

Design of CID-Cleavable Protein Cross-linkers: Identical Mass Modifications for Simpler Sequence Analysis

Wynne V. Kandur^a, Athit Kao^b, Danielle Vellucci^a, Lan Huang^b, and Scott D.
Rychnovsky^{a*}

^aDepartment of Chemistry, University of California, Irvine, CA 92697

^bDepartments of Physiology & Biophysics, University of California, Irvine, CA 92697

Supporting Information

Contents:

I.	General Methods.....	2
II.	Synthesis of Substrates.....	4
III.	IML 2 MS Data.....	9
IV.	References.....	10
V.	Spectral Data.....	11

General Methods

^1H NMR and ^{13}C NMR spectra were recorded at ambient temperature at 500 MHz and 125 MHz, respectively, on a Bruker DRX500 NMR instrument. ^1H and ^{13}C NMR data is reported as follows: chemical shifts are reported in ppm on a δ scale and referenced to internal tetramethylsilane or residual solvent (TMS: δ 0.00; CHCl_3 : δ 7.27(^1H), 77.23 (^{13}C)), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, qu = quintet, m = multiplet), coupling constants (Hz), and integration. Infrared (IR) spectra were obtained using a FT-IR spectrometer. Accurate mass spectra were acquired on a Waters LCT Premier quadrupole time-of-flight spectrometer and were obtained by peak matching. Gas chromatography/mass spectrometry (GC/MS) was performed with a Thermo-Finnigan Trace Mass Spectrometer Plus quadrupole system with a fused silica capillary column (30 m x 0.32 mm x 0.25 mm) wall-coated with DB-5 (J & W Scientific) using electron ionization (70 eV) or a Waters GCT Premier orthogonal acceleration time-of-flight spectrometer using chemical ionization. Melting points are uncorrected and were obtained using a Büchi 510 melting point apparatus. Analytical thin layer chromatography was performed on EMD Chemicals Inc. silica gel 60 F₂₅₄ plates. Liquid chromatography was performed using forced flow (flash chromatography) of the indicated solvent system on Sorbent Technologies silica gel (SiO_2) 60 (230–400 mesh). Unless otherwise noted, all reactions were carried out under an atmosphere of argon in flame-dried glassware. Solvents were distilled from CaH_2 or filtered through alumina before use.¹

General procedure 1: Ester Hydrolysis

Starting ester was dissolved in 2:1 or 4:1 THF:H₂O, depending on solubility. The mixture was cooled to 0 °C and lithium hydroxide (98%, 5 equiv.) dissolved in minimal

H₂O was added slowly. The reaction was monitored by thin layer chromatography, while stirring vigorously. Upon completion, the mixture was diluted with diethyl ether and water before the layers were separated. Diethyl ether was added to the resulting aqueous layer, followed by addition of HCl to pH 1. The layers were separated and the aqueous layer was extracted with diethyl ether (2x). Organics were combined, washed with brine, dried over Na₂SO₄, filtered and evaporated *in vacuo*.

General procedure 2: TBAF-Promoted Michael Additions

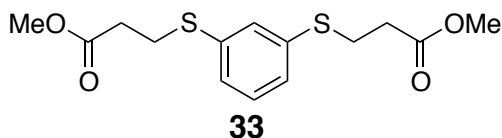
Acrylic acid (0.28 mmol), thiophenol (0.51 mmol), and THF (1 mL) were combined in a round bottom flask to which TBAF•3H₂O (0.06 mmol) was added at rt. The flask was fitted with a cold-water condenser and the mixture was heated to 50° C for 16 h. The mixture was cooled to rt, evaporated *in vacuo*, diluted with EtOAc and washed with 1N HCl (2x). The organic layer was washed with brine, dried over Na₂SO₄, filtered, and evaporated *in vacuo* to yield the crude product. In some cases, this reaction was run neat.²

General procedure 3: NHS Ester Preparation

Crude diacid (0.08 mmol) was dissolved in dry dichloromethane (0.62 mL) in a flame dried round bottom flask under argon. NHS•TFA (0.41 mmol) was added before a slow addition of triethylamine (0.49 mmol) at 0° C. The mixture was left to warm slowly overnight while stirring under argon. After 16 h, the mixture was diluted with dichloromethane and washed with water (2x). The organic layer was dried with Na₂SO₄, filtered, and evaporated *in vacuo*.

General procedure 4: *m*-CPBA-Oxidation to Sulfoxide

Di-NHS ester (0.026 mmol) was dissolved in CDCl_3 (1 mL) and *m*-CPBA (77% w/w, 0.026 mmol) was added slowly while monitoring by LRMS ESI + and ^1H NMR. When this reaction was run on larger scale than 30 mg starting material, the solution was cooled in an ice bath prior to *m*-CPBA addition. Once the reaction was complete, the solution was diluted with dichloromethane (2 mL) and washed with 10% sodium bicarbonate solution (3 x 2 mL). The organic layer was dried over Na_2SO_4 , filtered, and evaporated *in vacuo* to afford the product.

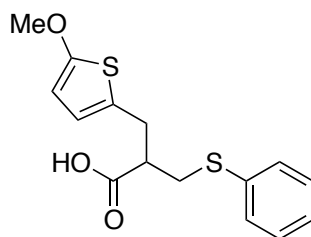


Dimethyl 3,3'-(1,3-phenylenebis(sulfanediyl))dipropionate (33): 1,3-benzenedithiol (0.500 g, 3.52 mmol) and methyl acrylate (0.632 mL, 7.02 mmol) were dissolved in THF (6 mL), and Et_3N (0.980 mL, 7.03 mmol) was added to the solution. After 5 h, the reaction mixture was partitioned between H_2O (10 mL) and EtOAc (10 mL). The EtOAc layer was collected, washed with brine (10 mL), dried over MgSO_4 , filtered, and concentrated under reduced pressure. The crude oil was purified by silica gel chromatography (5:1 hexanes:EtOAc) to afford 0.550 g (50%) of the sulfide as a clear, colorless oil: ^1H (500 MHz, CDCl_3) δ 7.38 (s, 1H), 7.29–7.22 (m, 3H), 3.73 (s, 6H), 3.22 (t, J = 7.5 Hz, 4H), 2.69 (t, J = 7.5 Hz, 4H); ^{13}C (125 MHz, CDCl_3) δ 172.2, 136.6, 130.5, 129.6, 127.0, 52.0, 34.2, 28.8; IR (neat) 2997, 2951, 2846, 1739, 1570 cm^{-1} ; Accurate Mass (ES+/MeOH) m/z calcd for $\text{C}_{14}\text{H}_{18}\text{O}_4\text{S}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 337.0544, found 337.0547.

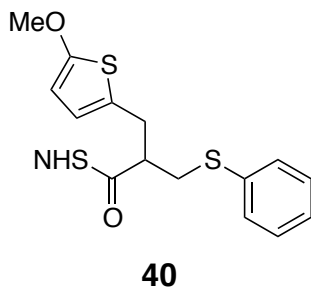
Bis(2,5-dioxopyrrolidin-1-yl) 3,3'-(1,3-phenylenedisulfinyl)dipropionate (2).

Bis(2,5-dioxopyrrolidin-1-yl) 3,3'-(1,3-phenylenebis(sulfanediyl))dipropionate (0.110 g, 0.229 mmol) was dissolved in CHCl_3 (7 mL), and the reaction mixture was cooled to 0°C.

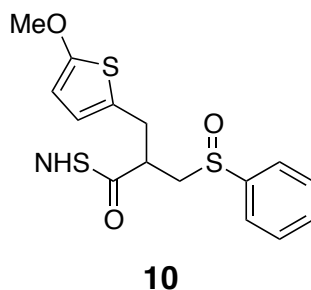
A solution of *m*-CPBA (0.105 g, 0.457 mmol) in CHCl₃ (2 mL) was added drop-wise, and the reaction mixture was stirred for 10 min. The reaction mixture was then partitioned between CHCl₃ (10 mL) and saturated aqueous NaHCO₃ (10 mL). The CHCl₃ layer was collected, dried over MgSO₄, filtered, and concentrated to afford 0.080 g (75%) of sulfoxide **2** as a white solid: ¹H (500 MHz, CDCl₃) δ 7.93–7.90 (m, 1H), 7.80–7.78 (m, 3H), 3.46–3.38 (m, 2H), 3.17–3.10 (m, 4H), 2.82 (brs, 10H); ¹³C (125 MHz, CDCl₃) δ 169.0, 167.0, 144.6, 144.5, 131.0, 127.1, 120.3, 49.7, 49.5, 25.7, 22.9, 22.7; IR (KBr) 2943, 2935, 1786, 1736, 1427 cm⁻¹; Accurate Mass (ES+/MeOH) *m/z* calcd for C₂₀H₂₀N₂₀O₁₀S₂Na [M+Na]⁺ 535.0457, found 535.0457.

**31**

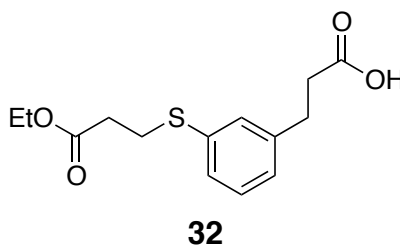
3-(5-methoxythiophen-2-yl)-2-((phenylthio)methyl)propanoic acid (31). Acrylic acid **29** (0.073 g, 0.25 mmol) and thiophenol (0.039 mL, 0.38 mmol) were subjected to general procedure 2 without solvent. The crude product was purified by flash column chromatography (30% EtOAc/hexanes) to yield desired product (0.018 g, 23%). ¹H NMR (500 MHz, CDCl₃) δ 7.36 (d, *J* = 7.9, 2H), 7.29 (t, *J* = 7.6, 2H), 7.22 (t, *J* = 7.6, 1H), 6.42 (d, *J* = 3.6, 1H), 5.99 (d, *J* = 3.8, 1H), 3.84 (s, 3H), 3.23 (dd, *J* = 7.7, 13.7, 1H), 3.16–3.04 (m, 3H), 2.88 (quintet, *J* = 6.8, 1H); ¹³C NMR (125 MHz, CDCl₃) δ 179.2, 165.3, 135.2, 130.5, 129.3, 127.0, 126.0, 123.8, 103.4, 60.4, 47.0, 34.6, 31.6; cm⁻¹; Accurate Mass ES- *m/z* calcd for C₁₅H₁₅O₃S₂ (M-H)⁻ 307.0463, found 307.0466.



NHS ester 40. Carboxylic acid **31** (0.018 g, 0.058 mmol) was combined with dry DMF (0.3 mL), NHS (0.013 g, 0.12 mmol), pyridine (0.019 mL, 0.23 mmol), and, lastly, trifluoroacetic anhydride (0.016 mL, 0.12 mmol). After 2.5 h stirring at rt, the mixture was diluted with dichloromethane and 1M HCl. After the layers were separated, the organic layer was washed with 1 M HCl (2x) and dilute NaHCO₃ solution (2x). The organic layer was dried with Na₂SO₄, filtered, and evaporated *in vacuo*. This procedure is adapted from a published procedure.³ The crude product was purified by flash column chromatography (2% EtOAc/CH₂Cl₂) to yield desired product (0.015 g, 63%). ¹H NMR (500 MHz, CDCl₃) δ 7.26-7.20 (m, 3H), 7.07 (d, *J* = 6.5, 1H), 4.15 (q, *J* = 7.2, 2H), 3.17 (t, *J* = 7.4, 2H), 3.04 (t, *J* = 7.5, 2H), 2.92 (t, *J* = 7.7, 2H), 2.85 (s, 4H), 2.62 (t, *J* = 7.4, 2H), 1.26 (t, *J* = 7.1, 3H); ¹³C NMR (125 MHz, CDCl₃) δ 172.0, 169.3, 168.0, 140.2, 136.0, 129.9, 129.5, 128.4, 126.7, 60.9, 34.6, 32.7, 30.5, 29.1, 25.8, 14.4; IR (liquid) 2928, 1814, 1784, 1737, 1591 cm⁻¹; Accurate Mass ES+ *m/z* calcd for C₁₉H₁₉O₅S₂NNa (M+Na)⁺ 428.0602, found 428.0591.

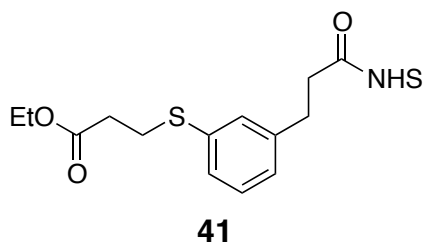


Sulfoxide 10. NHS ester **40** (0.012 g, 0.03 mmol) was subjected to general procedure 4. The crude product was purified by flash column chromatography (10–20% EtOAc/CH₂Cl₂) to yield desired product as a mixture of diastereomers (0.003 g, 26%). ¹H NMR (500 MHz, CDCl₃) δ 7.68–7.63 (m, 2H), 7.53 (m, 3H), 6.61 (d, *J* = 3.8, 0.6H), 6.49 (d, *J* = 3.7, 0.5H), 6.04 (d, *J* = 3.8, 0.6H), 5.98 (d, *J* = 3.8, 0.5H), 3.86 (s, 1.8H), 3.84 (s, 1.3H), 3.60–3.53 (m, 0.5H), 3.47–3.40 (m, 0.6H), 3.39 (s, 0.6H), 3.38 (s, 0.5H), 3.30–3.24 (m, 1.5H), 3.10 (dd, *J* = 7.2, 15.3, 0.5H), 3.03 (dd, *J* = 8.0, 13.2, 0.6H), 2.96 (dd, *J* = 4.0, 13.4, 0.5H), 2.87 (s, 2.1H), 2.84 (s, 2.3H); ¹³C NMR (125 MHz, CDCl₃) δ 168.7, 168.6, 168.3, 165.7, 165.6, 143.5, 143.0, 131.6, 131.4, 129.6, 129.5, 125.3, 124.8, 124.2, 124.0, 123.5, 123.2, 103.6, 103.5, 60.2, 60.2, 57.8, 56.6, 39.5, 39.0, 32.5, 31.0, 29.8, 25.7, 25.6; IR (thin film) 2924, 1810, 1782, 1739, 1507, cm⁻¹; Accurate Mass ES+ *m/z* calcd for C₁₉H₁₉NO₆S₂Na (M+Na)⁺ 444.0551, found 444.0548.

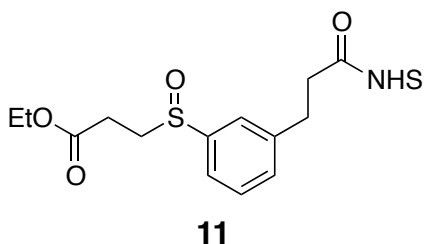


3-(3-((3-Ethoxy-3-oxopropyl)thio)phenyl)propanoic acid (32). Carboxylic acid **15** (0.10 g, 0.55 mmol) was combined with ethyl acrylate (0.088 mL, 0.82 mmol) in ethanol (5.5 mL) in a scintillation vial. Triethylamine (0.15 mL, 1.1 mmol) was added and the mixture was capped and stirred 14 days. The solvent was removed and the crude product was purified by flash column chromatography (50% EtOAc/hexanes) to yield desired product (0.022 g, 14%). ¹H NMR (500 MHz, CDCl₃) δ 7.23 (br s, 3H), 7.07 (d, *J* = 6.7, 1H), 4.16 (q, *J* = 7.2, 2H), 3.17 (t, *J* = 7.3, 2H), 2.95 (t, *J* = 7.6, 2H), 2.68 (t, *J* = 7.6, 2H), 2.61 (t, *J* = 7.5, 2H), 1.25 (t, *J* = 7.2, 3H); ¹³C NMR (125 MHz, CDCl₃) δ 177.7, 172.2, 141.3, 135.6, 130.0, 129.4, 128.2, 126.8, 61.1, 35.5, 34.7, 30.7, 29.2, 14.4; IR

(thin film) 3200, 2982, 2661, 2361, 1732 cm^{-1} ; Accurate Mass ES- m/z calcd for $\text{C}_{14}\text{H}_{17}\text{O}_4\text{S}$ ($\text{M}-\text{H}$) $^-$ 281.0847, found 281.0844.



NHS ester 41. Carboxylic acid **32** (0.022 g, 0.078 mmol) was combined with dry DMF (0.4 mL), NHS (0.018 g, 0.16 mmol), pyridine (0.025 mL, 0.31 mmol), and, lastly, trifluoroacetic anhydride (0.022 mL, 0.16 mmol). After 2.5 h stirring at rt, the mixture was diluted with dichloromethane and 1M HCl. After the layers were separated, the organic layer was washed with 1 M HCl (2x) and dilute NaHCO_3 solution (2x). The organic layer was dried with Na_2SO_4 , filtered, and evaporated *in vacuo*. This procedure is adapted from a published procedure.³ The crude product was purified by flash column chromatography (10% EtOAc/ CH_2Cl_2) to yield desired product (0.009 g, 30%). ^1H NMR (500 MHz, CDCl_3) δ 7.40 (d, J = 7.4, 2H), 7.31 (t, J = 7.5, 2H), 7.26 (t, J = 7.3, 1H), 6.50 (d, J = 3.7, 1H), 6.00 (d, J = 3.7, 1H), 3.85 (s, 3H), 3.33–3.05 (m, 5H), 2.84 (s, 4H); ^{13}C NMR (125 MHz, CDCl_3) δ 169.1, 169.0, 165.5, 134.6, 131.1, 129.4, 127.3, 124.7, 124.6, 103.6, 60.4, 44.9, 34.7, 31.4, 25.8; IR (liquid) 3057, 2944, 1810, 1737, 1562 cm^{-1} ; Accurate Mass ES+ m/z calcd for $\text{C}_{18}\text{H}_{21}\text{O}_6\text{SNa}$ ($\text{M}+\text{Na}$) $^+$ 402.0987, found 402.0979.



Sulfoxide 11. NHS ester **41** (0.007 g, 0.02 mmol) was subjected to general procedure 4 to yield desired product (0.005 g, 60%). ^1H NMR (500 MHz, CDCl_3) δ 7.54–7.46 (m,

3H), 7.39 (d, $J = 3.8$, 1H), 4.18–4.07 (m, 2H), 3.28–3.20 (m, 1H), 3.15 (t, $J = 7.5$, 2H), 3.00–2.93 (m, 3H), 2.90–2.79 (m, 5H), 2.56–2.45 (m, 1H), 1.25 (t, $J = 7.0$, 3H); ^{13}C NMR (125 MHz, CDCl_3) δ 171.5, 169.2, 167.8, 143.7, 140.8, 131.4, 129.9, 124.0, 122.8, 61.3, 51.3, 32.6, 30.6, 26.4, 25.8, 14.4; IR (thin film) 2930, 2360, 1813, 1783, 1737 cm^{-1} ; Accurate Mass ES+ m/z calcd for $\text{C}_{18}\text{H}_{21}\text{NO}_7\text{SNa}$ ($\text{M}+\text{Na}$) $^+$ 418.0937, found 418.0917.

IML 2 MS Data:

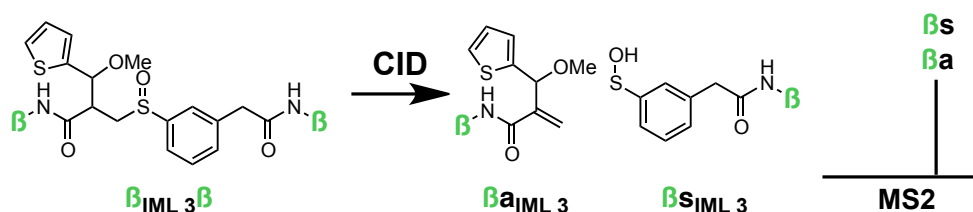


Figure 1. Expected fragmentation of IML 3 in single peptide crosslinking

Initial testing of IML 3 with synthetic peptide Ac-IR7 in the MS instrument showed significant loss of the methoxy group during CID (Figure 2). The MS^2 spectrum shows three peaks representing the IML 3-modified peptides. The first peak, m/z 496.73, represents a doubly charged IML 3-modified peptide that has lost a methoxy substituent during CID fragmentation (Figure 2, top left). This peak could only represent the alkene half of the linker, since it contains a labile methoxy in the benzylic position. When this peak was subjected to further CID, sequence data was acquired (Figure 2, top right). Similarly, the MS^2 peak with m/z 1024.48 represents a singly charged ion with the expected half IML modification. This peak could represent either the sulfenic acid-modified or the alkene-modified peptide, or it could represent a mixture of the two. Further CID of the m/z 1024.48 peak results in sequence data acquisition as well (Figure 2, lower right). The third peak in MS^2 , with m/z 595.75, represents a dead end-modified peptide that has lost the methoxy substituent during the first CID and has not undergone sulfoxide elimination. This dead end could be linked to the peptide on either end, so it is assumed the peak represents a mixture of both possible peptide linkages. When the m/z

595.75 peak is subjected to CID, the result is the expected products of sulfoxide fragmentation of the two differently-linked dead ends: one peak is the sulfenic acid modified peptide (1^+ with m/z 1024.48), and one is the alkene-modified peptide without a methoxy substituent (2^+ with m/z 496.73) (Figure 2, lower left). This data strongly supports the notion that the m/z 595.75 peak represents a mixture of both dead ends. With sulfoxide fragmentation occurring between MS^2 and MS^3 , sequence data could not be gathered for the m/z 595.75 peptides.

