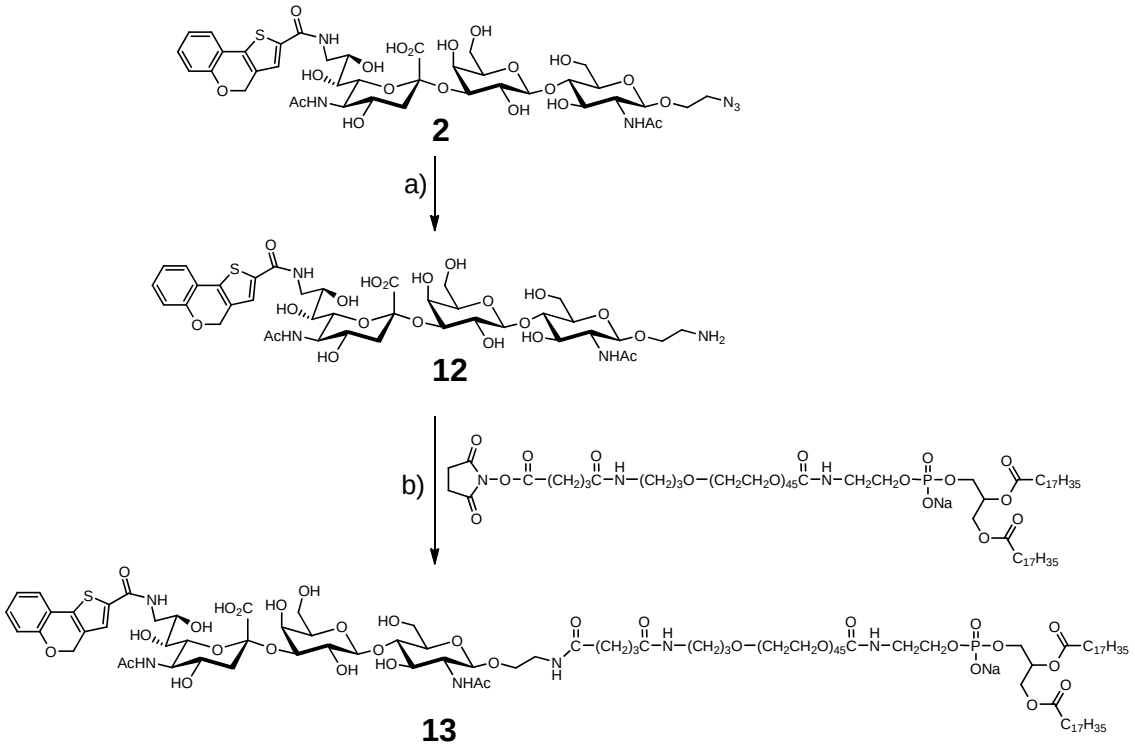
In silico aided design of a glycan ligand of sialoadhesin for *in vivo* targeting of macrophages.

Corwin M. Nycholat,‡ Christoph Rademacher,‡ Norihito Kawasaki and James C. Paulson*

Department of Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92037, USA.

Supporting Information

A. Supplemental Scheme 1.



Scheme S.1. Synthesis of ^{TCC}Neu5Acα2-3Galβ1-4GlcNAc pegylated lipid conjugate (**13**). a) trimethyl phosphine, H_2O, THF; b) DSPE-PEG-NHS, diisopropylethylamine, CH_2Cl_2, DMSO.

B. General methods.

Chemical reagents were obtained from Fisher Scientific or Sigma-Aldrich. The building block library used for the *in silico* screen and the carboxylic acids used for preparation of the sialoside analogues were obtained from Enamine Ltd. (Ukraine). 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC) and cholesterol were obtained from Avanti Polar Lipids (Alabaster, AL). 3-(*N*-succinimidyloxyglutaryl) aminopropyl polyethyleneglycol-carbamyl distearoyl phosphatidyl-ethanolamine (DSPE-PEG-NHS) and polyethyleneglycol-distearoyl phosphoethanolamine (DSPE-PEG) were obtained from NOF America Corporation (White Plains, NY). 5-Acetamido-9-amino-3,5,9-trideoxy-D-glycero-D-galacto-2-nonulopyranosyl-onate (9-amino-Neu5Ac) and Galβ1-4GlcNAc-ethyl azide (LacNAc-ethyl azide) were prepared as previously described.[1,2] The Sn ligand, 9-*N*-biphenylcarbamoyl-Neu5Acα2-3Galβ1-4GlcNAc-ethyl azide (**1**) was previously described.[3] Plasmids encoding *Pasteurella multocida* α2-3-sialyltransferase 1 (PmST1) was obtained from Prof. Peng Wu (Albert Einstein College of Medicine, NY) and expressed and purified as previously reported.[4] Other glycosyltransferases were prepared as previously described.[2,5] Enzymatic reactions were monitored by analytical thin-layer chromatography (TLC) performed on glass plates coated with Silica gel 60-F_{254} (E. Merck) and visualized under UV or by treatment with 10% sulfuric acid in ethanol followed by heating. Sialoadhesin-Fc chimera were generated as described previously.[6] Chinese hamster ovary (CHO) cell lines expressing human or murine siglecs were maintained as previously described.[7-9] ^{1}H NMR spectra were obtained on a *Bruker* DRX-600 (600 MHz) spectrometer and are reported in parts per million (δ) relative to HOD (4.78 ppm, D_2O). Coupling constants (*J*) are reported in Hertz. ^{13}C NMR spectra were obtained on a *Bruker* DRX-600 (150 MHz) spectrometer and are reported in parts per million (δ). Mass spectrometry data were acquired with an LC MSD TOF (Agilent Technologies, Foster City, California, USA) for ESI-TOF high resolution mass spectrometry (HRMS) data.

C. Construction of Siglec-H expressing BWZ cells

The extracellular domain of Siglec-H was subcloned into the pMXs-IRES-EGFP retroviral vector harboring the transmembrane region of mouse CD8a and intracellular region of mouse CD3z at the 3′ terminus. To generate the Siglec-H expressing BWZ cells, Plat-E cells were transfected with the pMXs-IG-Siglec-H-CD8a-CD3z ζ using Lipofectamine 2000 (Invitrogen, San Diego, CA) according to manufacturer's protocol. Two days after the transfection, the culture supernatant was harvested and added to BWZ.36 cells with 8 mg/mL polybrane. Established reporter cells with high levels of EGFP expression were enriched to more than 95% by cell sorting with the FACSVantage SE (BD Bioscience).

D. *In silico* screening

The *in silico* screen was performed using the coordinates of the co-crystal structure of murine sialoadhesin with methyl-α-9-*N*-(biphenyl-4-carbonyl)-amino-9-deoxy-Neu5Ac (1ODA.pdb).[10]

Initially, a commercial library of 8407 carboxylic acid building blocks (Enamine Ltd., Ukraine) was imported into Molecular Operation Environment (MOE, Chemical Computing Group, version 2010.10) using SMILES representations. Up to 250 low energy conformations were calculated for each acid using a fragment approach for conformer generation based on the MMFF94x force field. Conformations with a strain energy less than 4 kcal/mol were retained. Standard MOE force field setting for the stochastic conformer search were used, including a energy minimization with a 0.01 kcal/mol gradient and a limit 200 iterations. The generated conformers were individually conjugated *in silico* to the 9-NH_2-Neu5Ac in the protein binding pocket as follows. After the addition of hydrogens to the protein structure using Protonate3D[11] as implemented in MOE, the biphenyl-carbonyl substituent was removed and replaced with a hydrogen, forming methyl 9-NH_2-Neu5Ac and generating a connection point for the conformers. The carbonyl of the acid conformers was condensed on to the amine of methyl 9-NH_2-Neu5Ac

forming an amide. An aromatic ring pharmacophore was set using the coordinates of the first benzene ring of the biphenyl substituent. The atoms of both the Neu5Ac and protein were constrained allowing no flexibility, however, the acyl substituents were allowed to freely rotate around the newly formed amide bond. The generated Neu5Ac analogues were then assessed and ranked using a scoring function based on London dispersion without prior energy minimization. All computations in MOE were performed on a regular desktop computer.

From the MOE scoring, the top 3000 sialic acid derivatives were selected and inspected further with AutoDock 4.2.[12] Coordinates of the analogs were extracted from MOE and the input files for AutoDock were prepared using Raccoon[13] in combination with AutoDockTools (version 1.5.4). The affinity grid centered on the coordinates of the known inhibitor ^{BPC}Neu5Ac-OMe was calculated with 0.375 Å spacing. Rotatable bonds were automatically assigned, leaving the amide bond fixed. Docking was performed using the Lamarckian genetic algorithm with standard parameters, 30 runs per ligand, maximum number of energy evaluations of 2.5×10^6 and maximum number of generations being 27,000. The calculations were performed on the Garibaldi LINUX computer cluster (456 CPUs) at The Scripps Research Institute. The resulting docking poses were clustered using AutoDockTools and ranking was performed according to the largest energy cluster. A panel of six carboxylic acids were selected from the top 100 ranked docked structures for their structural diversity computed using MACCS fingerprints compared using the Tanimoto coefficient in MOE. Additionally, two acids were selected as 'non-ranked' controls that were excluded at the stage of fragment attachment prior to AutoDock evaluation. Administration of the ligands was done in Instant JChem (ChemAxon, version 2.5.3).

E. Competitive bead binding assay.

Magnetic beads (2 µl, $6.7x10^8$ beads/ml; Invitrogen Dynabeads M-280 Streptavidin, cat no. 112.06D) were coated overnight at 4 °C with Neu5Acα2-3Galβ1-4GlcNAc-PAA-biotin probe (2 µl, 1µg/µl, 30 kDa) in FACS buffer (200 µl, Hanks buffered saline solution, 0.5% bovine serum albumin). Beads were washed (FACS buffer, 3x 200µl) then re-suspended in FACS buffer. Stocks solutions of the 9-*N*-acyl substituted sialoside analogues (**1**-**9**) were prepared and quantified using the periodate-resorcinol assay.[14] For the inhibition assay, serial dilutions of the sialoside analogues (**1**-**9**) (50 µl) were first added to the wells in a 96-well plate. To each well was added $3.4x10^4$ Neu5Acα2-3Galβ1-4GlcNAc-coated beads followed by FITC-labeled goat anti-human IgG antibody (0.5 µl, 1.5 µg/µl, Jackson, 109-095-098). Sialoadhesin-Fc supernatant (40 µl) was then added to each well (total volume 100 µl) and incubated for 30 minutes at room temperature. Following incubation the supernatant was discarded and the beads were washed with FACS buffer (200µl). A bar magnet (Dynal MPC®-96S) fitted to the 96-well plate was used to prevent loss of the magnetic beads while discarding solutions. Finally, the beads were taken up with FACS buffer (200µl) and analyzed using a fluorescence-activated cell sorter (FACS). All titrations were performed in triplicate, and standard deviation is given from three independent measurements.

F. Liposome formulation and cell binding.

Liposomes were prepared as previously described.[8] Non-targeted naked-liposomes were composed of DSPC: cholesterol: PEG-DSPE in a 60:35:5 molar ratio. Sn-targeted ^{TCC}Neu5Ac-liposomes were composed of DSPC:cholesterol:PEG-DSPE:^{TCC}Neu5Ac-PEG-DSPE in a 60:35:4:1 molar ratio. Fluorescent labeled liposomes were prepared by incorporating 0.2 mol% of Alexa Fluor647-PEG-DSPE. For liposome preparation, lipids dissolved in $CHCl_3$ (DSPC, and cholesterol) and DMSO (^{TCC}Neu5Ac-PEG-DSPE, and PEG-DSPE) were mixed and concentrated

under reduced pressure followed by lyophilization. The lipid mixture was hydrated in nanopure H_2O to achieve a final liposome concentration of 1 mM (total phospholipids). Liposomes were extruded through polycarbonate membrane filters (Millipore) with controlled pore sizes of 0.4, 0.2, and 0.1 μM.

For the binding assay cells were suspended in mouse serum (2 mM EDTA) 10^6 cells/20 μL in the presence of fluorescently labeled liposomes at 100 μM (total phospholipids). After incubation at 37 °C for 30 minutes the cells were washed then analyzed using FACS.

