Electronic supplementary information for the manuscript

"Highly selective reactions of $C_{60}Cl_6$ with thiols for synthesis of functionalized [60]fullerene derivatives"

Ekaterina A. Khakina,[a] Anastasiya A. Yurkova,[a] Alexander S. Peregudov,[b] Sergey I. Troyanov,[c] Vyacheslav V. Trush,[d] Andrey I. Vovk,[d] Alexander V. Mumyatov,[a] Vyacheslav M. Martynenko,[a] Jan Balzarini[e] and Pavel A. Troshin[*]

Table of contents:

Selected spectroscopic data:

1a. ^{1}H NMR (600 MHz, $CDCl_3$, δ, ppm): 0.90 (t, 15H), 1.28 (br. s, 80H), 1.47-1.52 (m, 10H), 1.77-1.82 (m, 10H), 3.23-3.29 (m, 8H), 3.35 (t, 2H), 5.12 (s, 1H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 14.17 ($\underline{C}H_3$), 22.73 ($\underline{C}H_2$), 28.56 ($\underline{C}H_2$), 29.20 ($\underline{C}H_2$), 29.24 ($\underline{C}H_2$), 29.28 ($\underline{C}H_2$), 29.38 ($\underline{C}H_2$), 29.42 ($\underline{C}H_2$), 29.48 ($\underline{C}H_2$), 29.54 ($\underline{C}H_2$), 29.63 ($\underline{C}H_2$), 29.65 ($\underline{C}H_2$), 29.71 ($\underline{C}H_2$), 29.76 ($\underline{C}H_2$), 31.96 ($\underline{C}H_2$), 33.44 ($\underline{C}H_2$), 33.68 ($\underline{C}H_2$), 34.31 ($\underline{C}H_2$), 39.23 ($\underline{C}H_2$), 53.89 (sp^3 fullerene cage), 55.82 (sp^3 fullerene cage), 56.49 (sp^3 fullerene cage), 60.38 (sp^3 fullerene cage), 143.10, 143.16, 143.22, 143.27, 143.38, 143.90, 144.04, 144.10, 144.23, 144.36, 144.61, 145.08, 145.17, 146.51, 146.75, 146.80, 147.55, 147.90, 148.07, 148.09, 148.25, 148.46, 148.59, 148.71, 150.15, 151.02, 152.97, 154.17.

ESI MS: m/z=1727 ($[M-H]^-$).

1b. ^{1}H NMR (600 MHz, $CDCl_3$, δ, ppm): 3.73 (s, 6H), 3.79 (s, 9H), 4.06-4.11 (m, 4H), 4.12-4.20 (m, 4H), 4.26 (s, 2H), 5.56 (s, 1H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 35.12 ($\underline{C}H_3$), 35.38 ($\underline{C}H_3$), 36.30 ($\underline{C}H_3$), 52.71 ($\underline{C}H_2$), 52.78 ($\underline{C}H_2$), 52.86 ($\underline{C}H_2$), 53.69 (sp^3 fullerene cage), 55.74 (sp^3 fullerene cage), 56.78 (sp^3 fullerene cage), 60.02 (sp^3 fullerene cage), 142.20, 142.73, 142.88, 143.38, 143.50, 143.87, 143.96, 144.12, 144.23, 144.32, 144.37, 144.53, 144.73, 146.54, 146.75, 146.80, 147.54, 148.00, 148.12, 148.23, 148.32, 148.47, 148.78, 148.88, 149.26, 150.01, 153.02, 153.32, 169.80 ($\underline{C}OOCH_3$), 169.98 ($\underline{C}OOCH_3$), 170.27 ($\underline{C}OOCH_3$).

ESI MS: m/z=1245 ($[M-H]^-$).

1c. ^{1}H NMR (600 MHz, $CDCl_3$, δ, ppm): 1.22-1.25 (m, 6H), 1.26-1.34 (m, 9H), 3.92-4.04 (m, 2H), 4.06 (s, 2H), 4.09-4.12 (m, 2H), 4.15-4.19 (m, 6H), 4.20-4.40 (m, 8H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 14.12 ($\underline{C}H_3$), 14.17 ($\underline{C}H_3$), 14.21 ($\underline{C}H_3$), 35.34 ($\underline{C}H_2$), 35.61 ($\underline{C}H_2$), 36.44 ($\underline{C}H_2$), 53.73 (sp^3 fullerene cage), 55.78 (sp^3 fullerene cage), 56.79 (sp^3 fullerene cage), 60.08 (sp^3 fullerene cage), 61.82 ($O\underline{C}H_2$), 61.89 ($O\underline{C}H_2$), 61.95 ($O\underline{C}H_2$), 142.31, 142.82, 142.89, 143.37, 143.58, 143.97, 144.07, 144.12, 144.22, 144.30, 144.46, 144.52, 144.80, 146.54, 146.76, 146.81, 147.54, 147.99, 148.12, 148.22, 148.31, 148.46, 148.77, 148.87, 149.39, 150.15, 153.14, 153.37, 169.34 ($\underline{C}OOCH_2CH_3$), 169.53 ($\underline{C}OOCH_2CH_3$), 169.79 ($\underline{C}OOCH_2CH_3$).

ESI MS: m/z=1315 ($[M-H]^-$).

1d. ^{1}H NMR (600 MHz, $CDCl_3$, δ, ppm): 0.87-0.95 (m, 30H), 1.12-1.23 (m, 5H), 1.36-1.48 (m, 5H), 1.64-1.78 (m, 5H), 3.60 (s, 1H), 3.63 (d, 1H), 3.87-4.20 (m, 16H), 4.26 (s, 2H), 5.55(s, 1H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 11.22 ($\underline{C}H_3$), 11.27 ($\underline{C}H_3$), 16.37 ($\underline{C}H_3$), 16.41 ($\underline{C}H_3$), 25.94 ($\underline{C}H_2CH_3$), 25.97 ($\underline{C}H_2CH_3$), 34.04 ($\underline{C}H$), 34.07 ($\underline{C}H$), 34.08 ($\underline{C}H$), 35.38 ($S\underline{C}H_2$),

35.63 ($S\underline{C}H_2$), 36.43 ($S\underline{C}H_2$), 53.72 (sp^3 fullerene cage), 55.76 (sp^3 fullerene cage), 56.85 (sp^3 fullerene cage), 60.08 (sp^3 fullerene cage), 70.44 ($O\underline{C}H_2$), 70.51 ($O\underline{C}H_2$), 70.60 ($O\underline{C}H_2$), 142.34, 142.80, 142.94, 143.10, 143.35, 143.57, 143.95, 144.09, 144.21, 144.31, 144.44, 144.51, 144.80, 146.53, 146.76, 146.80, 147.51, 147.98, 148.09, 148.20, 148.29, 148.43, 148.75, 148.85, 149.40, 150.15, 153.17, 153.29, 169.46 ($\underline{C}OOCH_2CH$), 169.61 ($\underline{C}OOCH_2CH$), 169.96 ($\underline{C}OOCH_2CH$).

ESI MS: m/z=1526 ($[M-H]^-$).

