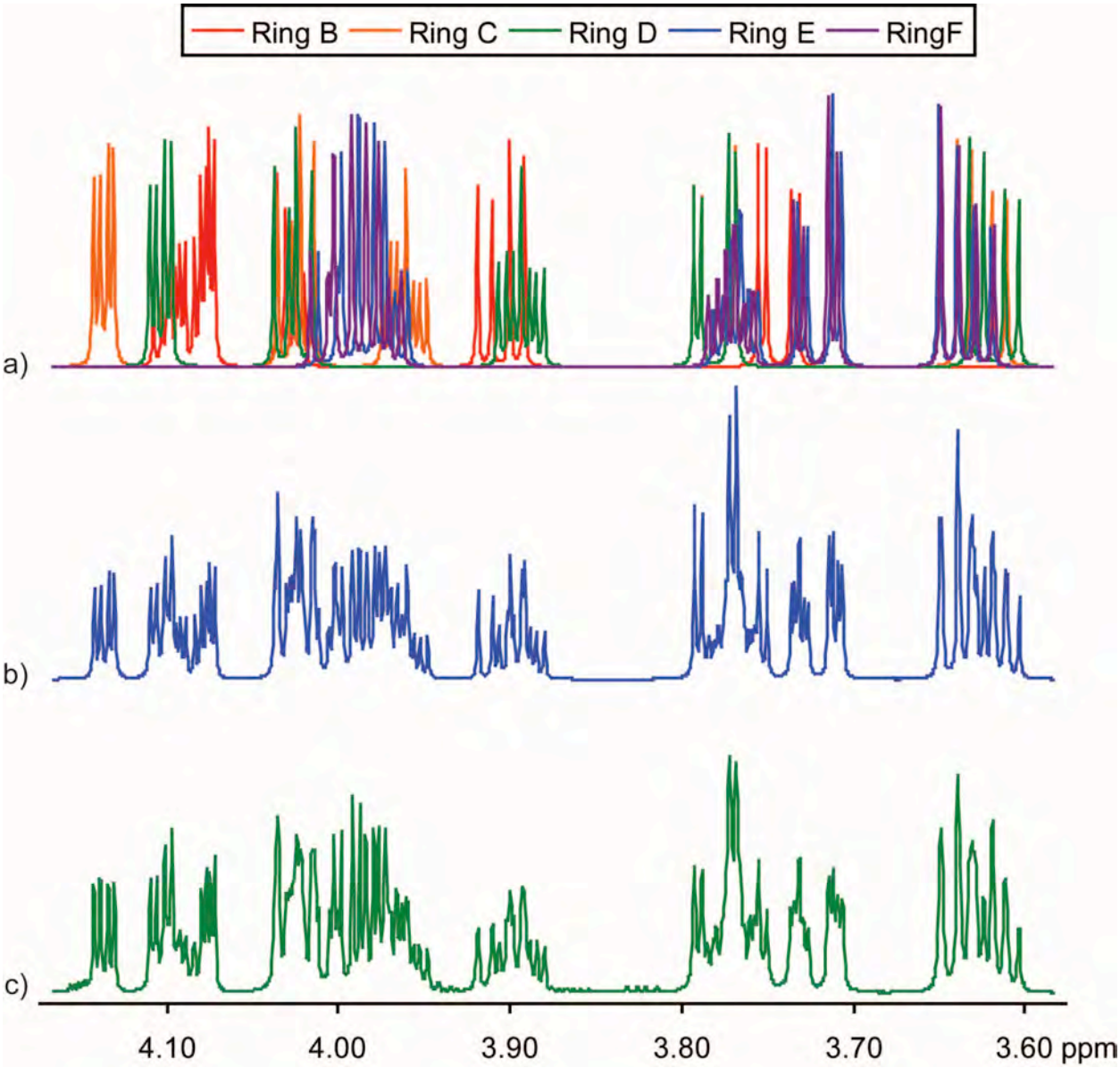


Figure S7. Simulated and acquired spectra for **4**. The spectra in *a* are the simulated spectra for each ring (an example for ring F is shown in Figure S6). These spectra can be summed to give a simulation of the 1D ^{1}H spectrum of **4** (spectrum *b*). Spectrum *c* is the acquired 1D ^{1}H spectrum for the same region (also shown in Figure S1). The correlation coefficient (i.e. Pearson correlation) between the *b* and *c* is r = 0.950.

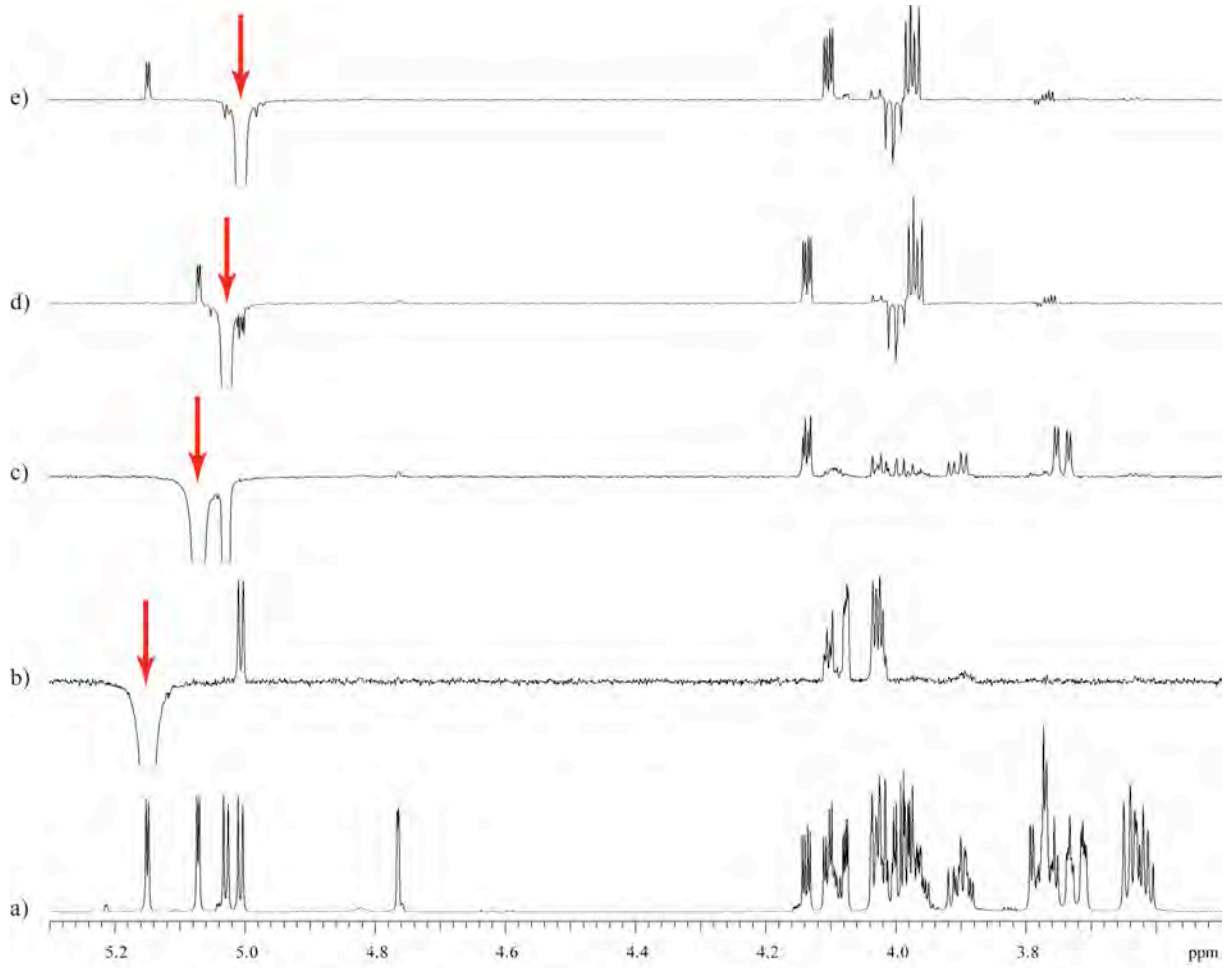


Figure S8. Chemical shift selective filtering NOESY spectra for **4** in CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *e* show 1D-ge-CSSF-DPFGSE-NOESY for rings C–F with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.15 ppm, Δ = 15 Hz; c) ring C, selectively irradiated at 5.07 ppm, Δ = 5.0 Hz; d) ring E, selectively irradiated at 5.03 ppm, Δ = 1.0 Hz; e) ring F, selectively irradiated at 5.00 ppm, Δ = 1.0 Hz. The mixing time is 0.4 s for spectra *b* through *e*. The NOE are graphically summarized below in Figure S9.

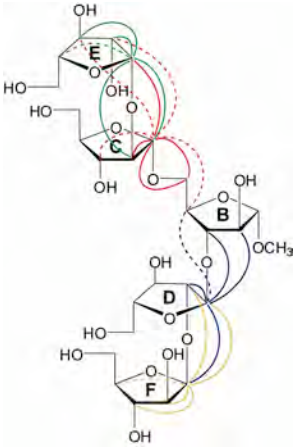


Figure S9. A graphical representation of the NOE for **4**, with weaker NOE shown as dashed lines. The green lines are NOE from irradiation at 5.03 ppm, the red lines from irradiation at 5.07 ppm, the blue lines from irradiation at 5.15 ppm, and the orange lines from irradiation at 5.00 ppm.

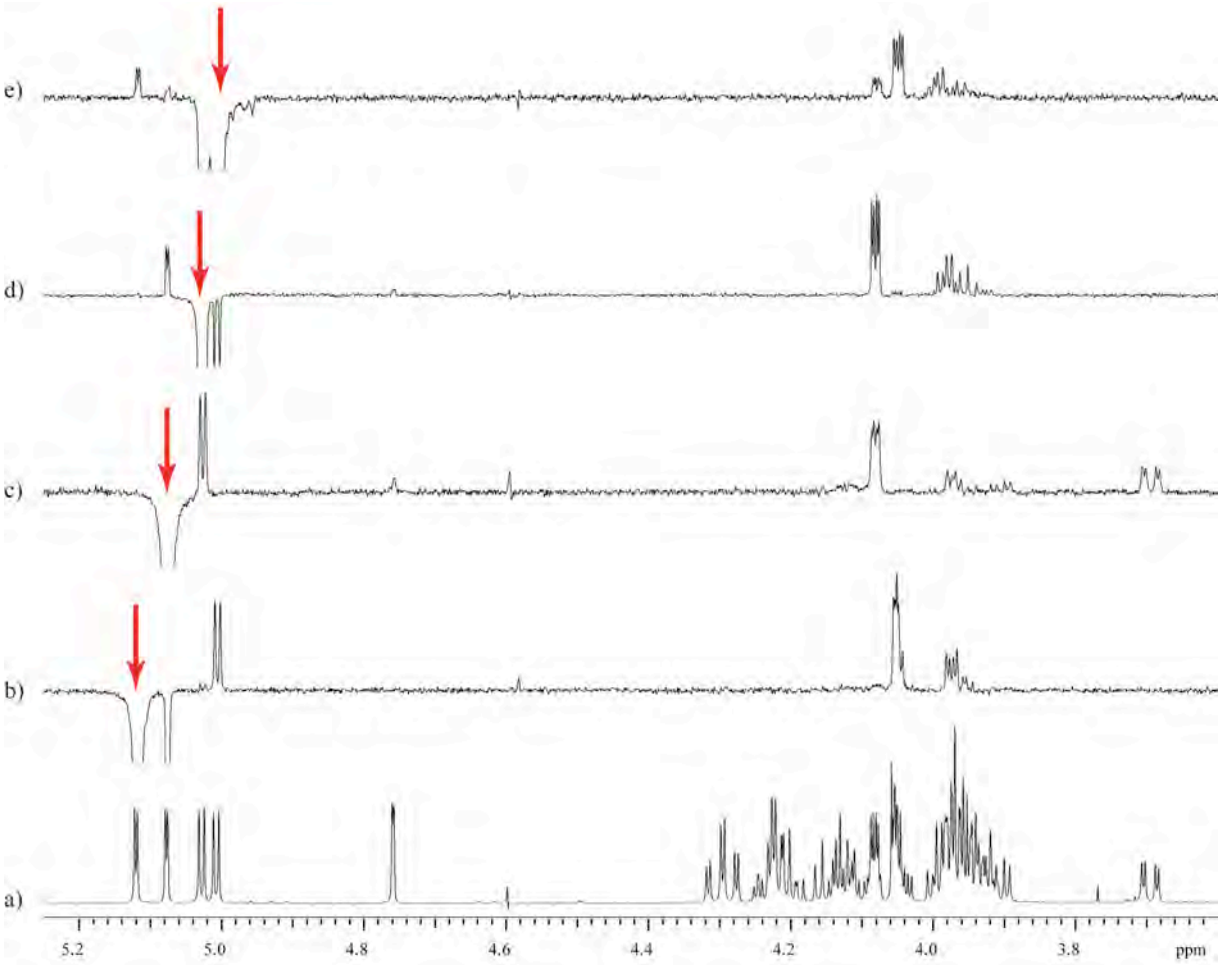


Figure S10. Chemical shift selective filtering NOESY spectra for **5** in 1:1 $CDCl_3$–CD_3OD. Spectrum *a* is an expansion of the 1D ^1H spectrum; spectra *b* through *e* show 1D-ge-CSSF-DPFGSE-NOESY for rings C–F with an arrow indicating the irradiation frequency, as follows: b) ring D, selectively irradiated at 5.12 ppm, mix = 1.6 s, Δ = 1.6 Hz; c) ring C, selectively irradiated at 5.08 ppm, mix = 1.6 s, Δ = 3.4 Hz; d) ring E, selectively irradiated at 5.03 ppm, mix =1.6 s, Δ = 1.0 Hz; e) ring F, selectively irradiated at 5.01 ppm, mix = 0.8 s, Δ = 1.0 Hz. The NOE are graphically summarized below in Figure S11.