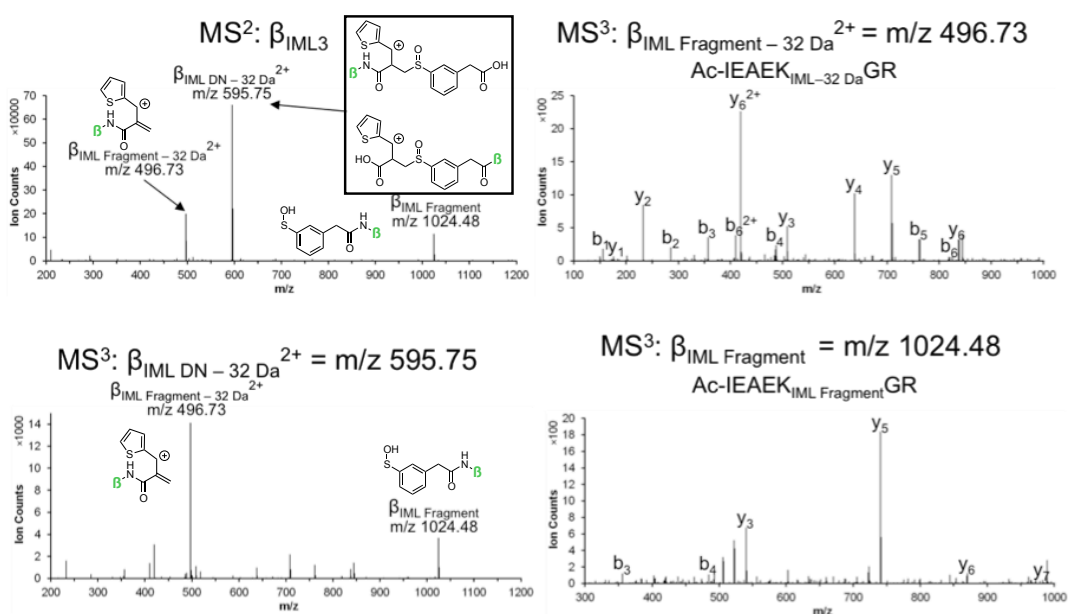
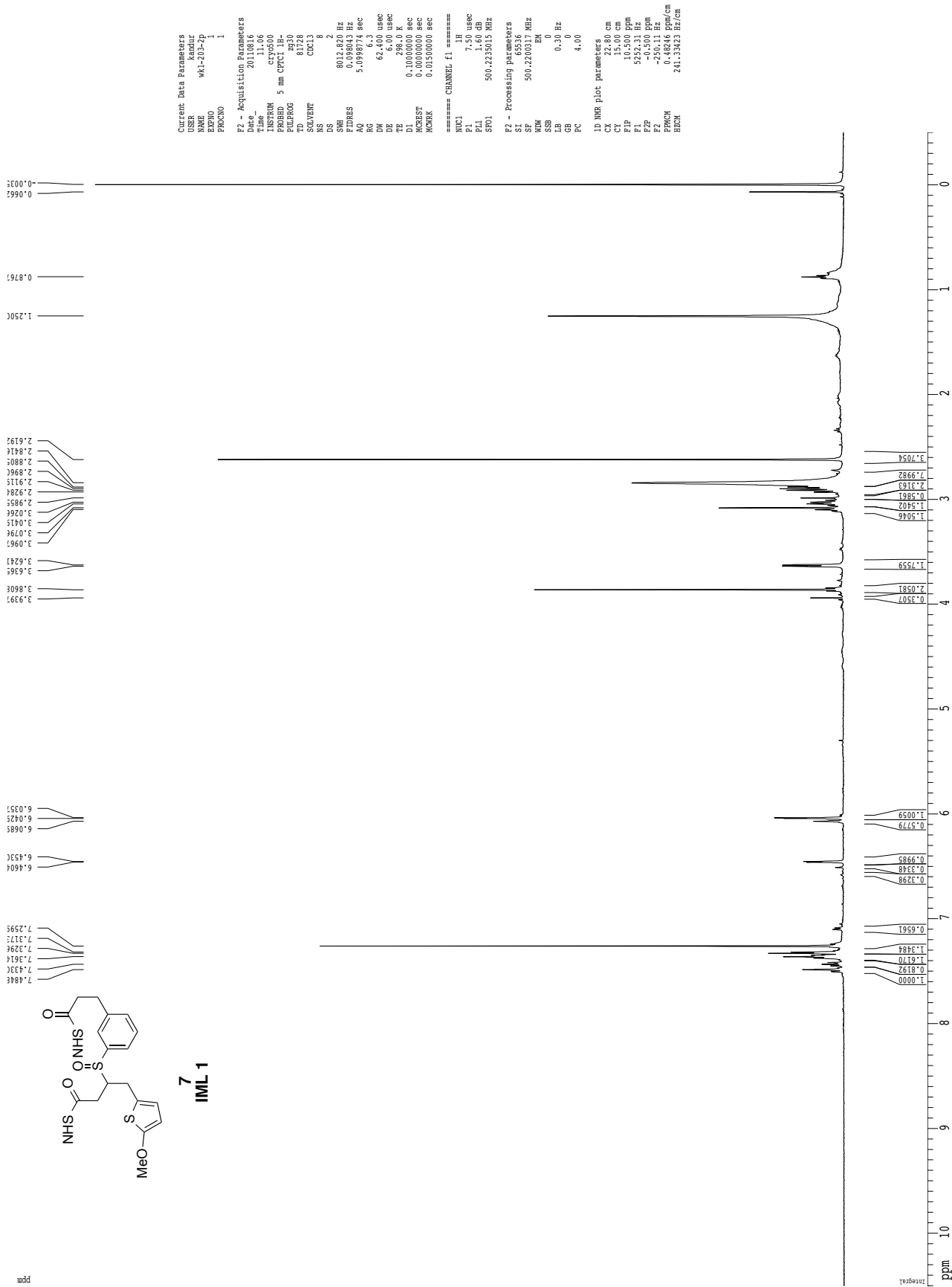


Figure 2. IML 2 Methoxy-loss in MS

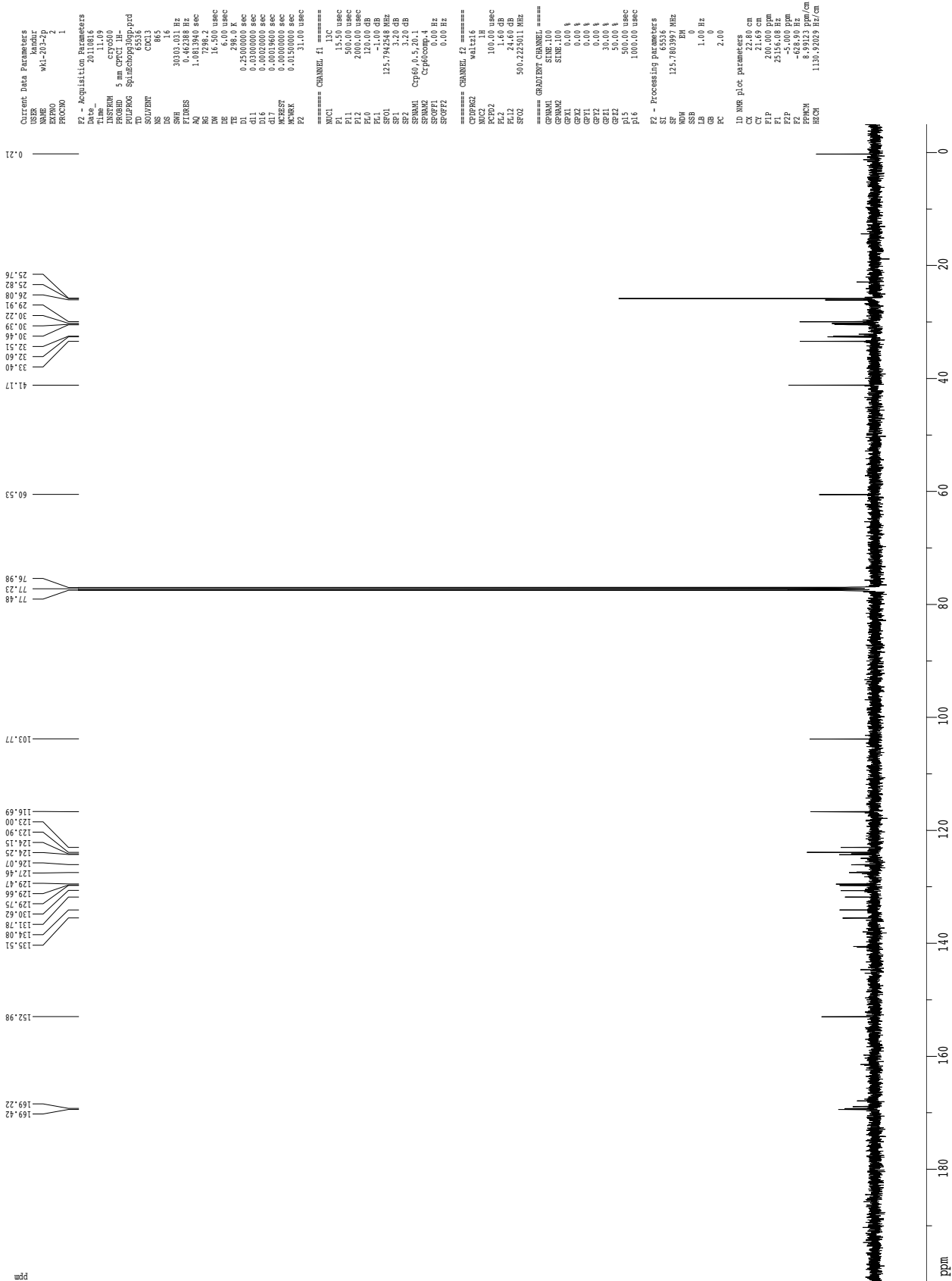
References:

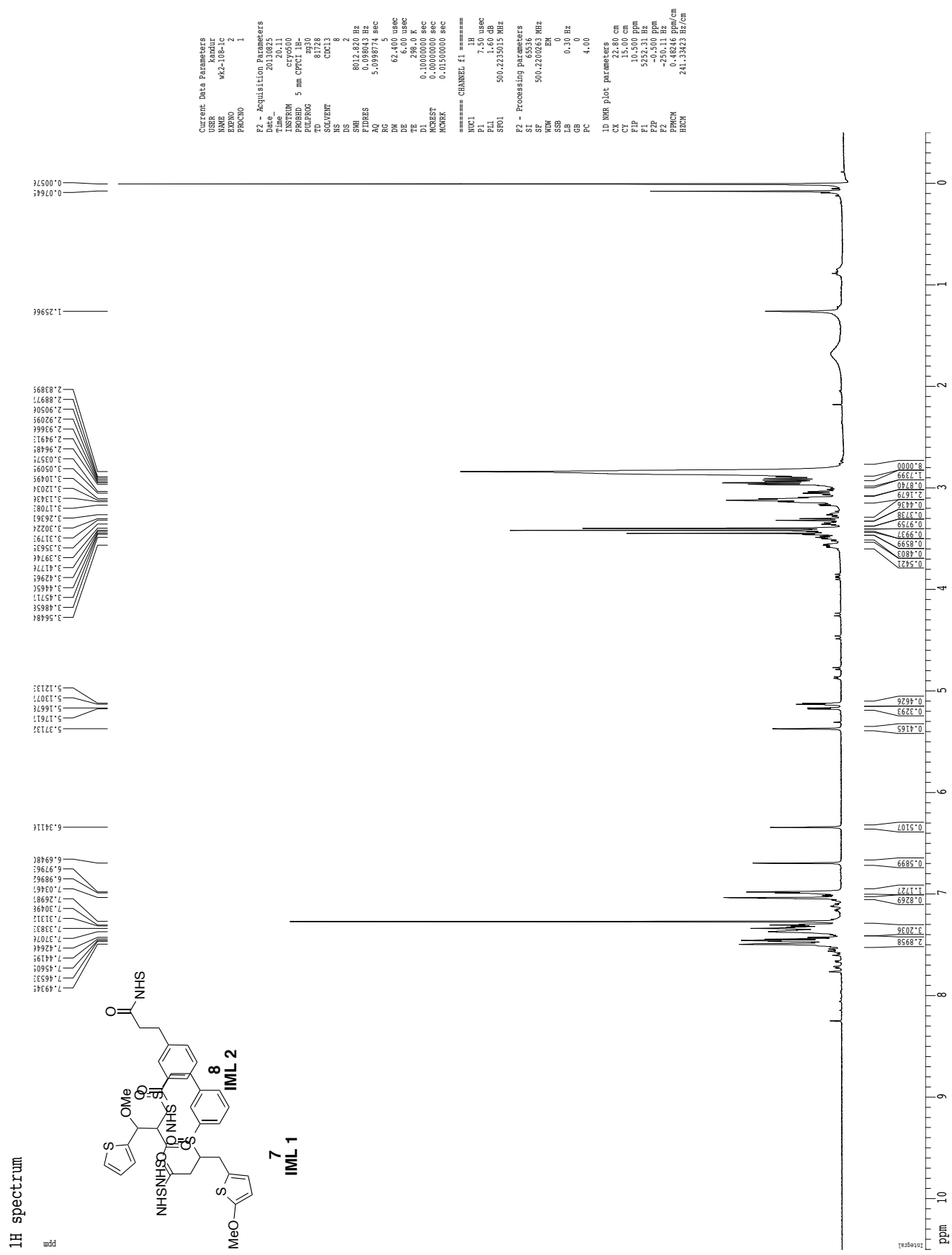
1. A. B. Pangborn, M. A. Giardello, R. H. Grubbs, R. K. Rosen, and F. J. Timmers, *Organometallics*, 1996, **15**, 1518–1520.
2. S. Gao, C. Tseng, C. H. Tsai, and C.-F. Yao, *Tetrahedron*, 2008, **64**, 1955–1961.
3. N. M. Leonard and J. Brunckova, *J. Org. Chem.*, 2011, **76**, 9169–9174.

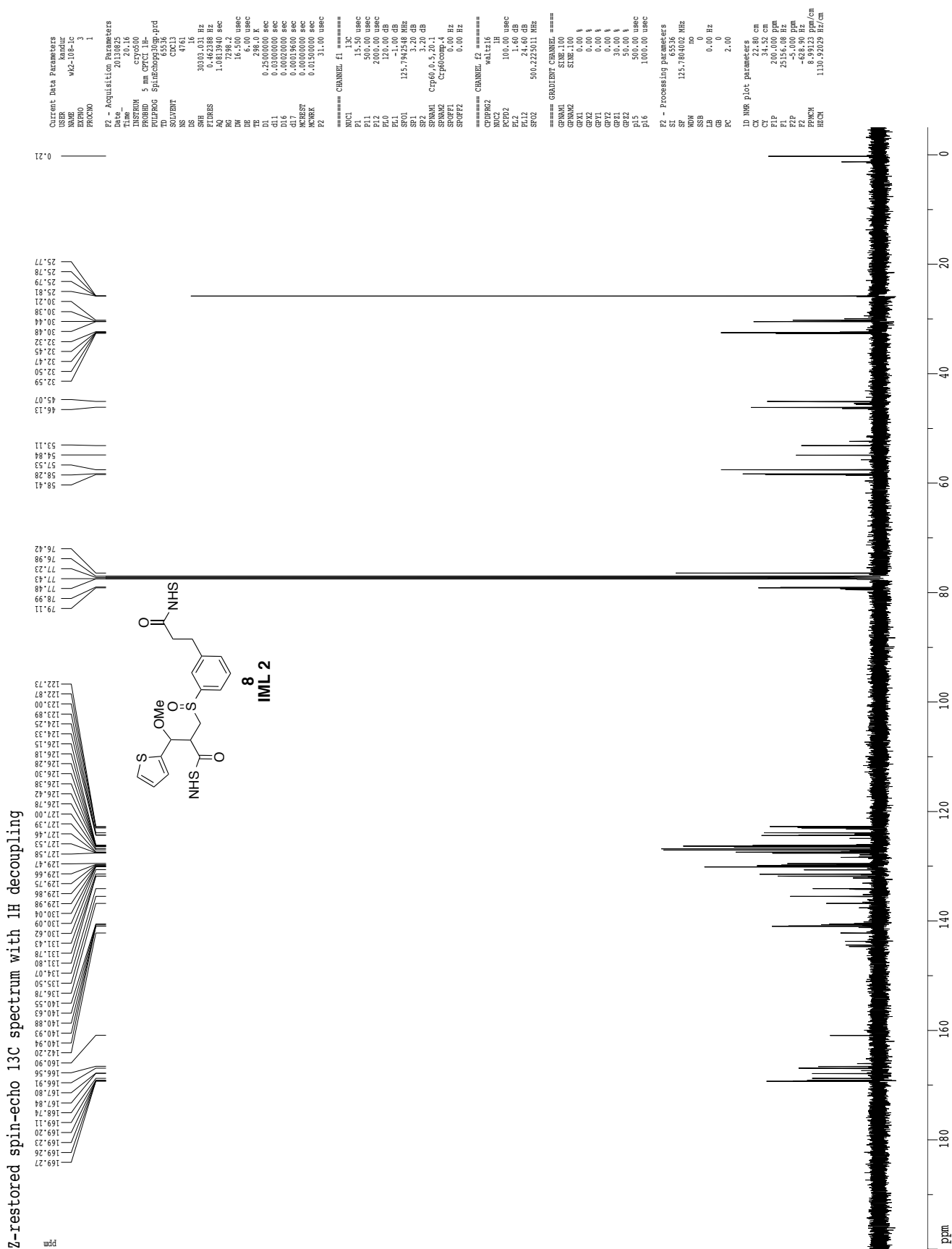
1H spectrum

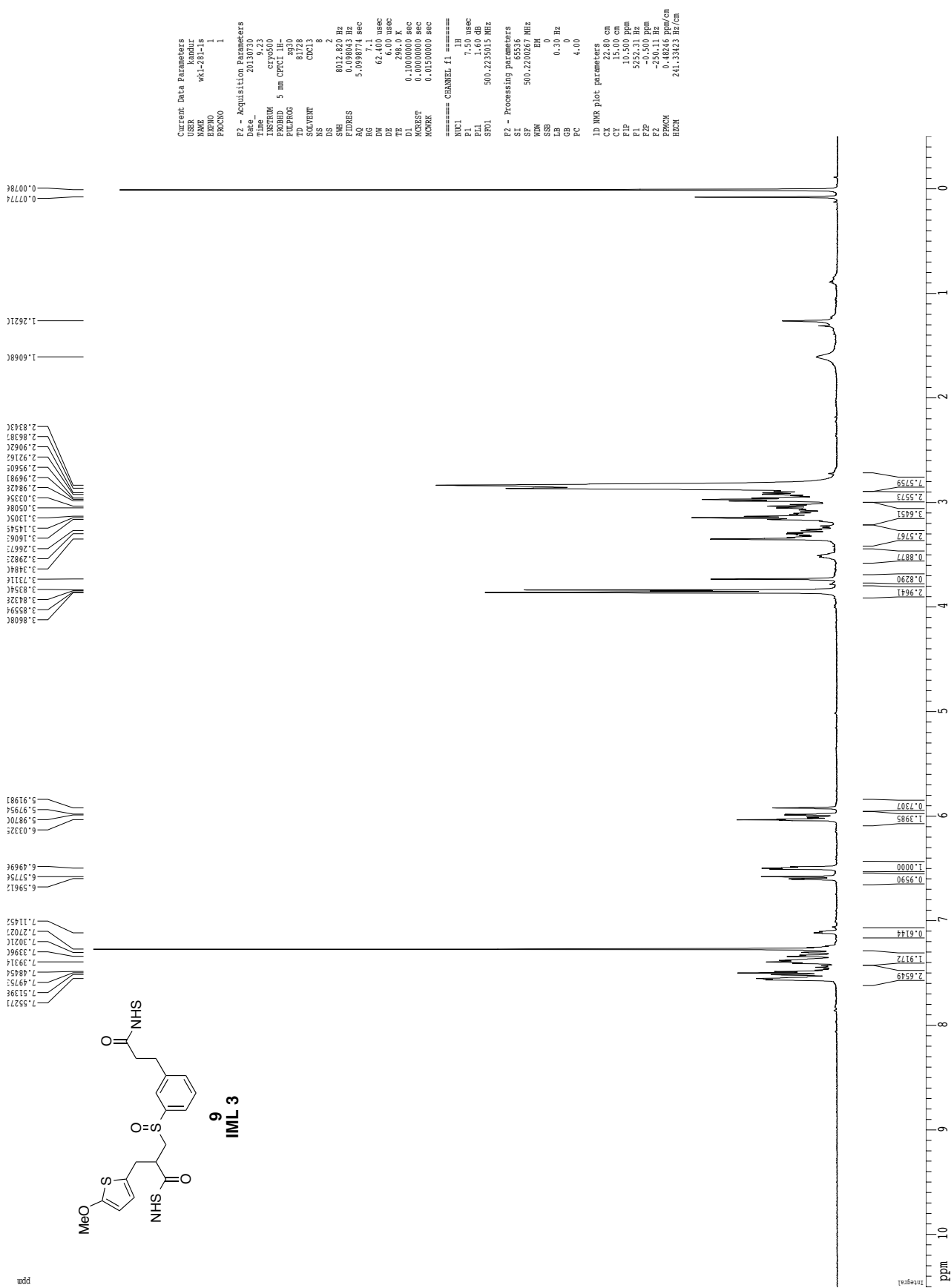
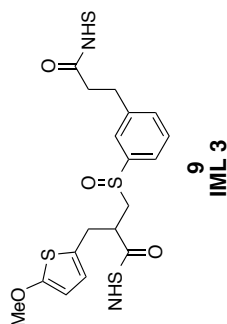


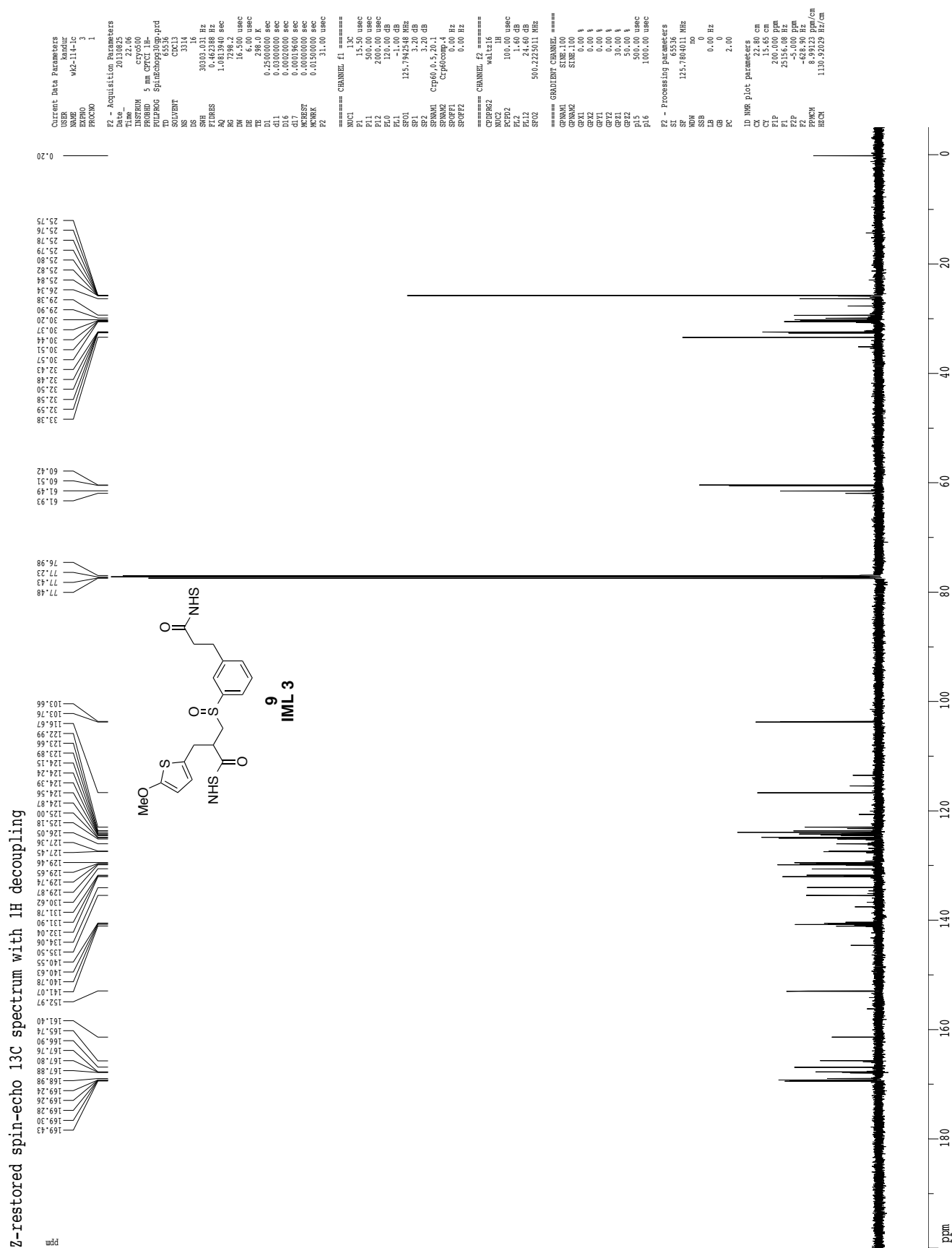
Z-restored spin-echo 13C spectrum with 1H decoupling