1e. ^1H NMR (600 MHz, $CDCl_3$, δ, ppm): 2.76 (t, 2H), 2.86 (m, 6H), 2.94 (t, 2H), 3.50 (m, 6H), 3.61 (t, 2H), 3.73 (m, 17H), 5.18 (s,1H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 27.95 ($\underline{C}H_2$), 28.19 ($\underline{C}H_2$), 33.08 ($\underline{C}H_2$), 33.90 ($\underline{C}H_2$), 34.41 ($\underline{C}H_2$), 34.95 ($\underline{C}H_2$), 51.93 ($O\underline{C}H_3$), 51.97 ($O\underline{C}H_3$), 51.99 ($O\underline{C}H_3$), 53.81 (sp^3 fullerene cage), 55.80 (sp^3 fullerene cage), 56.58 (sp^3 fullerene cage), 60.27 (sp^3 fullerene cage), 128.24, 129.05, 142.66, 142.90, 143.10, 143.66, 143.98, 144.23, 144.29, 144.36, 144.52, 144.55, 144.80, 146.50, 146.73, 146.79, 147.57, 147.97, 148.12, 148.22, 148.28, 148.48, 148.72, 148.84, 149.68, 150.57, 153.04, 153.70, 171.89 ($\underline{C}OOCH_3$), 171.91 ($\underline{C}OOCH_3$), 171.93 ($\underline{C}OOCH_3$).

ESI MS: m/z=1315 ($[M-H]^-$).

1f. ^1H NMR (600 MHz, $CDCl_3$, δ, ppm): 1.42-1.46 (m, 45H), 3.40-3.57 (m, 20H), 5.18 (s, 1H), 5.33-5.50 (m, 5H).

^{13}C NMR (150 MHz, $CDCl_3$, δ, ppm): 28.47 ($\underline{C}H_3$), 28.59 ($\underline{C}H_3$), 28.67 ($\underline{C}H_3$), 33.45 ($\underline{C}H_2$), 33.73 ($\underline{C}H_2$), 34.39 ($\underline{C}H_2$), 40.07 ($\underline{C}H_2$), 40.30 ($\underline{C}H_2$), 40.39 ($\underline{C}H_2$), 53.58 (sp^3 fullerene cage), 55.56 (sp^3 fullerene cage), 56.37 (sp^3 fullerene cage), 60.39 (sp^3 fullerene cage), 79.47 ($O\underline{C}(CH_3)_3$), 79.54 ($O\underline{C}(CH_3)_3$), 79.61 ($O\underline{C}(CH_3)_3$), 142.67, 142.89, 143.11, 143.38, 143.65, 143.99, 144.20, 144.27, 144.36, 144.47, 144.50, 144.80, 146.49, 146.73, 146.77, 147.54, 147.94, 148.09, 148.20, 148.28, 148.46, 148.71, 148.83, 149.74, 150.53, 153.14, 153.73, 155.68 ($\underline{C}OOC(CH_3)_3$), 155.79 ($\underline{C}OOC(CH_3)_3$), 155.85 ($\underline{C}OOC(CH_3)_3$).

ESI MS: m/z=1601 ($[M-H]^-$).

1g. ^1H NMR (600 MHz, $(CD_3)_2CO$:CS_2 (1:1), δ, ppm): 4.06-4.19 (m, 6H), 4.24-4.29 (m, 2H), 4.35 (s, 2H).

^{13}C NMR (150 MHz, $(CD_3)_2CO$:CS_2 (1:1), δ, ppm): 35,16 ($\underline{C}H_2$), 35.47 ($\underline{C}H_2$), 36.41 ($\underline{C}H_2$), 53.63 (sp^3 fullerene cage), 55.78 (sp^3 fullerene cage), 56.71 (sp^3 fullerene cage), 142.96, 143.13, 143.20, 143.34, 143.81, 144.09, 144.17, 144.21, 144.25, 144.49, 144.69, 145.12, 146.59, 146.80, 146.87, 147.02, 147.57, 148.00, 148.12, 148.32, 148.47, 148.67, 148.77, 150.03, 150.64, 153.54, 153.70, 169.60 ($\underline{C}OOH$), 169.88 ($\underline{C}OOH$), 170.01 ($\underline{C}OOH$).

ESI MS: m/z=1175 ($[M-H]^-$); 1083 ($[M-RSH]^-$).

1h. ^1H NMR (600 MHz, $(CD_3)_2CO$:CS_2 (1:1), δ, ppm): 2.69 (t, 1H), 2.78-2.86 (m, 8H), 2.92 (t, 1H), 3.46-3.55 (m, 7H), 3.59 (t, 1H), 3.65 (t, 2H), 5.43 (s, 1H).

^{13}C NMR (150 MHz, $(CD_3)_2CO:CS_2$ (1:1), δ, ppm): 33.49 ($\underline{C}H_2$), 33.53 ($\underline{C}H_2$), 34.04 ($\underline{C}H_2$), 34.48 ($\underline{C}H_2$), 34.57 ($\underline{C}H_2$), 53.98 (sp^3 fullerene cage), 55.98 (sp^3 fullerene cage), 56.82 (sp^3 fullerene cage), 60.35 (sp^3 fullerene cage), 143.16, 143.21, 143.37, 143.90, 144.21, 144.23, 144.27, 144.50, 144.63, 145.16, 145.19, 146.57, 146.79, 146.86, 147.64, 148.01, 148.15, 148.17, 148.34, 148.53, 148.65, 148.77, 150.35, 151.12, 153.55, 154.20, 171.91 ($\underline{C}OOH$), 171.99 ($\underline{C}OOH$), 172.02 ($\underline{C}OOH$).

ESI MS: m/z=1245 ($[M-H]^-$); 1140 ($[M-RSH]^-$); 622 ($[M-2H]^{2-}$), 613.5 ($[M-H_2O-H]^{2-}$).

1i. ^{1}H NMR (600 MHz, $(CD_3)_2CO:CS_2$ (1:1), δ, ppm): 1.26-1.36 (m, 50H), 1.54-1.60 (m, 20H), 1.80-1.87 (m, 10H), 2.25-2.27 (m, 10H), 3.23-3.34 (m, 10H), 5.15 (s, 1H).

^{13}C NMR (150 MHz, $(CD_3)_2CO:CS_2$ (1:1), δ, ppm): 25.19 ($\underline{C}H_2$), 25.22 ($\underline{C}H_2$), 29.52 ($\underline{C}H_2$), 29.63 ($\underline{C}H_2$), 29.80 ($\underline{C}H_2$), 30.00 ($\underline{C}H_2$), 30.05 ($\underline{C}H_2$), 30.09 ($\underline{C}H_2$), 33.6 ($\underline{C}H_2$), 33.71 ($\underline{C}H_2$), 33.87 ($\underline{C}H_2$), 53.99 (sp^3 fullerene cage), 55.91 (sp^3 fullerene cage), 56.59 (sp^3 fullerene cage), 60.50 (sp^3 fullerene cage), 143.19, 143.22, 143.28, 143.60, 144.01, 144.12, 144.17, 144.34, 144.39, 144.43, 144.69, 145.24, 145.34, 146.77, 146.85, 147.60, 147.95, 148.12, 148.30, 148.51, 148.59, 148.71, 150.35, 151.22, 153.15, 154.34, 174.20 ($\underline{C}OOH$), 174.26 ($\underline{C}OOH$).

ESI MS: m/z=1807 ($[M-H]^-$).