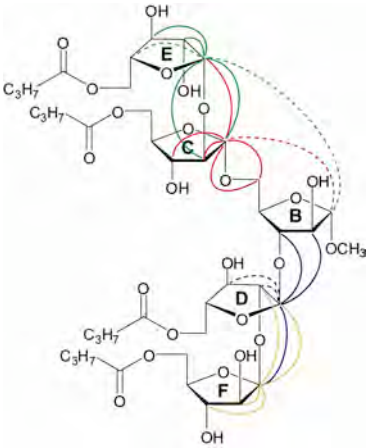


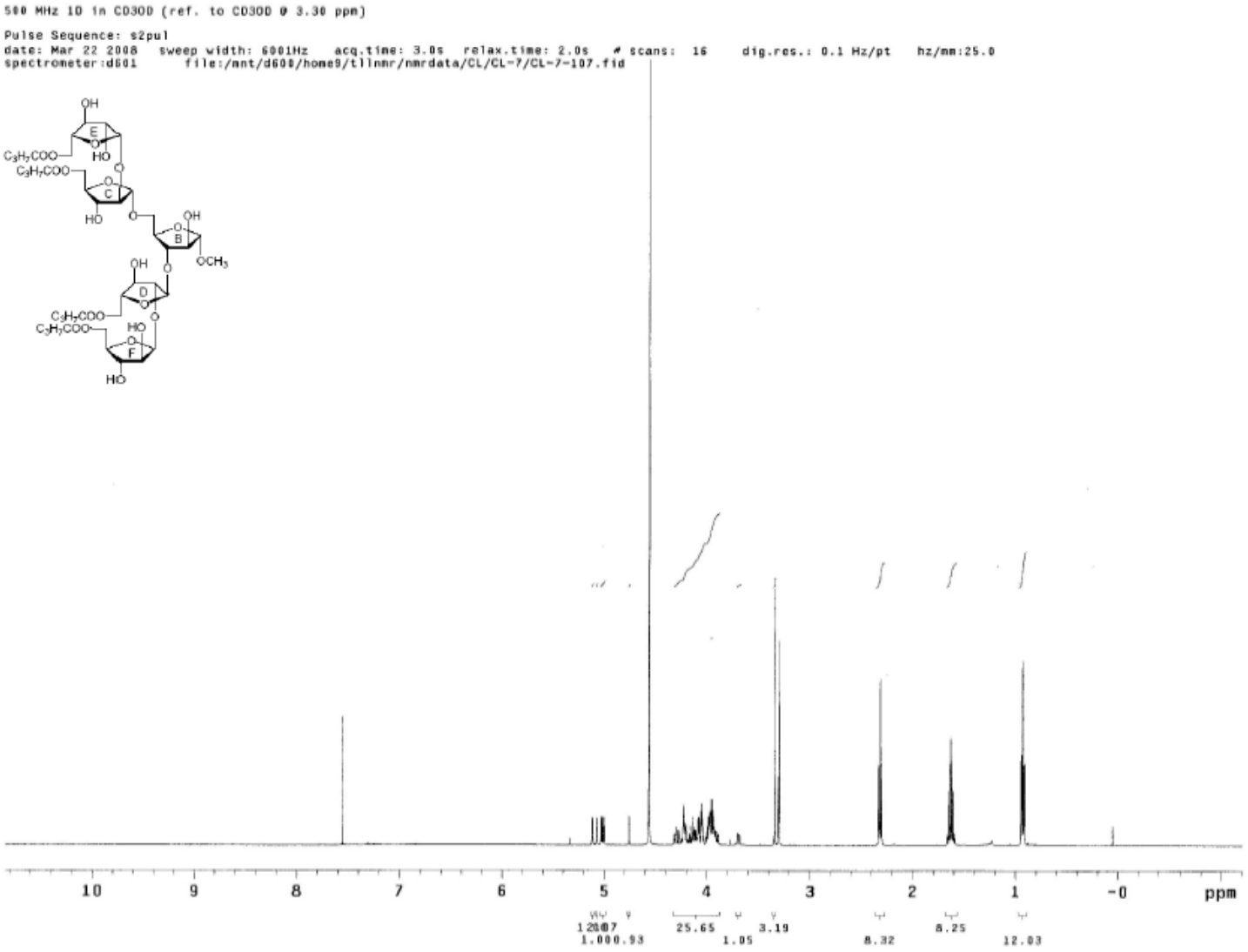
Figure S11. A graphical representation of the NOE for **5**, with weaker NOE shown as dashed lines. The green lines are NOE from irradiation at 5.03 ppm, the red lines from irradiation at 5.08 ppm, the blue lines from irradiation at 5.12 ppm, and the orange lines from irradiation at 5.01 ppm.

References for Supporting Information.

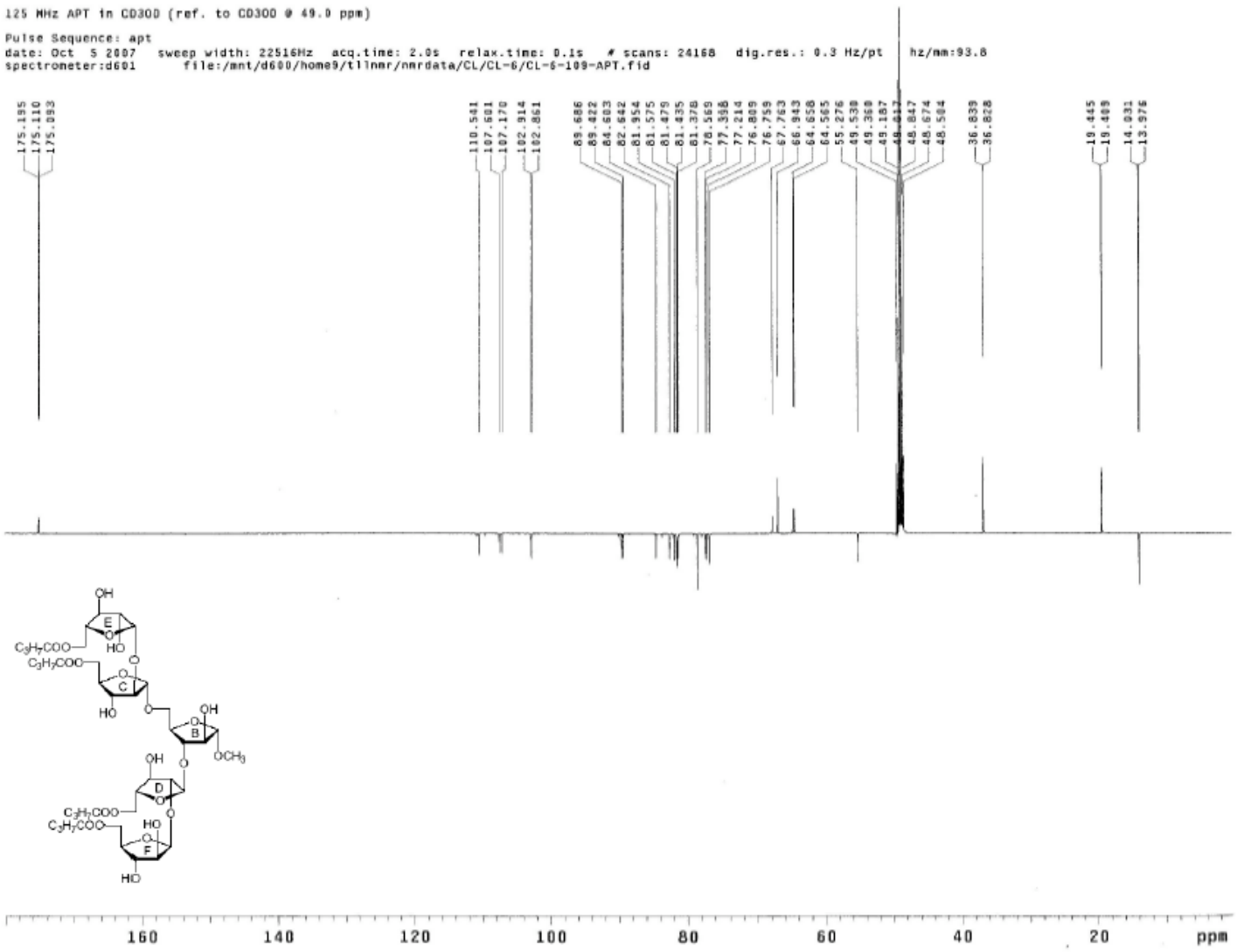
1. Robinson, P. T.; Pham, T. N.; Uhrin, D. *J. Magn. Reson.* **2004**, *170*, 97–103.
2. Duncan, S. J.; Lewis, R.; Bernstein, M. A.; Sandor, P. *Magn. Reson. Chem.* **2007**, *45*, 283–288.
3. Reich, H. J. *WinDNMR: Dynamic NMR Spectra for Windows*, Journal of Chemical Education Software, 3D2: Madison, WI, USA, 1996.
4. Taha, H. A.; Castillo, N.; Roy, P.-N.; Lowary, T. L. *J. Chem. Theory Comput.* **2009**, *5*, 430–438.
5. Wu, G. D.; Serianni, A. S.; Barker, R. *J. Org. Chem.* **1983**, *48*, 1750–1757.
6. Serianni, A. S.; Barker, R. *Can. J. Chem.* **1979**, *57*, 3160–3167.
7. Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870–2871.
8. Taha, H. A.; Castillo, N.; Sears, D. N.; Wasylishen, R. E.; Lowary, T. L. *J. Chem. Theory Comput.* **2010**, *6*, 212–222.
9. Liu, C.; Richards, M. R.; Lowary, T. L. *J. Org. Chem.* **2010**, DOI: 10.1021/jo100575a

.