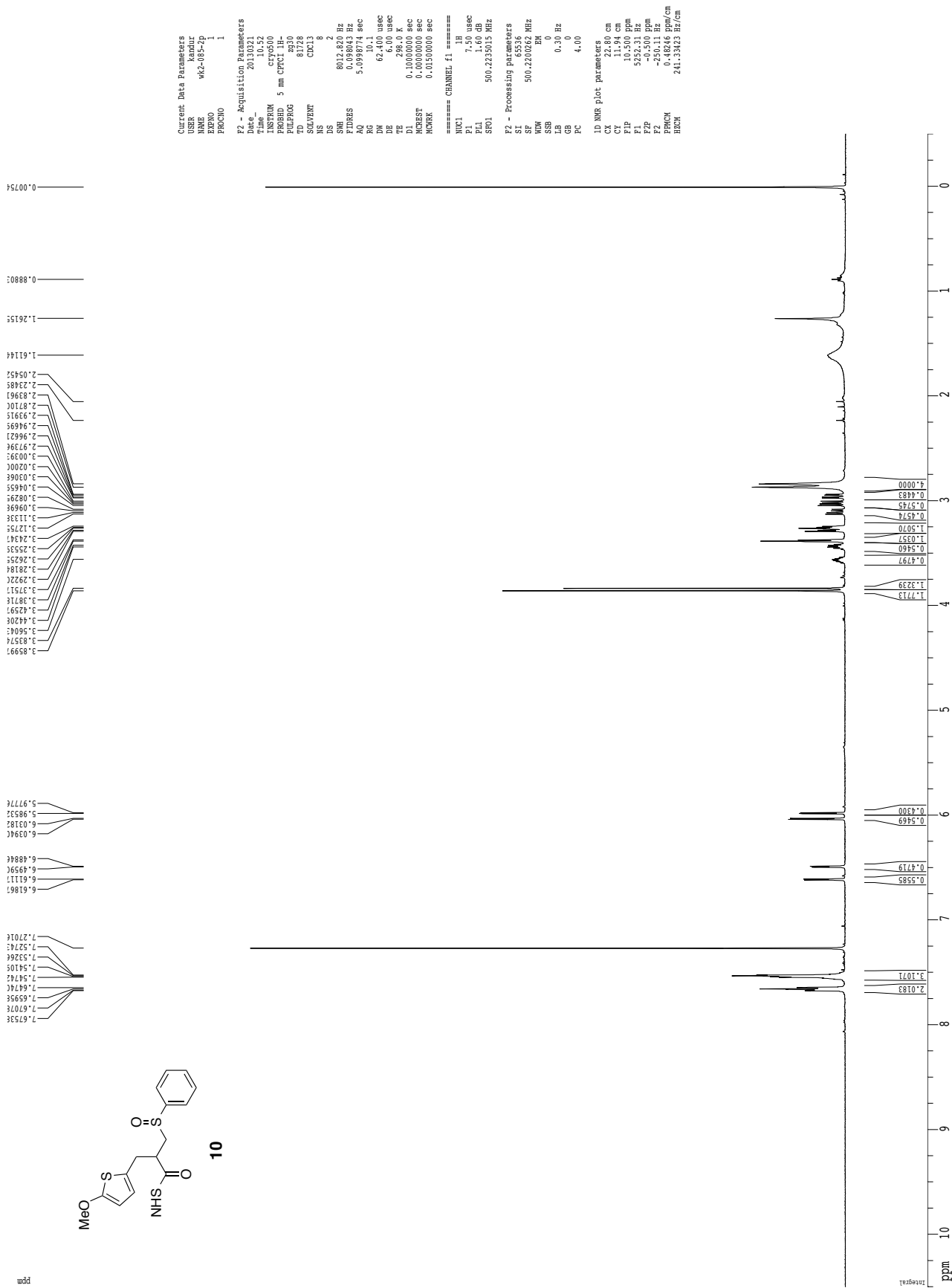




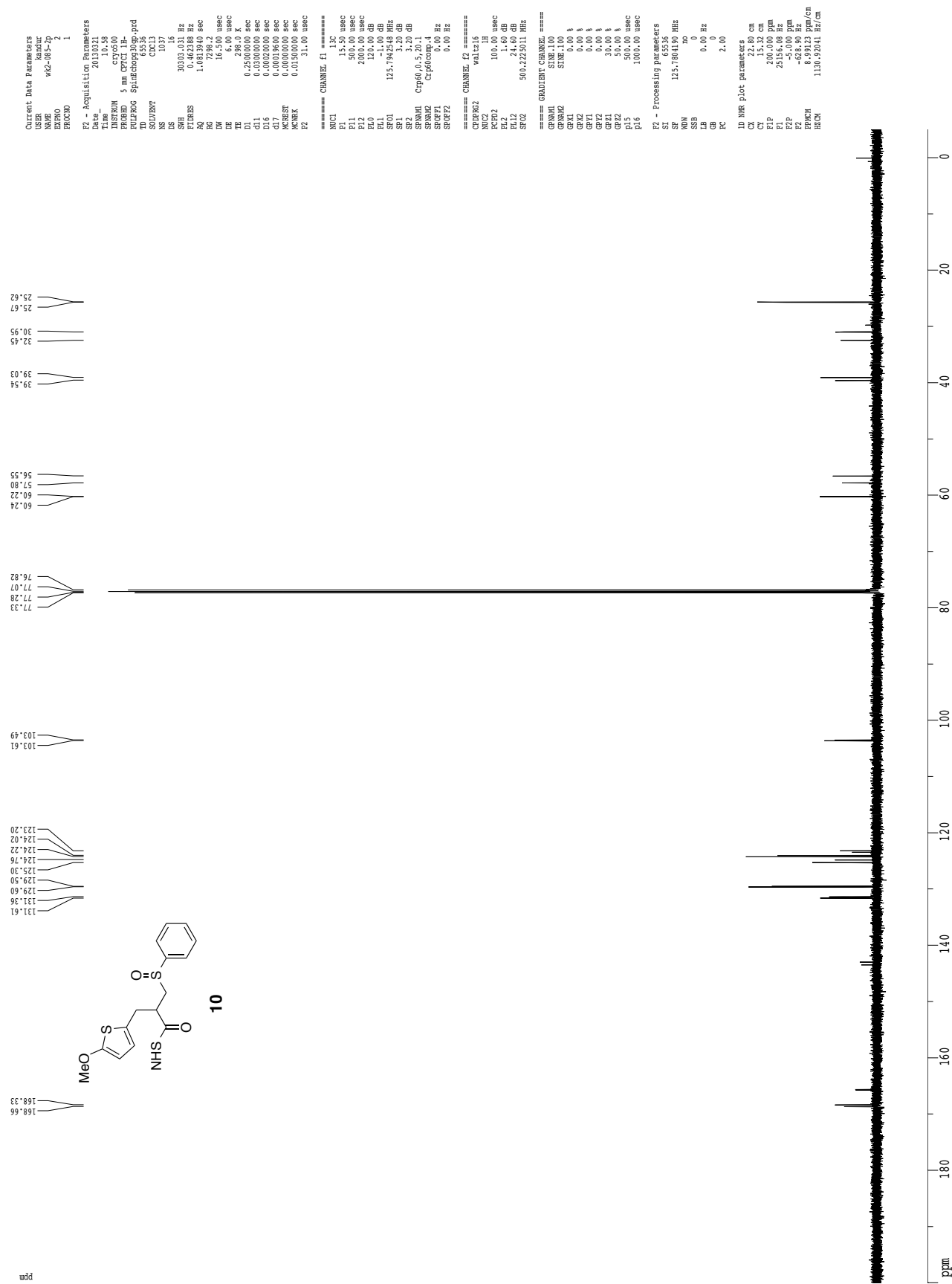




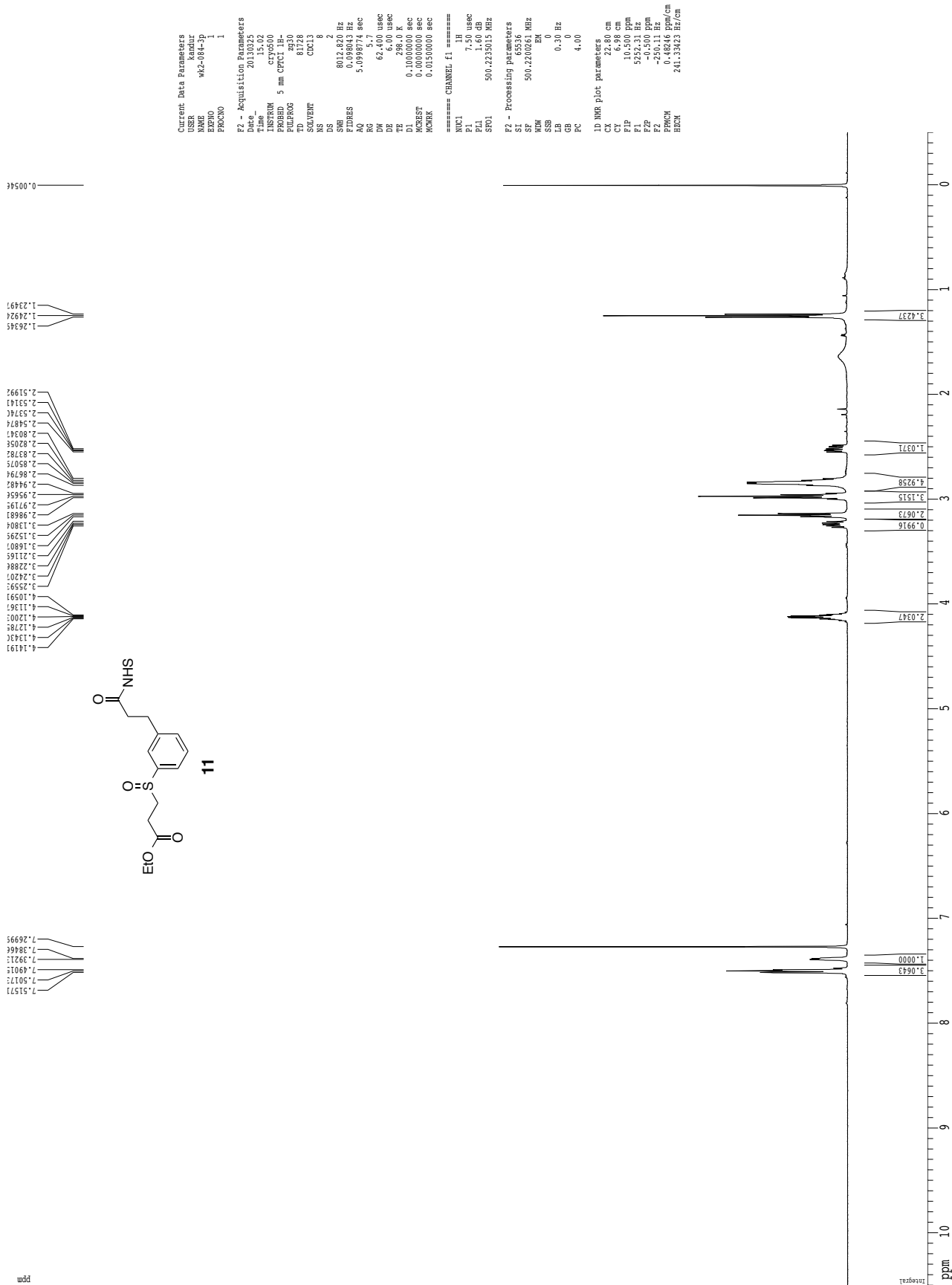
¹H spectrum



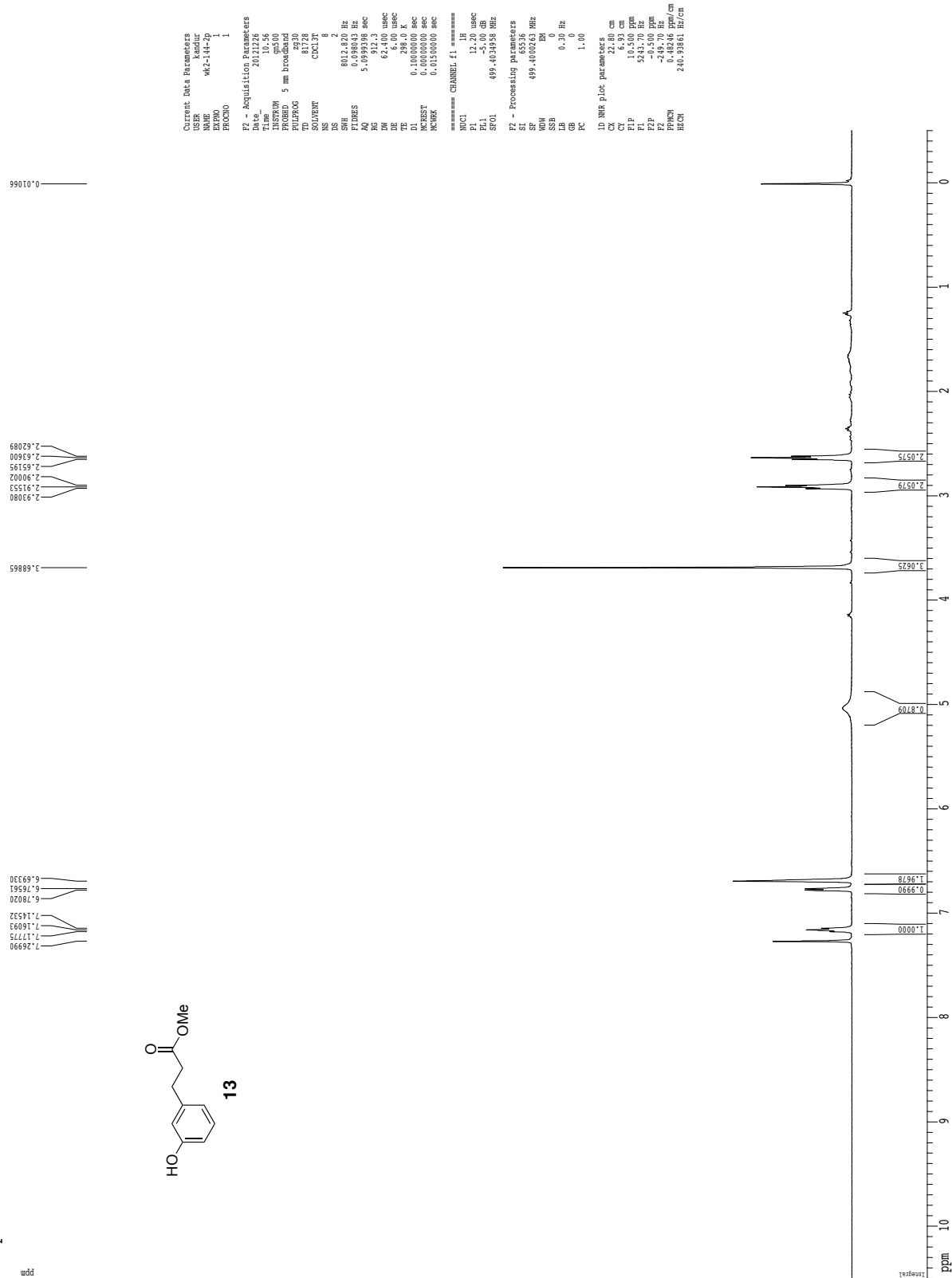
Z-restored spin-echo 13C spectrum with 1H decoupling