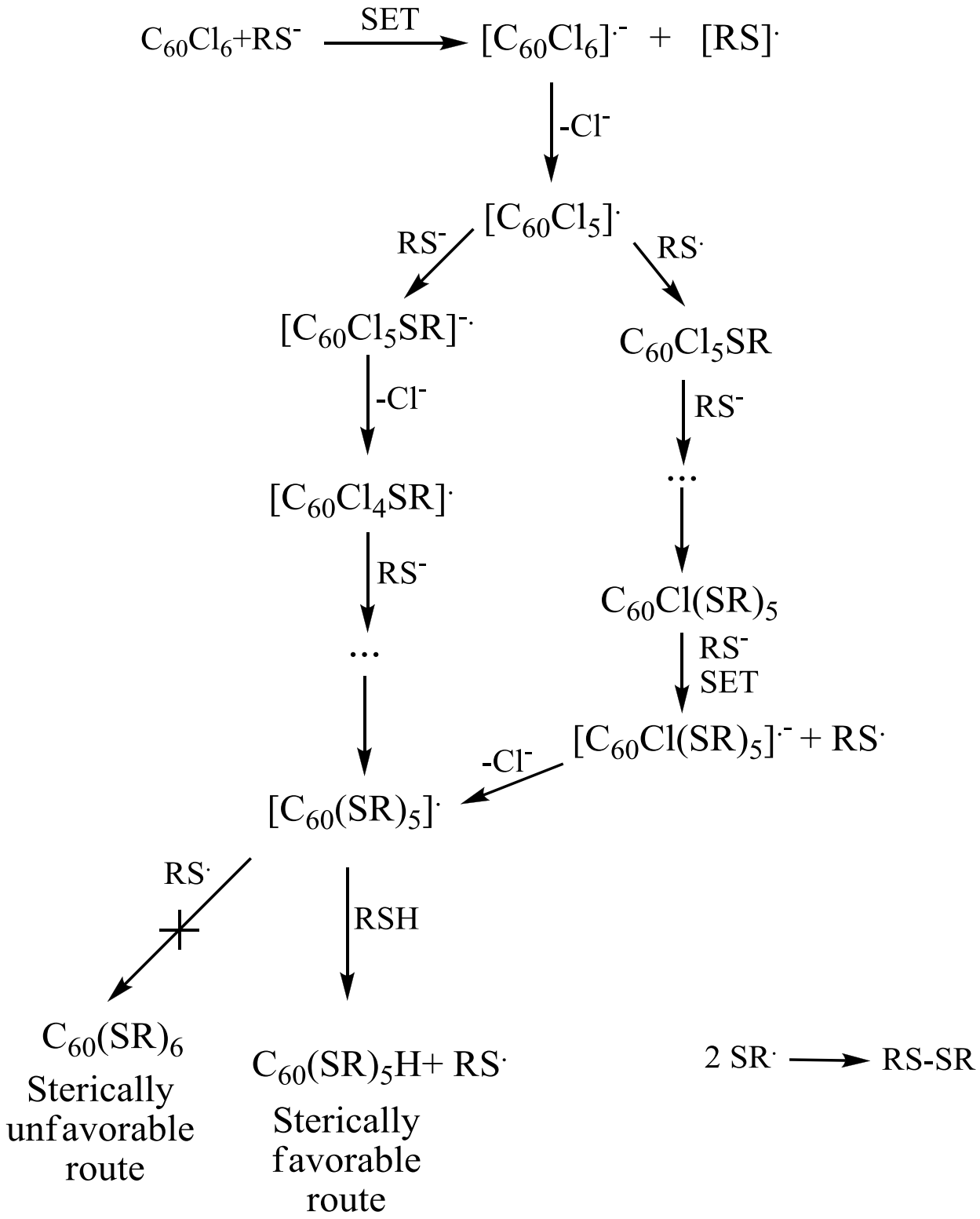
2e. ^{1}H NMR (600 MHz, $CDCl_3$, δ, ppm): 3.05-3.11 (m, 4H), 3.82 (s, 6H), 3.84-3.92 (m, 4H).

^{13}C NMR (150 MHz, $(CD_3)_2CO:CS_2$ (1:1), δ, ppm): 28.06 ($\underline{C}H_2$), 35.38 ($\underline{C}H_2$), 51.58 ($O\underline{C}H_3$), 57.21 (sp^3 fullerene cage), 128.44, 137.80, 138.87, 141.21, 141.78, 142.06, 142.52, 142.74, 142.82, 143.29, 143.49, 143.51, 143.55, 143.97, 144.14, 144.41, 144.44, 144.49, 144.51, 144.61, 144.69, 145.36, 145.79, 145.96, 147.07, 147.17, 147.35, 147.69, 148.83, 150.06, 154.06, 170.58 ($\underline{C}OOCH_3$).

ESI MS: m/z=984 ($[M+CN]^-$).

X-ray crystallography for 1c·

Data collection for single crystal of **1c** (0.52 × 0.33 × 0.16 mm^3) was carried out with an IPDS diffratometer (Stoe) at 100 K (λ = 0.71073 Å). The structure was solved using direct methods (SHELXS97) and anisotropically refined against $|F^2|$ with SHELXL97. Absorption correction was not applied. Crystal data for **1c**: $C_{80}H_{36}O_{10}S_5$, M = 1317.39, triclinic, $P\ \bar{1}$, a = 10.3484(3), b = 12.9522(3), c = 20.1312(5) Å, α = 90.823(2), β = 92.058(2), γ = 91.539(2)°, V = 2695.25(12) $Å^3$, Z = 2, D_{calc} = 1.623 g cm^{-3}. Anisotropic refinement with 14465 reflections and 911 parameters yielded a conventional R_1 = 0.053 for 12359 reflections with $I > 2\sigma(I)$ and wR_2 = 0.141 for all reflections. All methylene and methyl hydrogen atoms were placed into geometrically calculated positions and refined in the riding mode. The hydrogen atom attached to the fullerene cage was found in a difference Fourier map and refined isotropically. One sulfide group and one OEt group are orientationally disordered between two positions. For more details see CCDC – 875844.



Scheme 1. Tentative mechanism of the reaction of $C_{60}Cl_6$ with thiols.