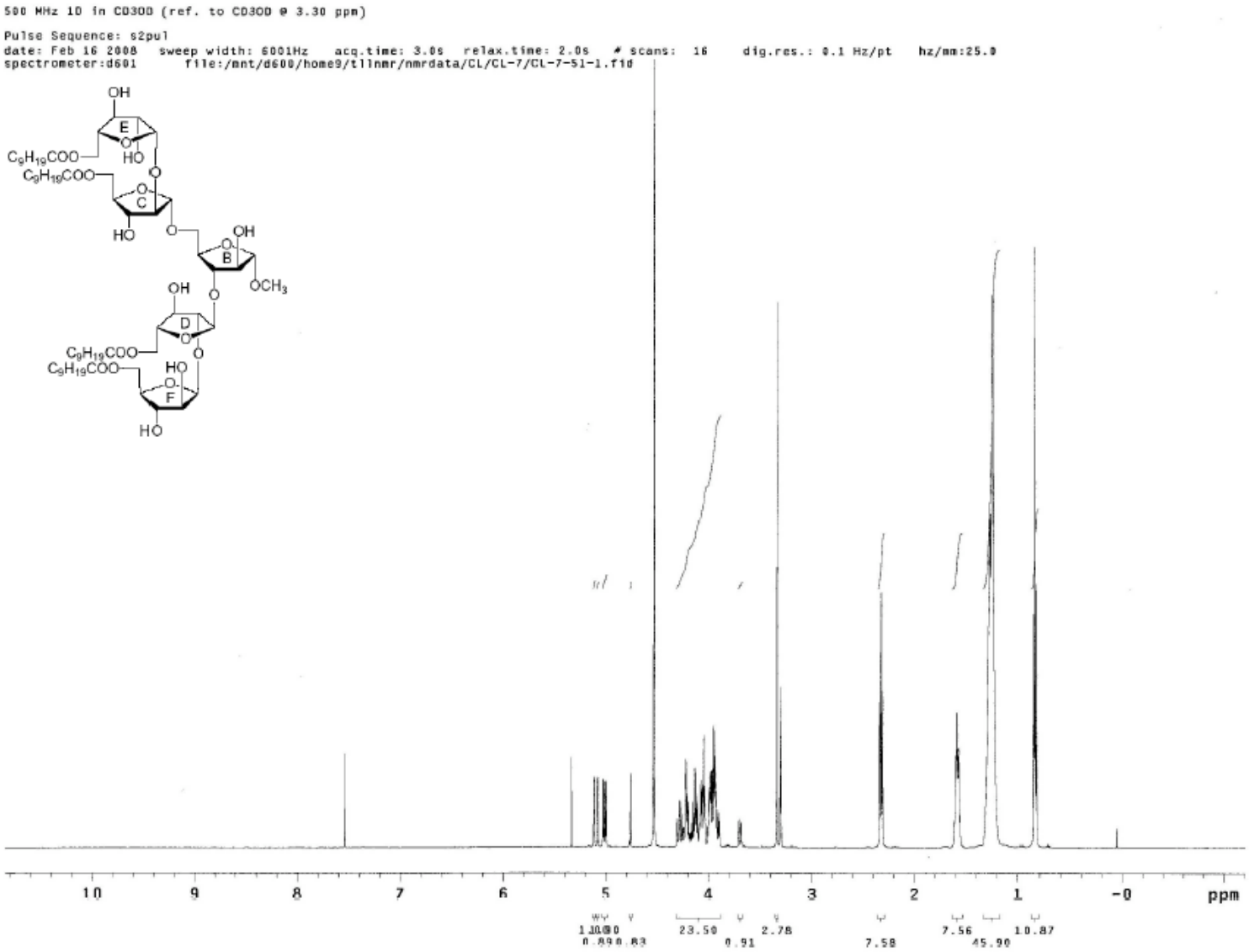
^{1}H NMR spectrum of **5**



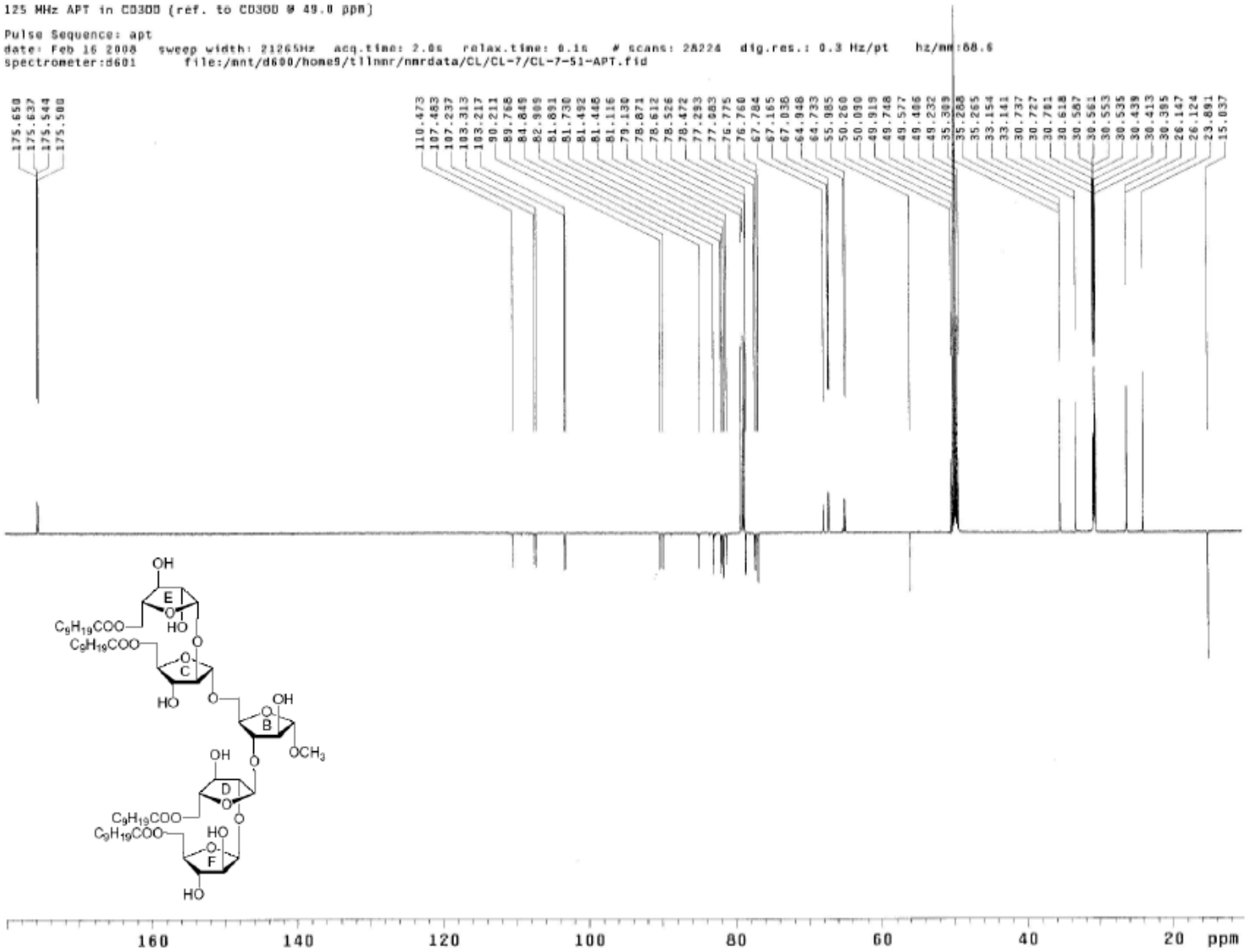
^{13}C NMR spectrum of **5**



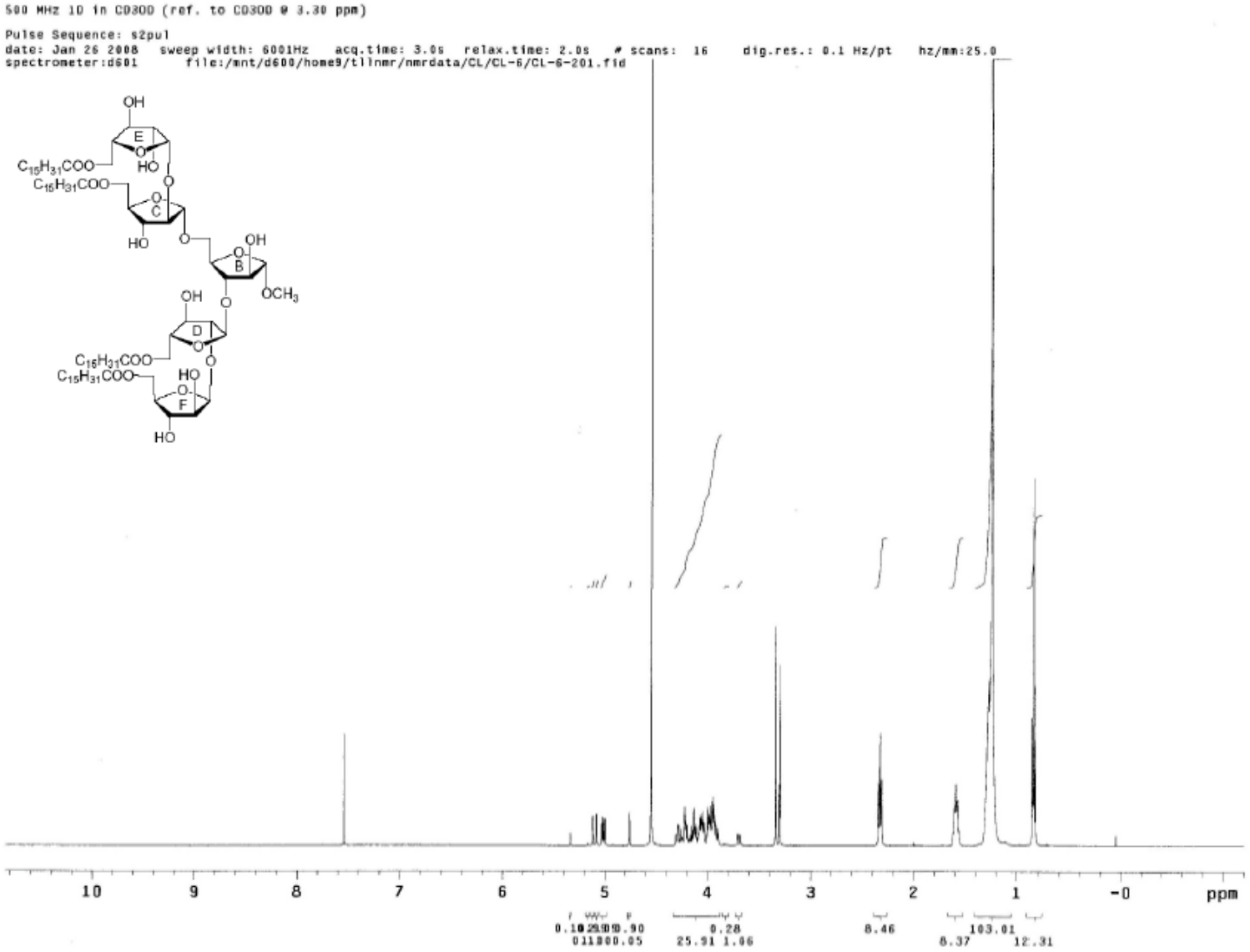
^{1}H NMR spectrum of **6**



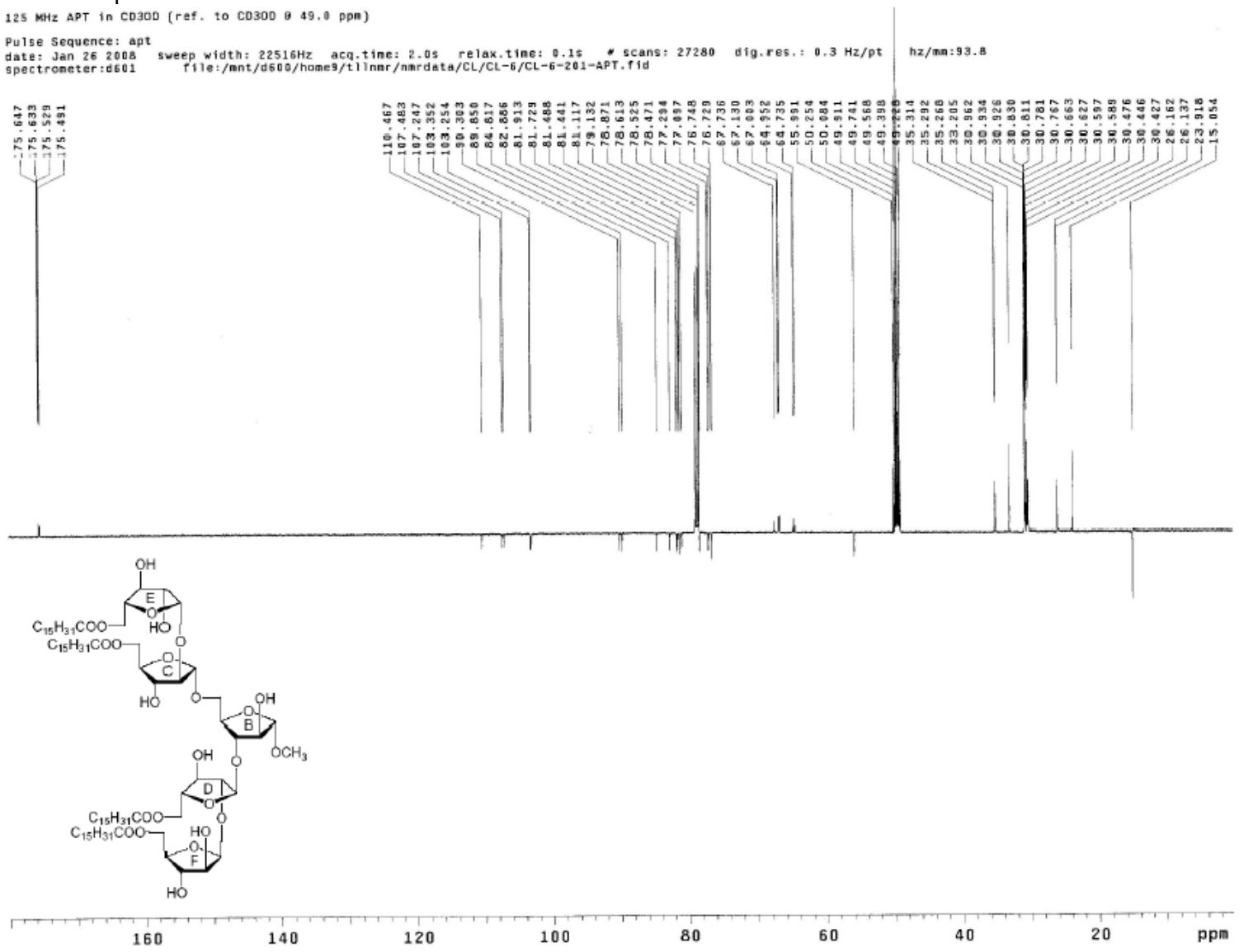
^{13}C NMR spectrum of **6**



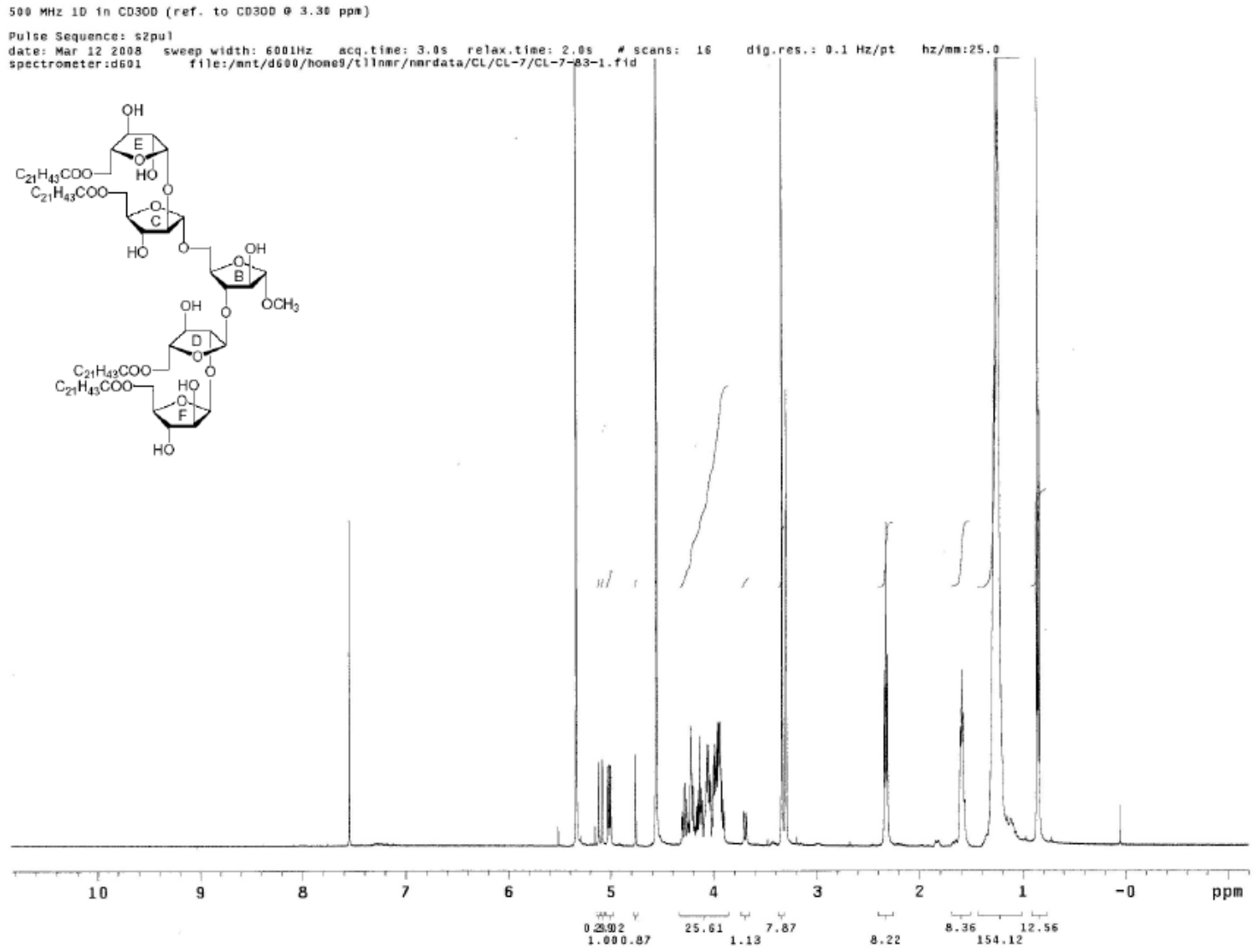