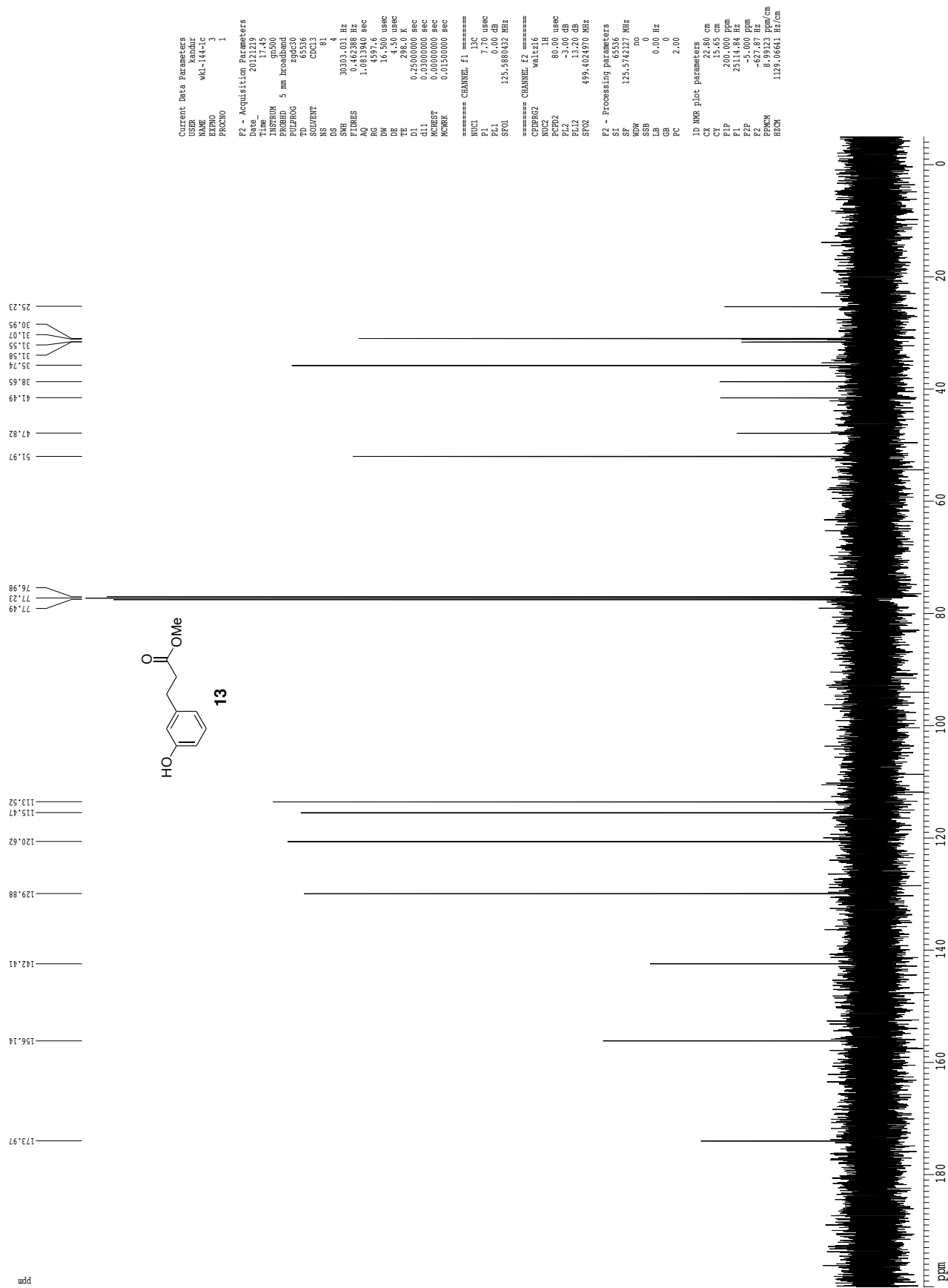
¹H spectrum



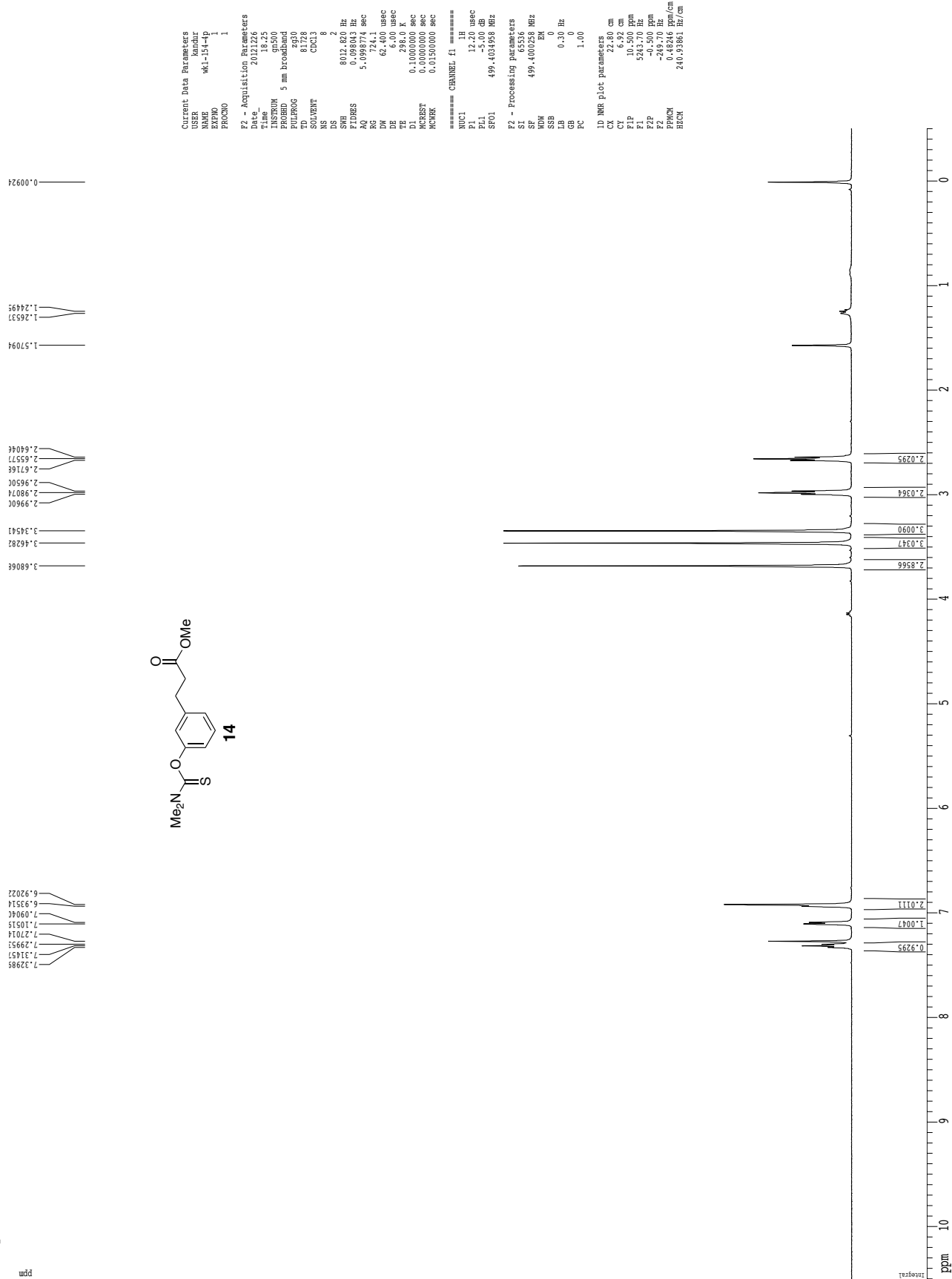
¹H spectrum

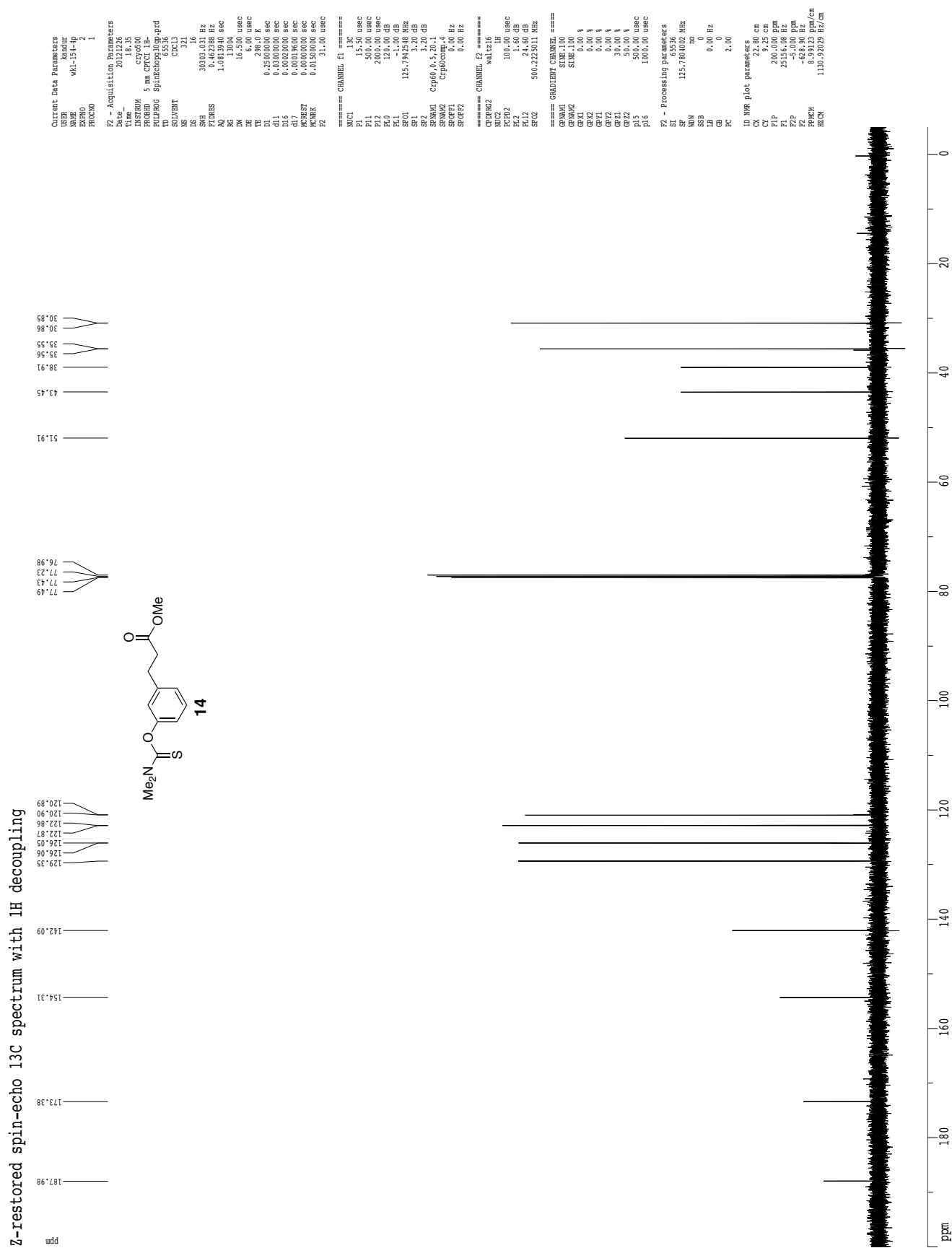


¹³C spectrum with ¹H decoupling

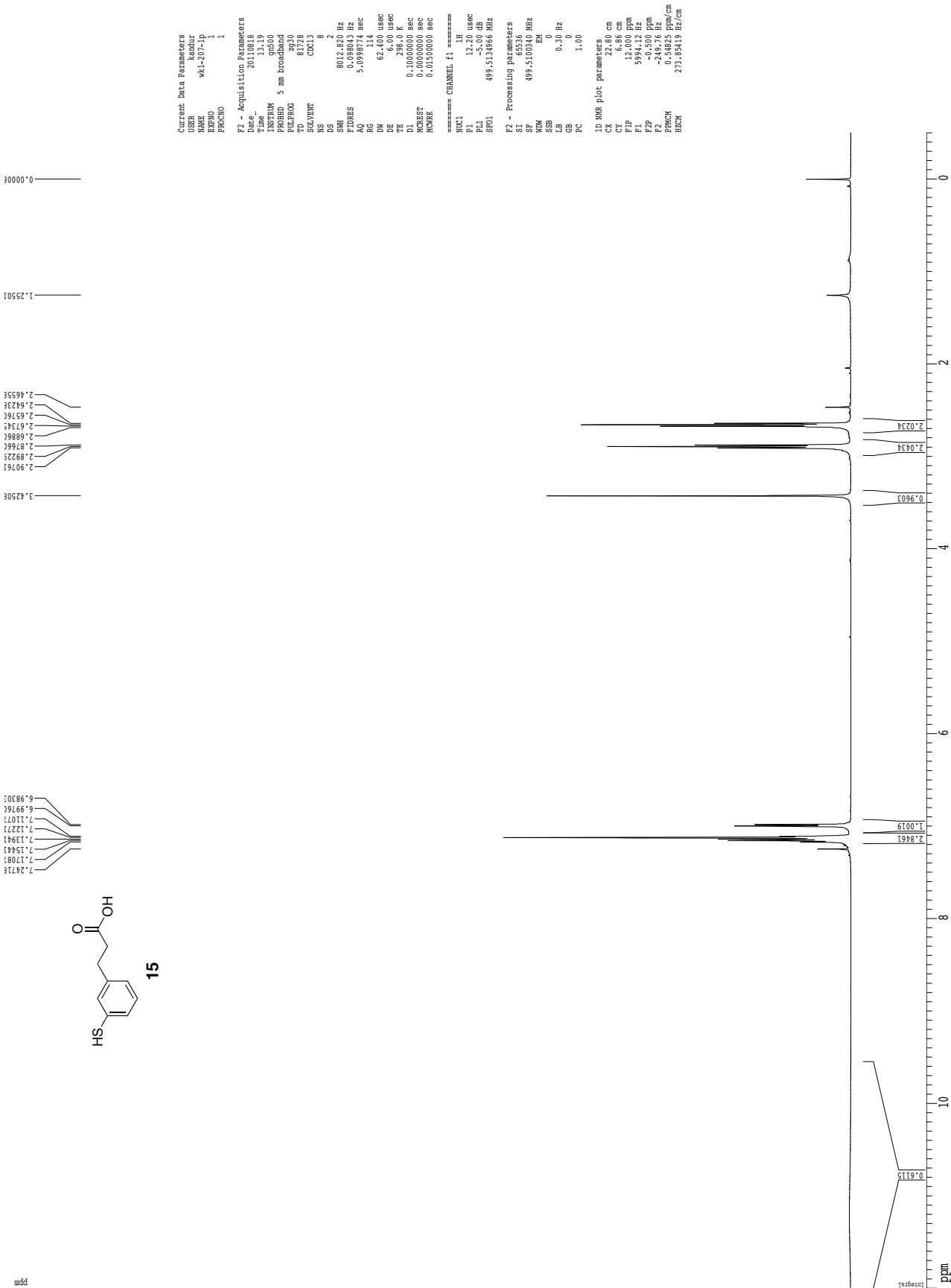


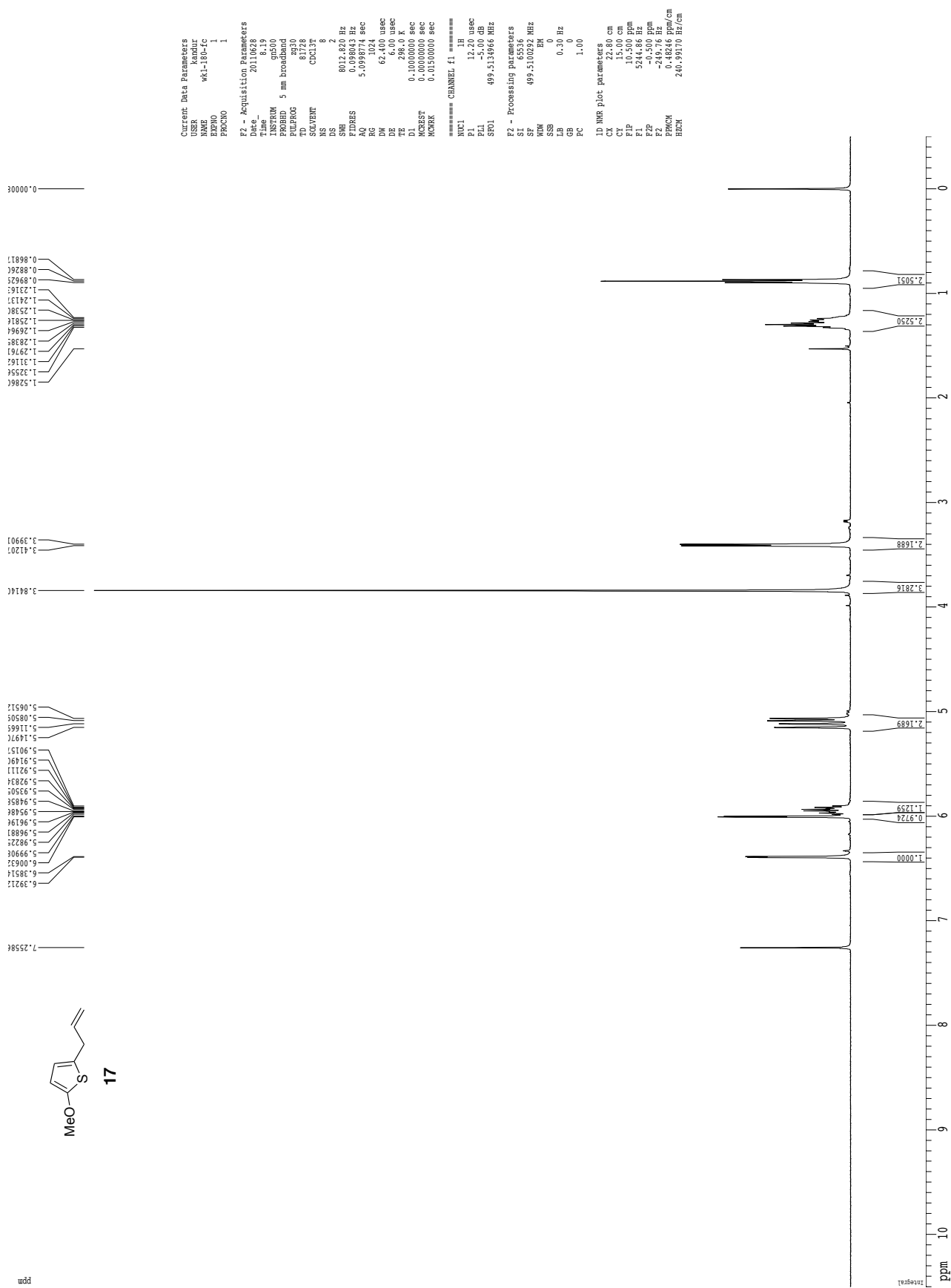
¹H spectrum



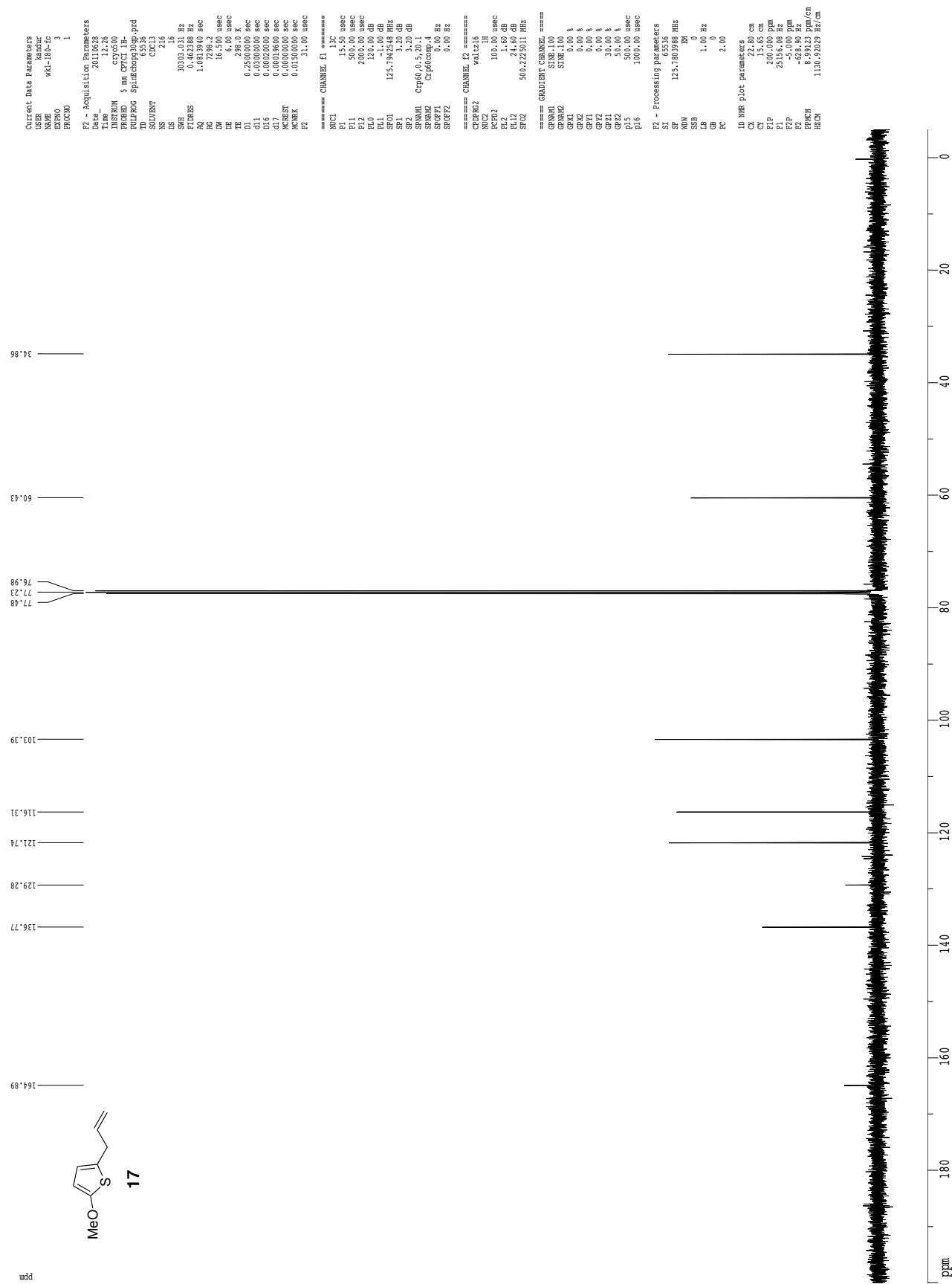


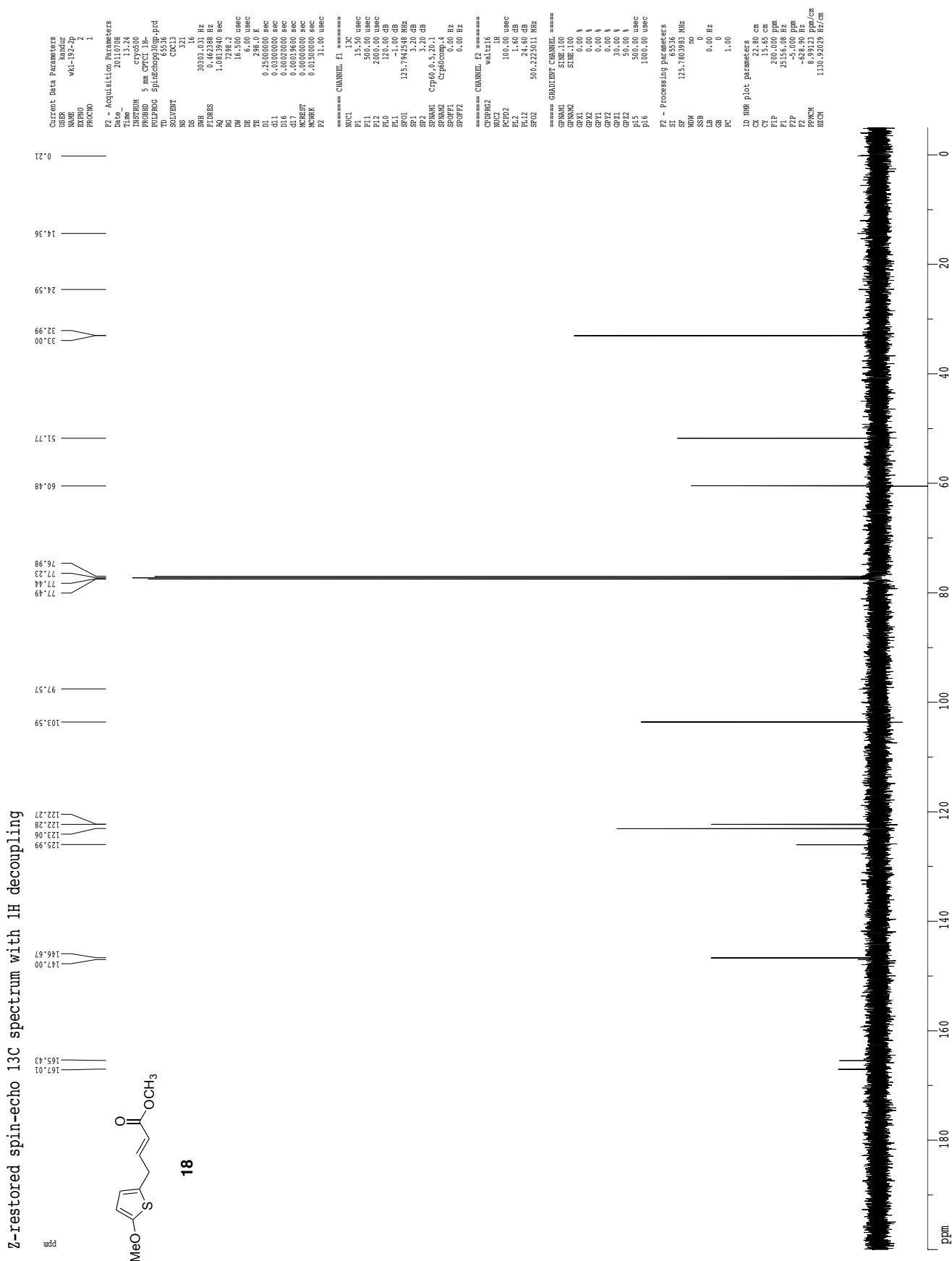
¹H spectrum



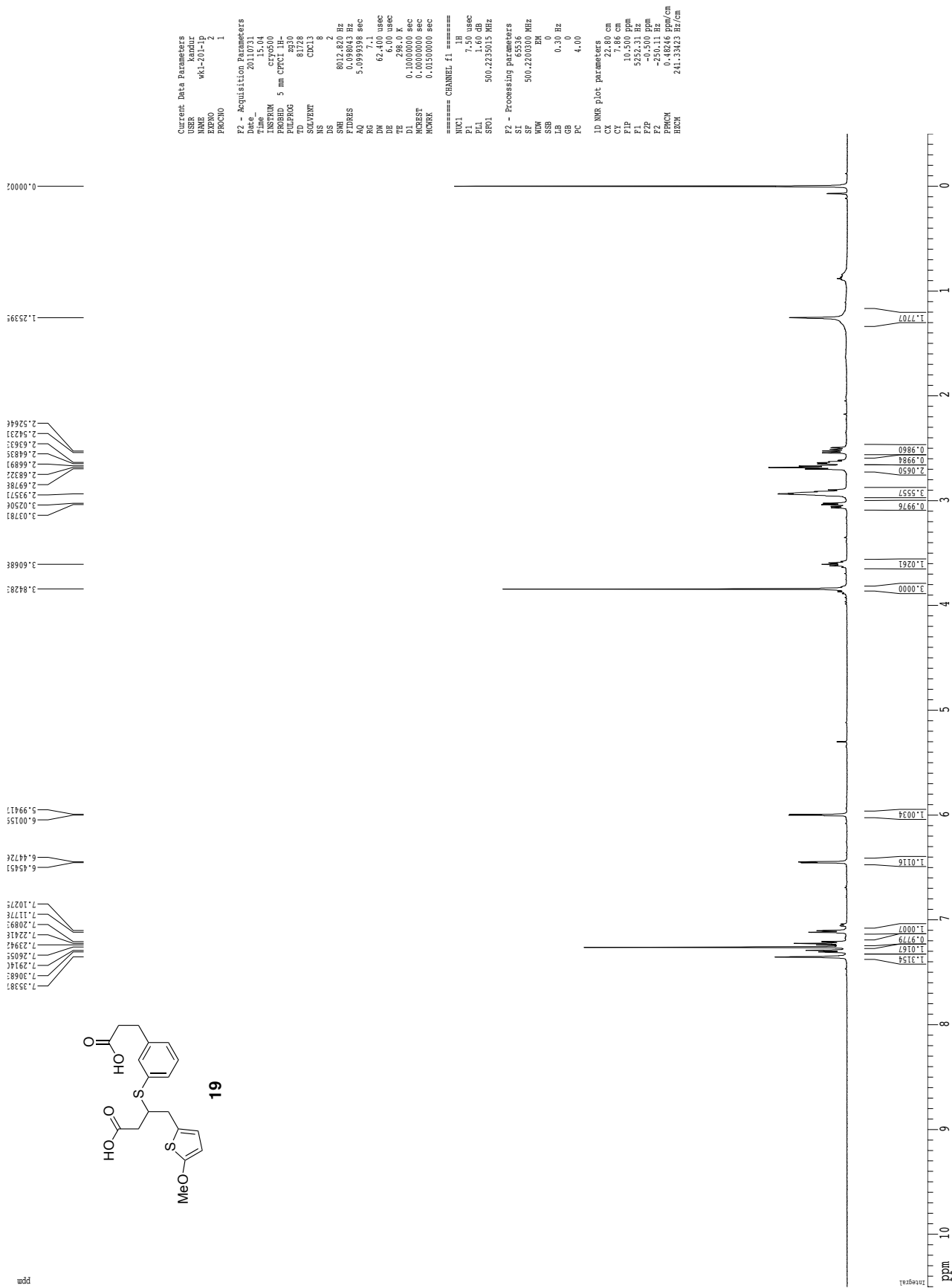
¹H spectrum

Z-restored spin-echo ¹³C spectrum with ¹H decoupling