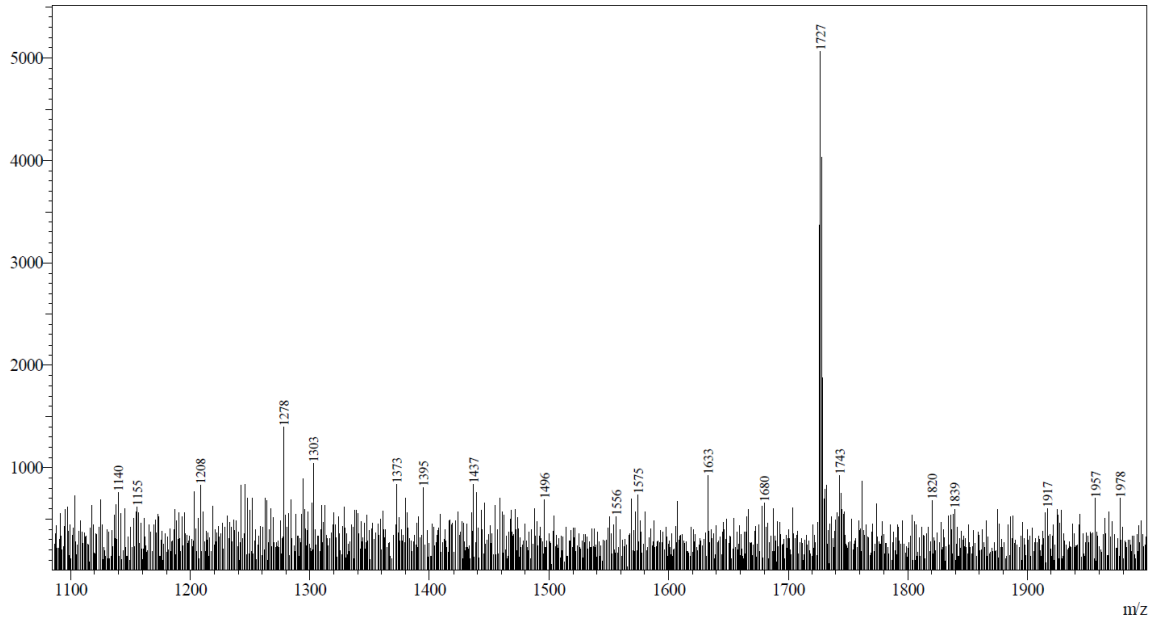


Fig. S1. ESI MS spectrum of compound **1a**

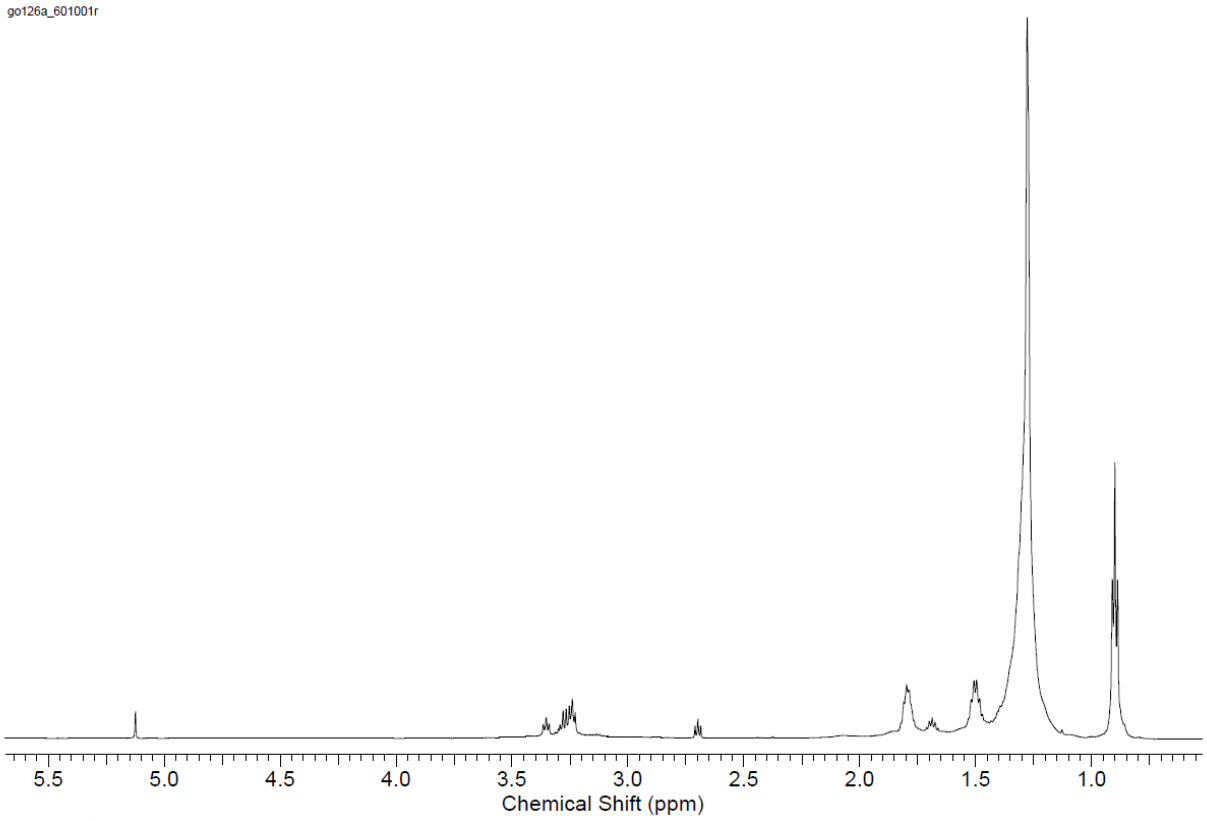


Fig. S2. ^{1}H NMR spectrum of compound **1a**

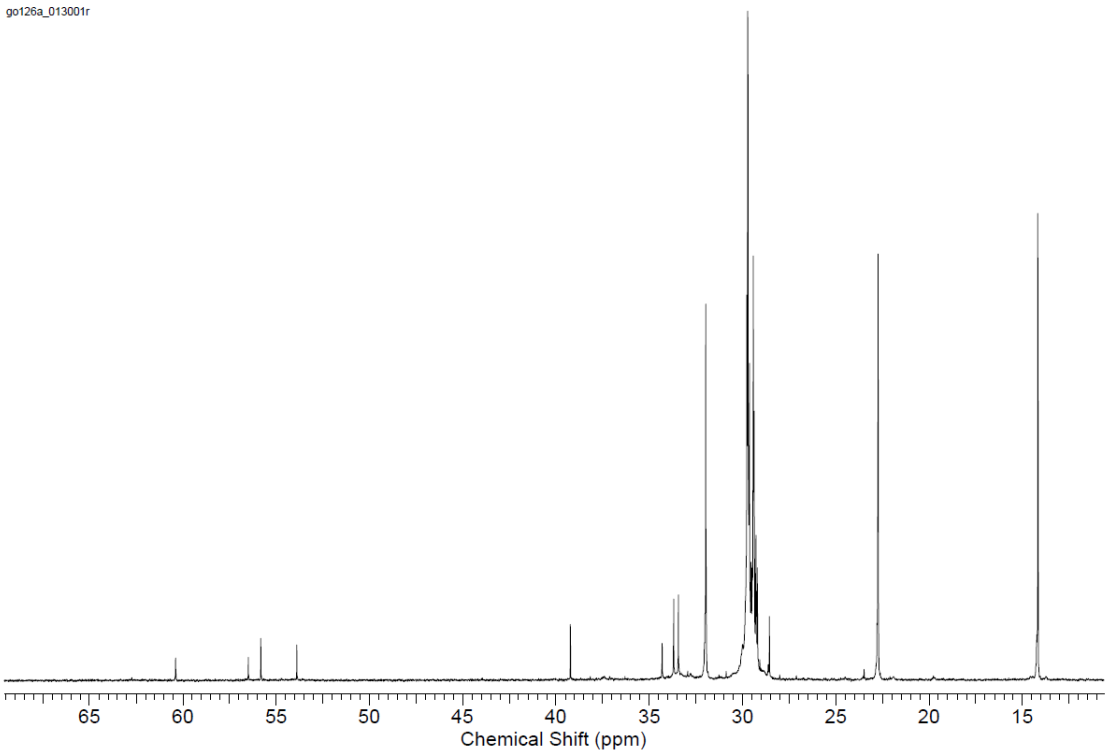


Fig. S3. High-field part of the ^{13}C NMR spectrum of compound **1a**