^{1}H NMR spectrum of **7**



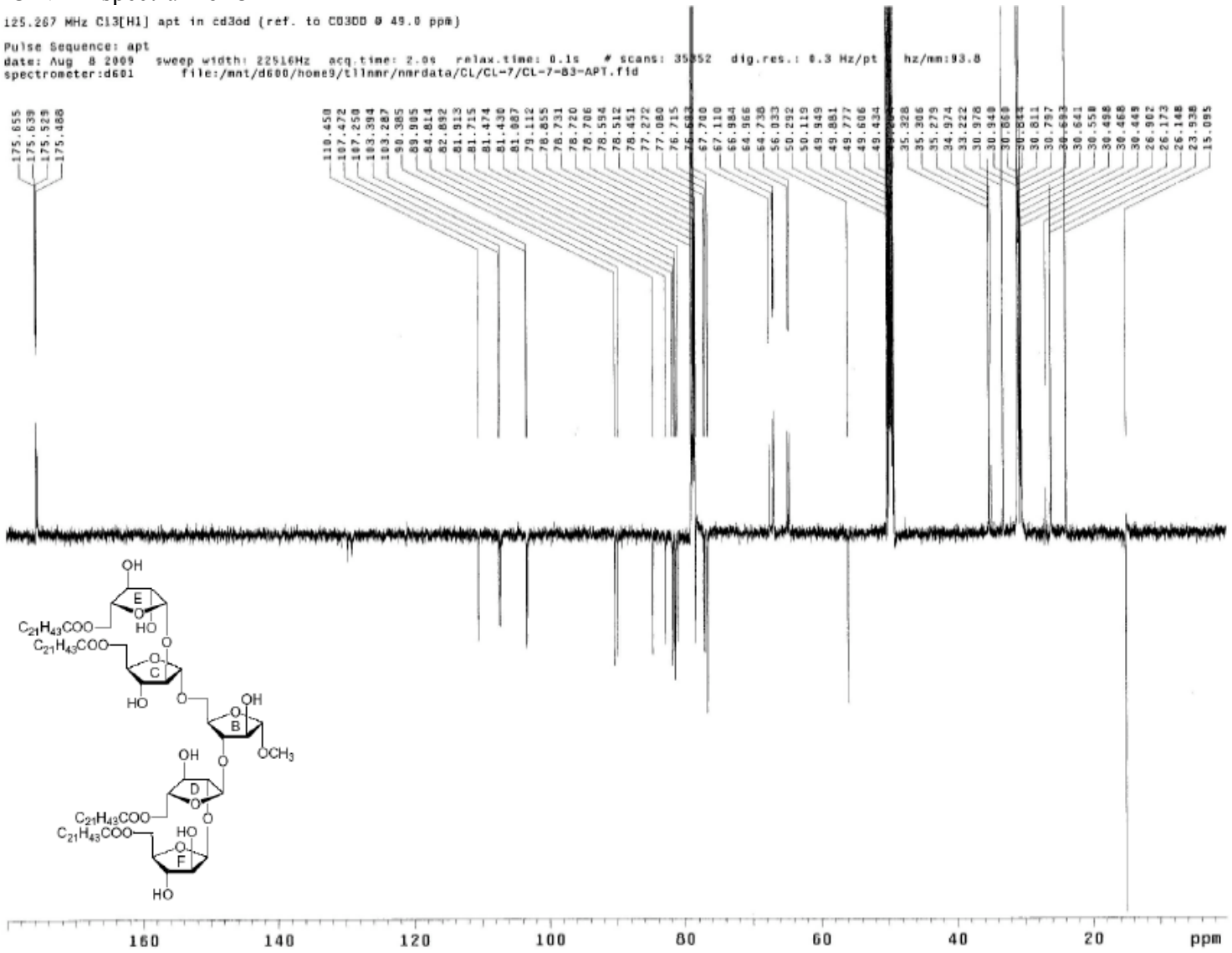
^{13}C NMR spectrum of **7**



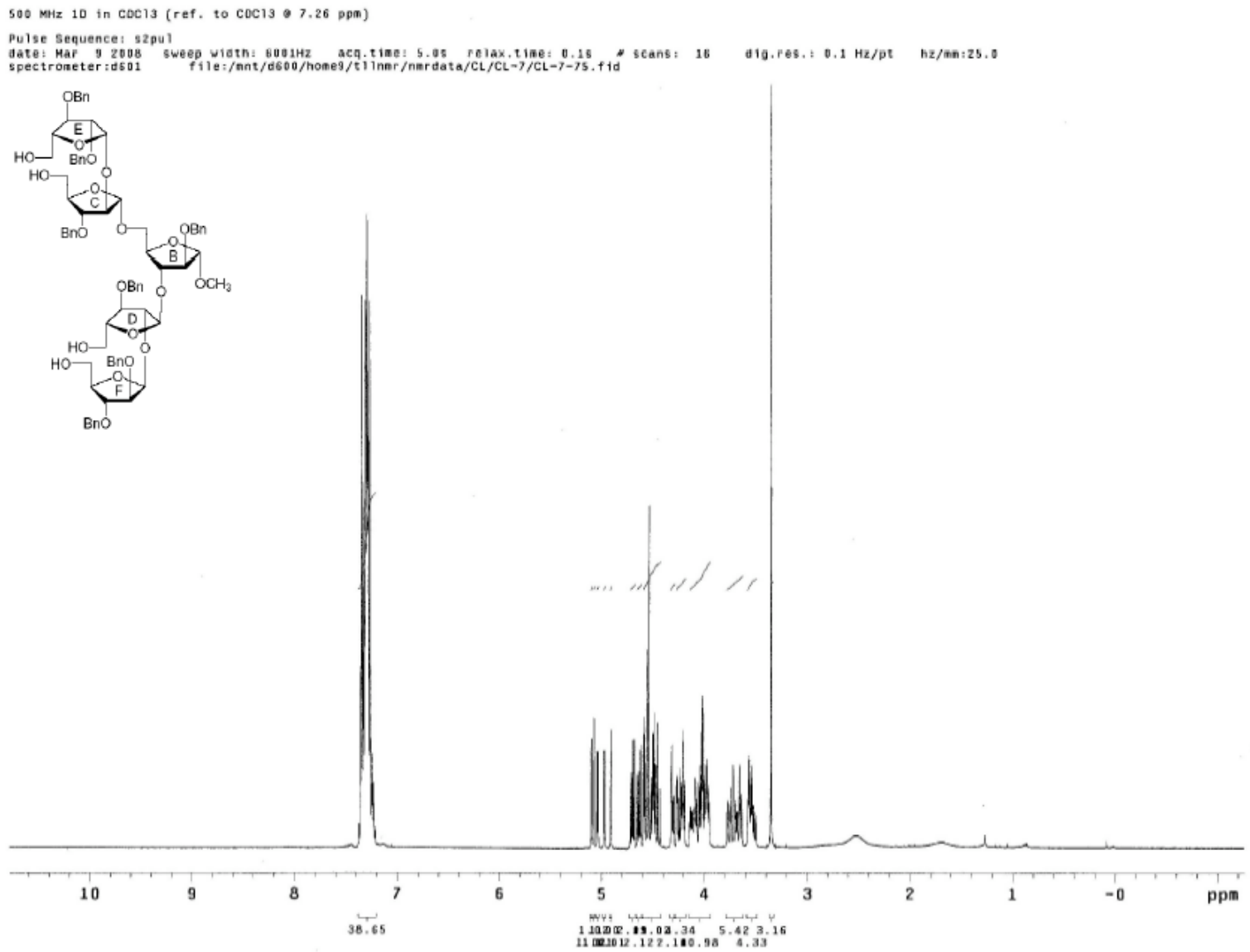
^{1}H NMR spectrum of **8**



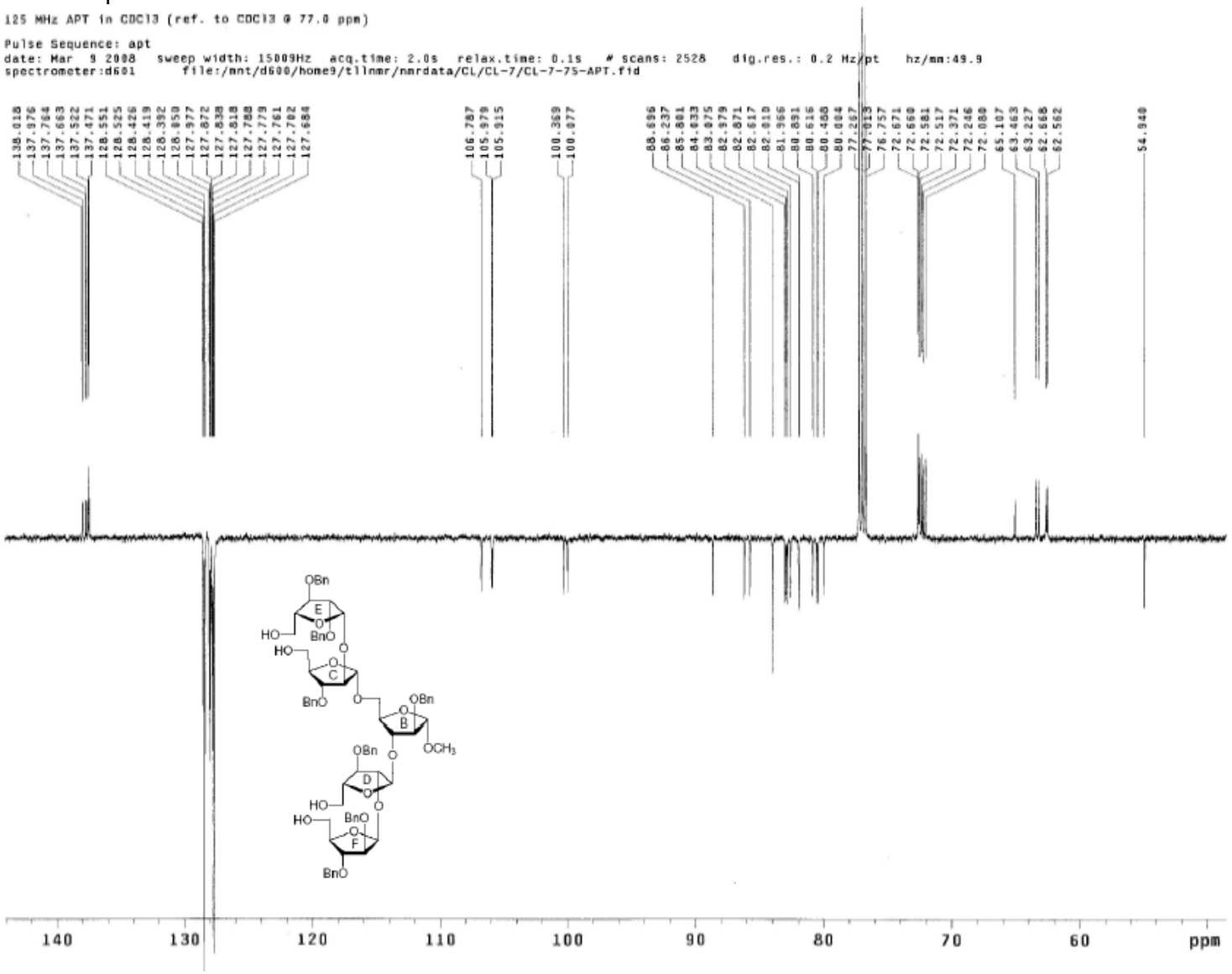
^{13}C NMR spectrum of **8**



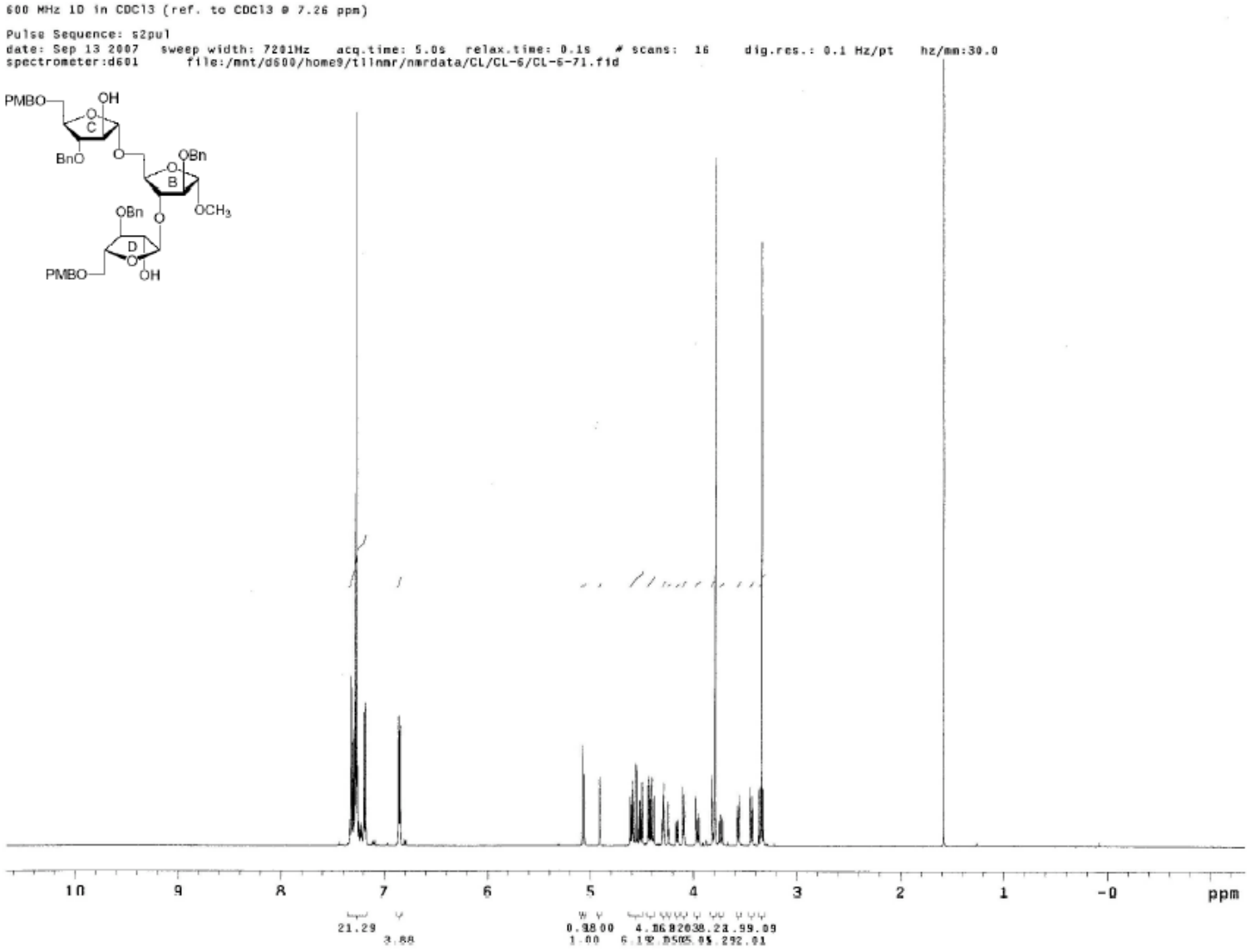
^{1}H NMR spectrum of **9**