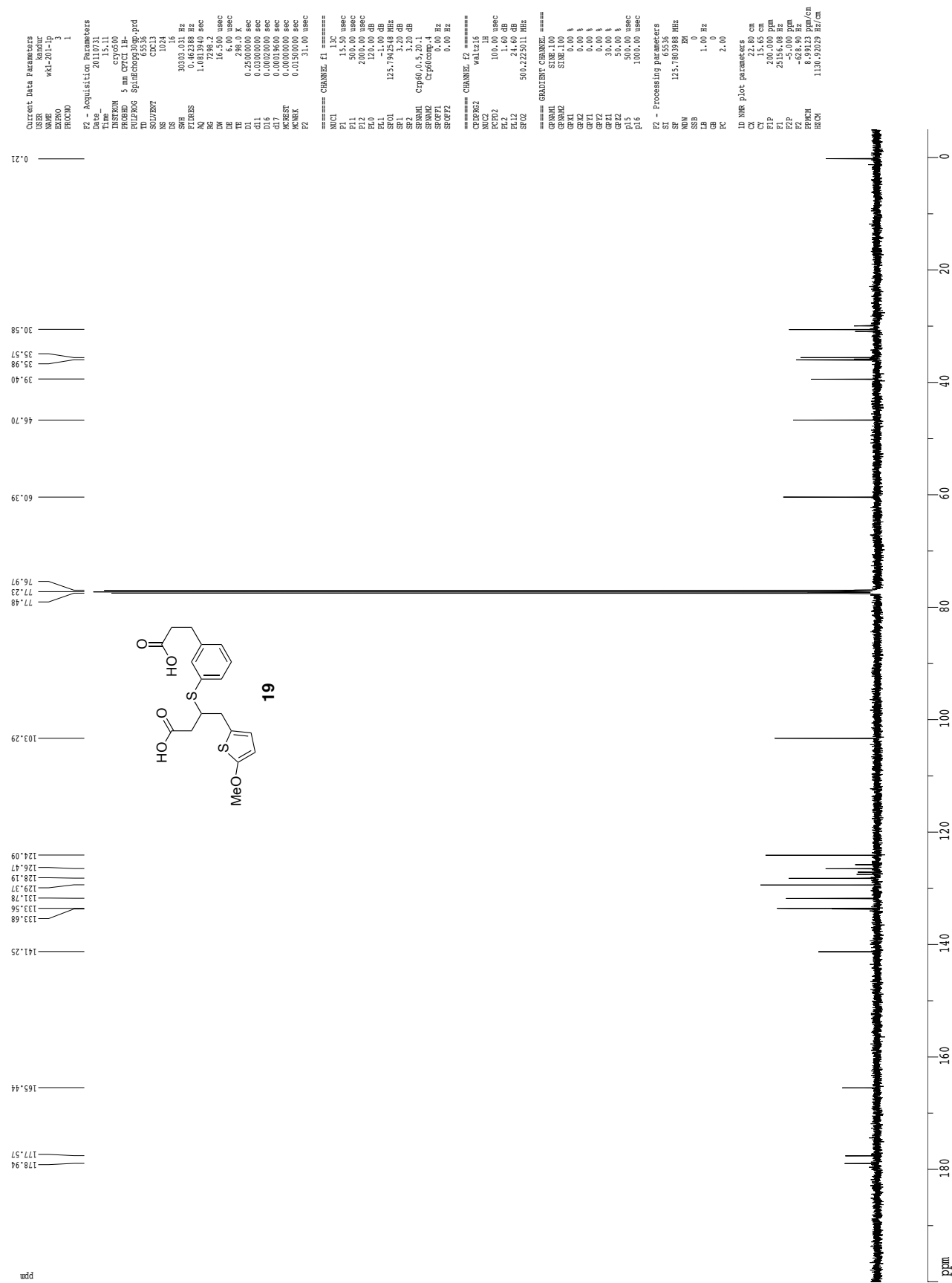




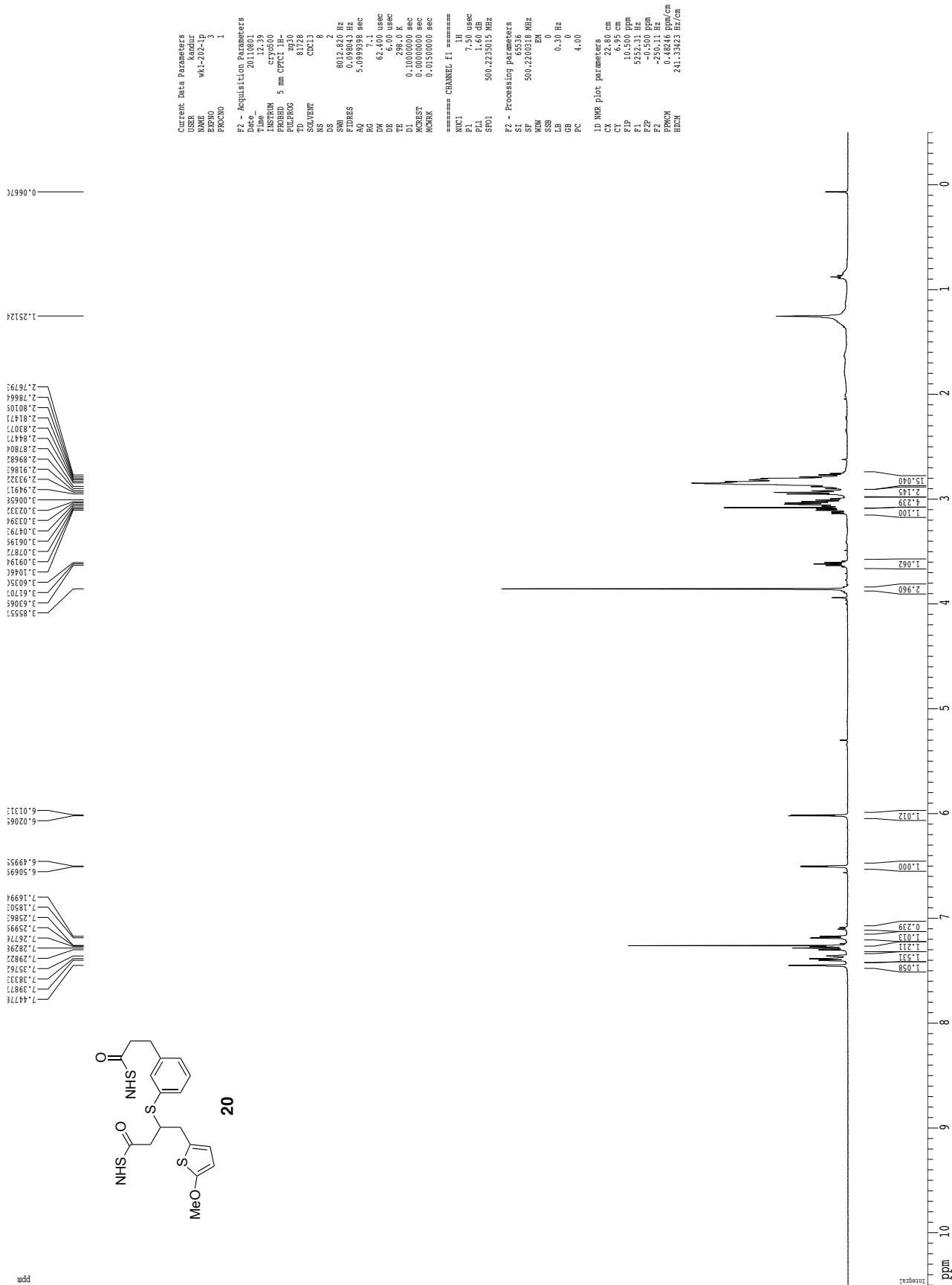
¹H spectrum



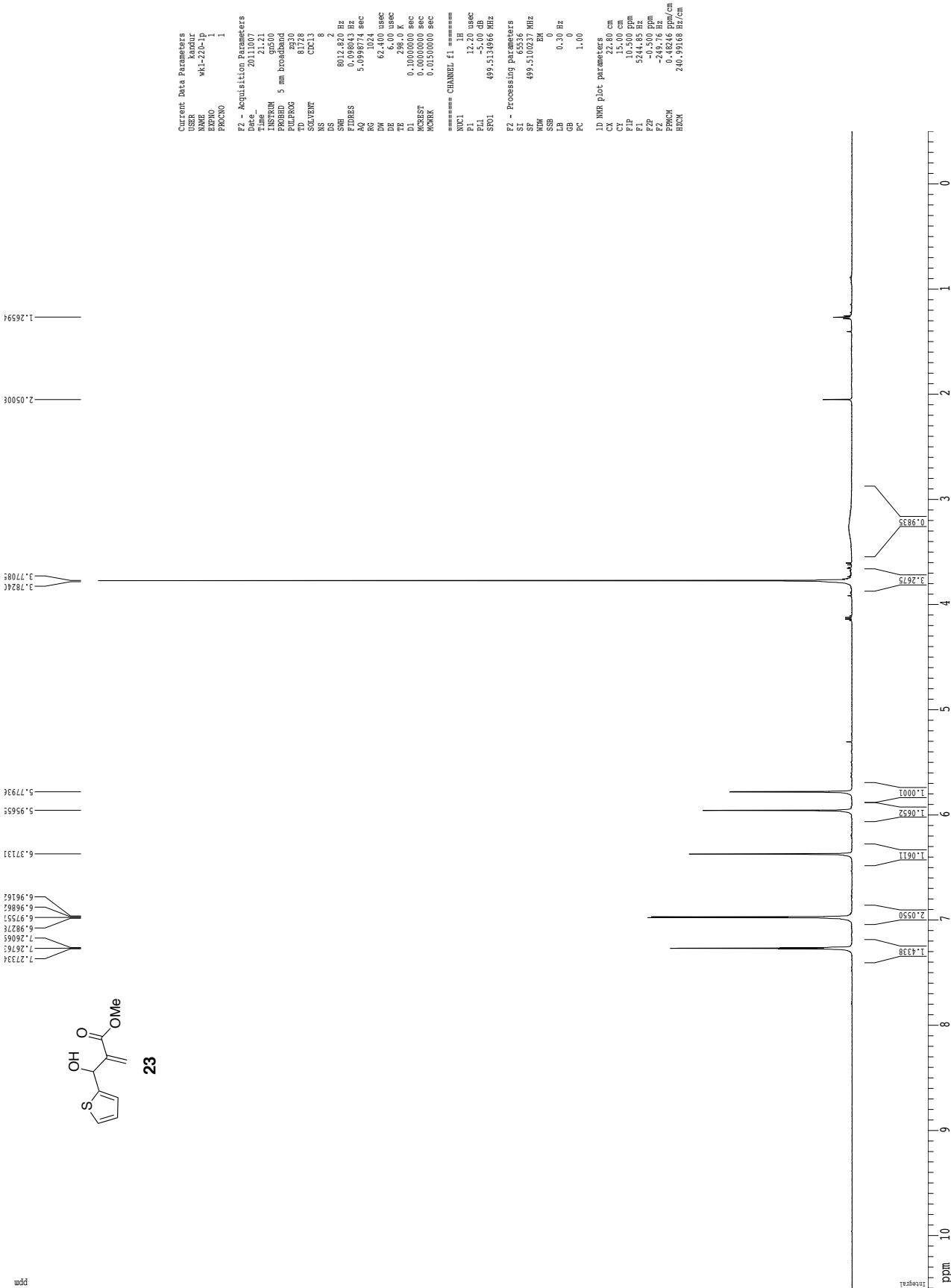
Z-restored spin-echo 13C spectrum with 1H decoupling



1H spectrum

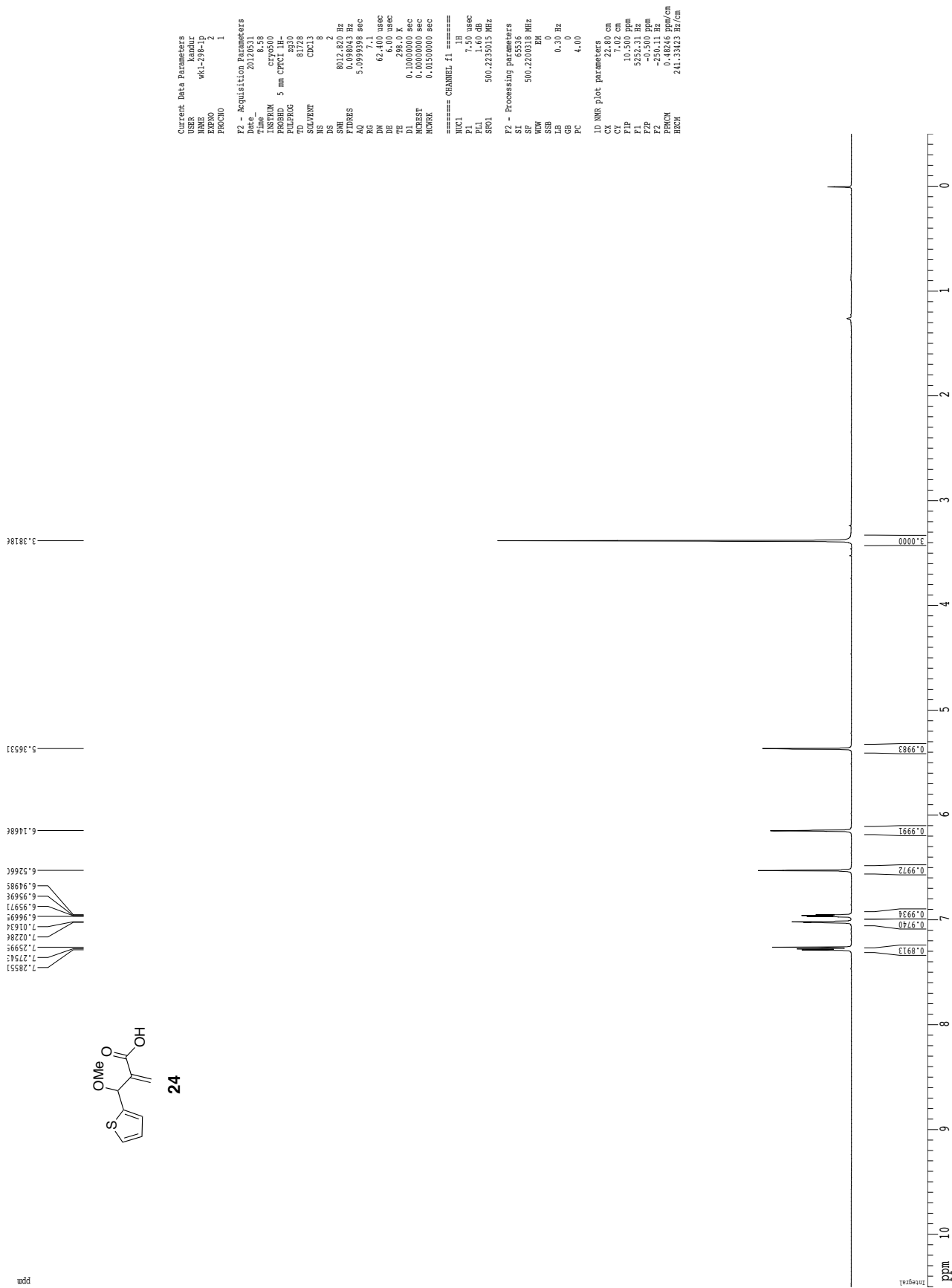


¹H spectrum

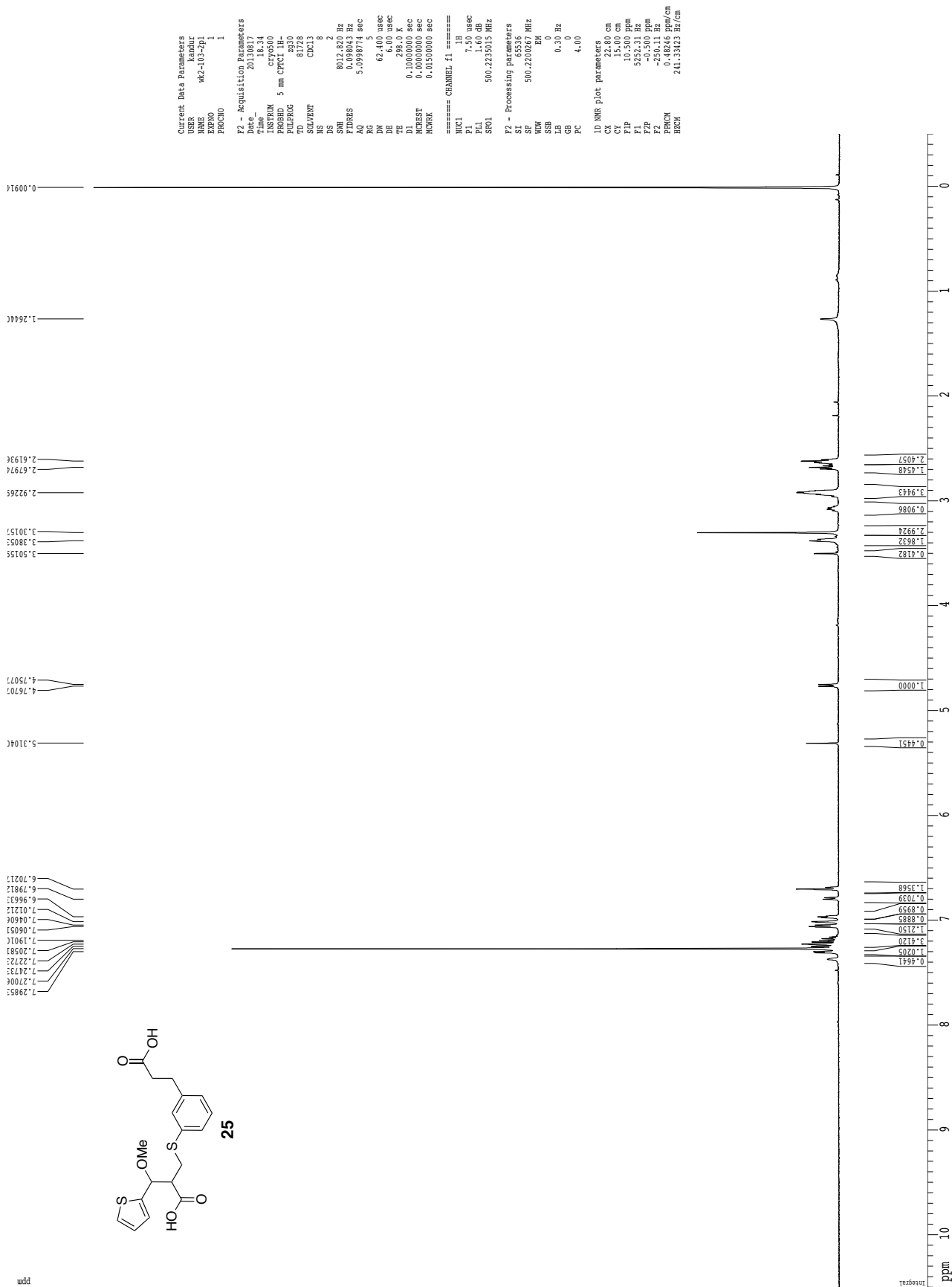


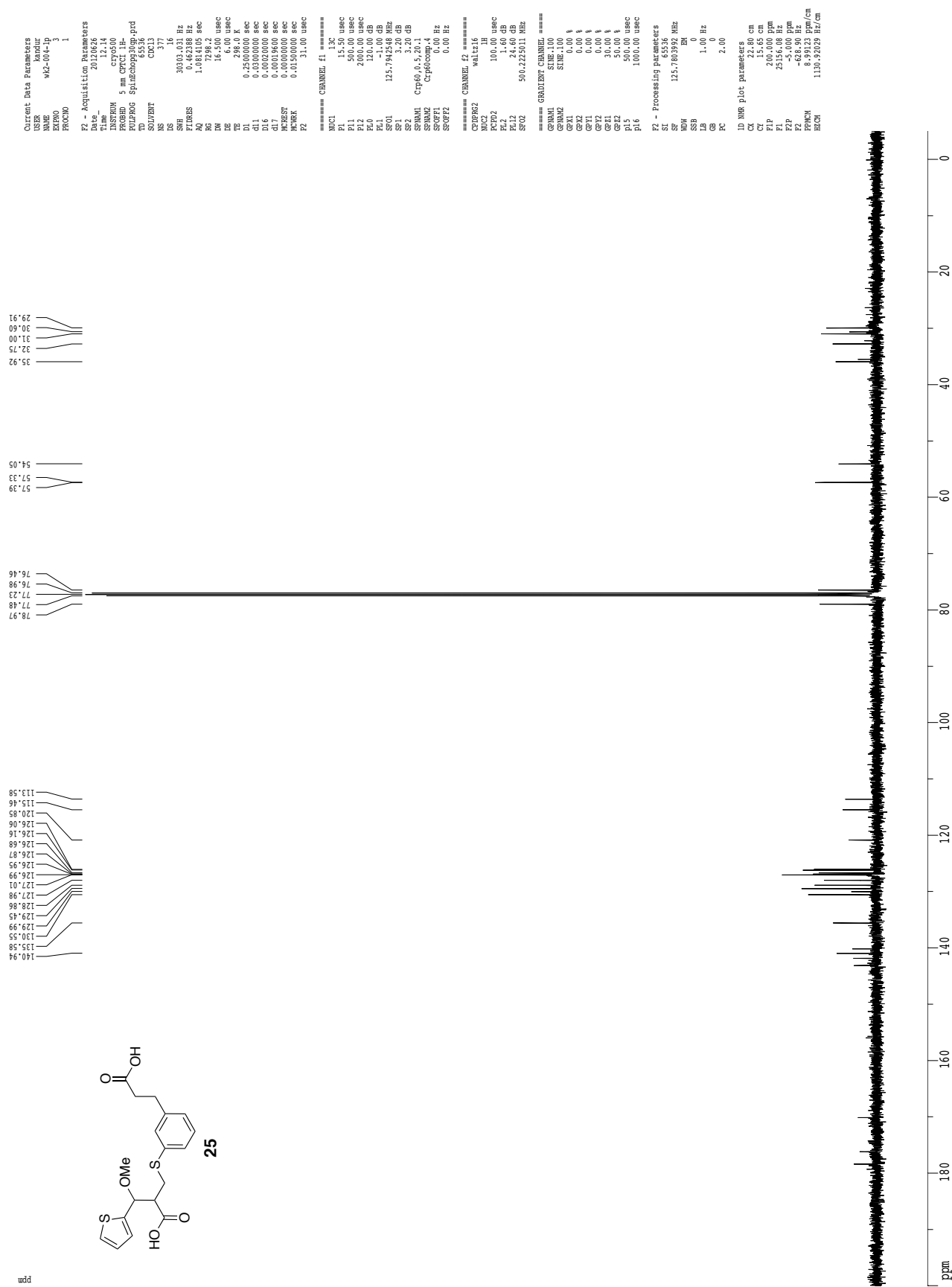
23

1H spectrum



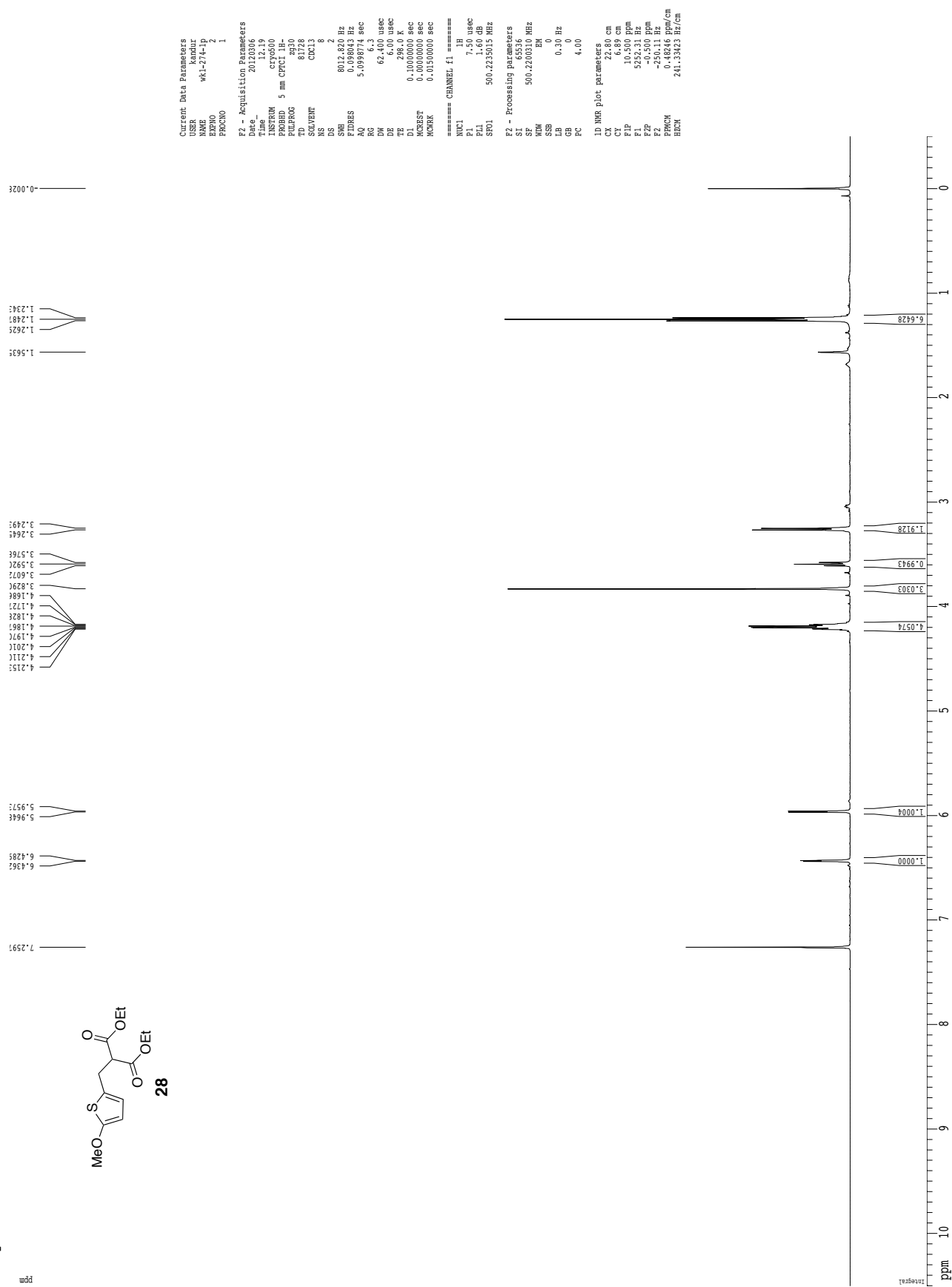
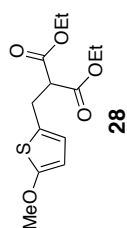
¹H spectrum