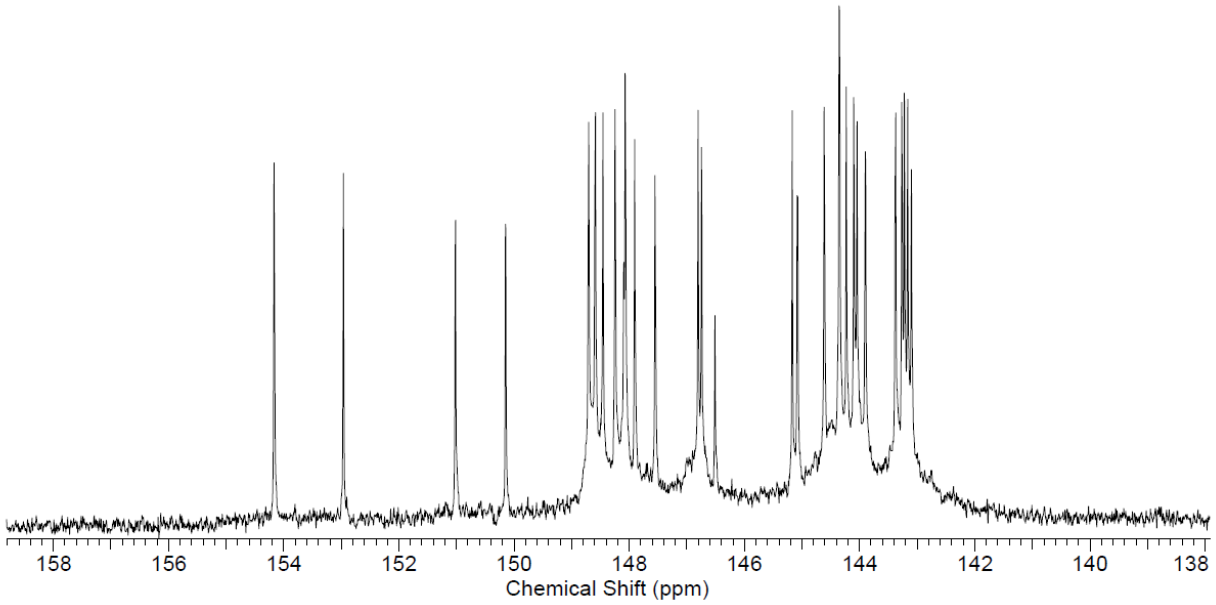


Fig. S4. Low-field part of the ^{13}C NMR spectrum of compound **1a**

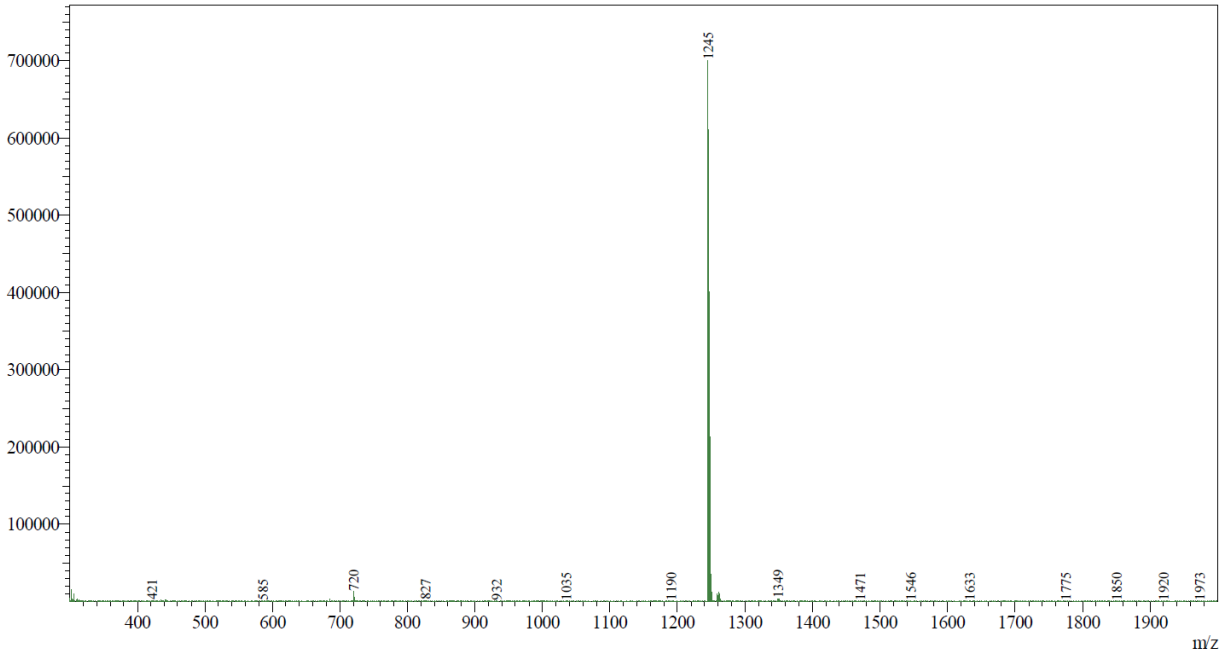


Fig. S5. ESI mass spectrum of compound **1b**

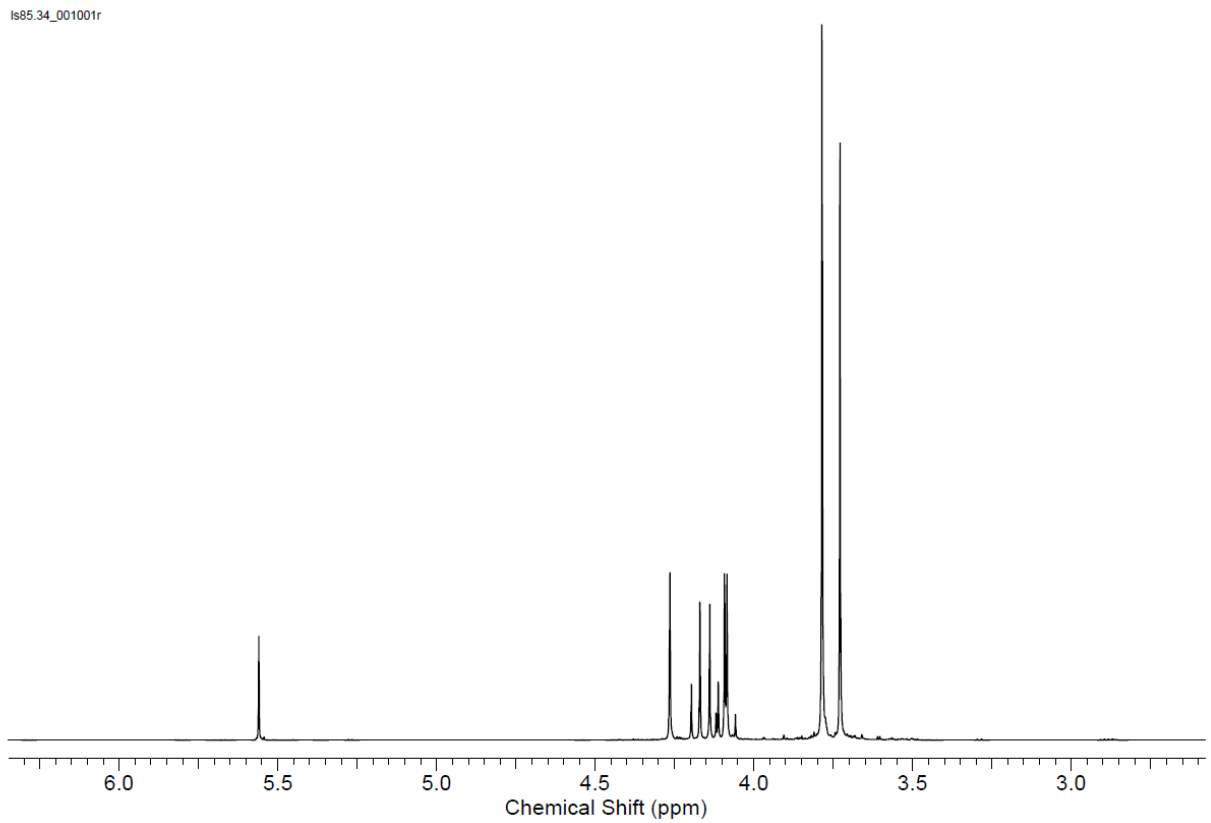


Fig. S6. ^{1}H NMR spectrum of compound **1b**