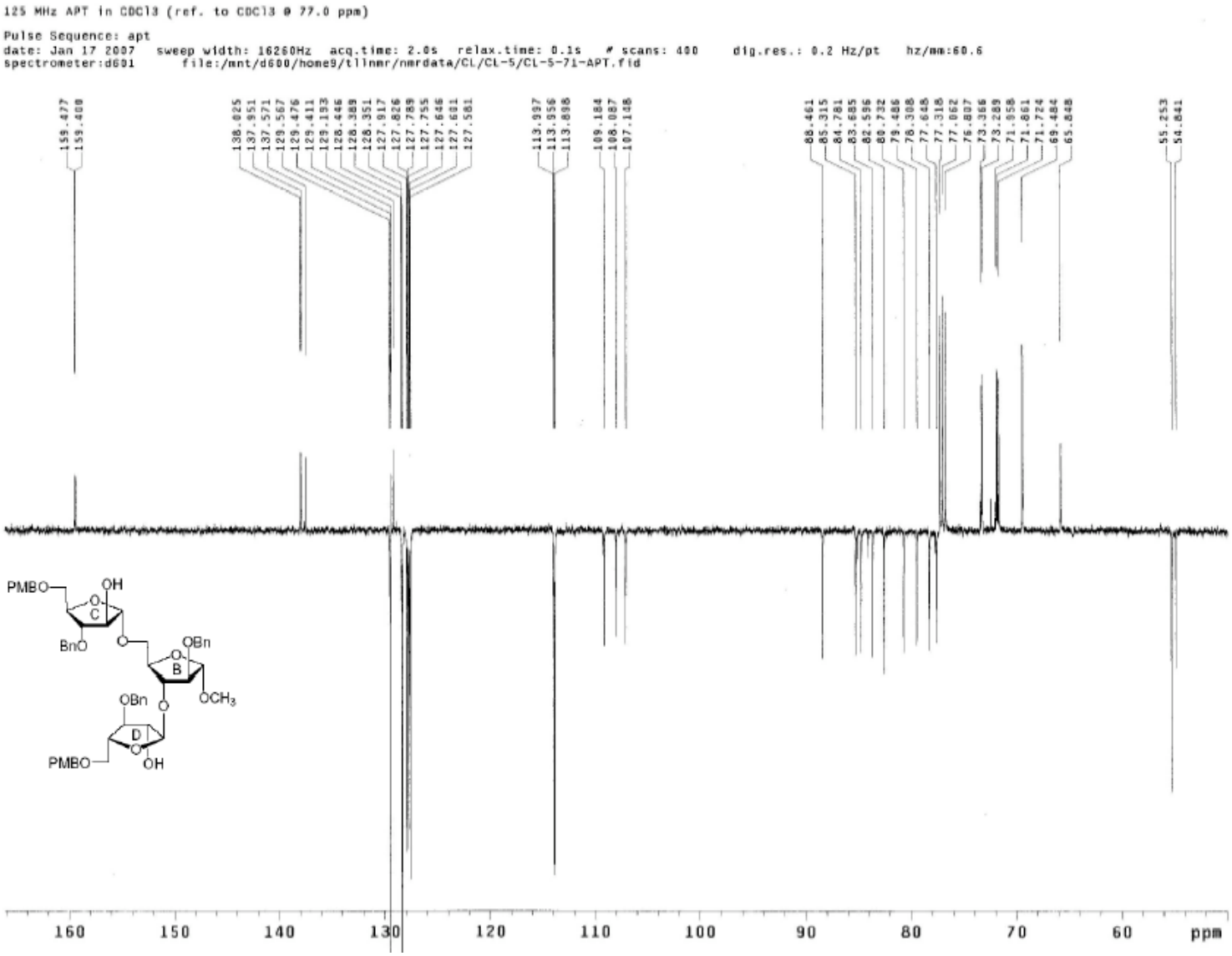
^{13}C NMR spectrum of **9**



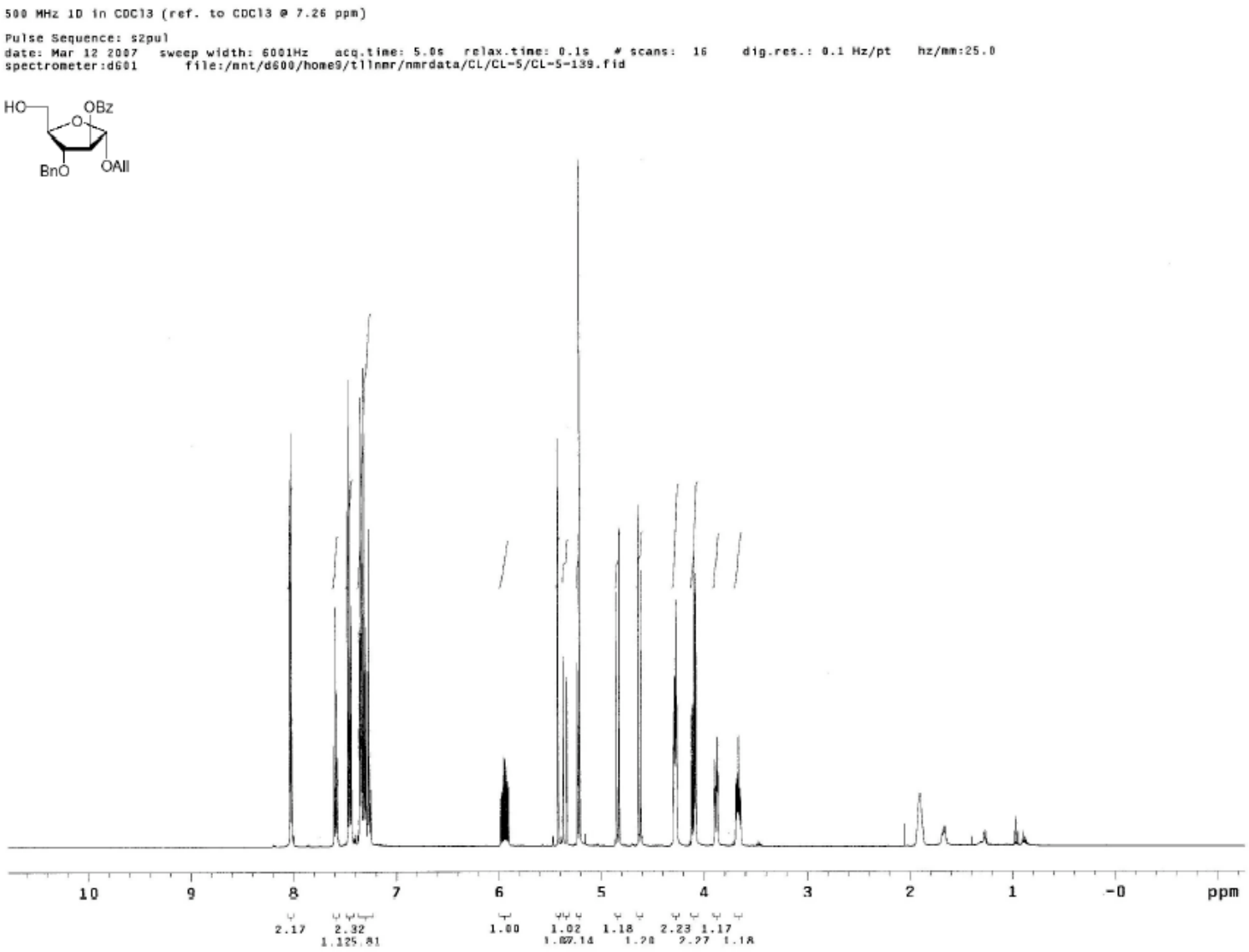
^{1}H NMR spectrum of **11**



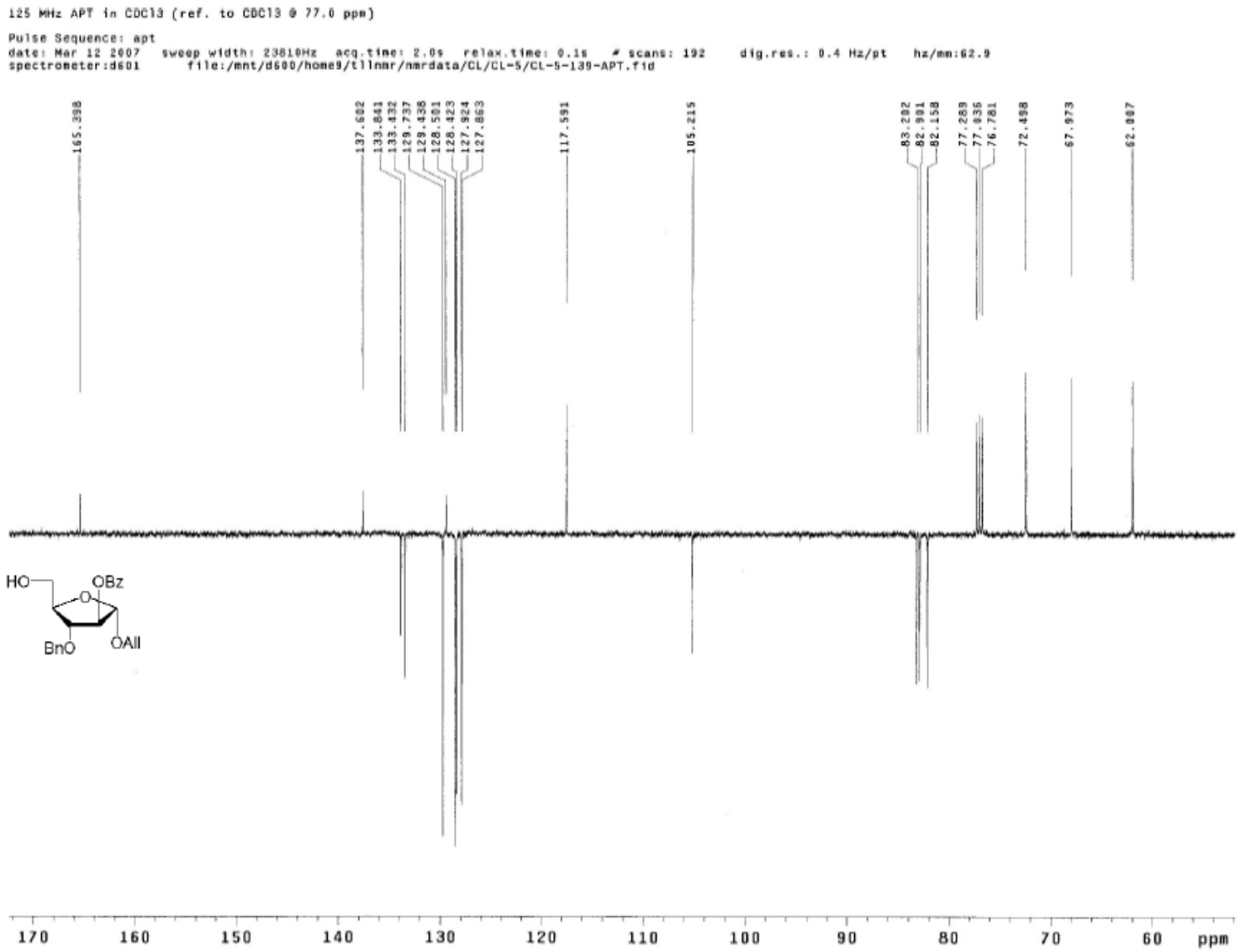
^{13}C NMR spectrum of **11**



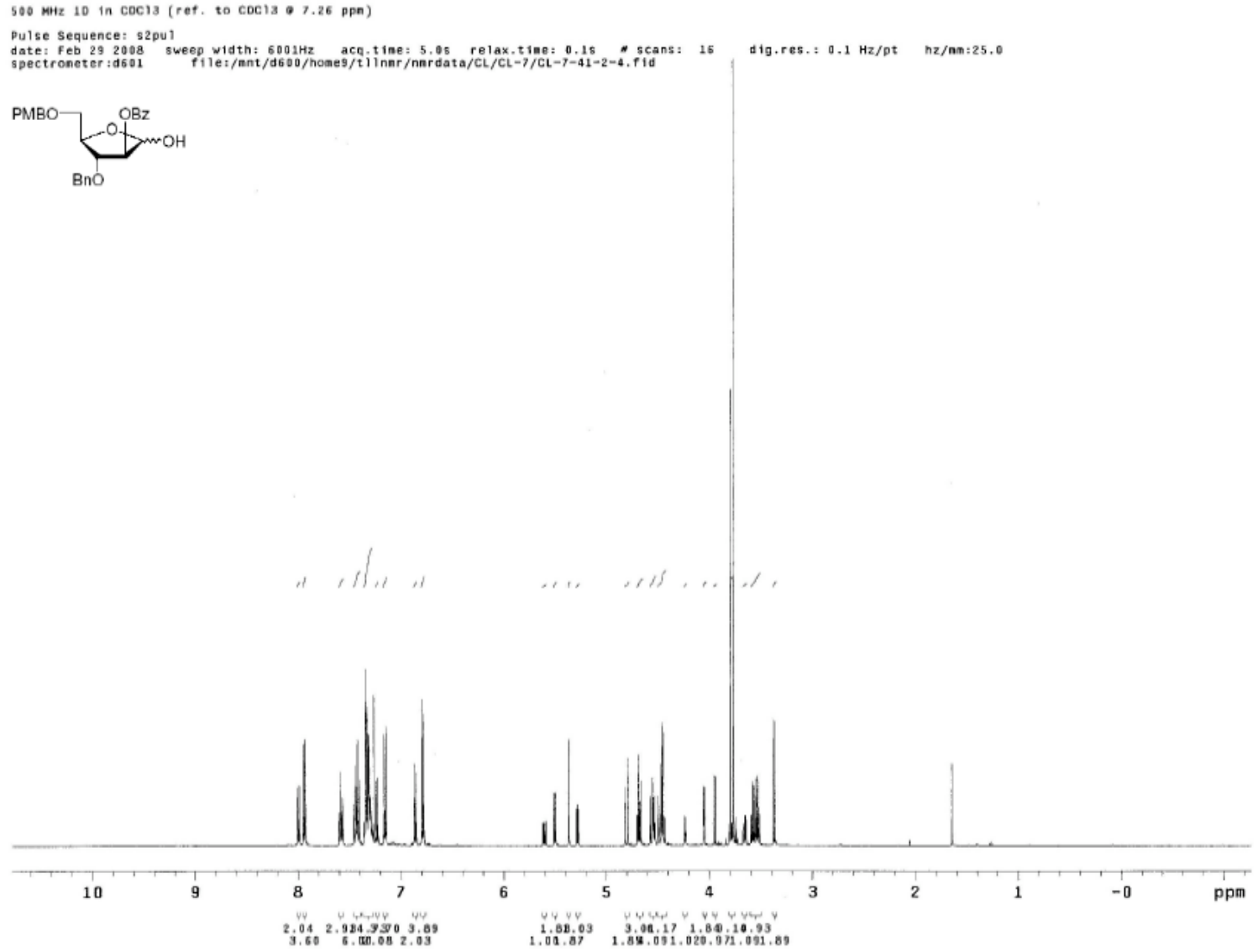
^{1}H NMR spectrum of **16**