G. *In vivo* binding

C57BL/6J (wild type) and Siglec1 KO mice (a generous gift from Dr. Paul Crocker, University of Dundee, UK) were maintained in pathogen-free conditions at The Scripps Research Institute breeding facility and were used in accordance to the guidelines of the Institutional Animal Care Committee at the National Institutes of Health. For the *in vivo* binding assay, wild type or Siglec-1 KO mice were subcutaneously injected on the back with Alexa-647 labeled ^{TCC}Neu5Ac liposomes. Neighboring lymph nodes were harvested 2.5 hours after injection. Mononuclear cells were stained with anti-TCRβ (H57-597, Biolegend), NK1.1 (PK136, Biolegend), and CD19 (3D6, Biolegend) to identify T, NK, and B cells, respectively. Myeloid cells were identified using anti-CD11b (M1/70, Biolegend), and anti-Siglec-1 (3D6.112, AbD serotec, Oxford, UK). The stained cells were washed then analyzed by LSR II (BD Bioscience, San Diego, CA) with the exclusion of dead cells using 1 μg/mL of propidium iodide staining.

H. Synthetic methods.

General procedure A: Synthesis of 9-*N*-acyl-subsituted sialoside analogs (2-9)

Carboxylic acids (**2-9**, 0.021 mmol) and *N*-hydroxy succinimide (0.032 mmol) were dissolved in anhydrous CH_2Cl_2 (0.5 ml). 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (0.032 mmol) was added followed by *N*,*N*-diisopropylethylamine (0.032 mmol). The reaction was mixed at room temperature under an atmosphere of N_2. When the reaction was complete as indicated by tlc (CH_2Cl_2-MeOH, 95:5) the mixture was concentrated under reduced pressure. The remaining syrup was subjected to silica gel chromatography (glass pipette column, 5 cm; hexanes-EtOAc, 1:1). Fractions containing pure product were combined and concentrated under reduced pressure. The remaining solids were used directly in the next reaction. Compound **10** (3.0 mg, 0.0039 mmol) was dissolved in deionized distilled H_2O (0.5 ml). A solution of the *N*-hydroxy succinimide activated carboxylic acid (2.5 eq) in THF was added followed by *N*,*N*-diisopropylethylamine (1 µl, 0.0059 mmol, 1.5 eq). The reaction was mixed vigorously at room temperature. When the reaction was complete as indicated by tlc (EtOAc-MeOH-acetic acid-H_2O, 6:3:3:2) the mixture was concentrated under reduced pressure. The remaining solid was re-dissolved in H_2O (1 ml) then passed through a pipette column containing Amberlite™ Resin IR-120 H^+ (3 cm). The fractions containing product were combined and lyophilized. The residue was then re-dissolved in H_2O and subjected to size exclusion chromatography (Bio-Gel® P-2 Gel Fine, 1x45 cm) with H_2O elution. Fractions containing product were combined and the sample lyophilized. Typical isolated yields for the 9-*N*-acyl substituted analogs (**2-9**) were 60-70%.

2-azidoethyl (sodium 5-Acetamido-9-(4*H*-thieno[3,2-c]chromene-2-carboxamido)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (2)

Prepared using general procedure A. R_f 0.57 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 7.29 (1H, s), 7.28 (1H, d, *J* = 1.5 Hz), 7.14 (1H, ddd, *J* = 1.7, 8.2, 8.2 Hz), 6.91 (1H, ddd, *J* = 1.1, 7.5, 7.5 Hz), 6.84 (1H, d, *J* = 8.3 Hz), 5.11 (2H, s), 4.29 (1H, d, *J* = 7.9 Hz), 4.28 (1H, d, *J* = 7.7 Hz), 3.91 (1H, dd, *J* = 3.1, 9.8 Hz), 3.85 (1H, ddd, *J* = 2.5, 8.7, 8.7 Hz), 3.78 (1H, d, *J* = 2.9 Hz), 3.77-3.73 (2H, m), 3.63-3.40 (14H, m), 3.33 (1H, ddd, *J* = 2.4, 7.7, 11.3 Hz), 3.26 (1H, m), 3.21-3.14 (2H, m), 3.07 (1H, ddd, *J* = 2.6, 5.4, 13.6 Hz), 2.64 (1H, dd, *J* = 4.6, 12.4 Hz), 1.89 (3H, s), 1.87 (3H, s), 1.65 (1H, dd, *J* = 12.2 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 175.8, 175.3, 174.3, 164.6, 152.6, 137.7, 137.0, 133.2, 131.3, 127.3, 124.6, 123.6, 120.0, 117.4, 103.2, 101.7, 100.9, 78.4, 77.1, 76.2, 75.4, 73.8, 73.1, 71.8, 70.7, 69.9, 69.1, 69.0, 68.3, 66.5, 61.8, 60.4, 55.7, 52.5, 51.0, 43.7, 40.9, 23.0, 22.8; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{39}H_{52}N_6O_{20}S$ 957.3030, obs 957.3022.

2-azidoethyl (sodium 5-Acetamido-9-(5-(2,3-dihydro-1,4-benzodioxin-6-yl)-thiophene-2-carboxamido)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (3)

Prepared using general procedure A. R_f 0.49 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 7.53 (1H, d, *J* = 4.0 Hz), 7.20 (1H, d, *J* = 4.0 Hz), 7.13-7.11 (2H, m), 6.82 (1H, d, *J* = 8.8 Hz), 4.32 (1H, d, *J* = 7.9 Hz), 4.27 (1H, d, *J* = 8.1 Hz), 4.20 (4H, bs), 3.90 (1H, dd, *J* = 3.1, 9.8 Hz), 3.85 (1H, ddd, *J* = 2.5, 8.7, 8.7 Hz), 3.78 (1H, d, *J* = 2.9 Hz), 3.77-3.73 (2H, m), 3.65-3.20 (16H, m), 3.18 (1H, ddd, *J* = 2.8, 7.9, 13.7 Hz), 3.08 (1H, ddd, *J* = 2.8, 5.3, 13.7 Hz), 2.64 (1H, dd, *J* = 4.4, 12.4 Hz), 1.89 (3H, s), 1.87 (3H, s), 1.65 (1H, dd, *J* = 12.3 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 174.9, 174.5, 173.3, 164.1, 148.8, 143.8, 143.3, 135.3, 130.6, 126.8, 123.6, 119.6, 117.9, 114.5, 102.3, 100.9, 100.1, 77.3, 76.4, 75.4, 74.6, 73.0, 72.2, 71.0, 69.9, 69.1, 68.4, 68.1, 67.5, 64.6, 64.5, 61.0, 59.4, 55.0, 51.6, 50.2, 42.9, 40.1, 22.2, 22.0; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{40}H_{54}N_6O_{21}S$ 987.3135, obs 987.3123.

2-azidoethyl (sodium 5-Acetamido-9-(3-ethoxy-4-(2-phenoxy-ethoxy)-benzoylamino)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (4)

Prepared using general procedure A. R_f 0.53 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 7.34 (1H, dd, *J* = 2.1, 8.4 Hz), 7.31 (1H, d, *J* = 2.0 Hz), 7.26-7.23 (2H, m), 7.05 (1H, d, *J* = 8.5 Hz), 6.94 (1H, ddd, *J* = 0.7, 7.4, 7.4 Hz), 6.91-6.90 (2H, m), 4.38 (1H, d, *J* = 8.4 Hz), 4.36 (2H, m), 4.34 (1H, d, *J* = 7.9 Hz), 4.32 (2H, m), 4.02 (2H, q, *J* = 7.0 Hz), 3.95 (1H, dd, *J* = 3.2, 9.9 Hz), 3.89 (1H, ddd, *J* = 2.8, 8.4, 8.4 Hz), 3.85 (1H, ddd, *J* = 3.0, 5.6, 11.4 Hz), 3.80 (1H, d, *J* = 3.1 Hz), 3.74 (1H, dd, *J* = 10.2 Hz), 3.69 (1H, dd, *J* = 2.8, 14.2 Hz), 3.62-3.50 (11H, m), 3.44-3.33 (4H, m), 3.23 (1H, ddd, *J* = 2.9, 5.4, 13.7 Hz), 2.63 (1H, dd, *J* = 4.6, 12.4 Hz), 1.89 (3H, s), 1.86 (3H, s), 1.66 (1H, dd, *J* = 12.3 Hz), 1.24 (3H, t, 7.0); **^{13}C NMR** (D_2O, 150 MHz) δ 175.7, 175.4, 174.5, 170.9, 158.7, 151.5, 148.5, 130.7, 127.7, 122.6, 121.9, 115.8, 114.1, 113.2, 103.3, 101.8, 100.8, 78.8, 76.6, 76.1, 75.5, 73.7, 73.2, 71.4, 70.6, 70.1, 69.4, 69.0, 68.5, 68.3, 67.5, 66.0, 61.8, 60.6, 55.8, 52.5, 51.1, 43.6, 40.6, 23.0, 22.8, 14.6; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{44}H_{62}N_6O_{22}$ 1027.3990, obs 1027.3993.

2-azidoethyl (sodium 5-Acetamido-9-(2-benzyl-5-methyl-4-oxo-3,4-dihydrothieno[2,3-d]pyrimidine-6-carboxamido)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (5)

Prepared using general procedure A. R_f 0.51 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 7.31-7.22 (5H, m), 4.39 (1H, d, *J* = 7.8 Hz), 4.38 (1H, d, *J* = 8.1 Hz), 3.99 (1H, dd, *J* = 3.1, 9.8 Hz), 3.96 (2H, bs), 3.92 (1H, ddd, *J* = 3.0, 7.7, 8.6 Hz), 3.84 (1H, d, *J* = 3.1 Hz), 3.81 (1H, ddd, *J* = 2.8, 5.4, 11.3 Hz), 3.77 (1H, d, *J* = 10.2 Hz), 3.73 (1H, dd, *J* = 2.9, 14.1 Hz), 3.69 (1H, dd, *J* = 2.1, 12.3 Hz), 3.65-3.53 (10H, m), 3.48-3.37 (4H, m), 3.30 (1H, ddd, *J* = 2.9, 7.8, 13.8 Hz), 3.22 (1H, ddd, *J* = 2.9, 5.6, 13.8 Hz), 2.66 (1H, dd, *J* = 4.6, 12.4 Hz), 2.60

(3H, s), 1.93 (3H, s), 1.92 (3H, s), 1.69 (1H, dd, *J* = 12.1 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 174.9, 174.5, 173.4, 165.5, 165.3, 136.9, 128.93, 128.91, 127.3, 121.3, 102.4, 100.9, 99.9, 77.9, 75.8, 75.2, 74.7, 72.8, 72.2, 70.3, 69.8, 69.2, 68.6, 68.2, 67.4, 61.0, 59.9, 54.9, 51.6, 50.2, 42.8, 41.4, 39.8, 22.2, 22.0, 14.7; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{42}H_{56}N_8O_{20}S$ 1025.3404, obs 1025.3420.

2-azidoethyl (sodium 5-Acetamido-9-(4-{[(6-oxo-1,6-dihydropyridin-3-yl)-carbonyl]-amino}-benzoylamino)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (6)

Prepared using general procedure A. R_f 0.32 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 8.18 (1H, d, *J* = 2.7 Hz), 7.93 (1H, dd, *J* = 2.7, 9.4 Hz), 7.72 (2H, m), 7.55 (2H, m), 6.48 (1H, d, *J* = 9.4 Hz), 4.38 (1H, d, *J* = 8.4 Hz), 4.33 (1H, d, *J* = 7.9 Hz), 3.95 (1H, dd, *J* = 3.1, 9.8 Hz), 3.89 (1H, ddd, *J* = 2.8, 8.5, 8.5 Hz), 3.82 (1H, ddd, *J* = 3.0, 5.4, 11.3 Hz), 3.80 (1H, d, *J* = 3.2 Hz), 3.75 (1H, dd, *J* = 10.2 Hz), 3.72 (1H, dd, *J* = 2.8, 14.1 Hz), 3.61-3.41 (13H, m), 3.37-3.32 (2H, m), 3.26 (1H, ddd, *J* = 3.2, 7.8, 13.9 Hz), 3.19 (1H, ddd, *J* = 3.0, 5.5, 13.8 Hz), 2.63 (1H, dd, *J* = 4.7, 12.6 Hz), 1.89 (3H, s), 1.88 (3H, s), 1.66 (1H, dd, *J* = 12.2 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 174.9, 174.5, 173.6, 170.1, 166.3, 140.6, 140.0, 129.9, 128.2, 121.5, 117.2, 115.5, 102.6, 100.9, 99.9, 78.1, 75.9, 75.2, 74.6, 72.8, 72.4, 70.7, 69.7, 69.2, 86.6, 68.2, 67.4, 61.0, 59.8, 54.9, 51.6, 50.3, 42.8, 39.8, 22.2, 22.0; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{40}H_{54}N_8O_{21}$ 983.3476, obs 983.3473.