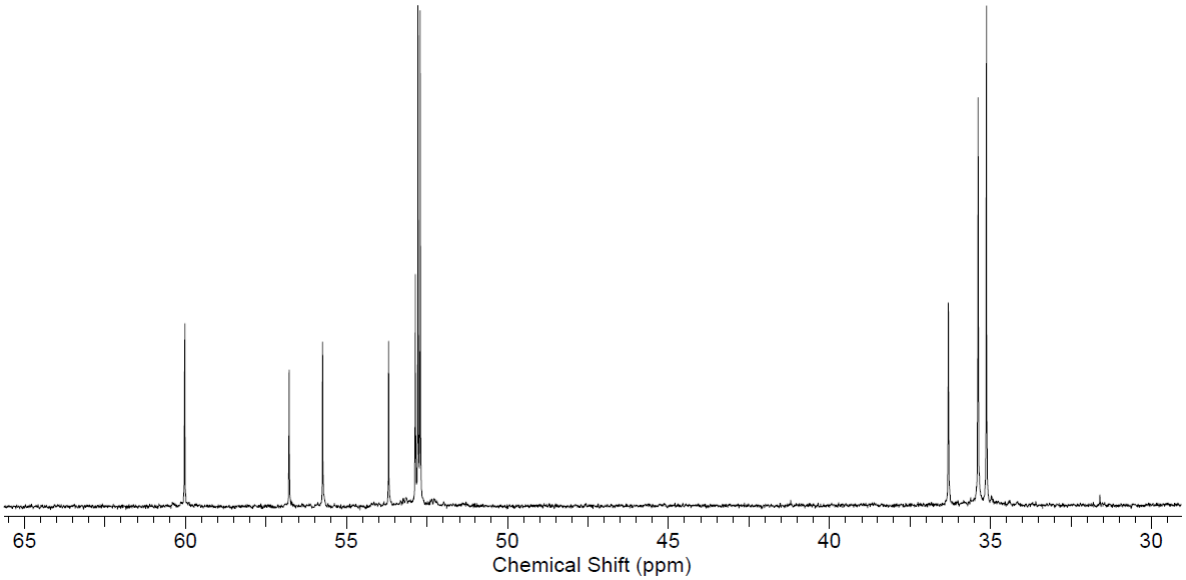


Fig. S7. High-field part of the ^{13}C NMR spectrum of compound **1b**

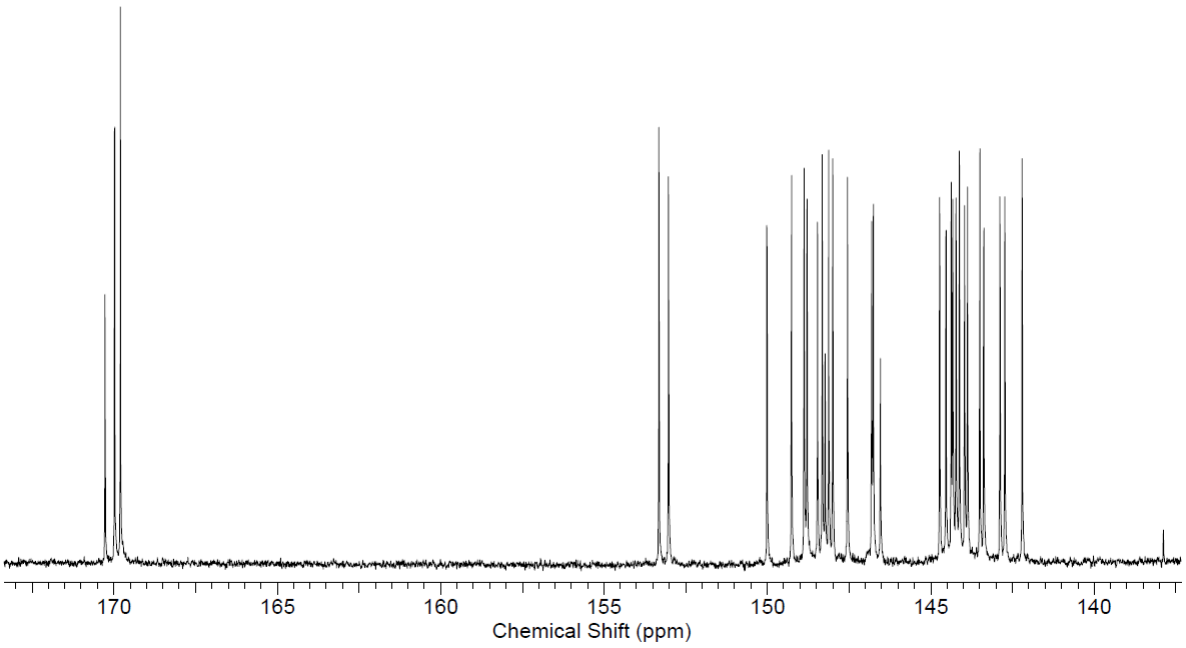


Fig. S8. Low-field part of the ^{13}C NMR spectrum of compound **1b**

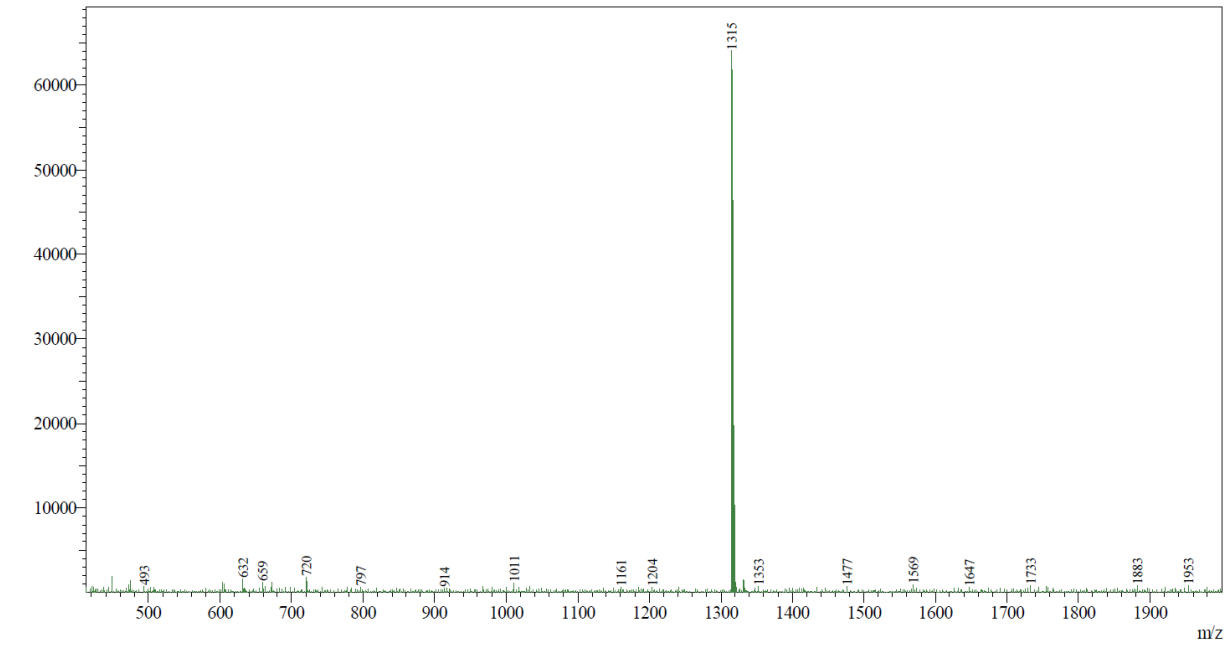