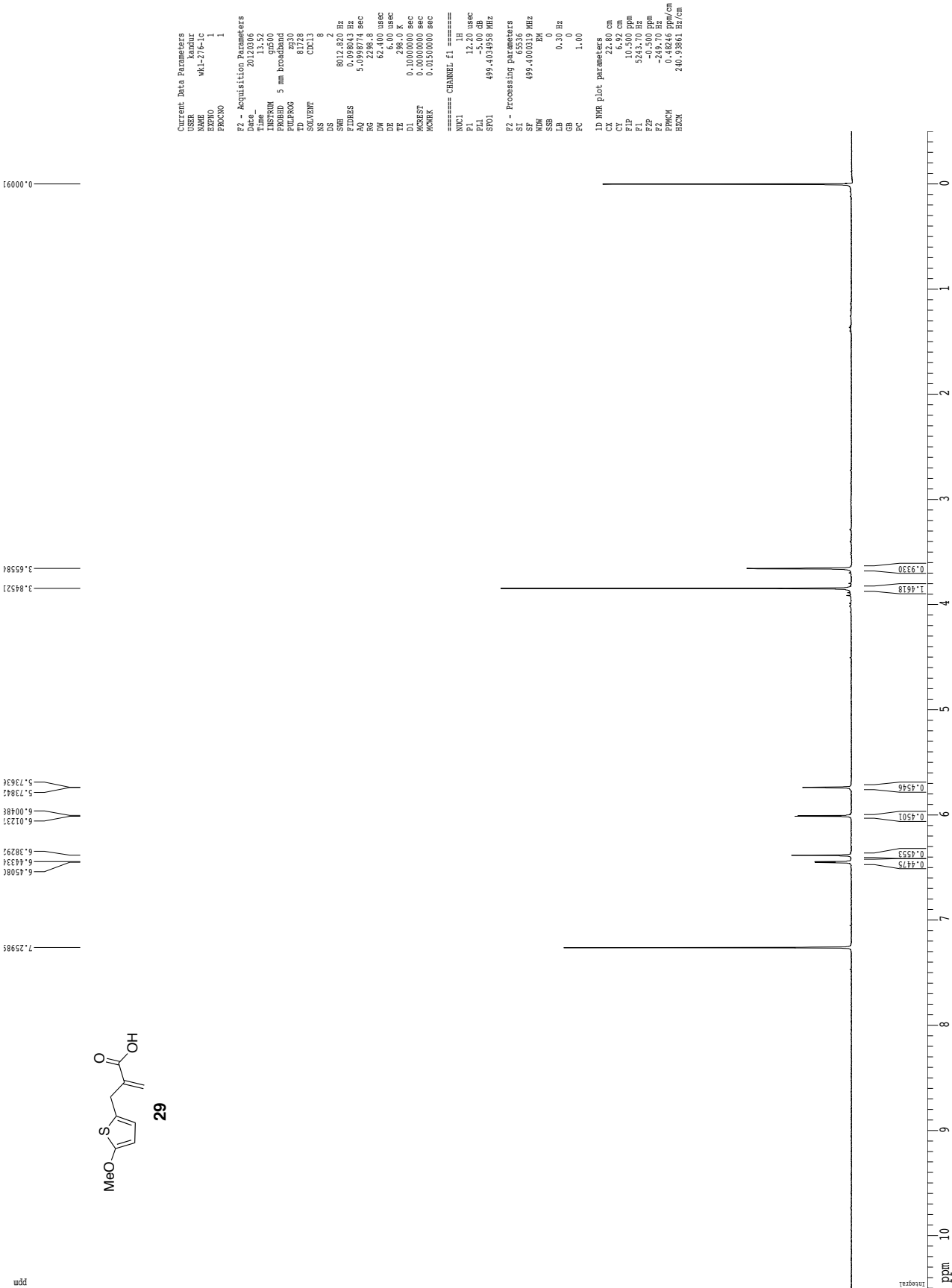
Z-restored spin-echo ^{13}C spectrum with ^1H decoupling



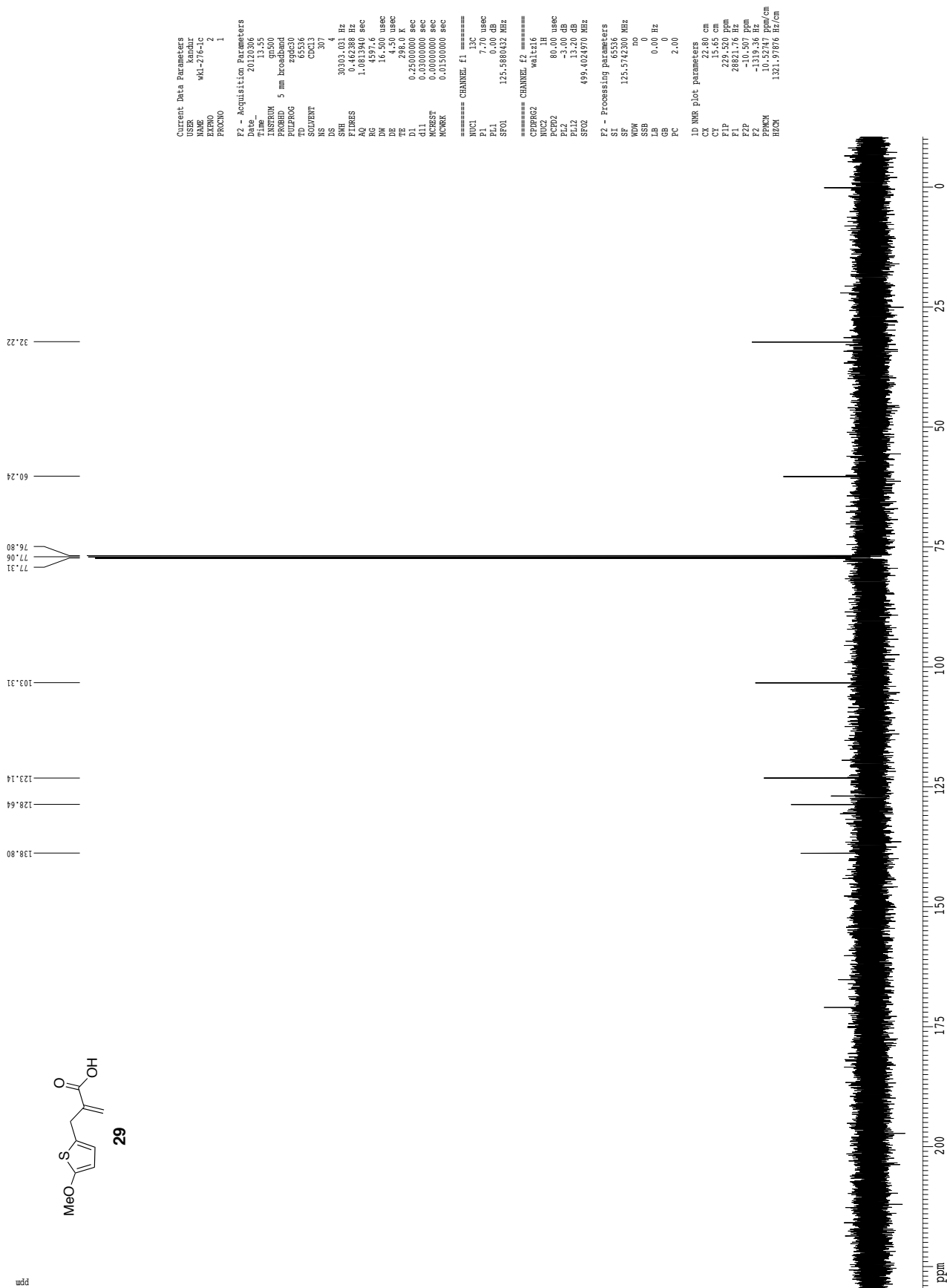
wdd

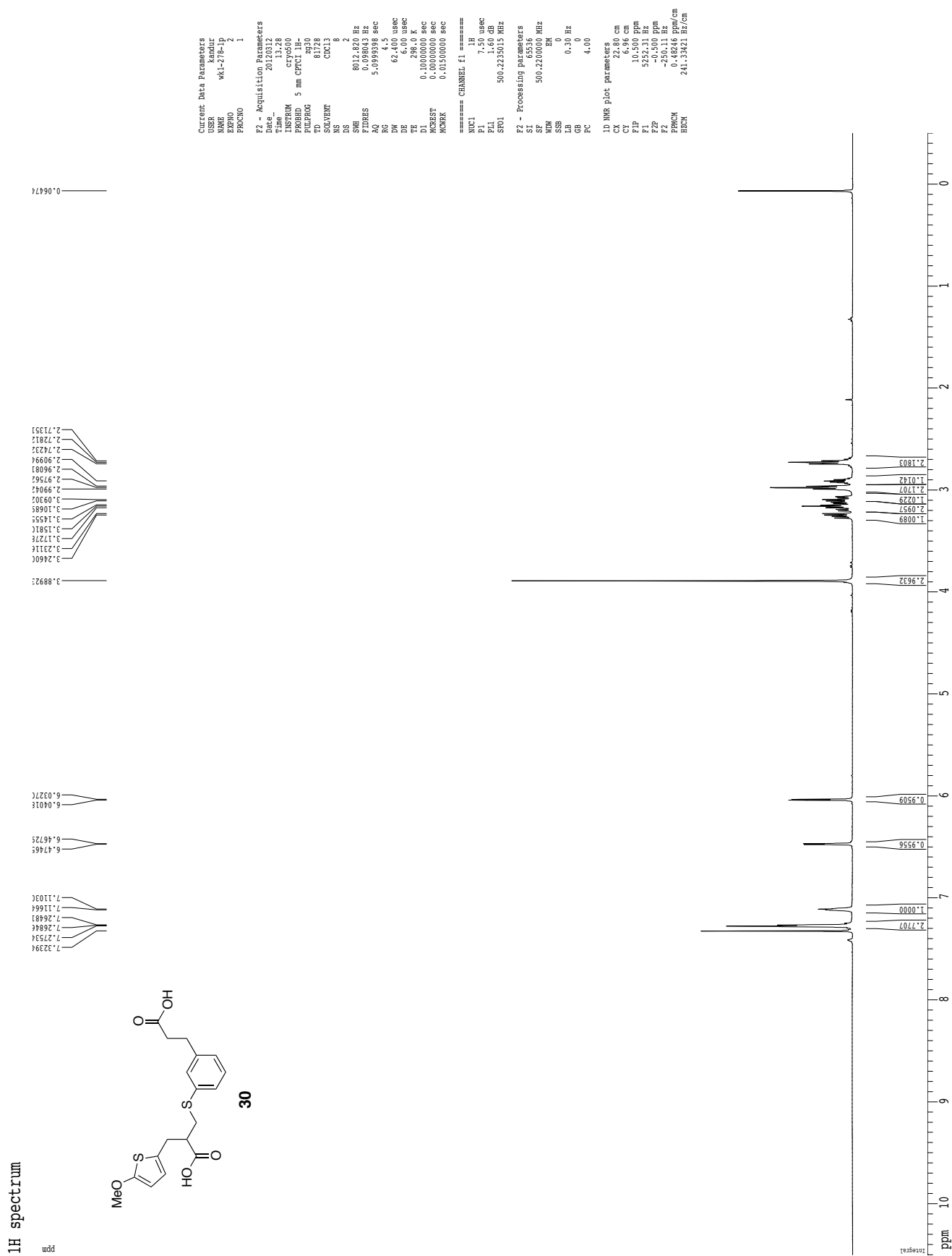


1H spectrum

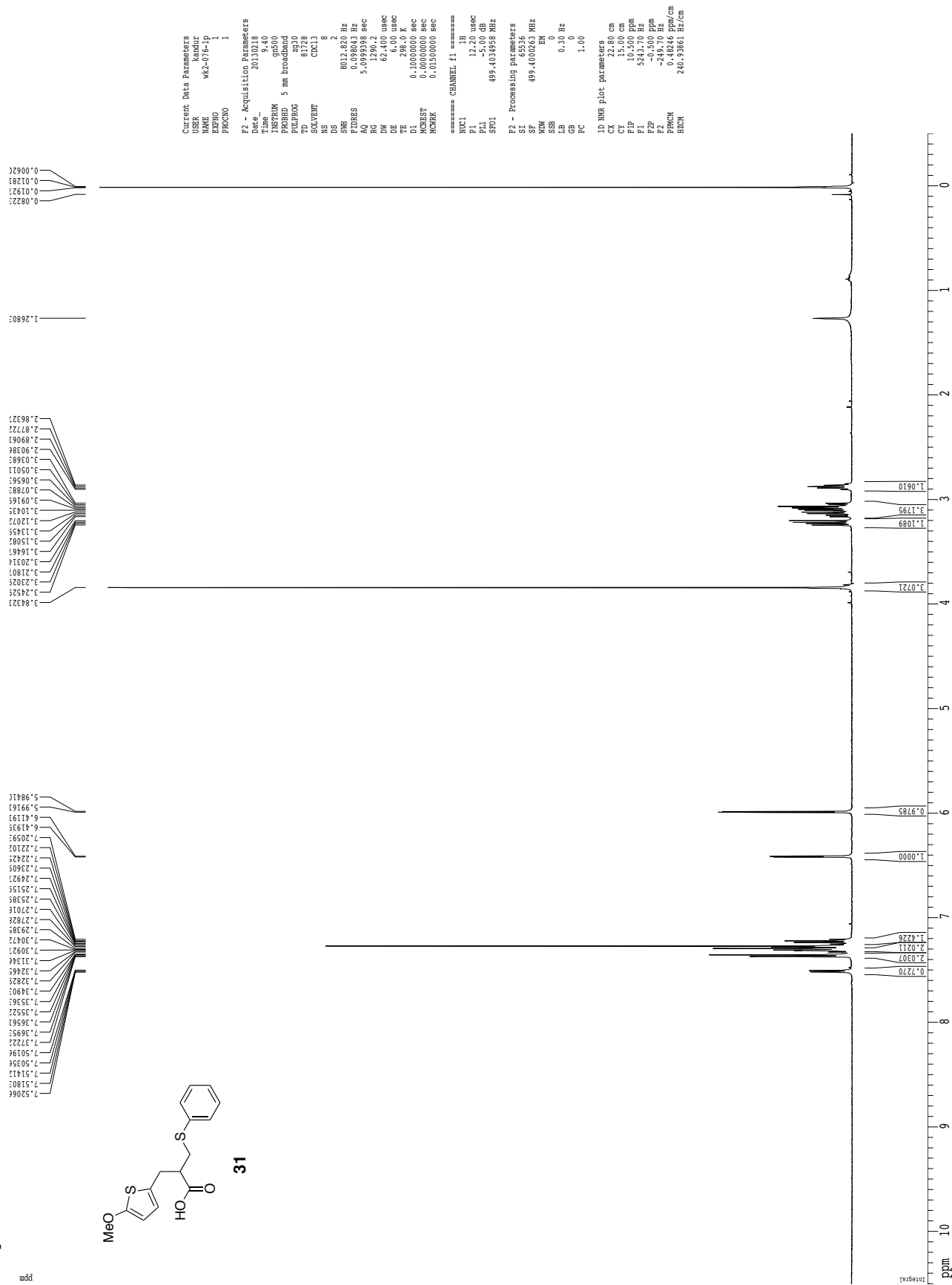


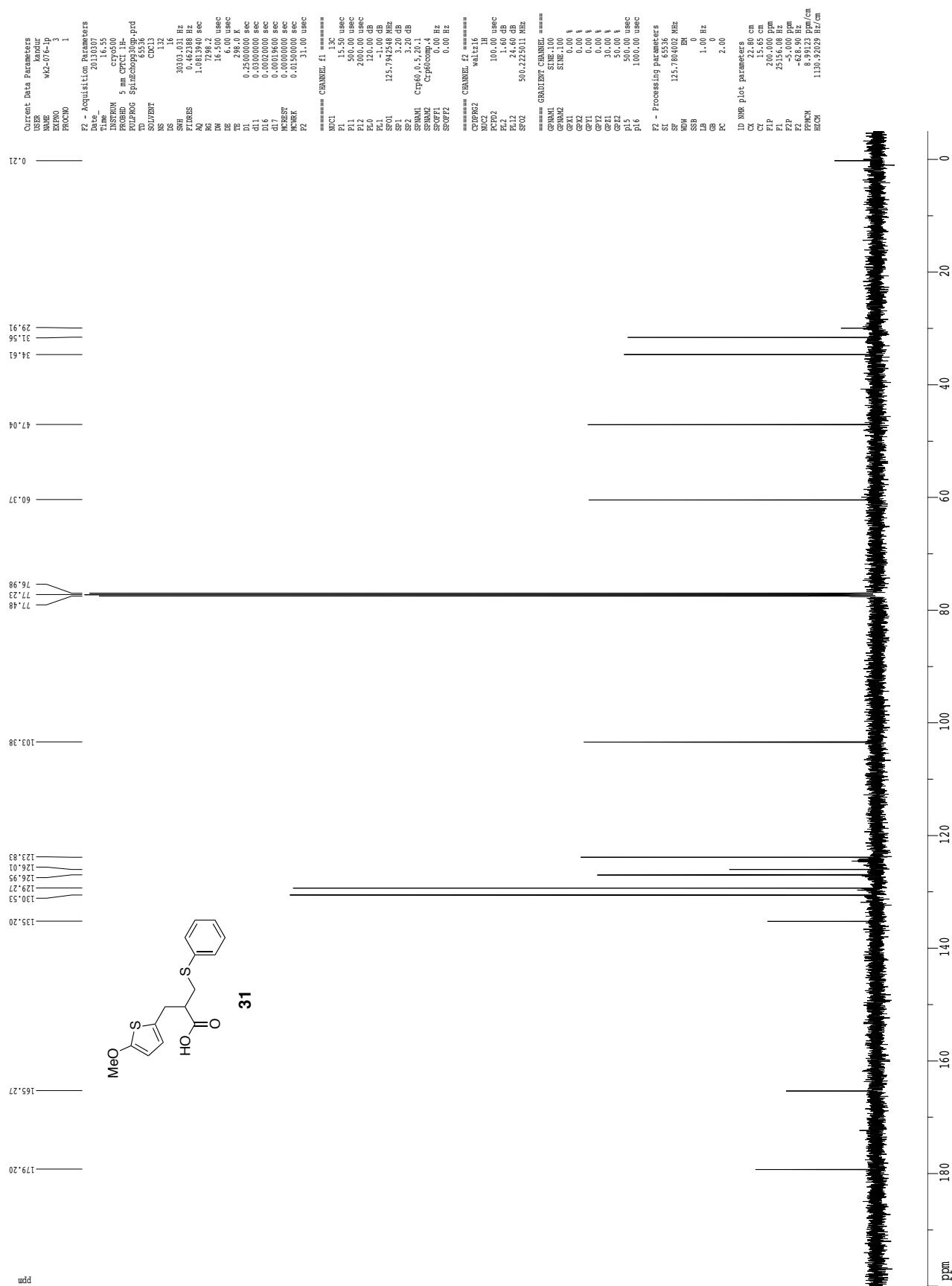
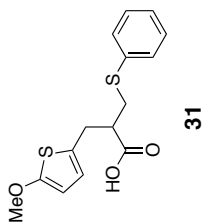
13C spectrum with 1H decoupling