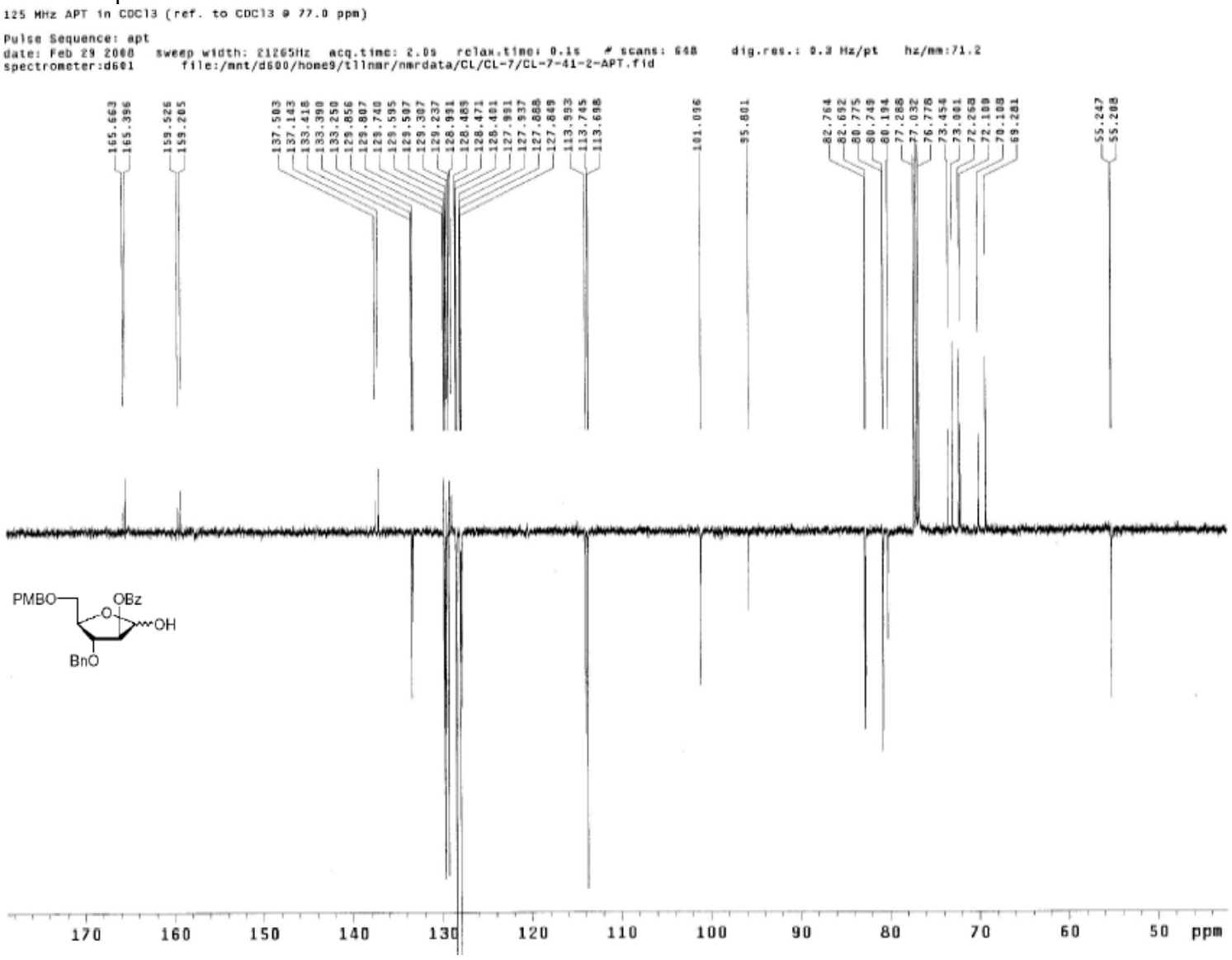
^{13}C NMR spectrum of **16**



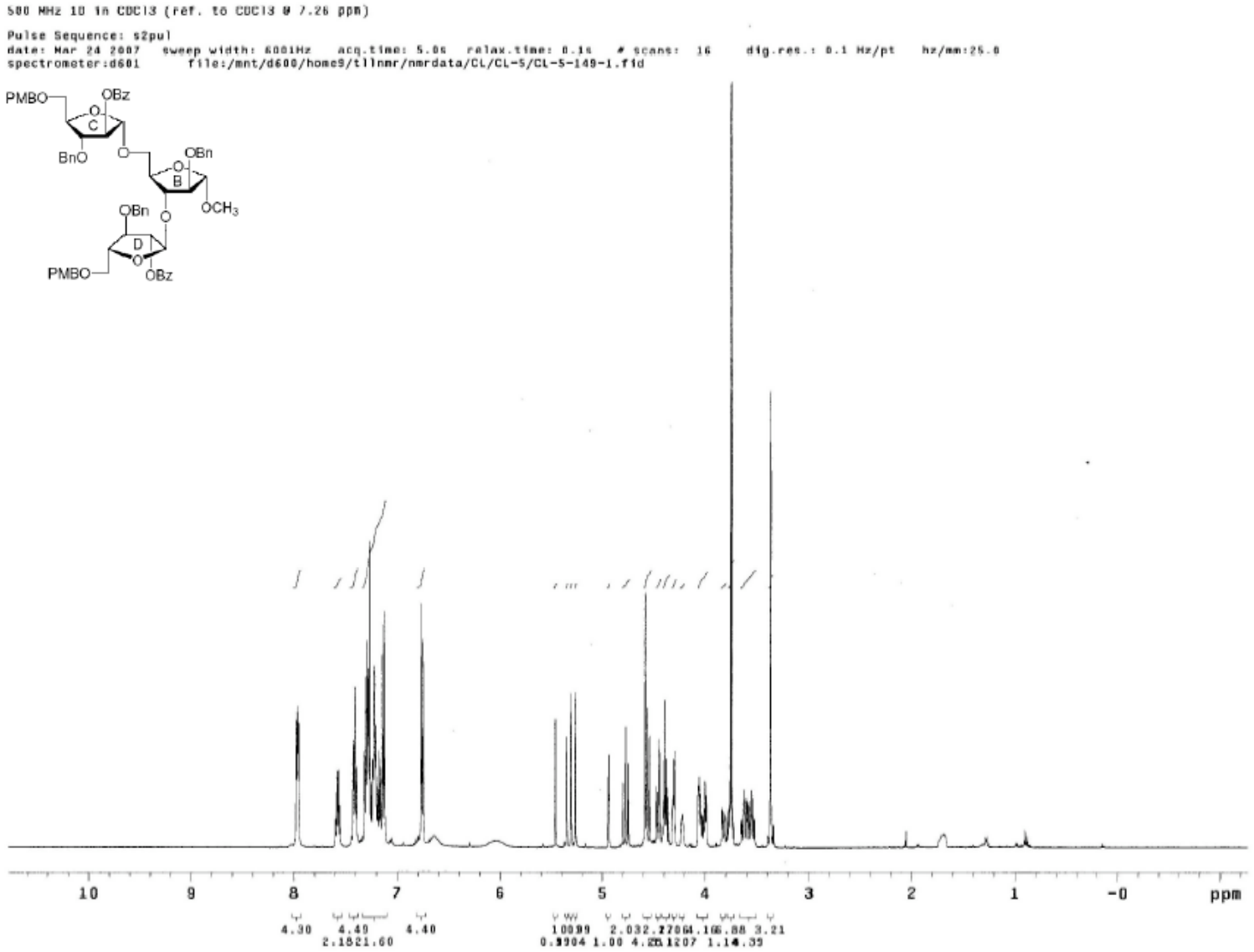
^{1}H NMR spectrum of **18**



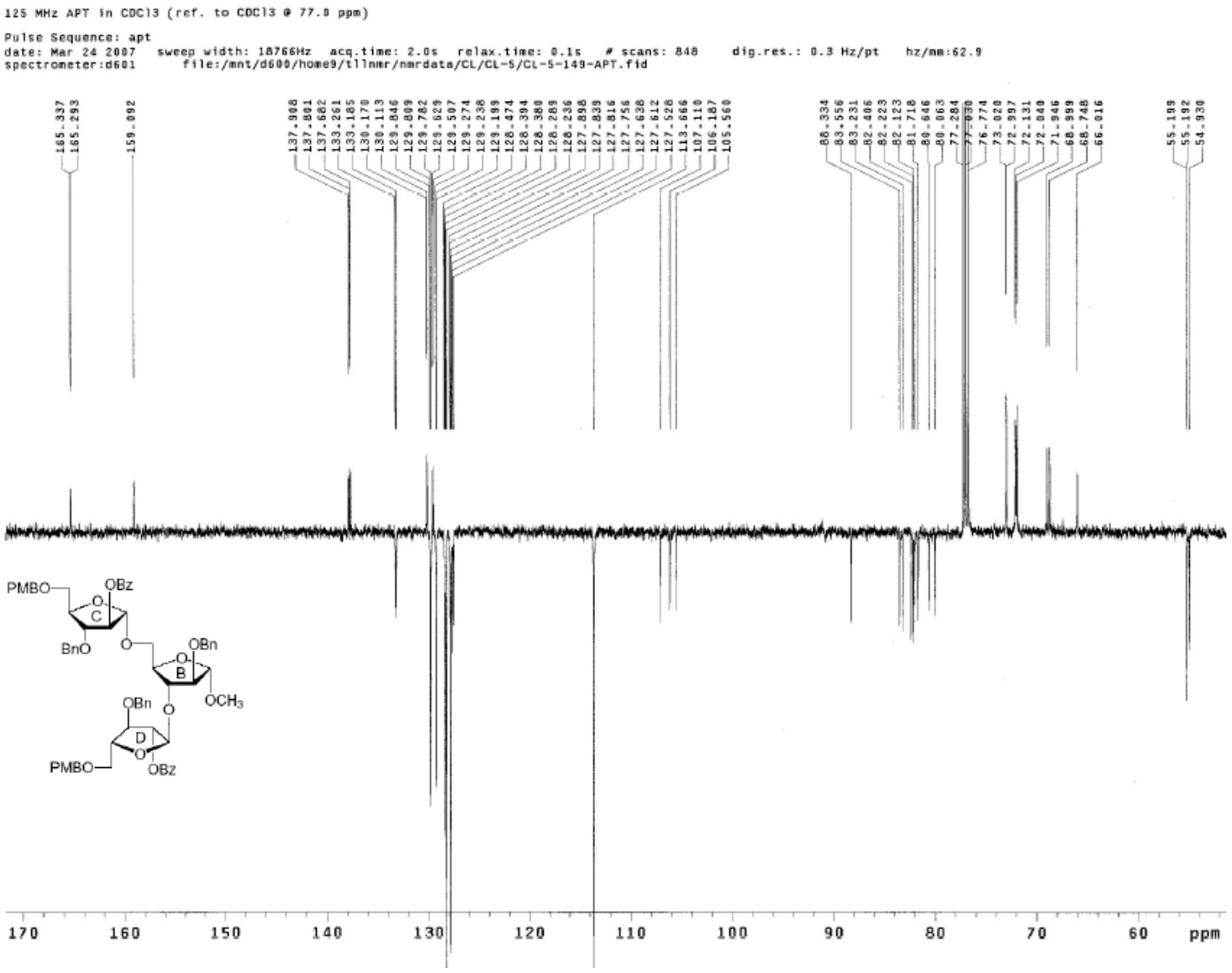
^{13}C NMR spectrum of **18**



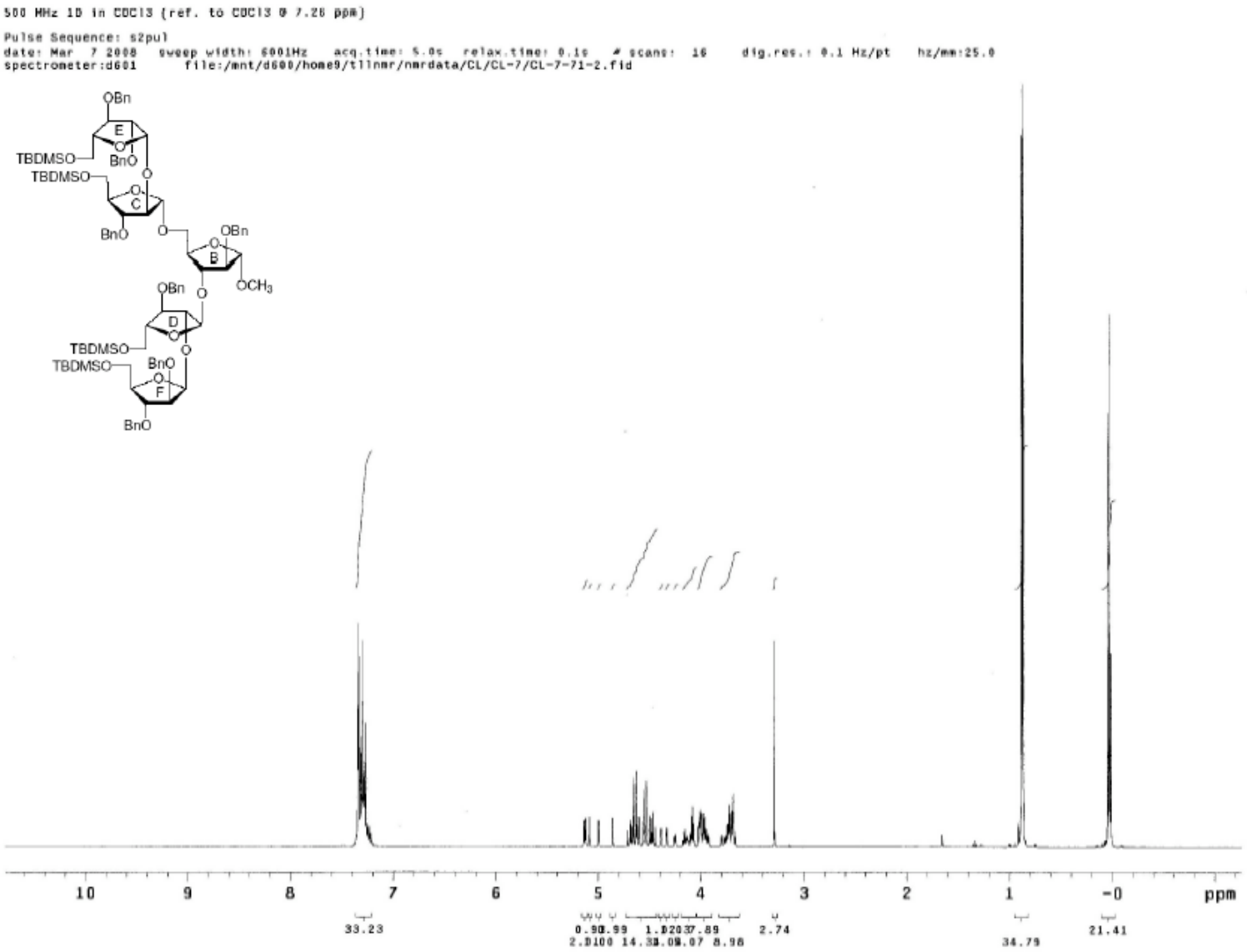
^{1}H NMR spectrum of **19**



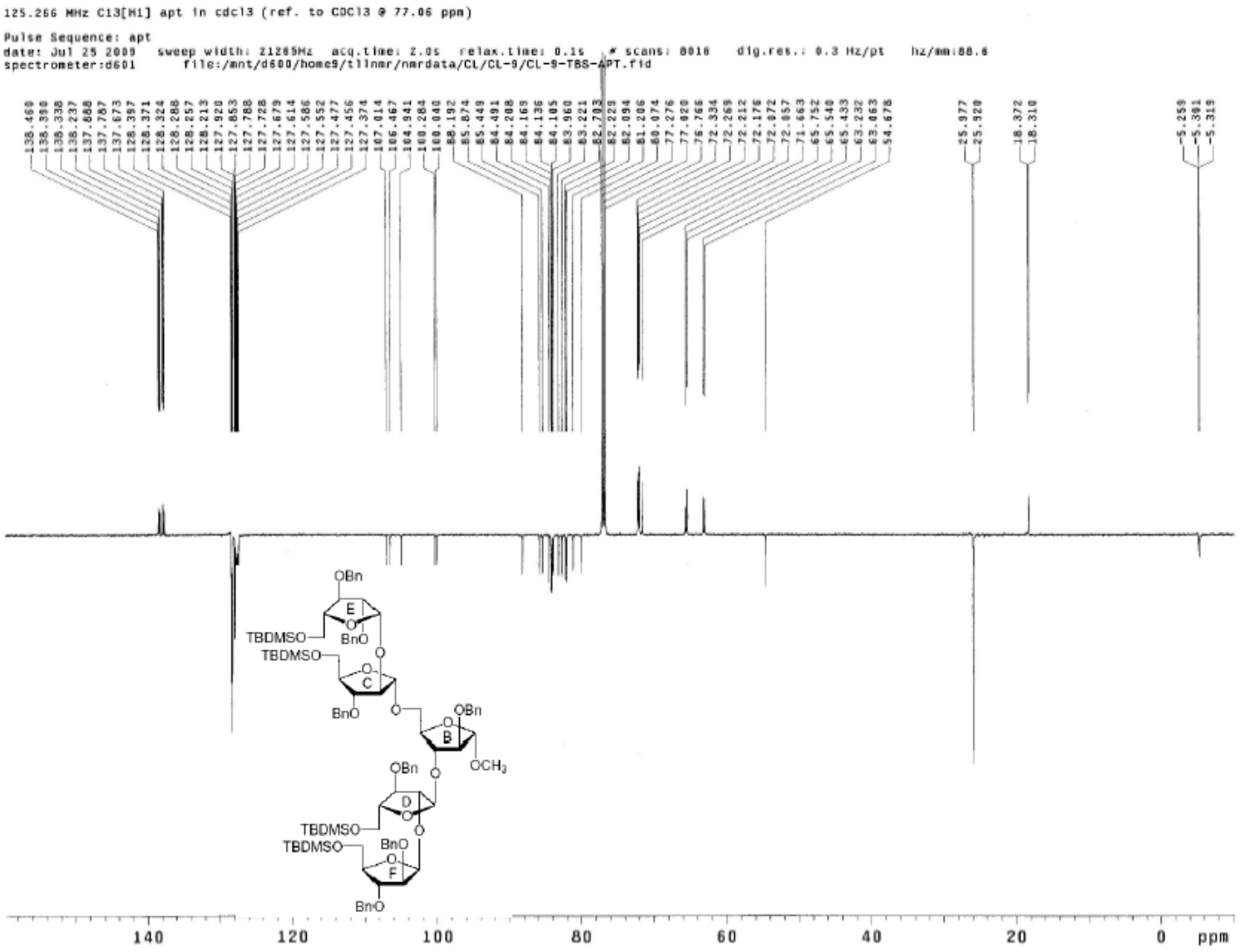