Fig. S9. ESI mass spectrum of compound **1c**

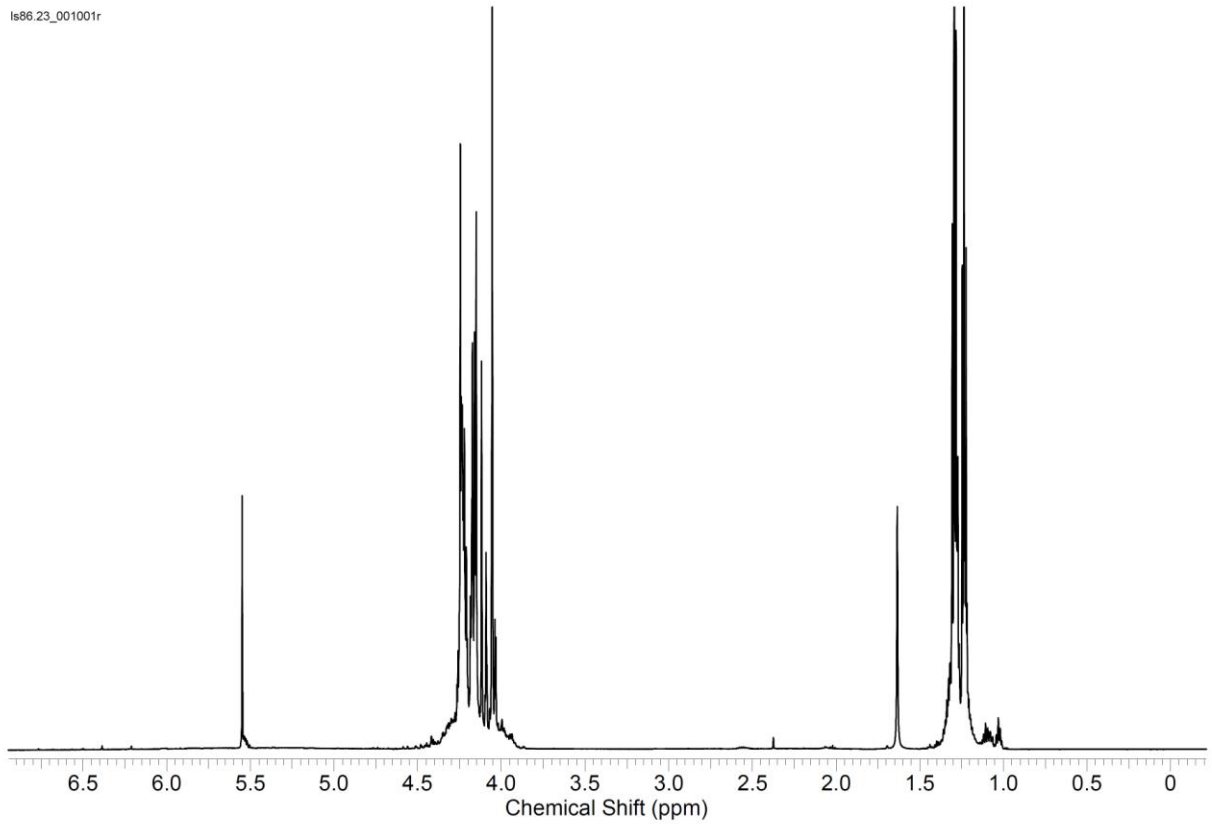


Fig. S10. ^{1}H NMR spectrum of compound **1c**

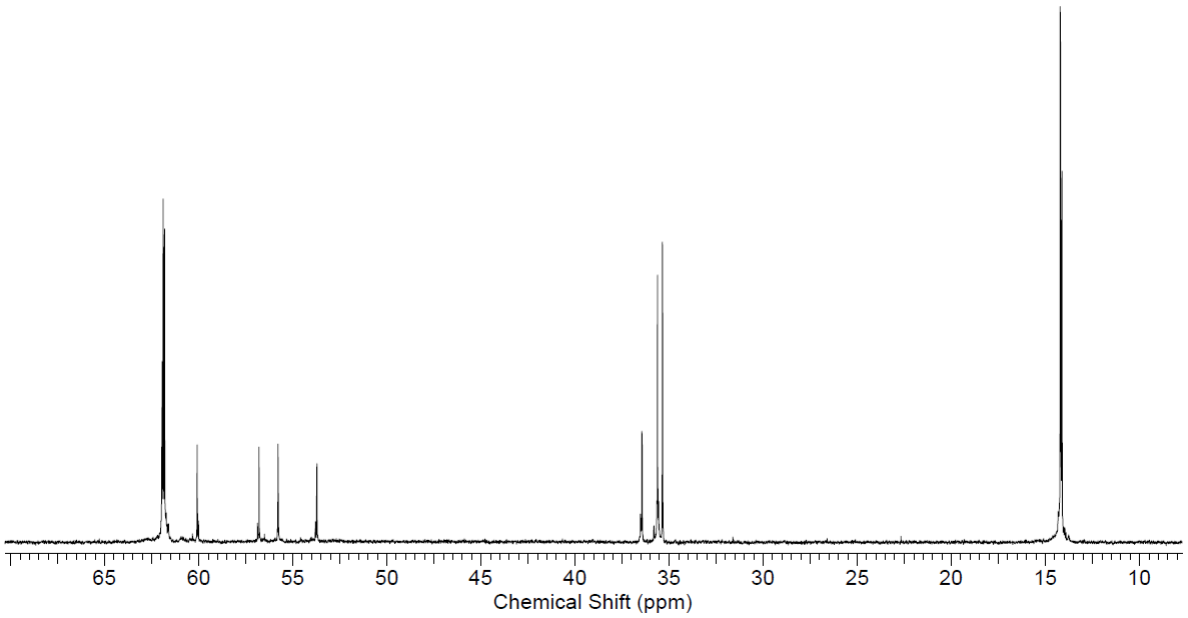


Fig. S11. High-field part of the ^{13}C NMR spectrum of compound **1c**