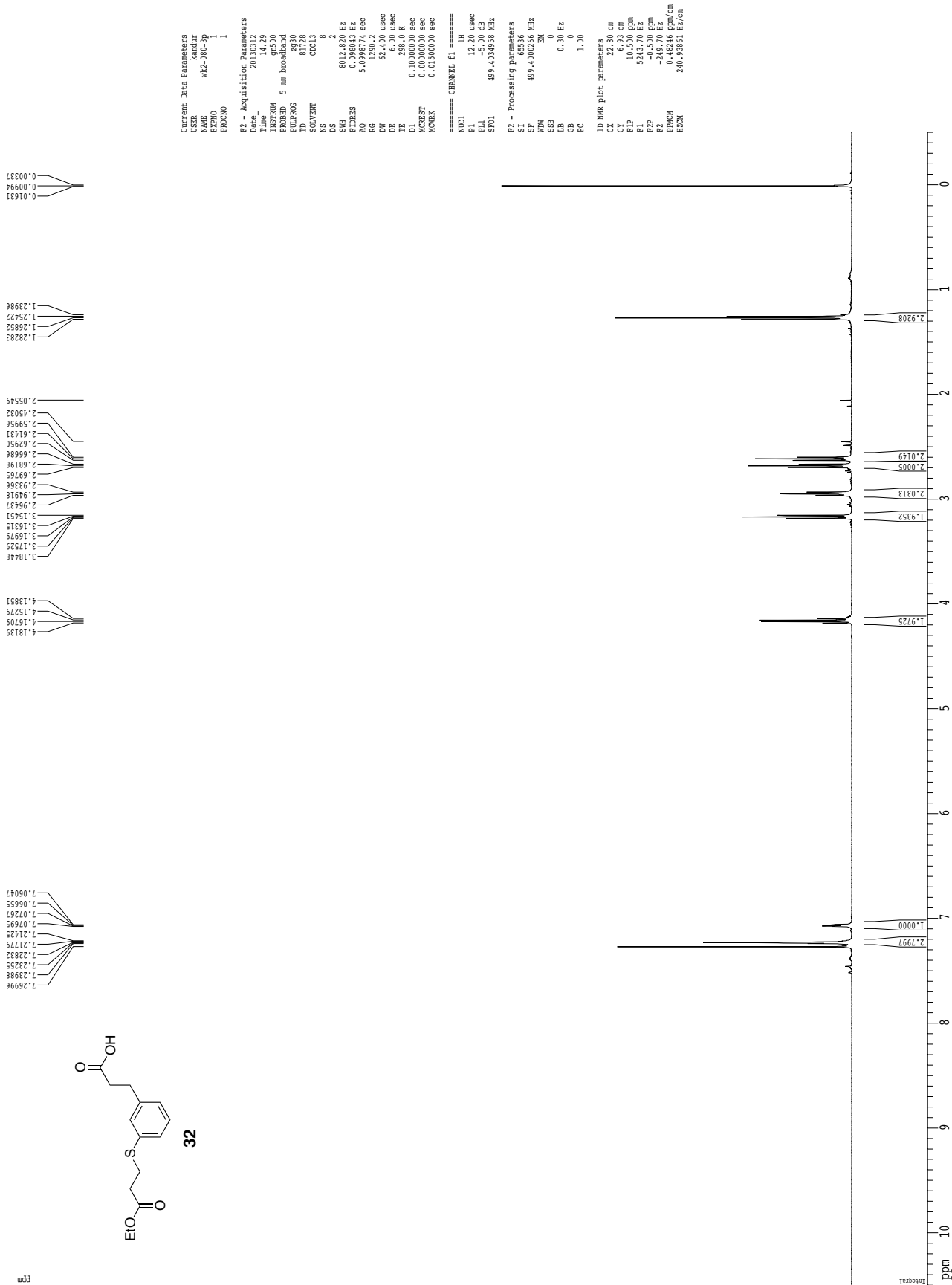


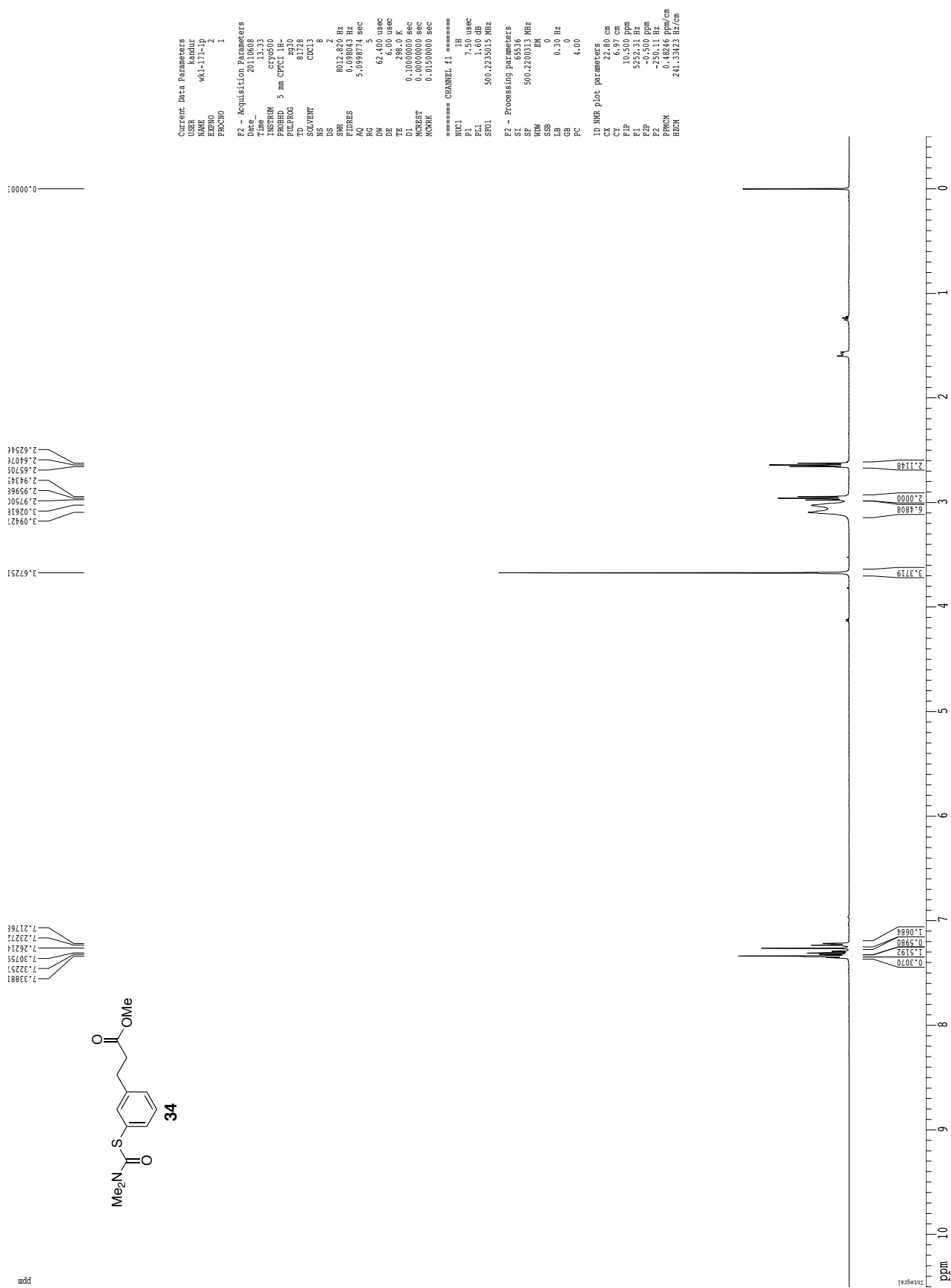
¹H spectrum



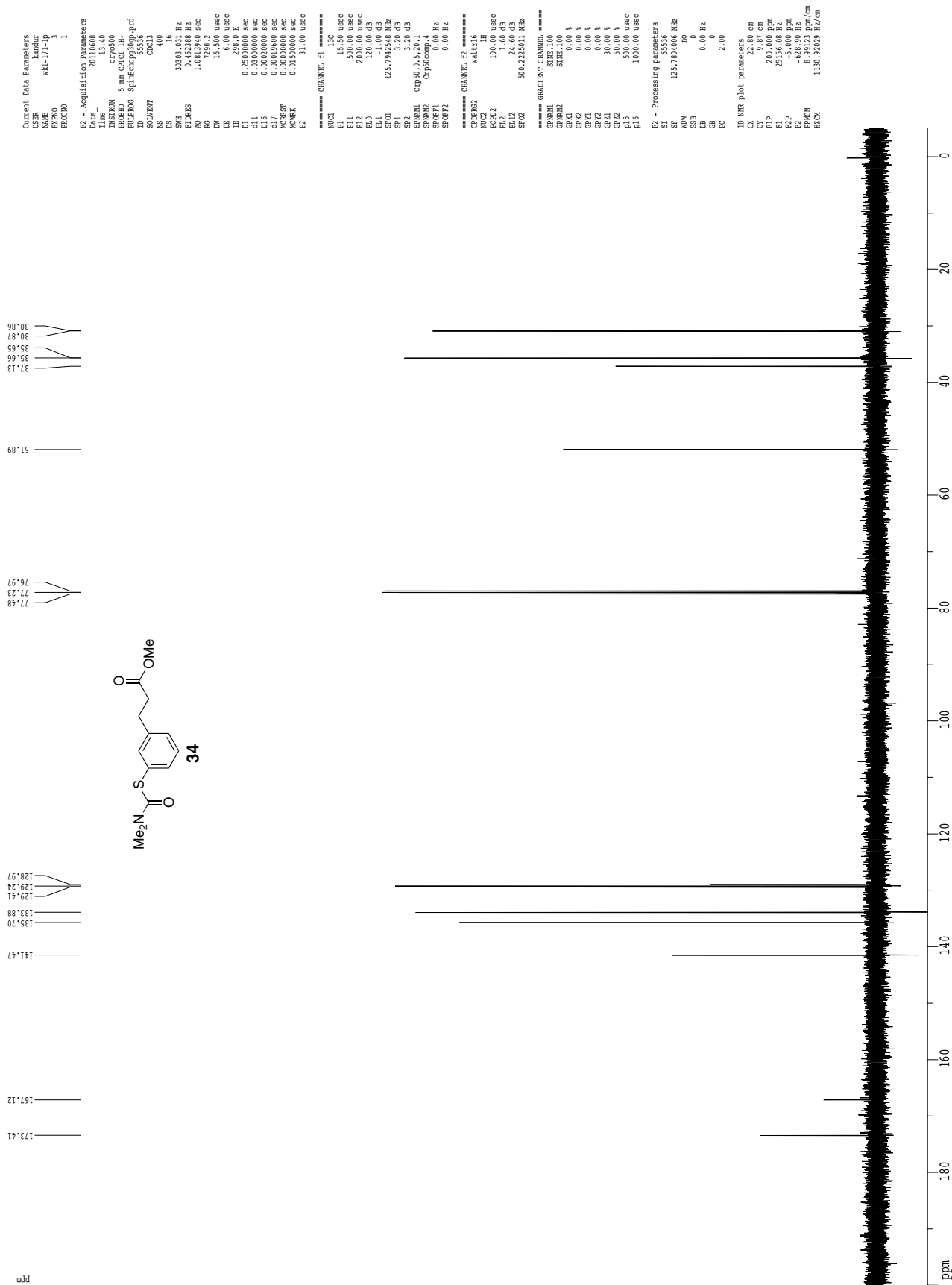
Z-restored spin-echo ^{13}C spectrum with ^1H decoupling

¹H spectrum

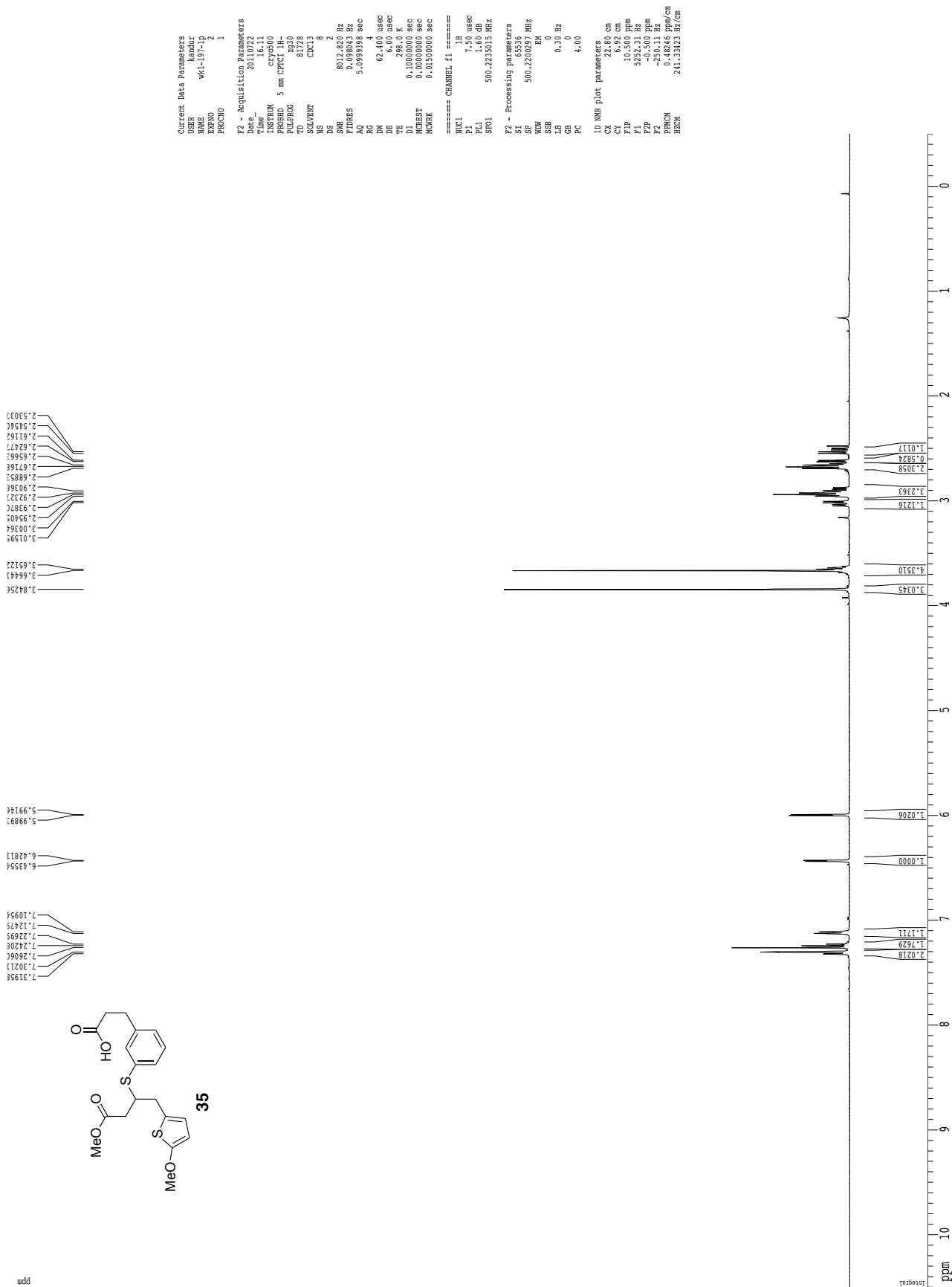


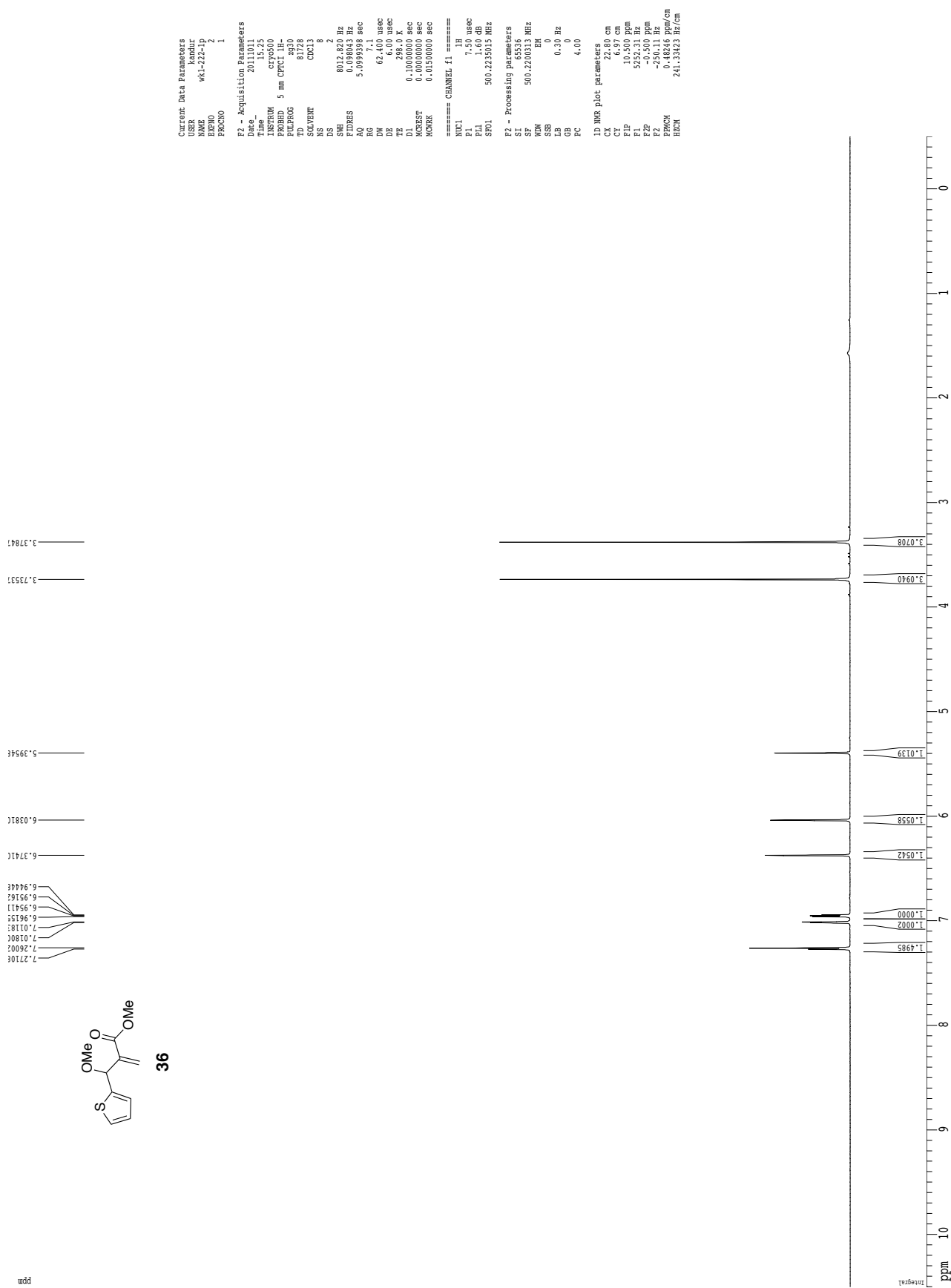
¹H spectrum

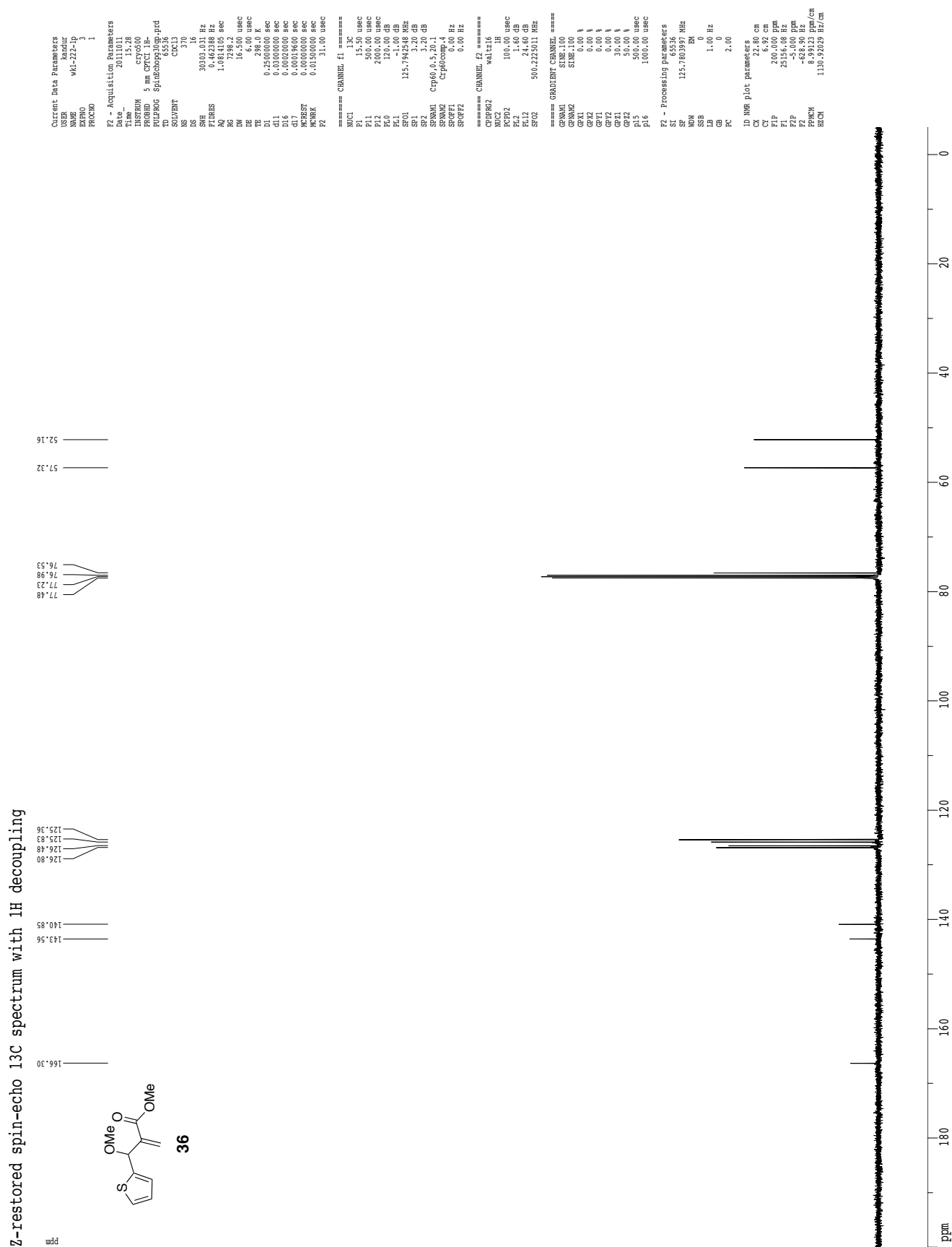
Z-restored spin-echo 13C spectrum with 1H decoupling