2-azidoethyl (sodium 5-Acetamido-9-(7-methoxycinnoline-3-carboxamido)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (7)

Prepared using general procedure A. R_f 0.37 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 8.47 (1H, s), 7.87 (1H, d, *J* = 9.1 Hz), 7.56 (1H, s), 7.41 (1H, dd, *J* = 2.1, 9.0 Hz), 4.35-4.33 (2H, m), 4.02-3.97 (2H, m), 3.94 (3H, s), 3.89 (1H, dd, *J* = 2.8, 14.1 Hz), 3.81 (1H, d, *J* = 3.0 Hz), 3.77 (1H, dd, *J* = 10.1 Hz), 3.76-3.73 (1H, m), 3.63 (1H, dd, *J* = 1.6, 10.4 Hz), 3.59-3.40 (13H, m), 3.32-3.20 (3H, m), 2.64 (1H, dd, *J* = 4.6, 12.5 Hz), 1.873 (3H, s), 1.871 (3H, s), 1.67 (1H, dd, *J* = 12.2 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 174.9, 174.5, 173.6, 165.6, 162.9, 152.3, 145.0, 129.5, 126.6, 124.3, 122.7, 104.6, 102.6, 100.8, 99.9, 78.2, 75.7, 75.2, 74.6, 72.8, 72.3, 70.6, 69.9, 69.2, 68.5, 68.3, 67.4, 61.0, 59.8, 56.1, 54.8, 51.7, 50.3, 42.5, 39.8, 22.2, 22.0; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{37}H_{52}N_8O_{20}$ 929.3370, obs 929.3345.

2-azidoethyl (sodium 5-Acetamido-9-(3-(4-fluoro-phenylsulfamoyl)-4-methyl benzoylamino)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (8)

Prepared using general procedure A. R_f 0.53 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 8.10 (1H, d, *J* = 1.9 Hz), 7.79 (1H, dd, *J* = 1.8, 7.9 Hz), 7.41 (1H, d, *J* = 8.1 Hz), 6.96-6.86 (4H, m), 4.40 (1H, d, *J* = 8.4 Hz), 4.36 (1H, d, *J* = 7.9 Hz), 3.95 (1H, dd, *J* = 3.1, 9.8 Hz), 3.88 (1H, ddd, *J* = 2.9, 8.3, 8.3 Hz), 3.85 (1H, ddd, *J* = 3.1, 5.7, 11.5 Hz), 3.80 (1H, d, *J* = 3.1 Hz), 3.73 (1H, d, *J* = 10.2 Hz), 3.68 (1H, dd, *J* = 2.8, 14.1 Hz), 3.63-3.52 (11H, m), 3.43 (1H, dd, *J* = 7.9, 9.7 Hz), 3.40 (1H, dd, *J* = 1.7, 8.7 Hz), 3.38-3.30 (3H, m), 3.26 (1H, ddd, *J* = 3.1, 5.6, 13.8 Hz), 2.63 (1H, dd, *J* = 4.6, 12.4 Hz), 2.49 (3H, s), 1.89 (3H, s), 1.86 (3H, s), 1.66 (1H, dd, *J* = 12.1 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 175.0, 174.6, 173.6, 168.8, 161.3, 159.7, 141.7, 136.2, 133.4, 131.9, 131.9, 128.5, 124.63, 124.57, 116.2, 116.1, 102.4, 100.9, 99.9, 77.9, 75.7, 75.2, 74.7, 72.8, 72.3, 70.3, 69.8, 69.3, 68.6, 68.2, 67.4, 60.9, 59.8, 54.9, 51.6, 50.3, 42.9, 39.8, 22.2, 22.0, 19.4; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{41}H_{56}FN_7O_{21}S$ 1034.3307, obs 1034.3299.

2-azidoethyl (sodium 5-Acetamido-9-(3-(4-methylphenyl)-thiophene-2-carboxamido)-3,5,9-trideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside) (9)

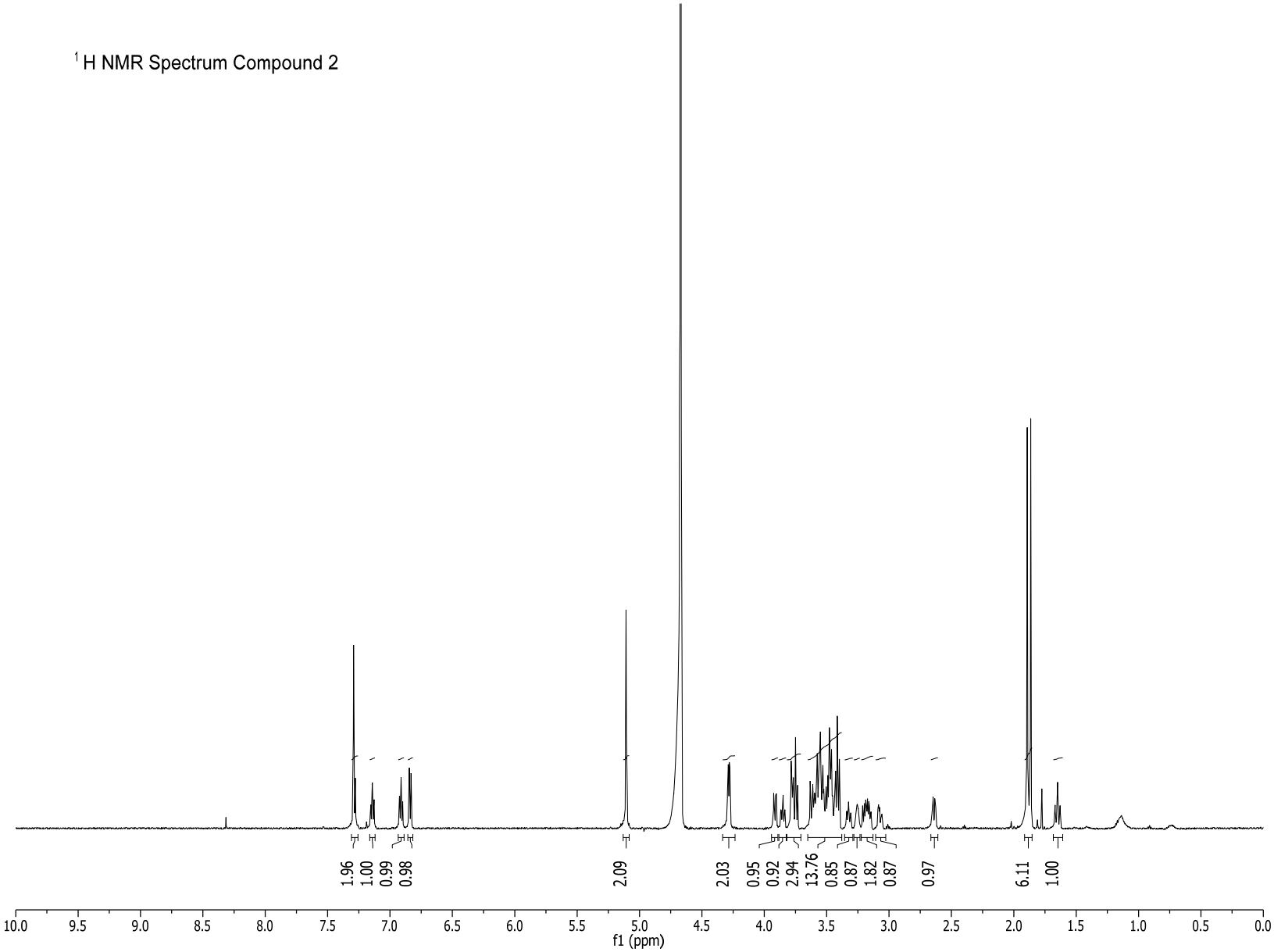
Prepared using general procedure A. R_f 0.57 (6:3:3:2, EtOAc-MeOH-AcOH-H_2O); **^{1}H NMR** (D_2O, 600 MHz) δ 7.58 (1H, d, *J* = 5.1 Hz), 7.31-7.30 (2H, m), 7.26-7.25 (2H, m), 7.11 (1H, d, *J* = 5.1 Hz), 4.42 (1H, d, *J* = 8.4 Hz), 4.34 (1H, d, *J* = 7.8 Hz), 3.96 (1H, dd, *J* = 3.2, 9.8 Hz), 3.93 (1H, ddd, *J* = 3.0, 5.6, 11.4 Hz), 3.84 (1H, d, *J* = 3.0 Hz), 3.79-3.73 (2H, m), 3.71 (1H, dd, *J* = 10.1 Hz), 3.67-3.52 (10H, m), 3.50 (1H, dd, *J* = 1.7, 10.5 Hz), 3.45 (1H, dd, *J* = 7.9, 9.7 Hz), 3.42 (1H, ddd, *J* = 2.0, 5.0, 9.2 Hz), 3.38 (1H, ddd, *J* = 3.0, 7.6, 13.6 Hz), 3.31 (1H, ddd, *J* = 2.9, 5.6, 13.9 Hz), 3.30 (1H, dd, *J* = 1.4, 8.7 Hz), 3.22 (1H, dd, *J* = 7.7, 14.1 Hz), 2.66 (1H, dd, *J* = 4.7, 12.5 Hz), 2.30 (3H, s), 1.93 (3H, s), 1.92 (3H, s), 1.67 (1H, dd, *J* = 12.1 Hz); **^{13}C NMR** (D_2O, 150 MHz) δ 175.0, 174.6, 173.5, 165.5, 143.4, 138.9, 131.7, 130.8, 130.7, 129.7, 129.0, 128.5, 102.5, 100.9, 99.8, 78.2, 75.7, 75.3, 74.7, 72.8, 72.3, 69.8, 69.5, 69.3, 68.7, 68.1, 67.4, 61.0, 60.0, 54.9, 51.7, 50.3, 42.3, 39.8, 22.2, 21.9, 20.4; **HRMS** (ESI-TOF high acc.) *m/z* $[M + H]^+$ calcd for $C_{39}H_{54}N_6O_{19}S$ 943.3237, obs 943.3240.