^{13}C NMR spectrum of **19**



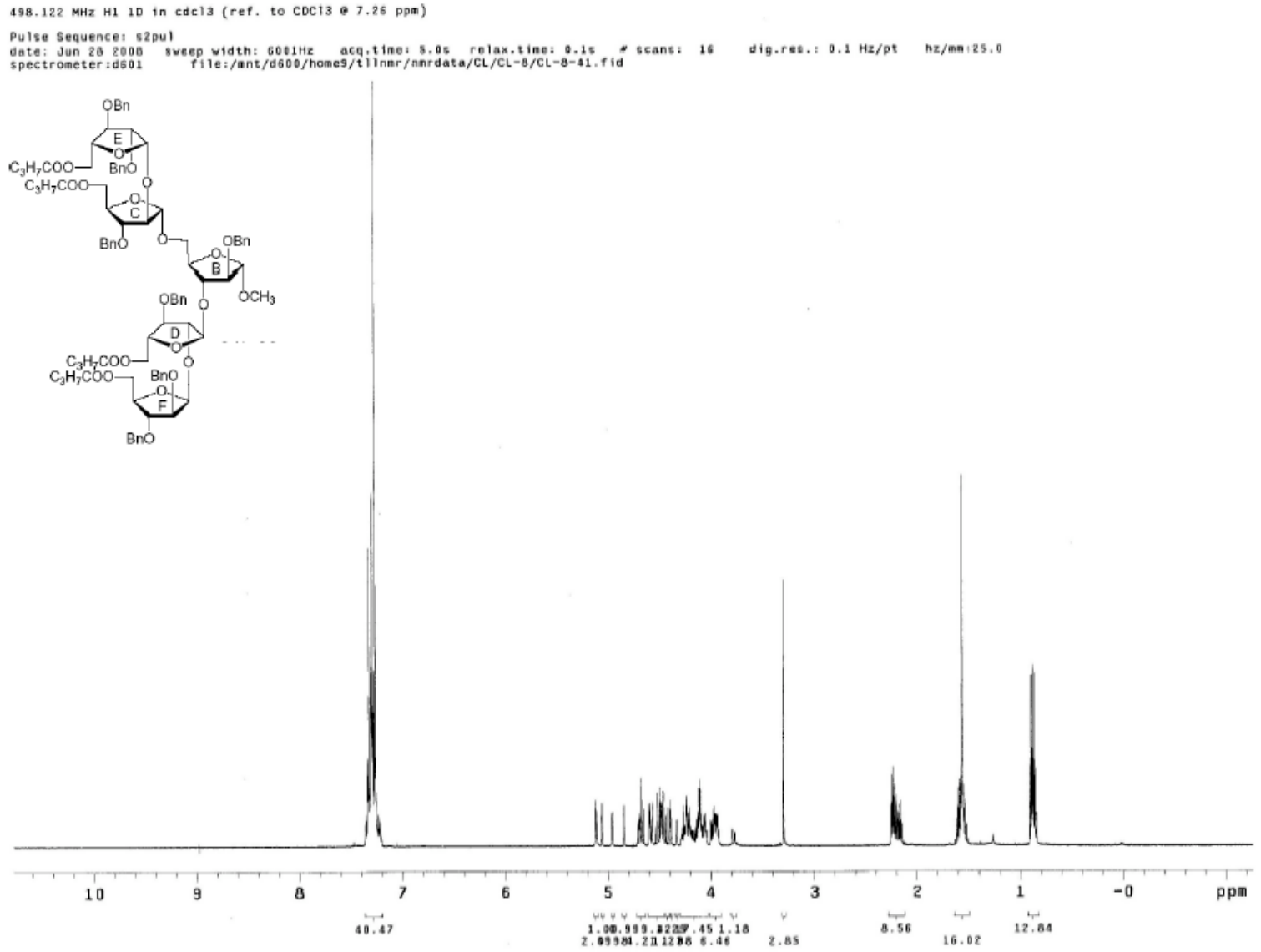
^{1}H NMR spectrum of **20**



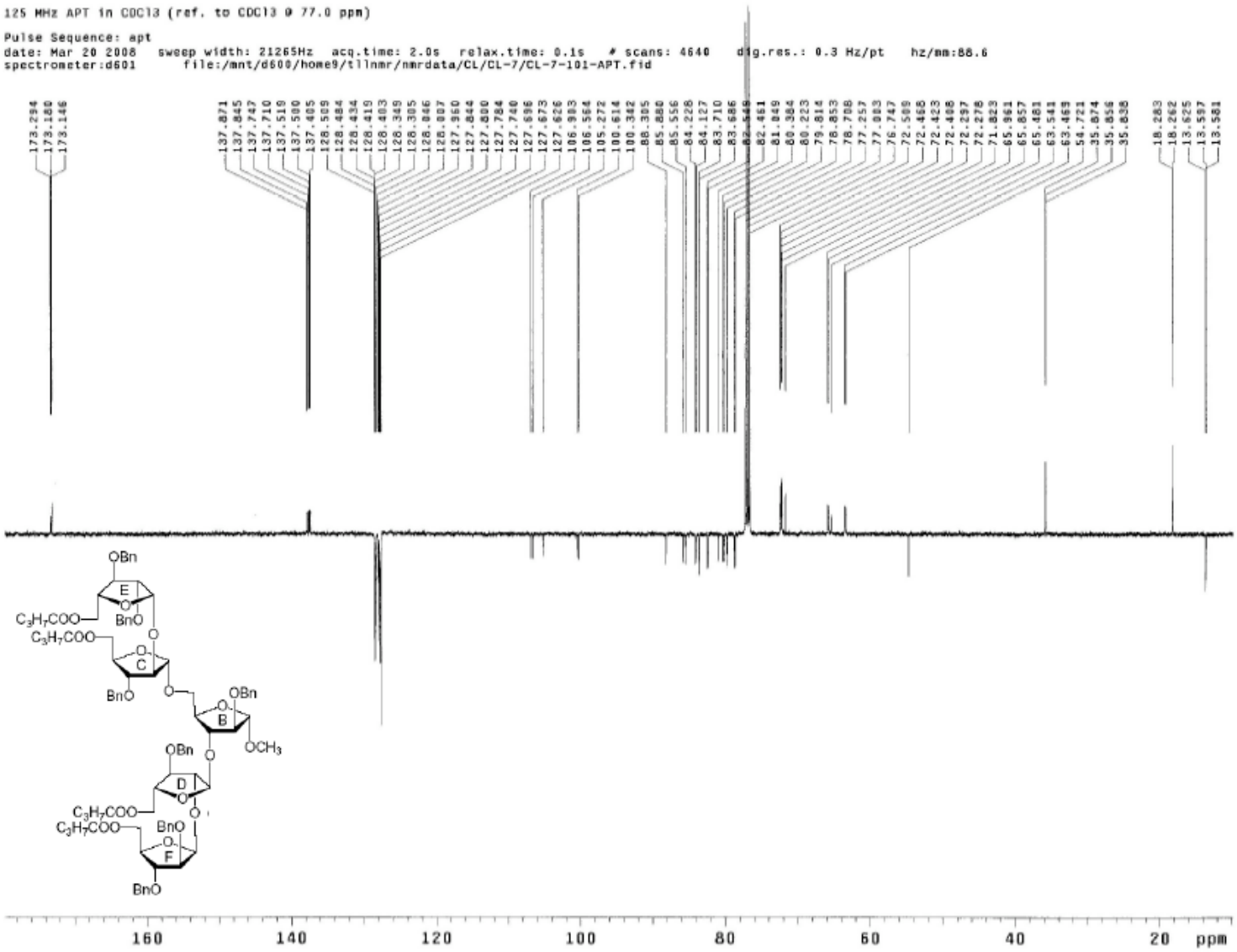
^{13}C NMR spectrum of **20**



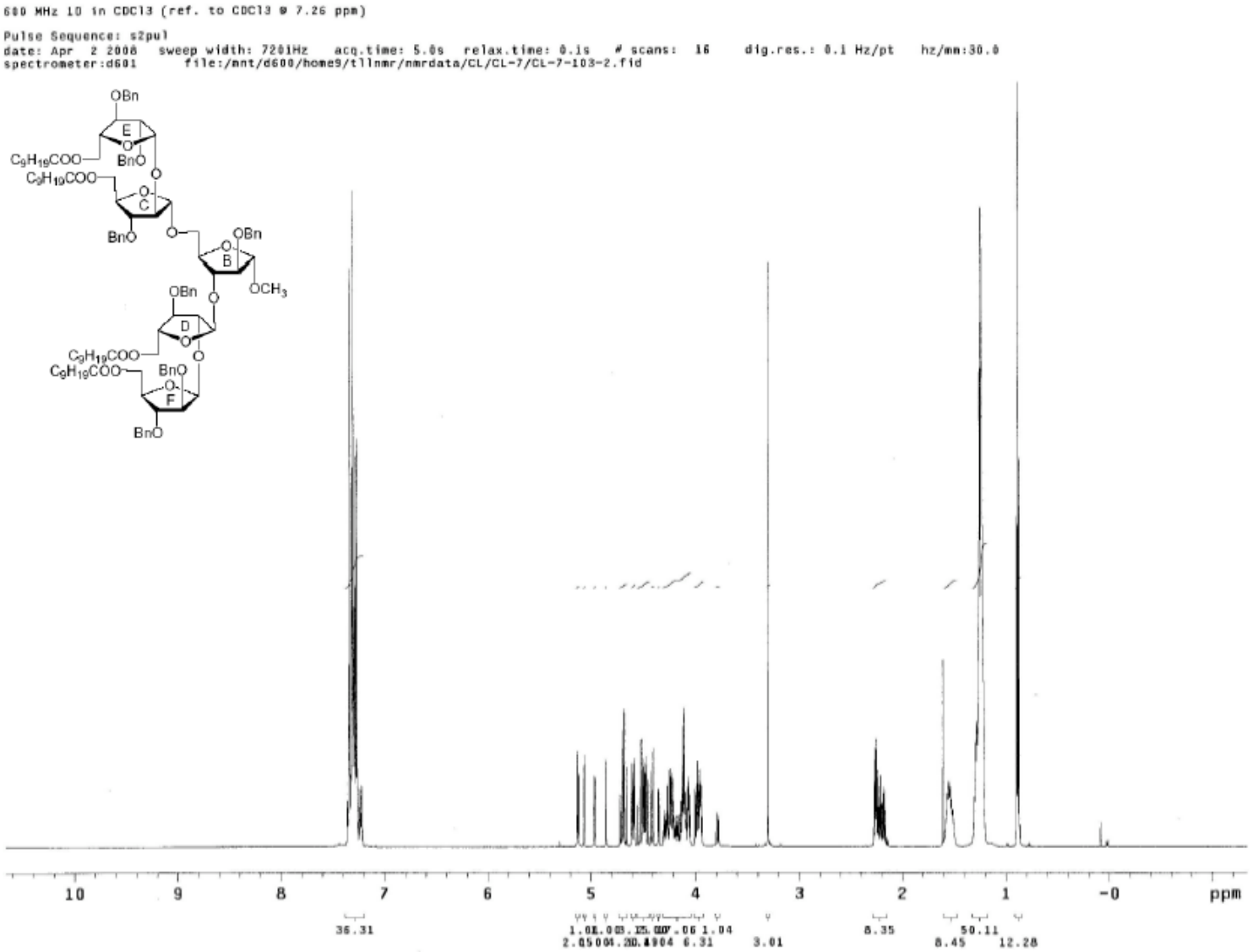
^{1}H NMR spectrum of **21**



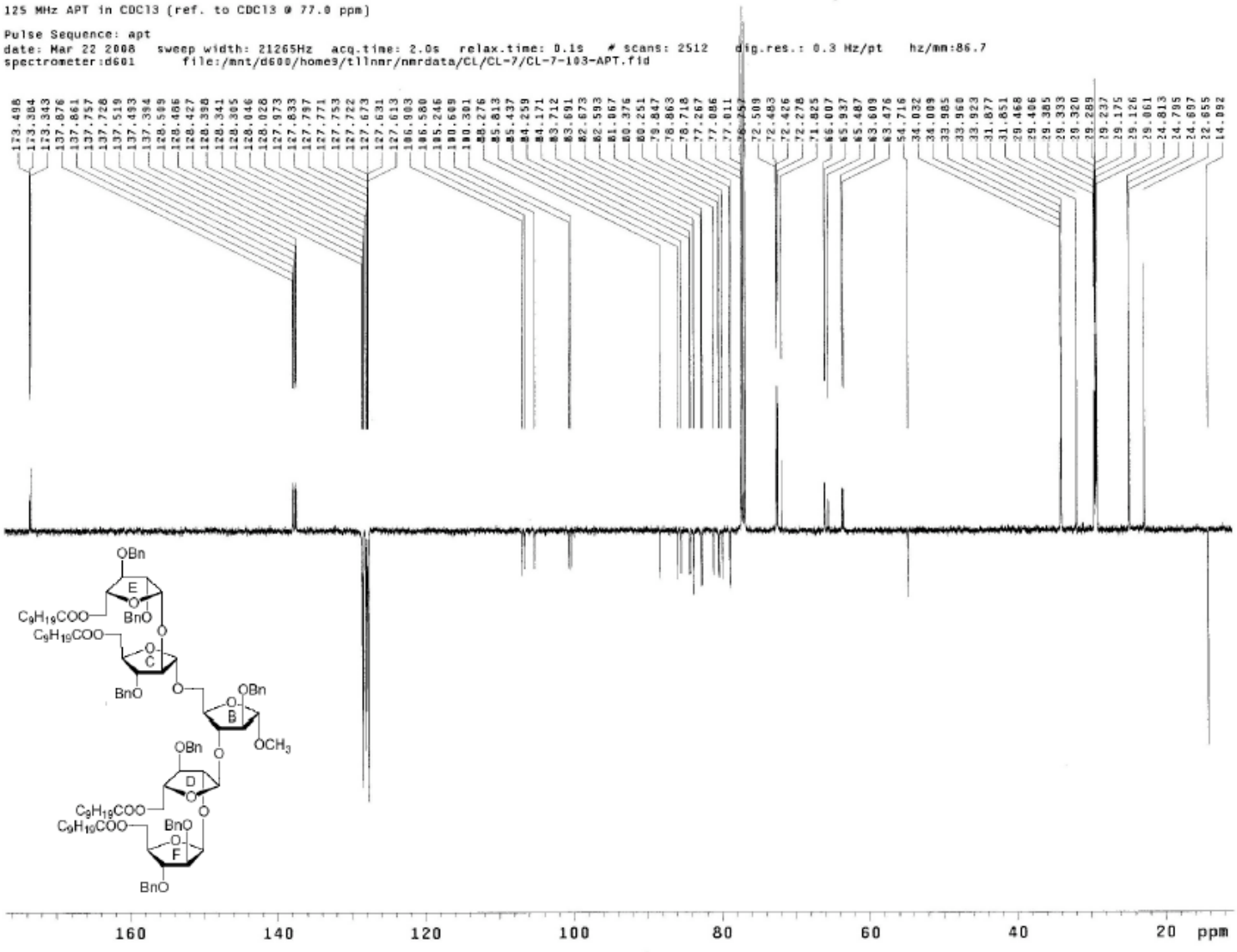
^{13}C NMR spectrum of **21**