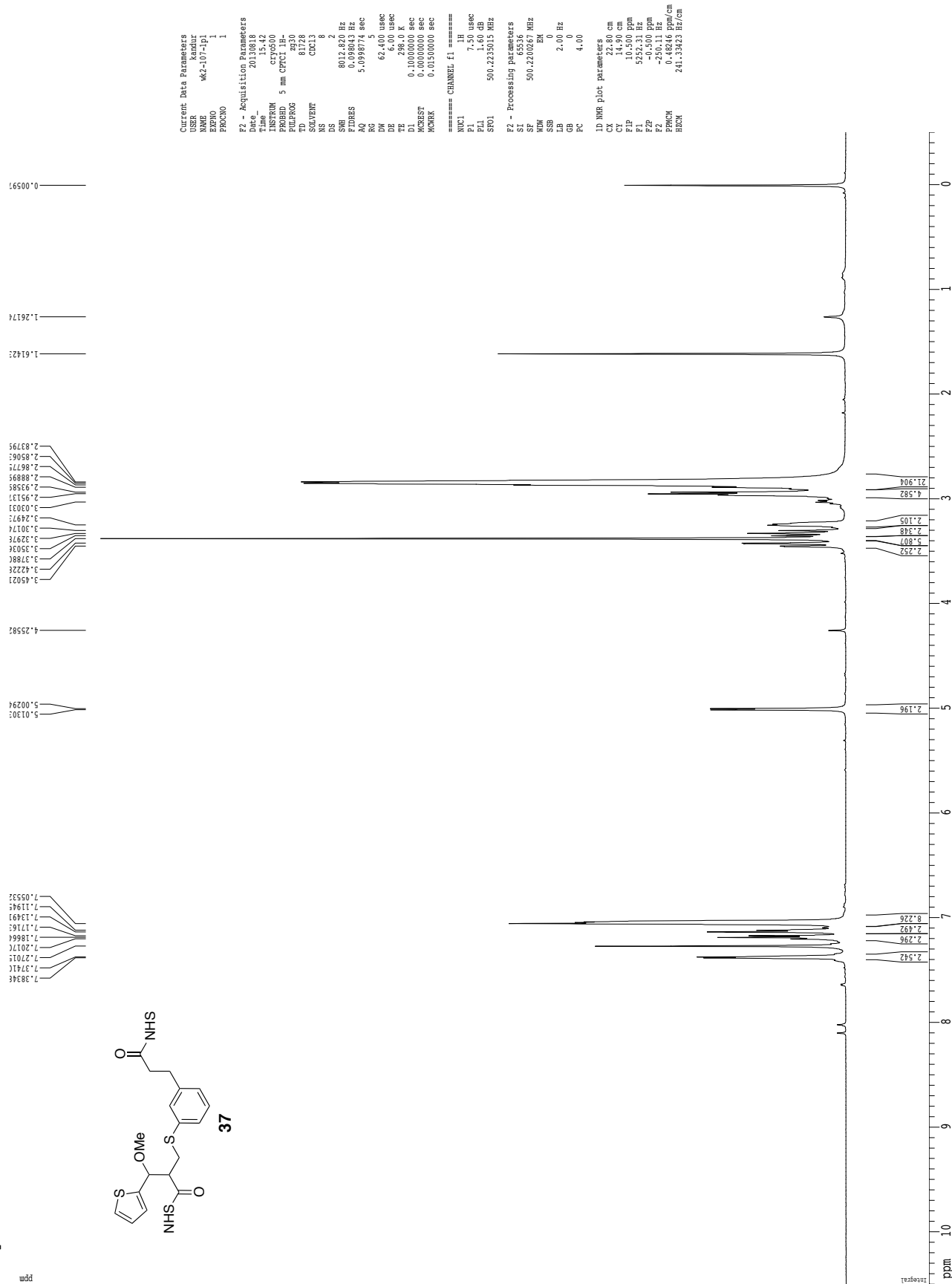
¹H spectrum



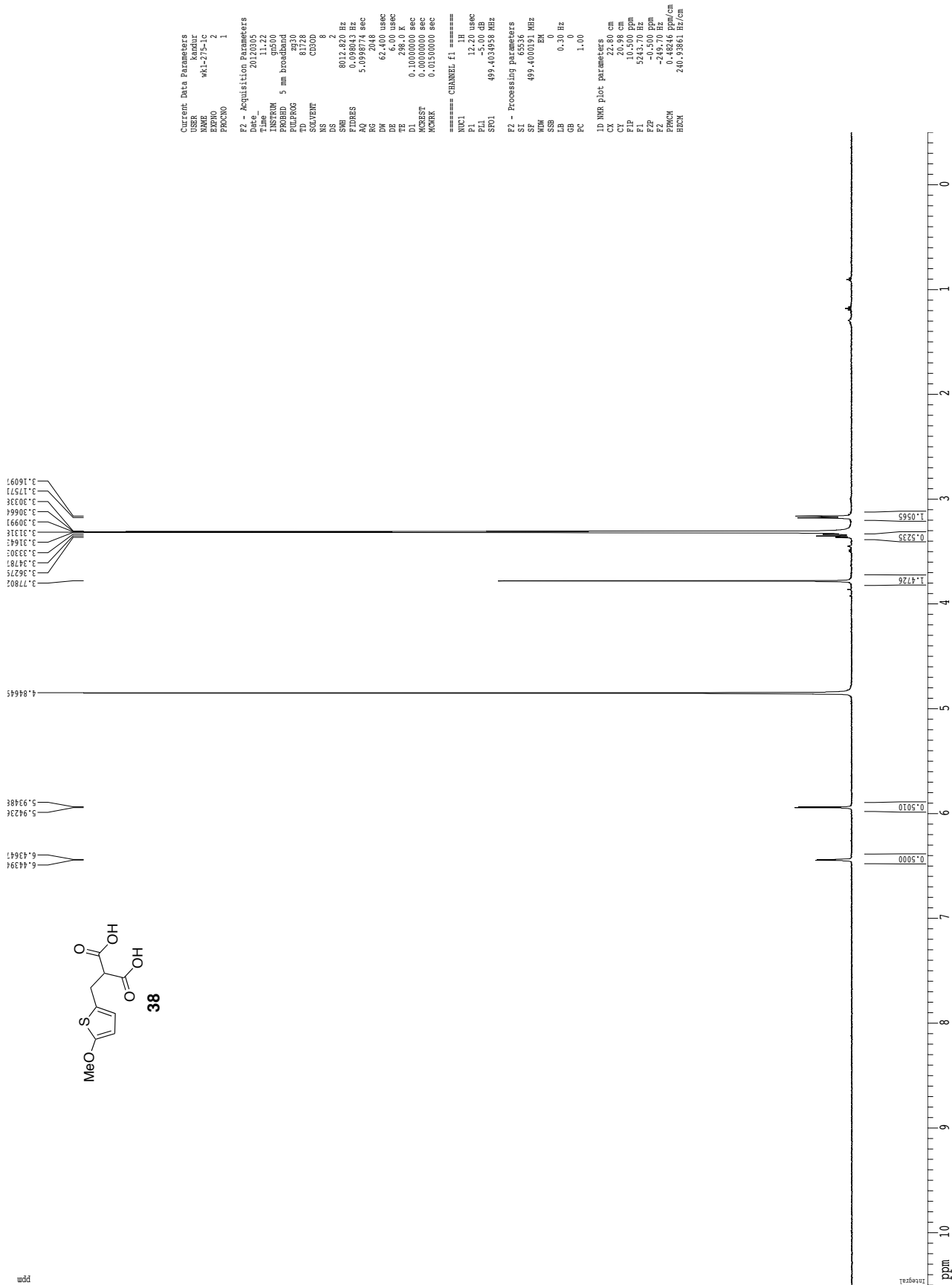
¹H spectrum

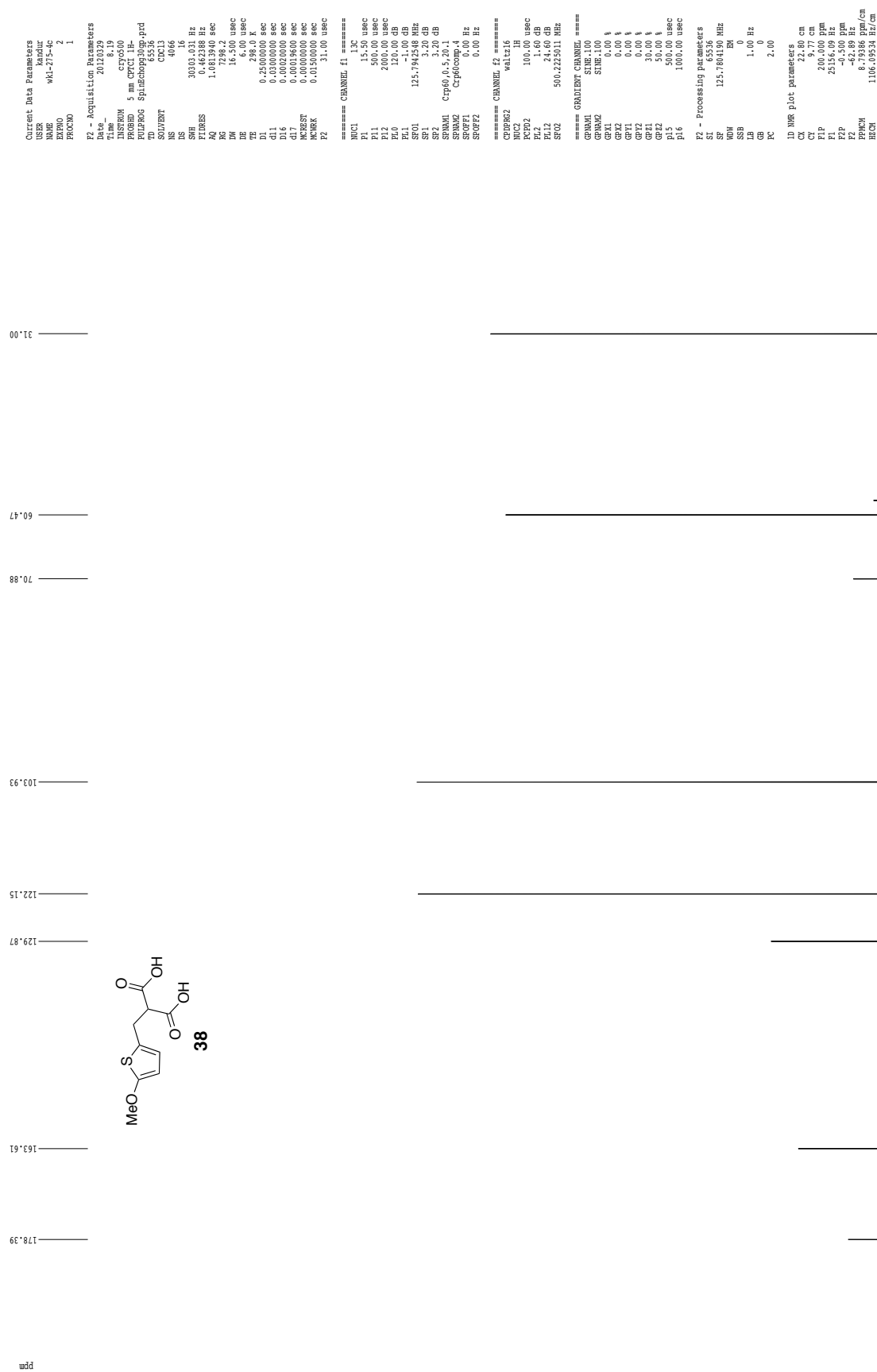
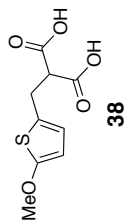


¹H spectrum

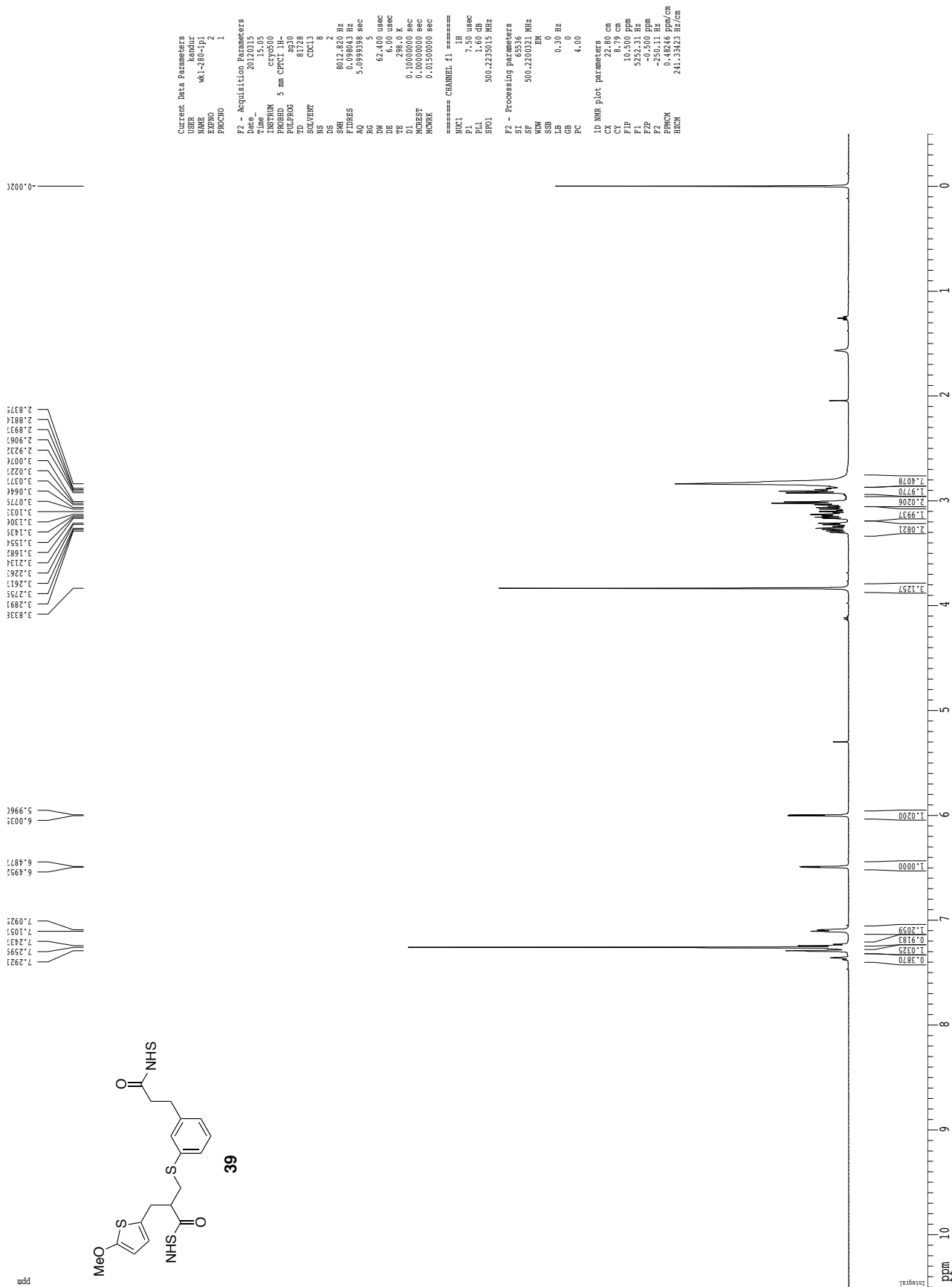


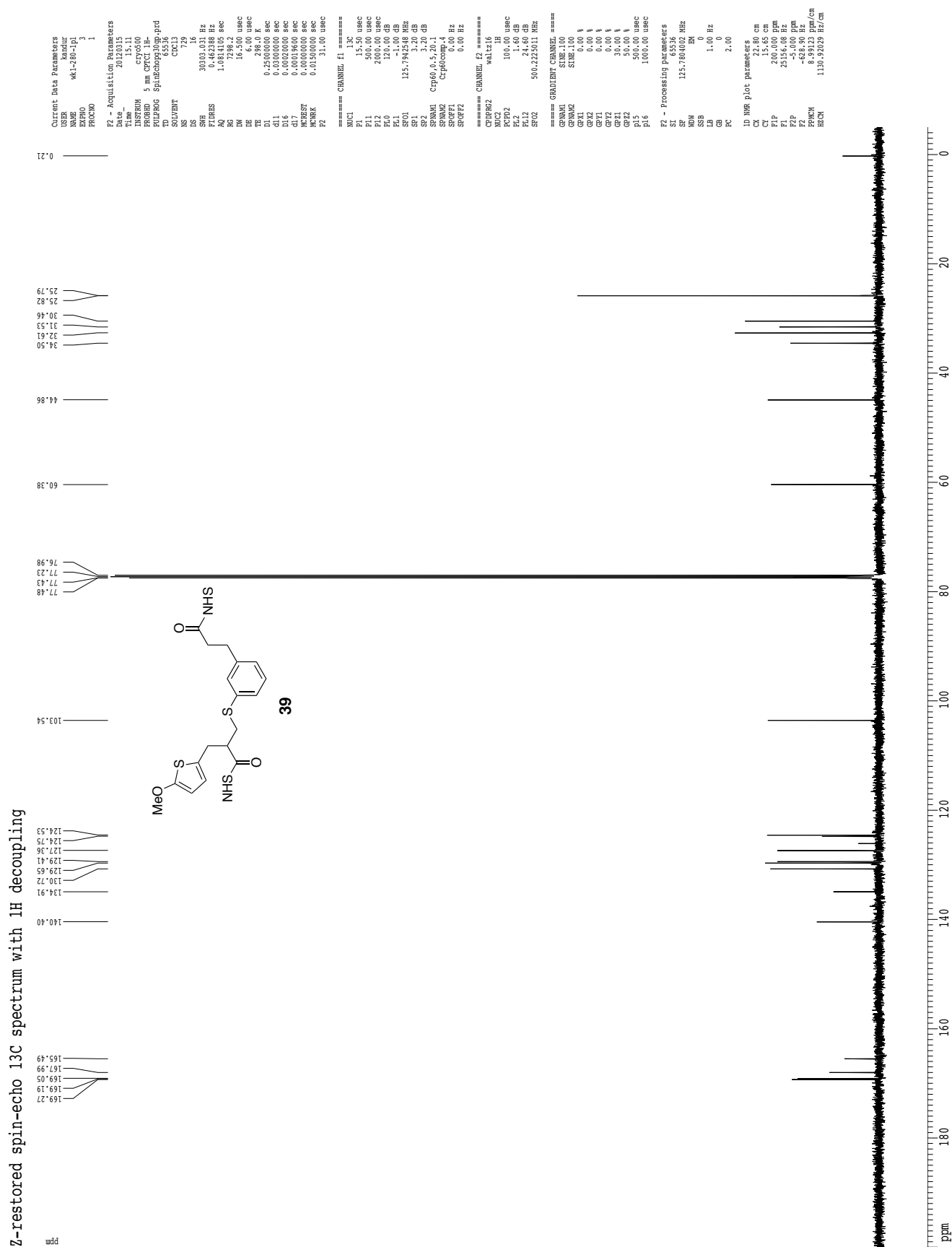
¹H spectrum



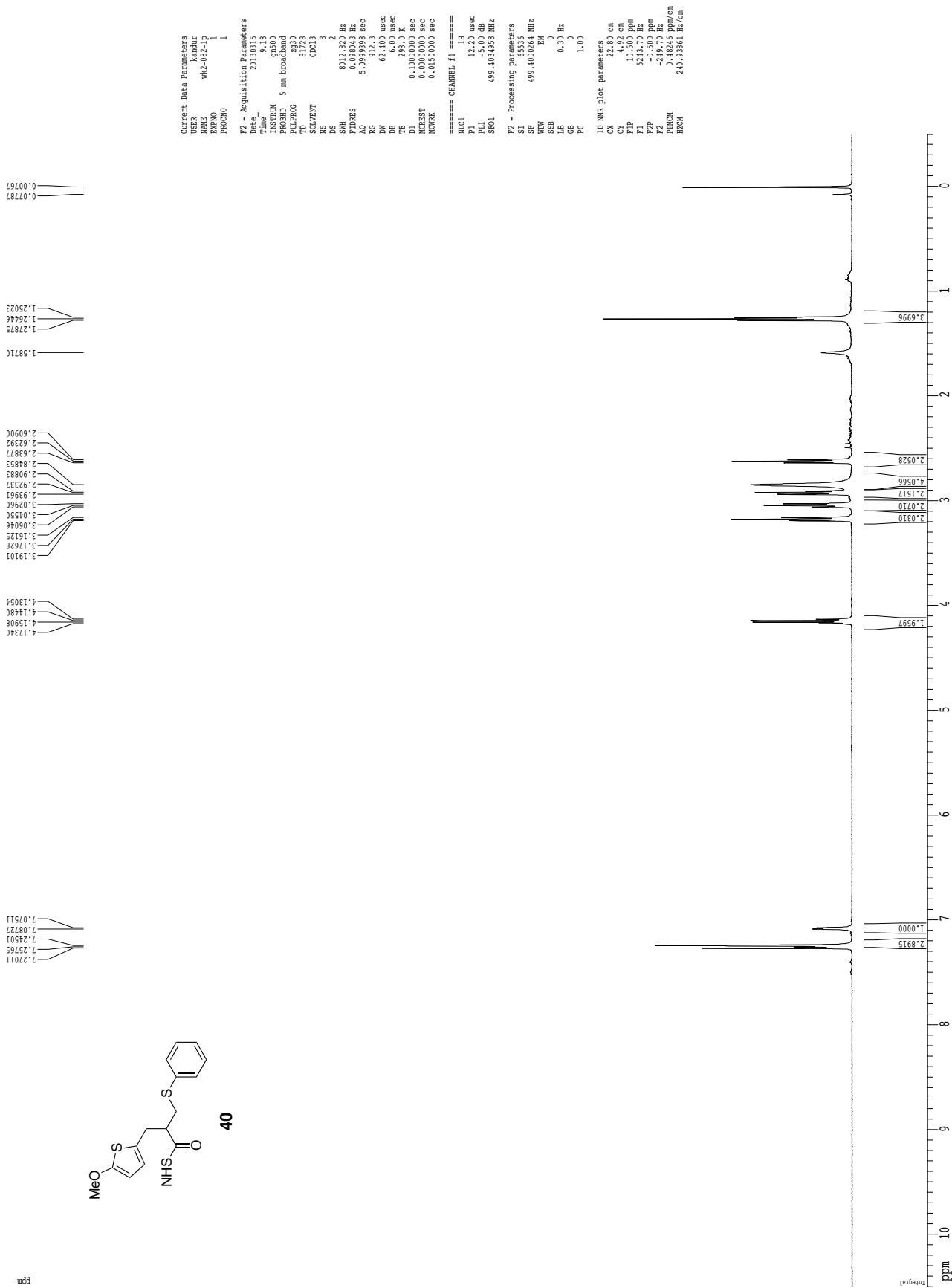
Z-restored spin-echo ^{13}C spectrum with ^1H decoupling

¹H spectrum

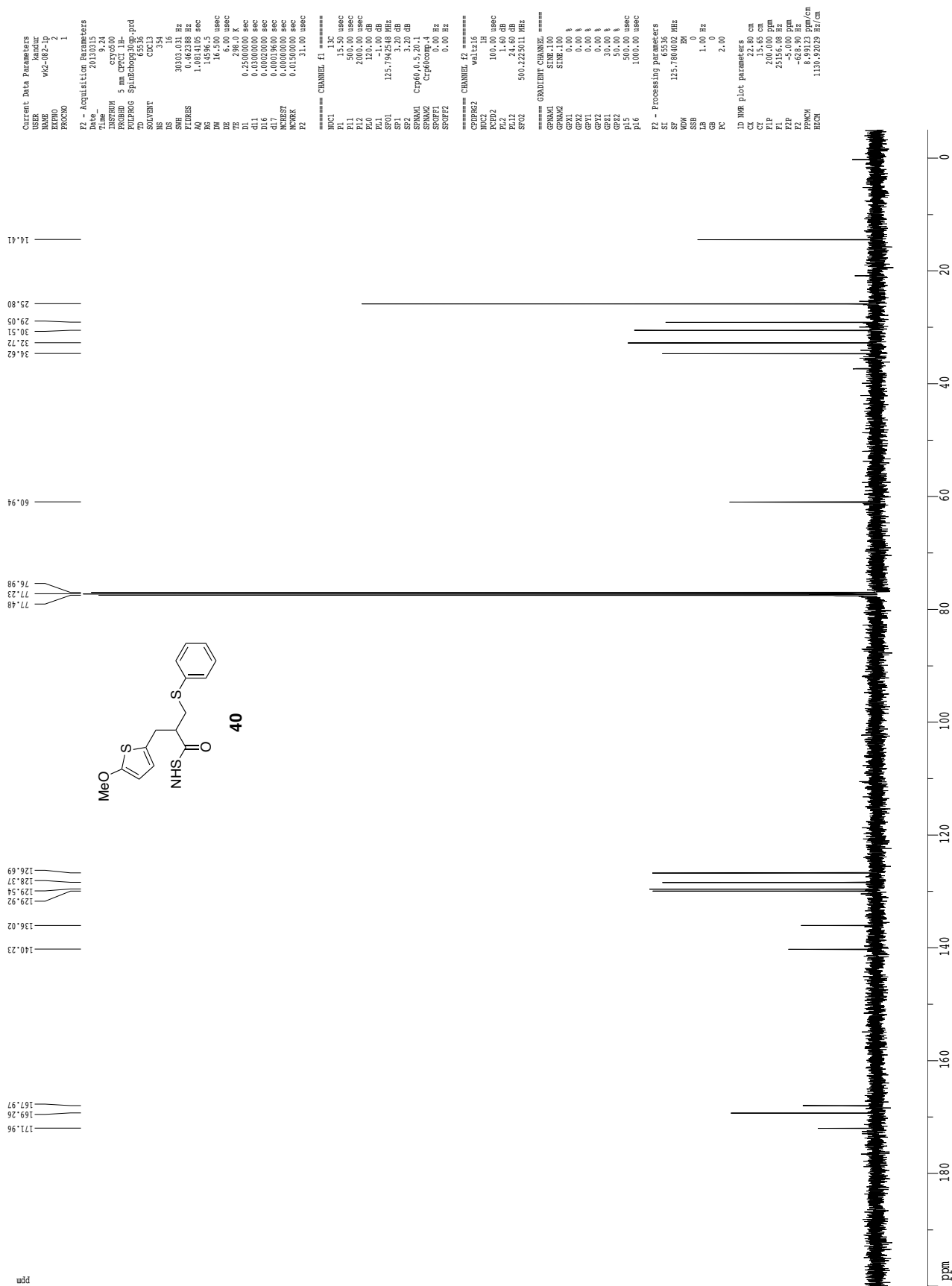




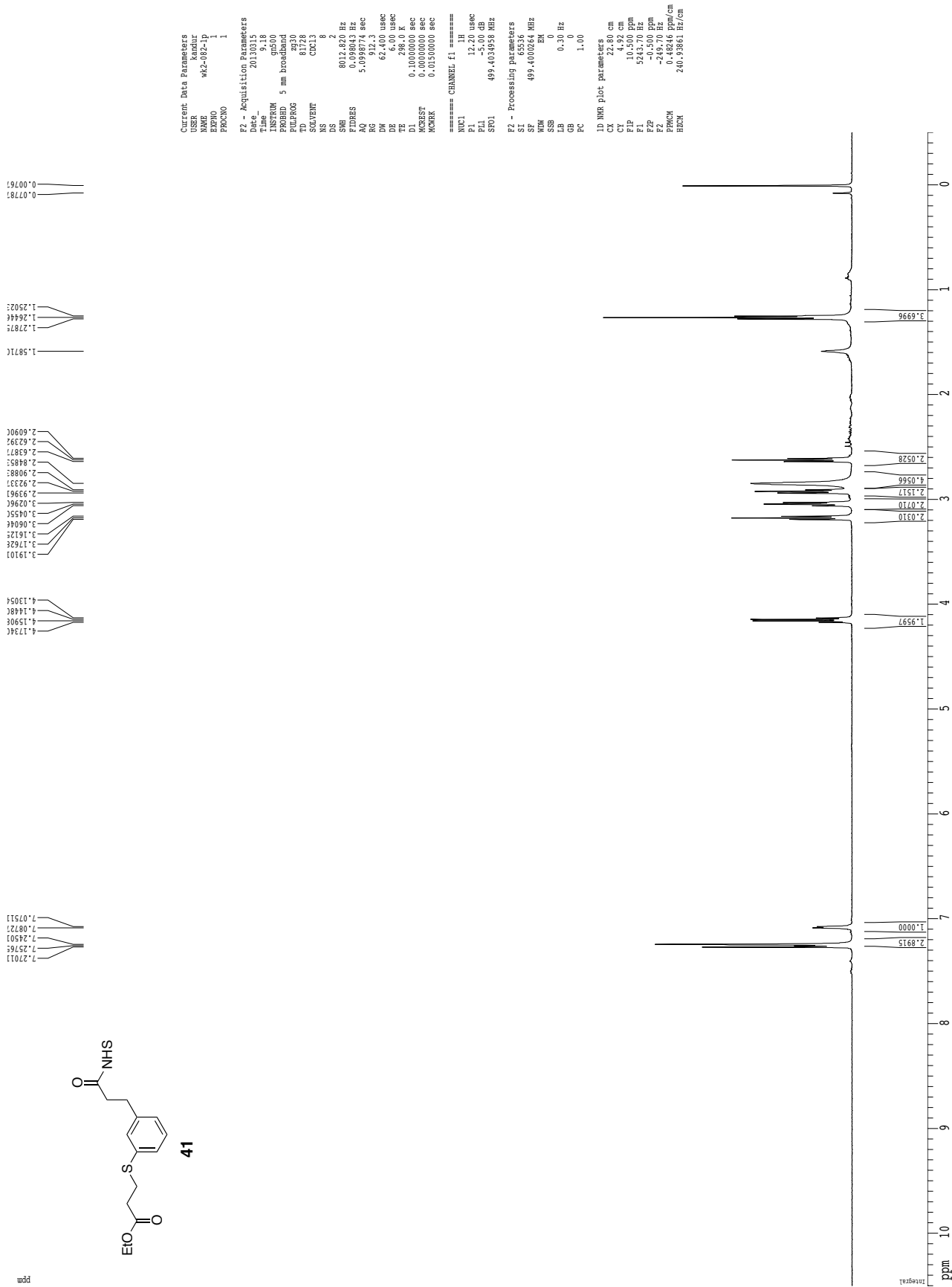
¹H spectrum



Z-restored spin-echo 13C spectrum with 1H decoupling



¹H spectrum



Z-restored spin-echo 13C spectrum with 1H decoupling