Synthesis of ^{TCC}Neu5Ac-pegylated lipid conjugate (13)

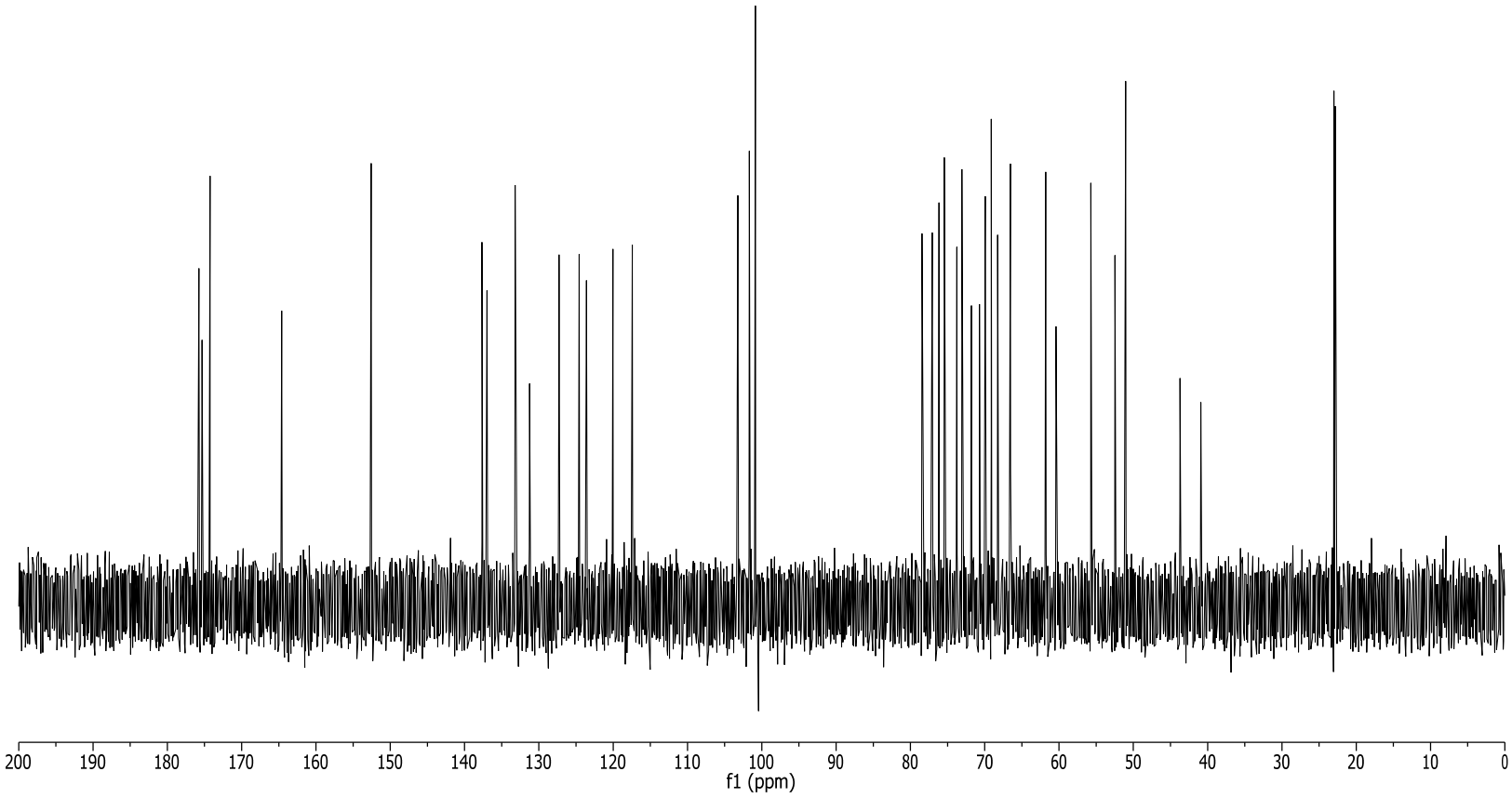
9-*N*-(4*H*-thieno[3,2-c]chromene-2-carbamoyl)-Neu5Acα2-3Galβ1-4GlcNAc-ethyl amine **11** (8.4 mg, 9.0 μmol) and DSPE-PEG-NHS (13.8 mg, 4.5 μmol) were dissolved in an anhydrous mixture of CH_2Cl_2 (1 ml) and DMSO (1 ml). *N*,*N*-Diisopropylethylamine (7.9 μl, 45 μmol) was added then the reaction was mixed at room temperature. The reaction was monitored by tlc (EtOAc-MeOH-AcOH-H_2O, 6:3:3:2) using an I_2/silica chamber for staining. After 2 hours, the reaction appeared complete by tlc and was concentrated under reduced to pressure using a rotary evaporator to remove CH_2Cl_2. During concentration the bath temperature was kept below 40 °C. The remaining DMSO solution was then diluted with H_2O then lyophilized. The remaining solid

was re-dissolved in H_2O (10 ml) and dialyzed overnight against H_2O (1 l) using a Thermo Scientific Slide-A-Lyzer Dialysis cassette 10000 MWCO. The aqueous solution was removed from the cassette then lyophilized to a white amorphous solid (96% yield). The product was used directly for the formulation of liposomes.

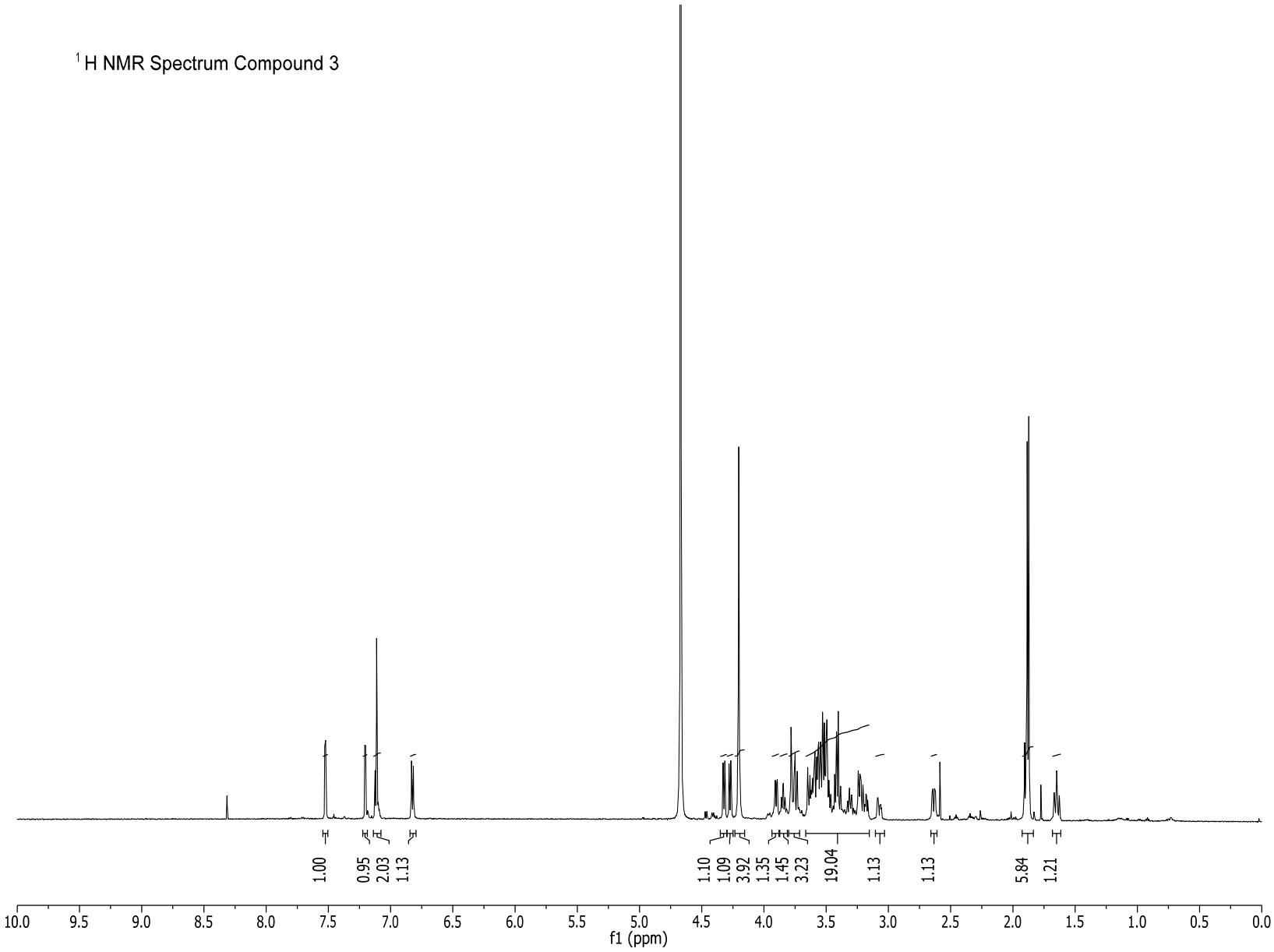
H. NMR Data


1 H NMR Spectrum Compound 2
1.96
1.00
0.99
0.98
2.09
2.03
0.95
0.92
2.94
13.76
0.85
0.87
1.82
0.87
0.97
6.11
1.00
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