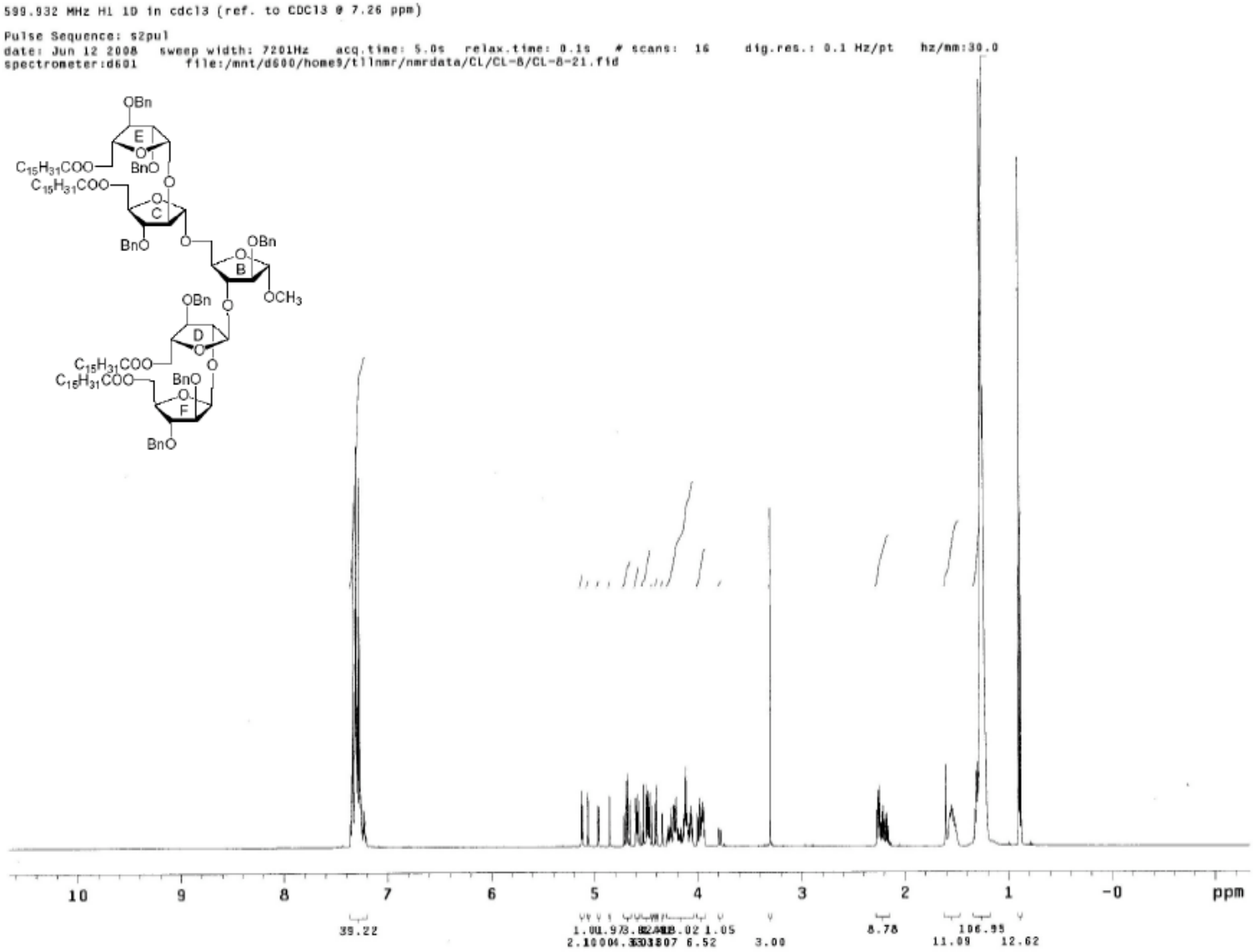
^{1}H NMR spectrum of **22**



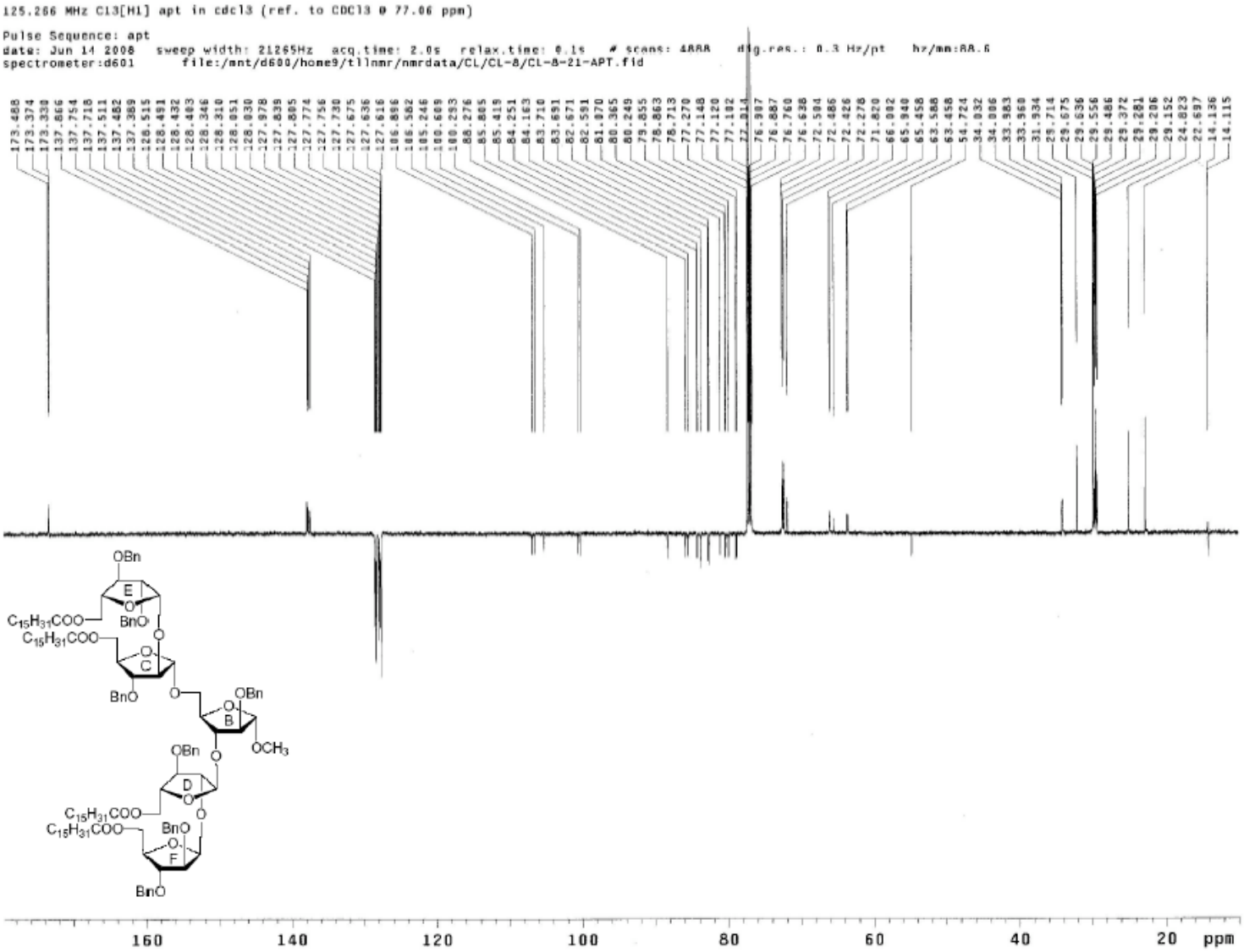
^{13}C NMR spectrum of **22**



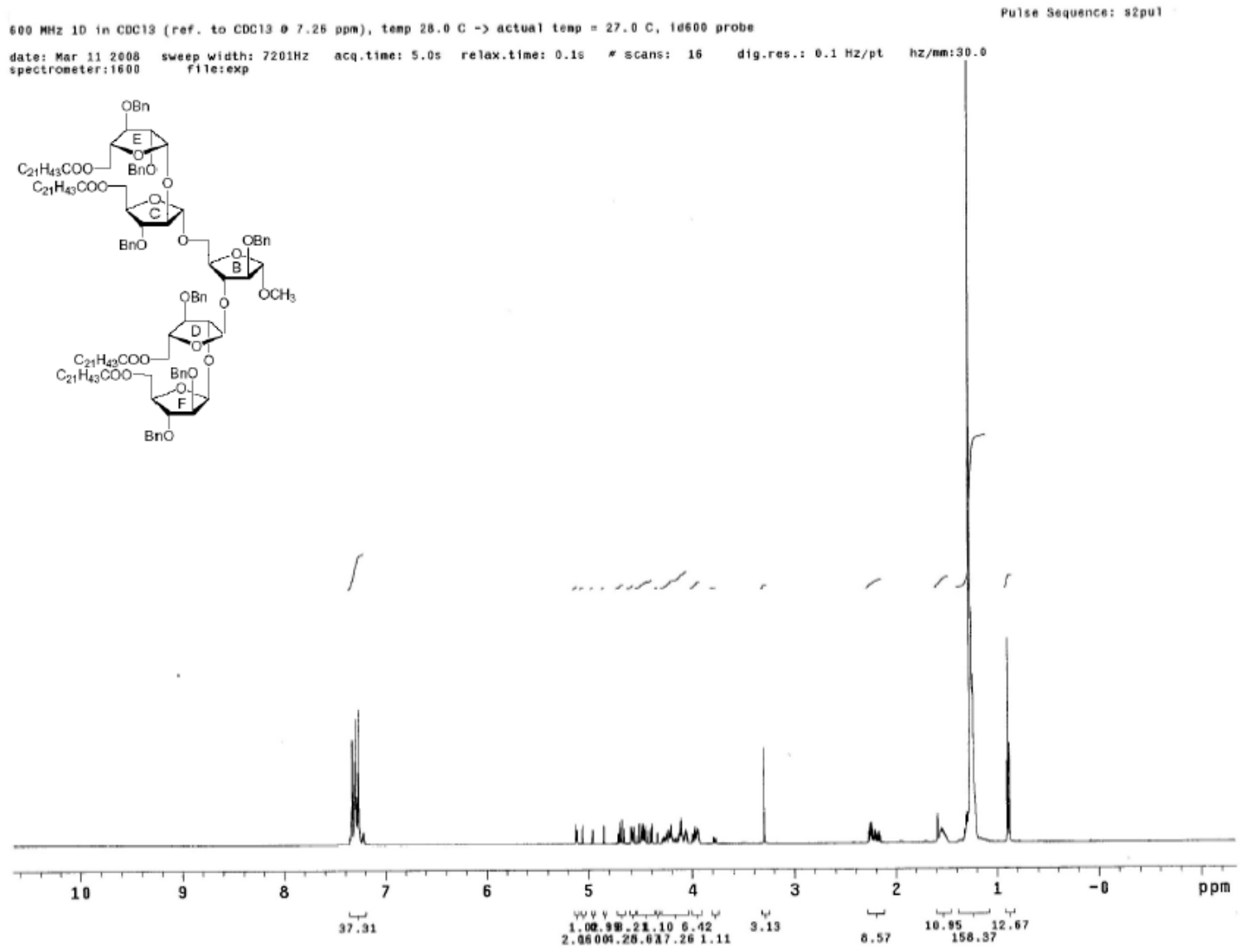
^{1}H NMR spectrum of **23**



^{13}C NMR spectrum of **23**



^{1}H NMR spectrum of **24**



^{13}C NMR spectrum of **24**