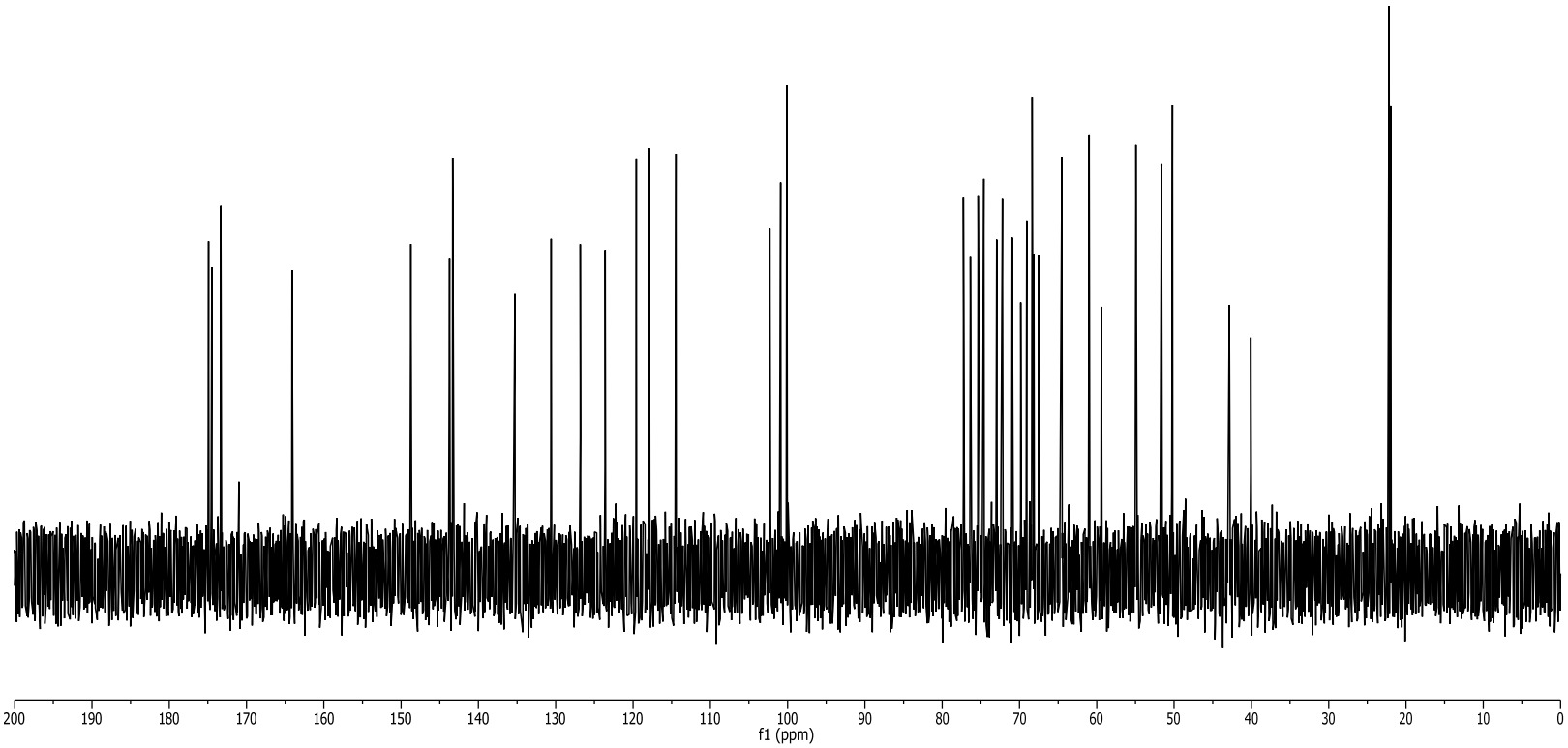
^{13} C NMR Spectrum Compound 2



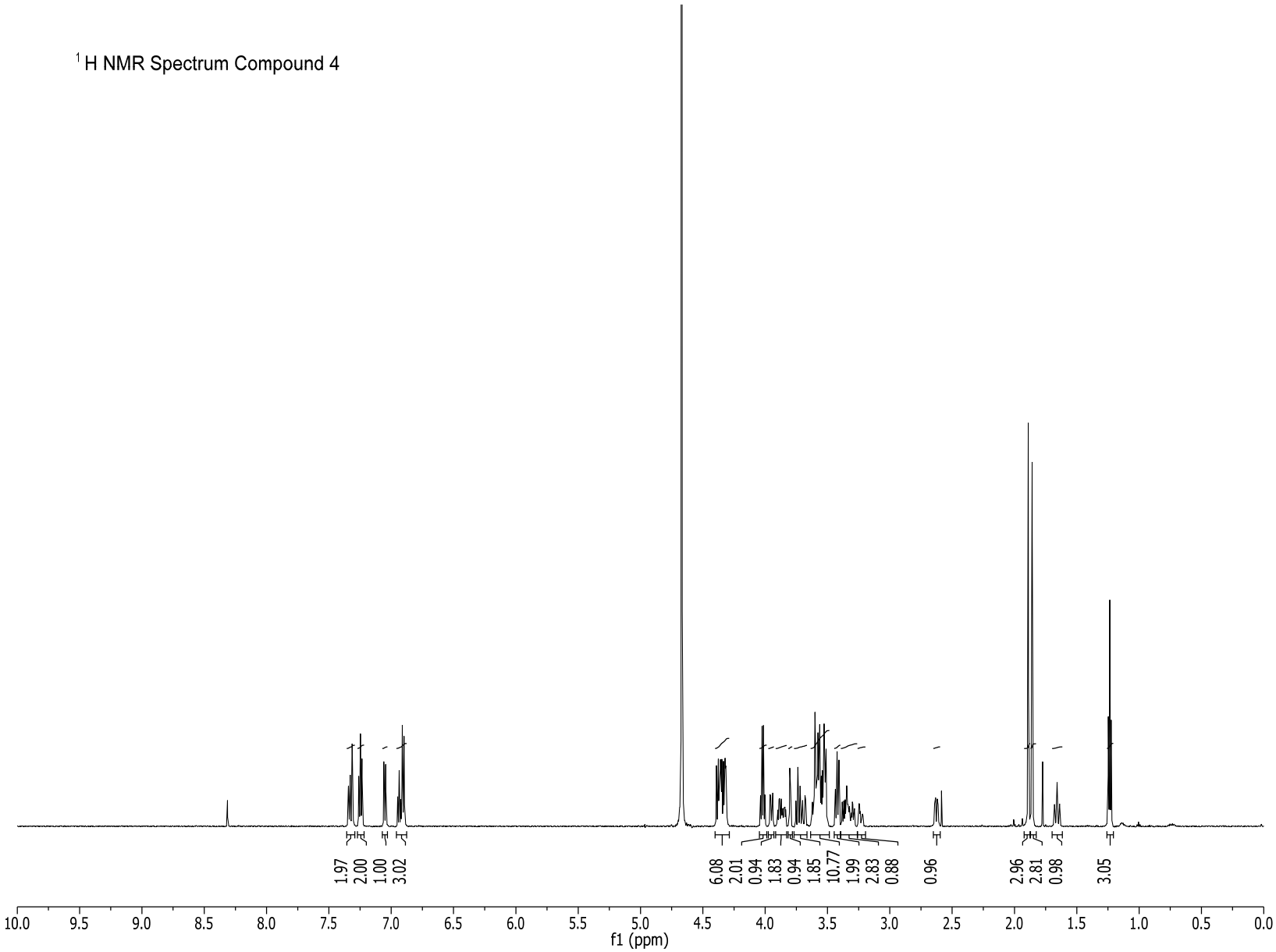
^{1}H NMR Spectrum Compound 3



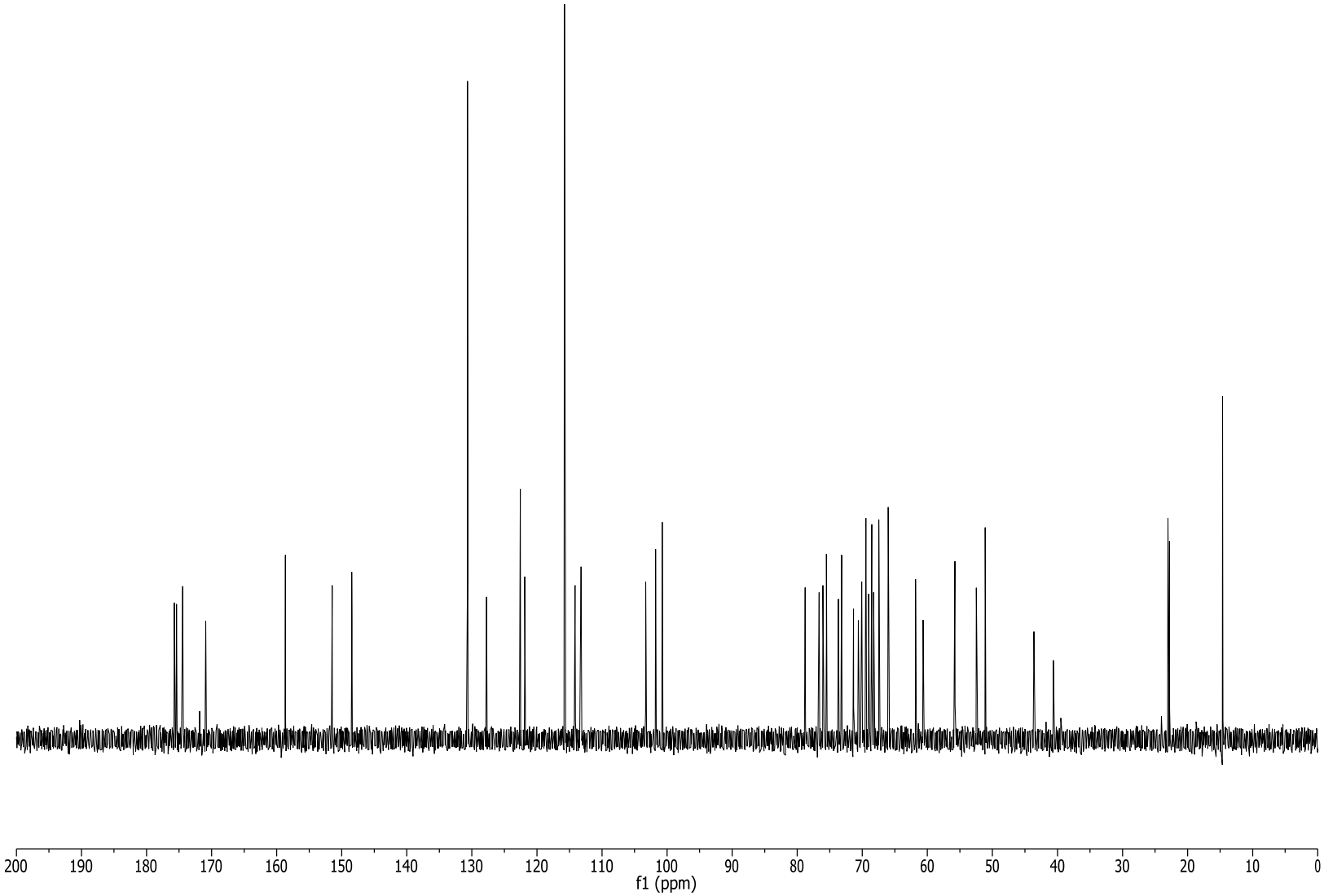
^{13} C NMR Spectrum Compound 3



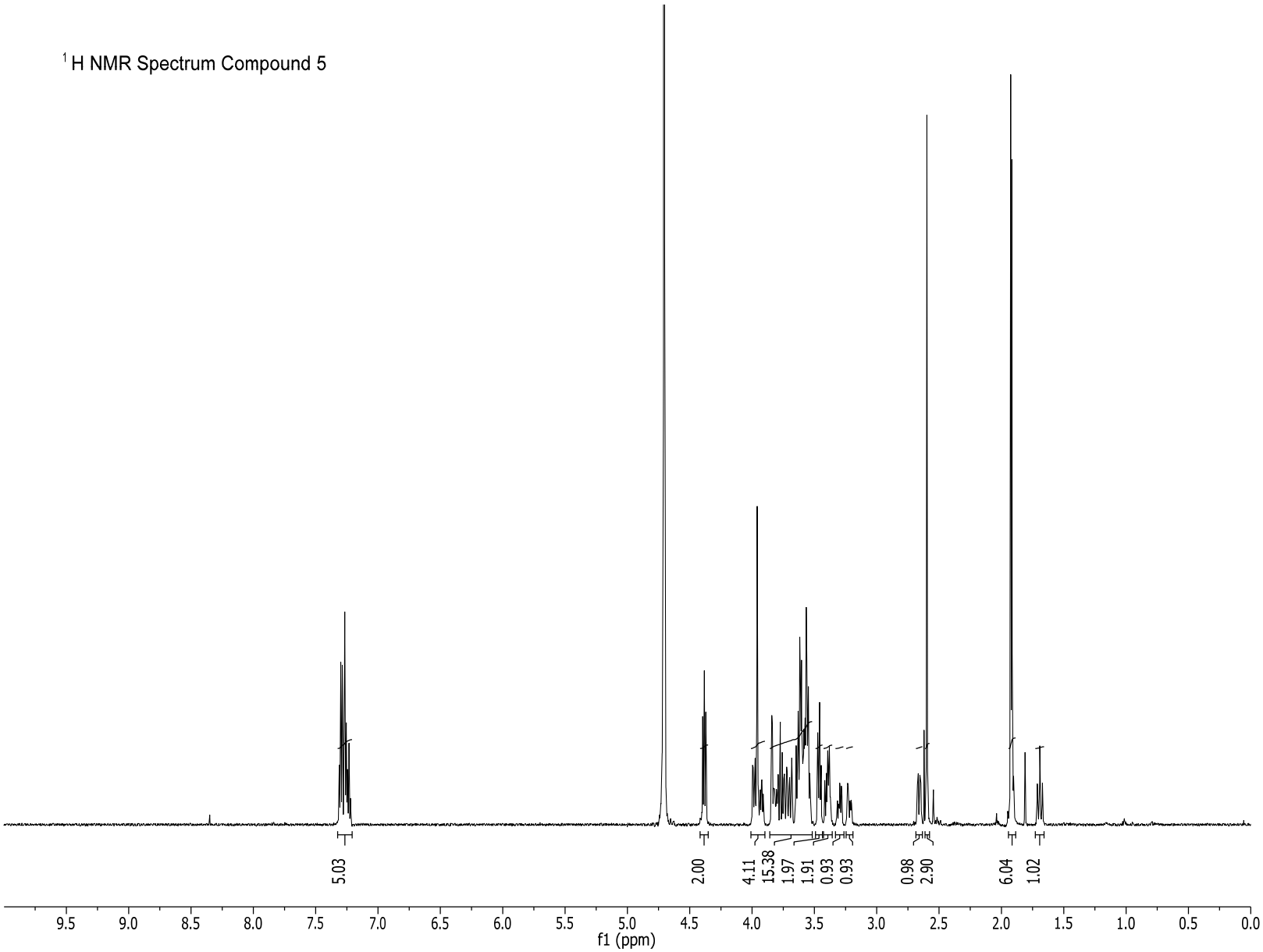
^{1}H NMR Spectrum Compound 4



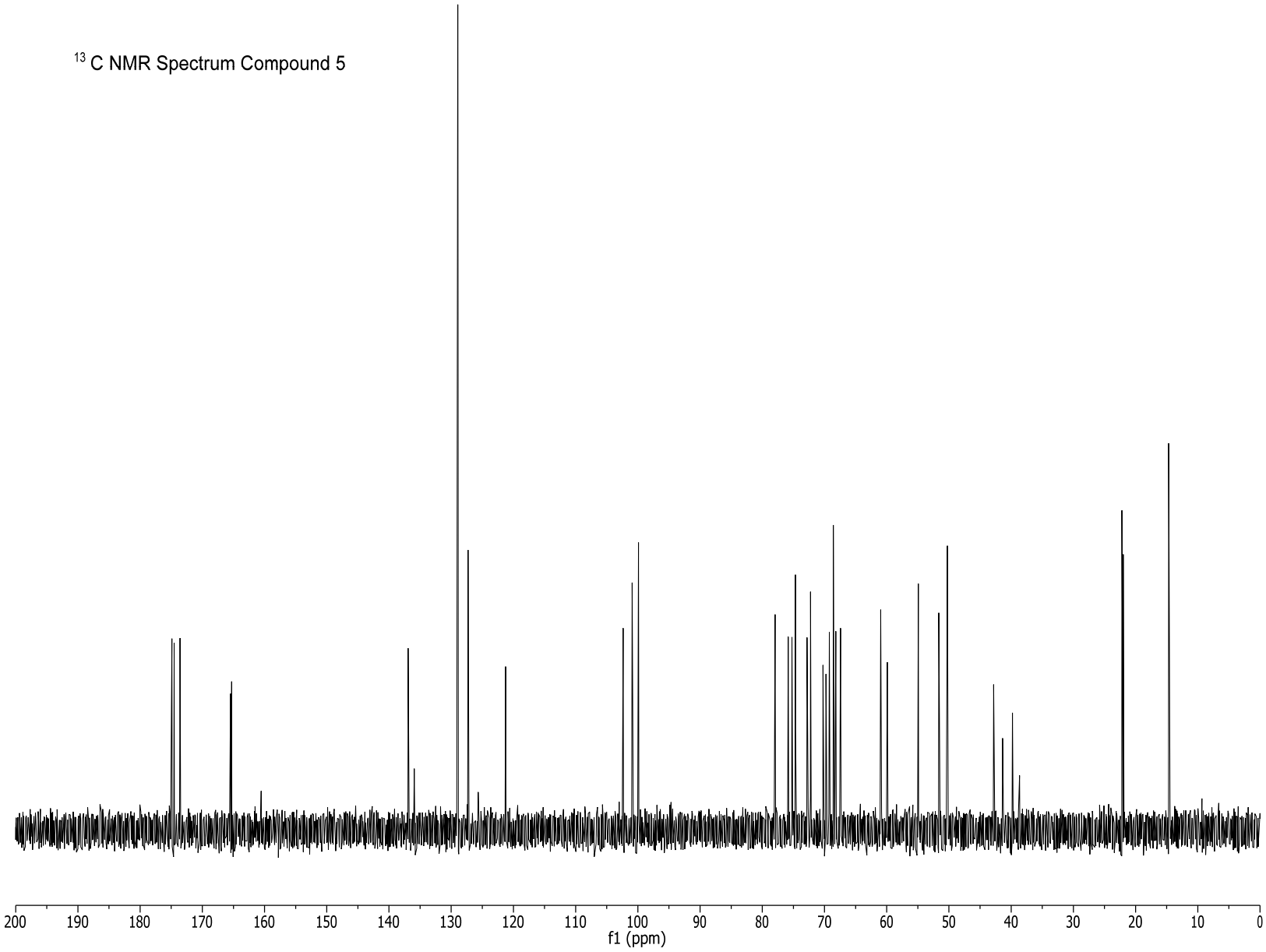
^{13} C NMR Spectrum Compound 4

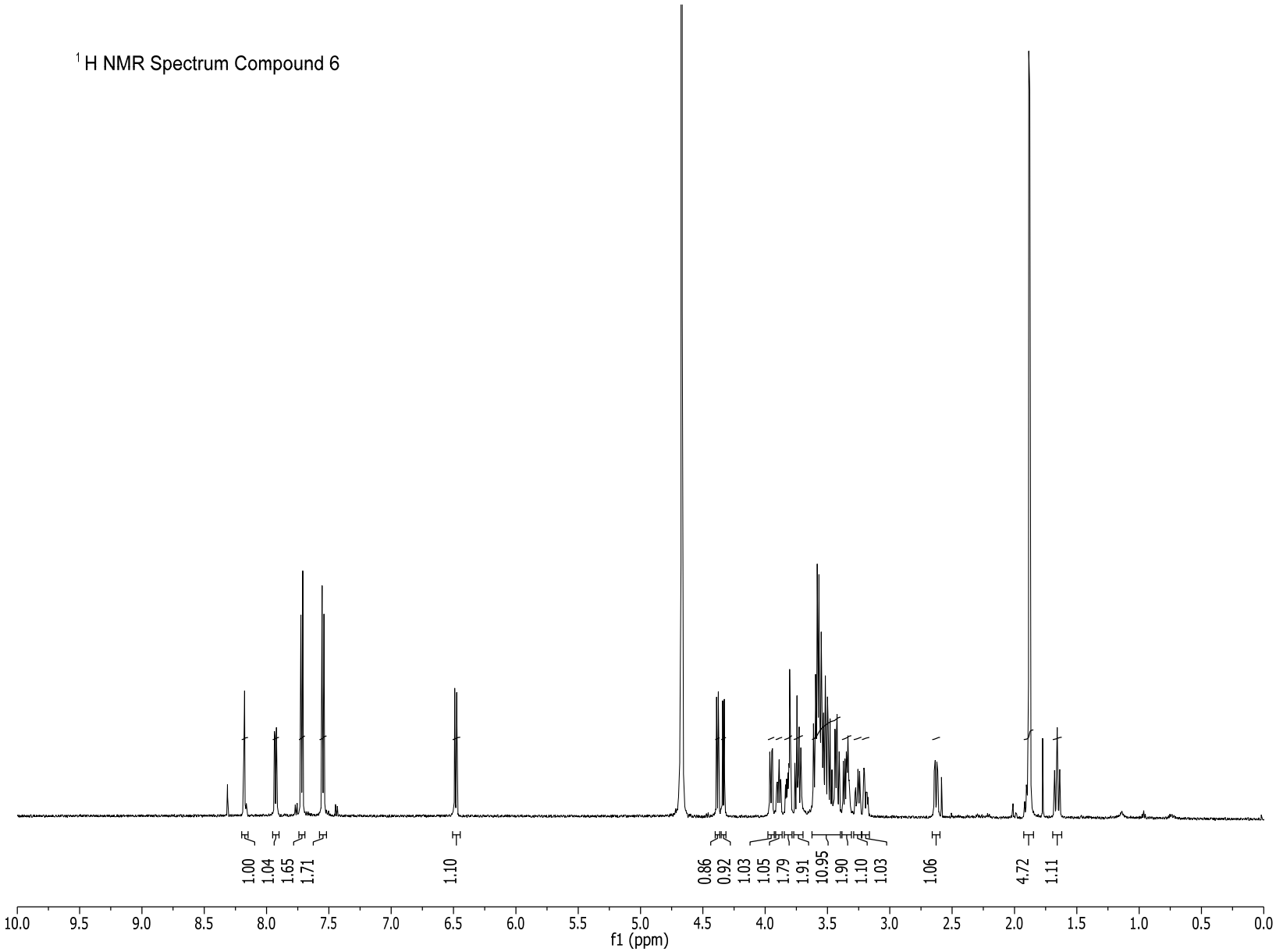


^{1}H NMR Spectrum Compound 5

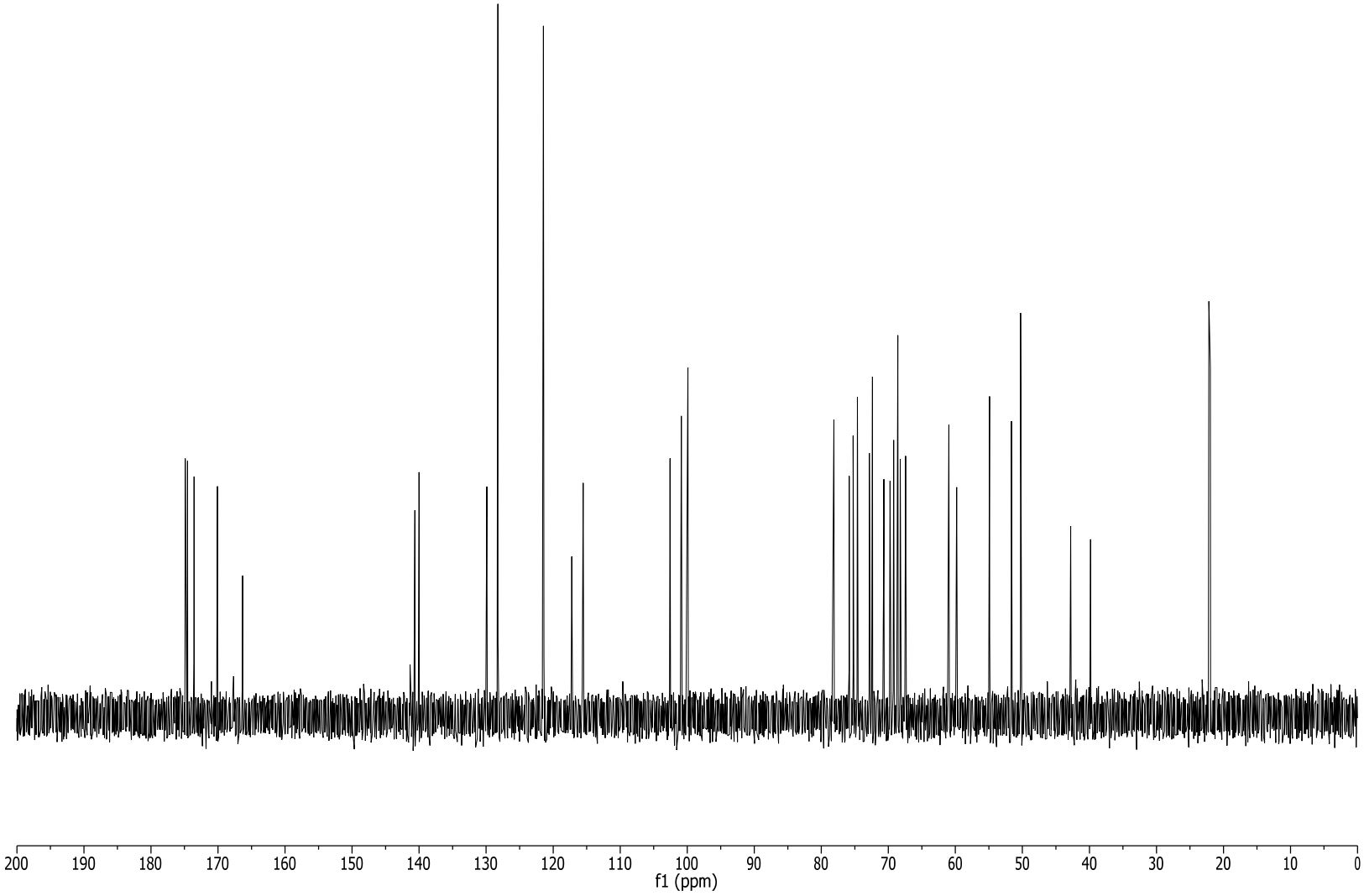


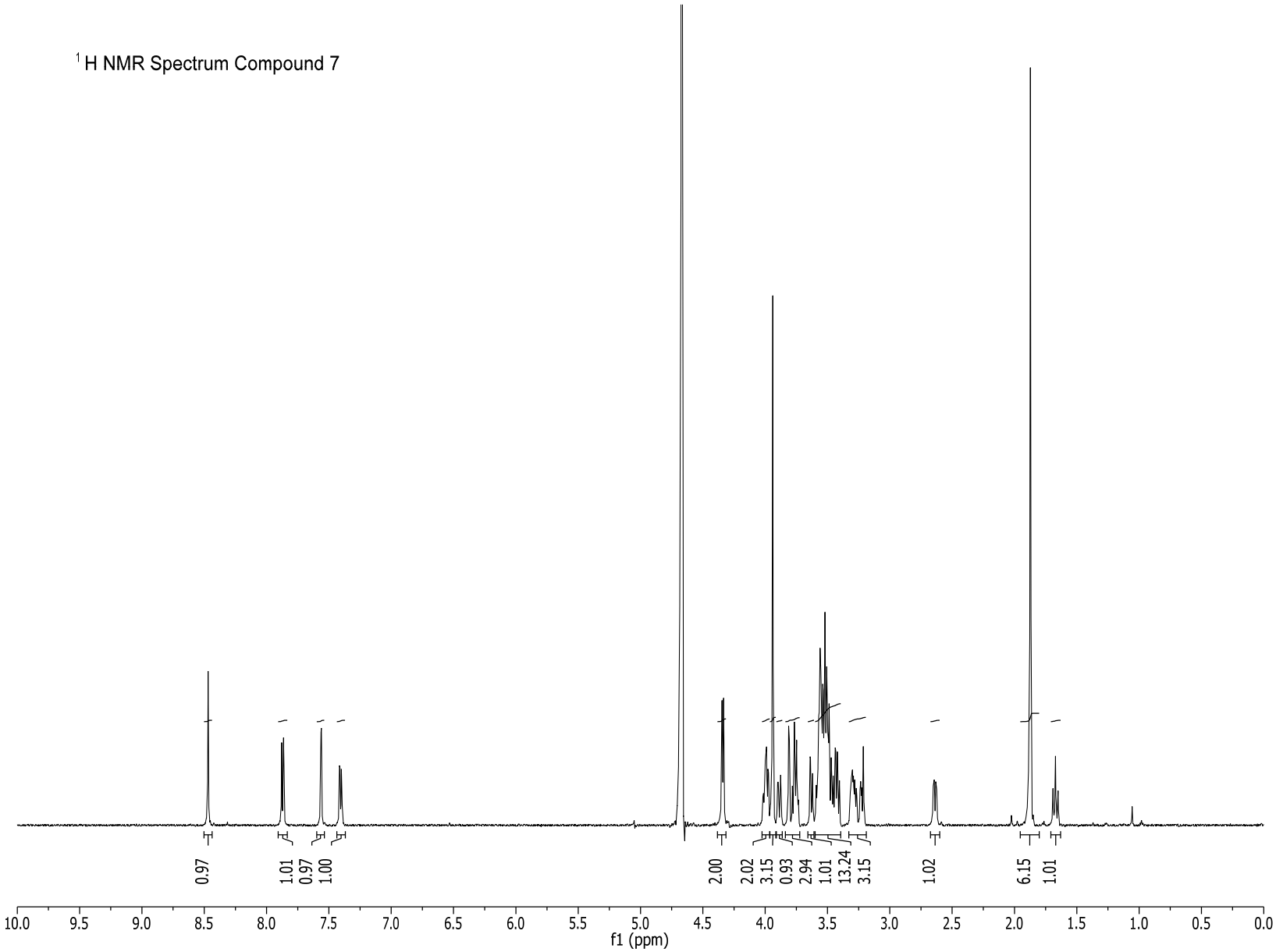
^{13} C NMR Spectrum Compound 5



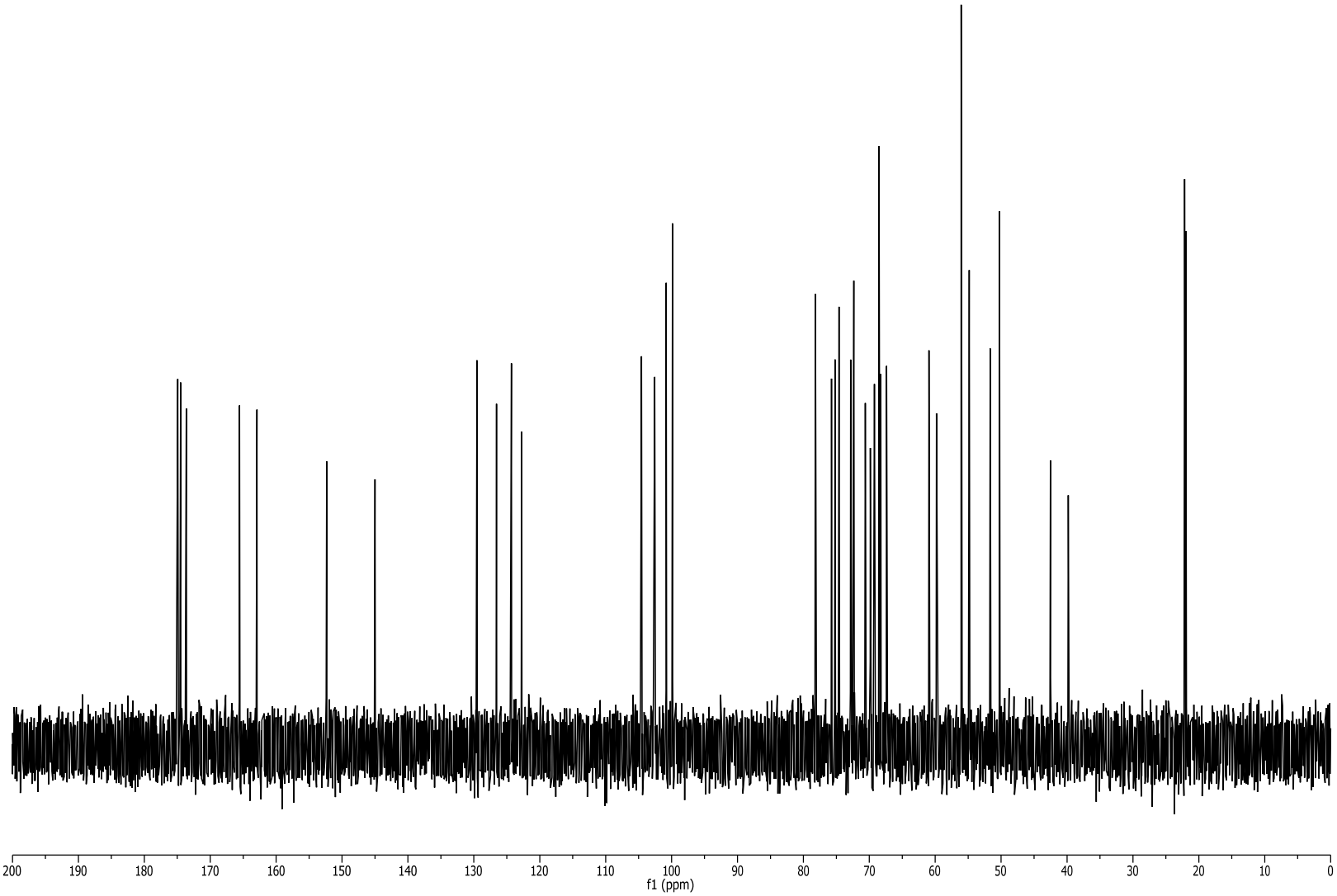

^{1}H NMR Spectrum Compound 6
1.00
1.04
1.65
1.71
1.10
0.86
0.92
1.03
1.05
1.79
1.91
10.95
1.90
1.10
1.03
1.06
4.72
1.11
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

^{13} C NMR Spectrum Compound 6

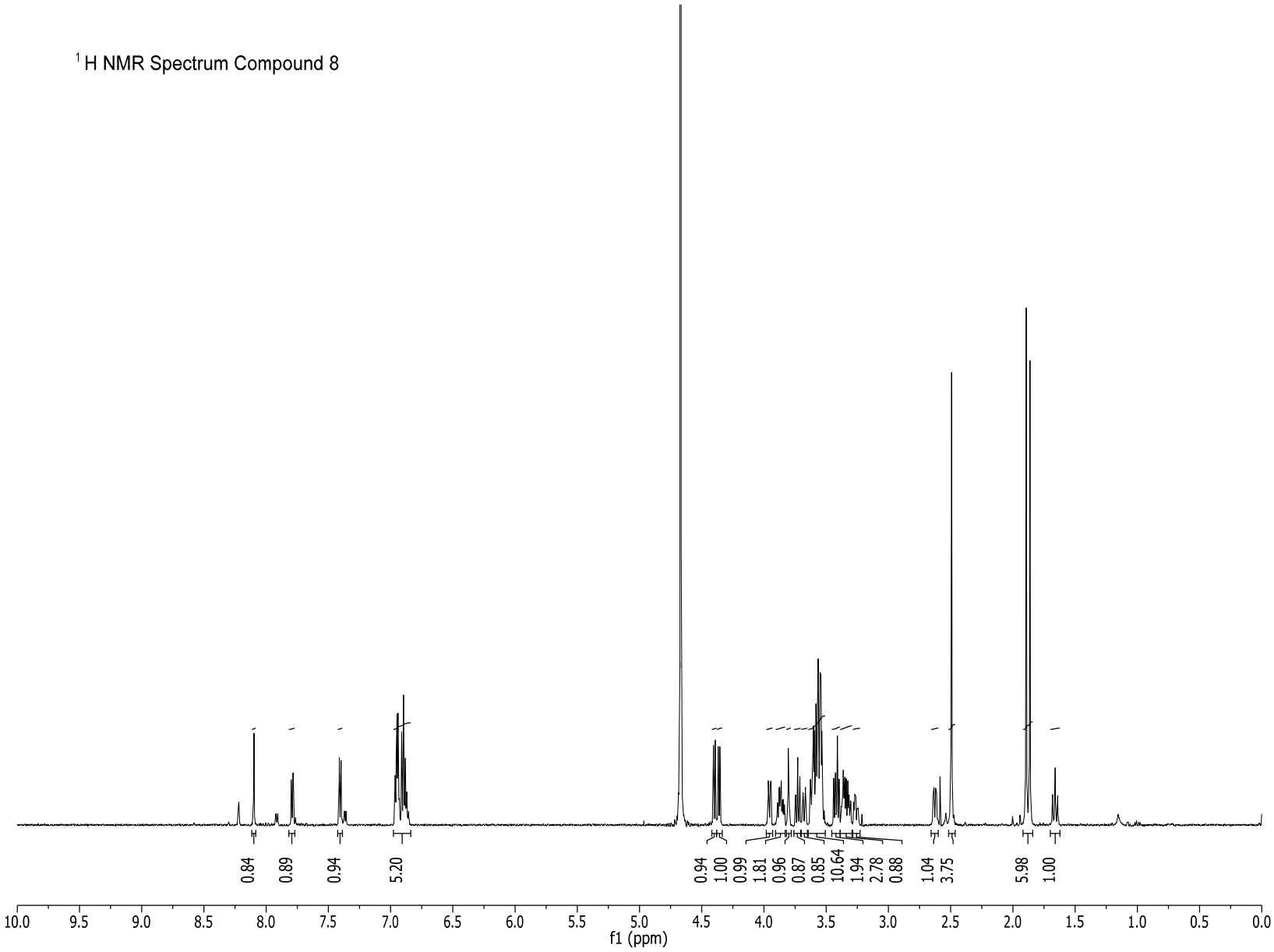



^{1}H NMR Spectrum Compound 7
0.97
1.01
0.97
1.00
2.00
2.02
3.15
0.93
2.94
1.01
13.24
3.15
1.02
6.15
1.01
10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

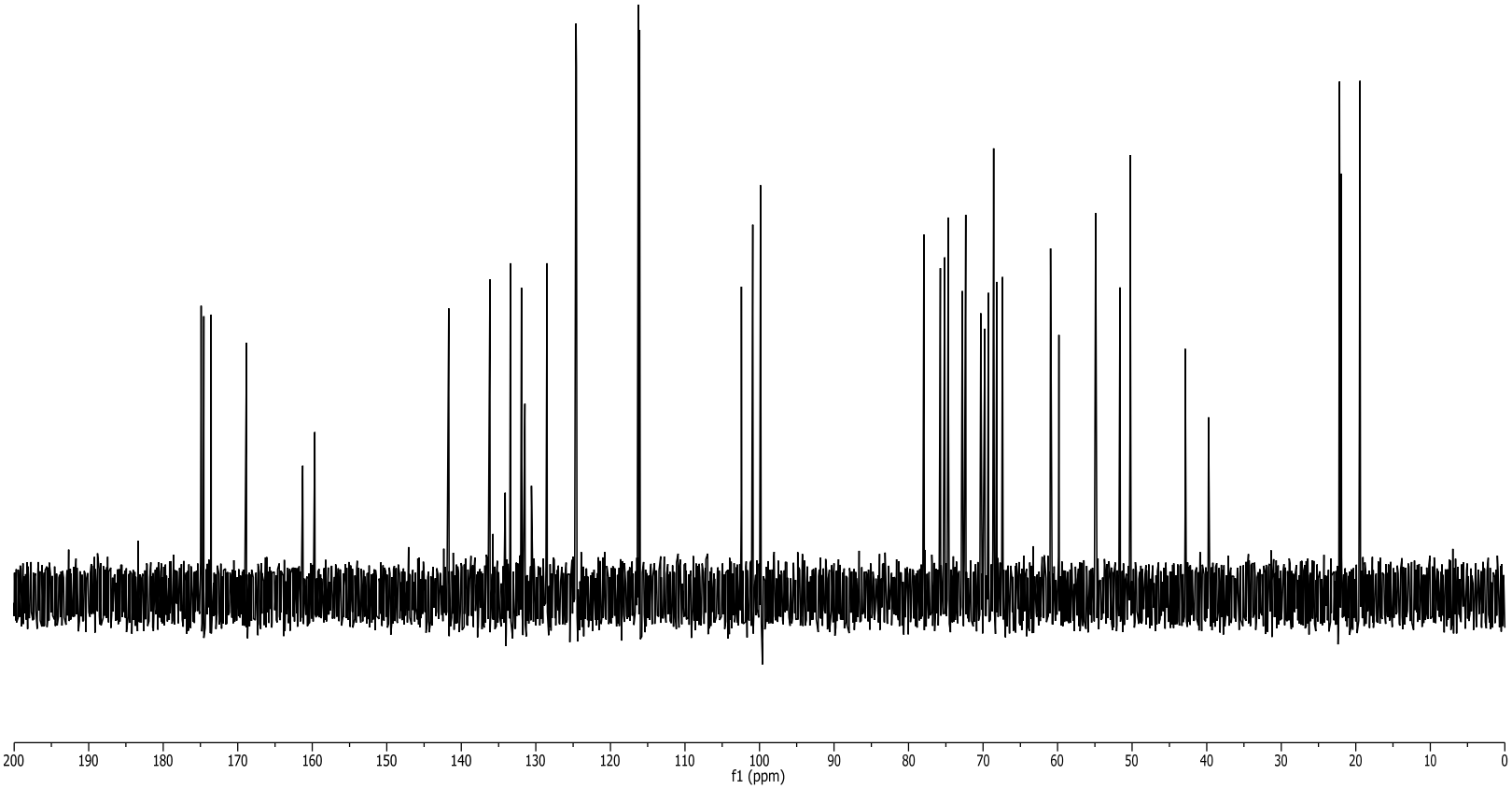
^{13} C NMR Spectrum Compound 7



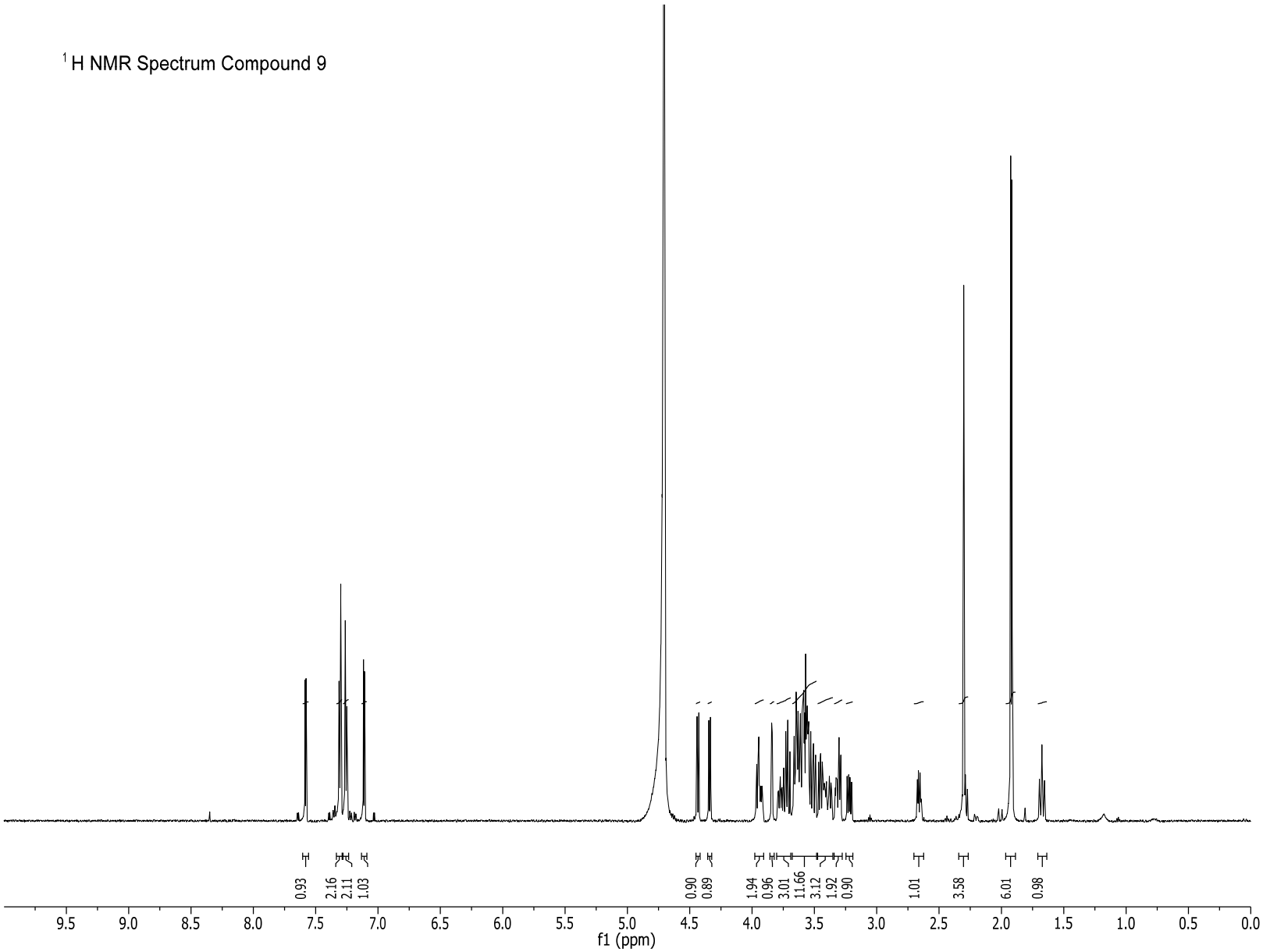
^{1}H NMR Spectrum Compound 8



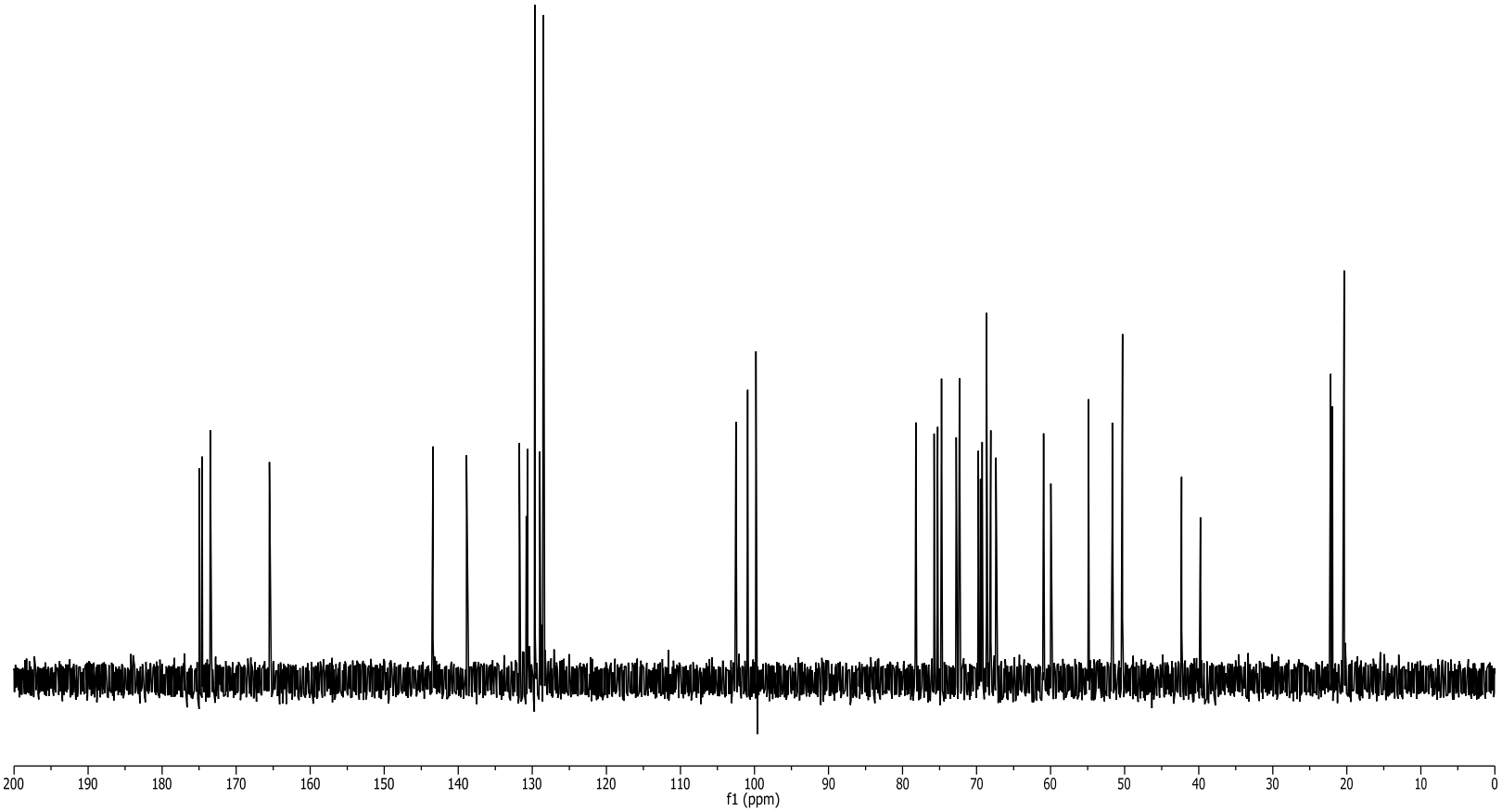
^{13}C NMR Spectrum Compound 8



^{1}H NMR Spectrum Compound 9



^{13} C NMR Spectrum Compound 9



I. **References**

(1) Han, S.; Collins, B. E.; Bengtson, P.; Paulson, J. C. *Nat Chem Biol* **2005**, *1*, 93.

(2) Blixt, O.; Brown, J.; Schur, M. J.; Wakarchuk, W.; Paulson, J. C. *J Org Chem* **2001**, *66*, 2442.

(3) Blixt, O.; Han, S. F.; Liao, L.; Zeng, Y.; Hoffmann, J.; Futakawa, S.; Paulson, J. C. *J Am Chem Soc* **2008**, *130*, 6680.

(4) Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B. Y.; Zhang, J. B.; Zhang, Y. X.; Jia, Q.; Chen, X. *J Am Chem Soc* **2005**, *127*, 17618.

(5) Karwaski, M. F.; Wakarchuk, W. W.; Gilbert, M. *Protein Expres Purif* **2002**, *25*, 237.

(6) Kelm, S.; Pelz, A.; Schauer, R.; Filbin, M. T.; Tang, S.; Debellard, M. E.; Schnaar, R. L.; Mahoney, J. A.; Hartnell, A.; Bradfield, P.; Crocker, P. R. *Curr Biol* **1994**, *4*, 965.

(7) Tateno, H.; Li, H. Y.; Schur, M. J.; Bovin, N.; Crocker, P. R.; Wakarchuk, W. W.; Paulson, J. C. *Mol Cell Biol* **2007**, *27*, 5699.

(8) Chen, W. C.; Completo, G. C.; Sigal, D. S.; Crocker, P. R.; Saven, A.; Paulson, J. C. *Blood* **2010**, *115*, 4778.

(9) Chen, W. H. C.; Kawasaki, N.; Nycholat, C. M.; Han, S. F.; Pilotte, J.; Crocker, P. R.; Paulson, J. C. *Plos One* **2012**, *7*.

(10) Zaccai, N. R.; Maenaka, K.; Maenaka, T.; Crocker, P. R.; Brossmer, R.; Kelm, S.; Jones, E. Y. *Structure* **2003**, *11*, 557.

(11) Labute, P. *Proteins* **2009**, *75*, 187.

(12) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J Comput Chem* **2009**, *30*, 2785.

(13) Forli, S. "Raccoon|AutoDock VS. an automated tool for preparing AutoDock virtual screenings", http://autodock.scripps.edu/resources/raccoon (accessed 02/24/2010)

(14) Jourdian, G. W.; Dean, L.; Roseman, S. *J Biol Chem* **1971**, *246*, 430.