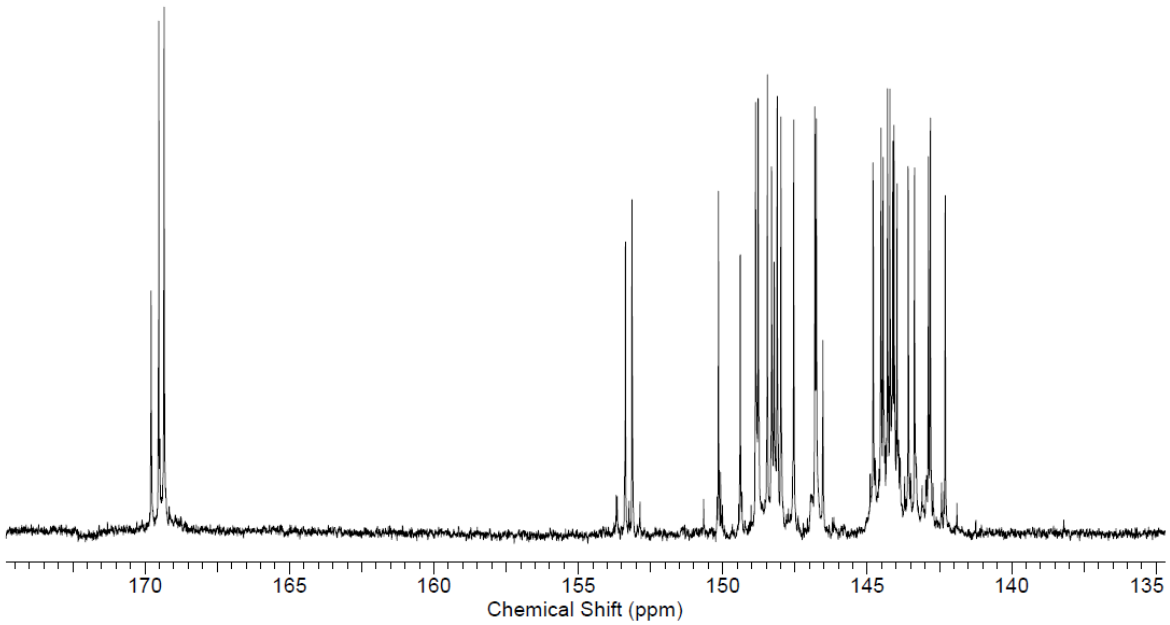


Fig. S12. Low-field part of the ^{13}C NMR spectrum of compound **1c**

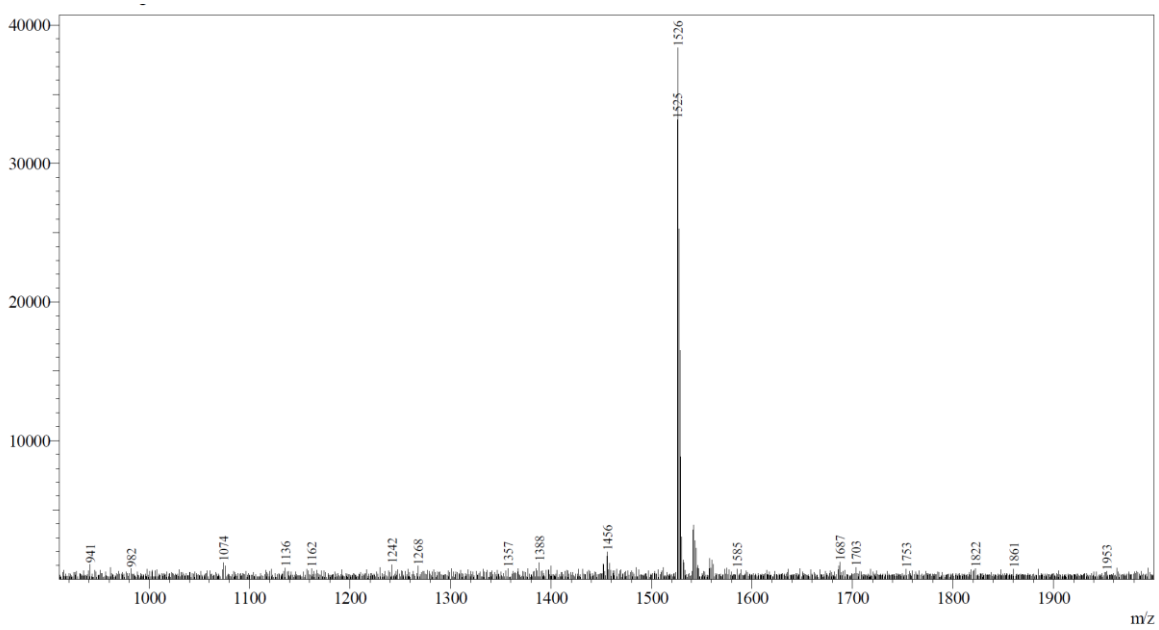


Fig. S13. ESI mass spectrum of compound **1d**

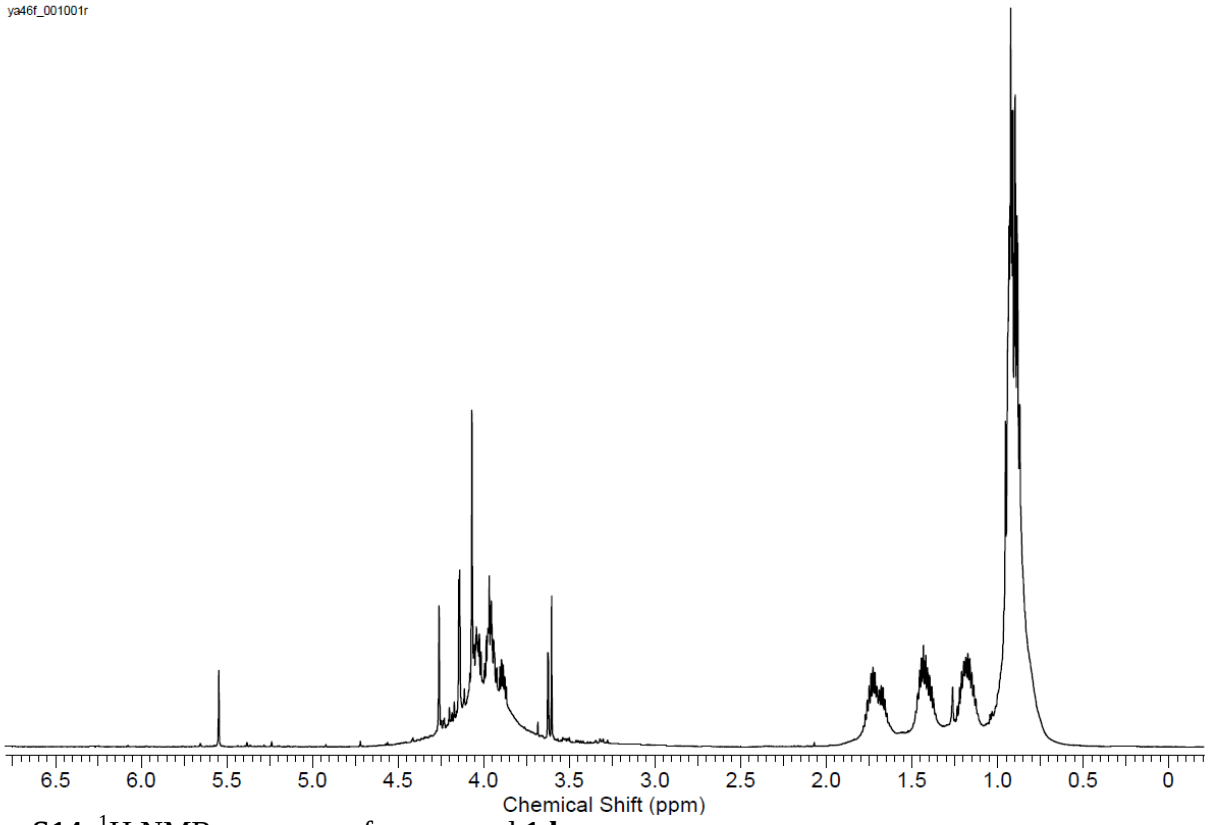


Fig. S14. ^{1}H NMR spectrum of compound **1d**

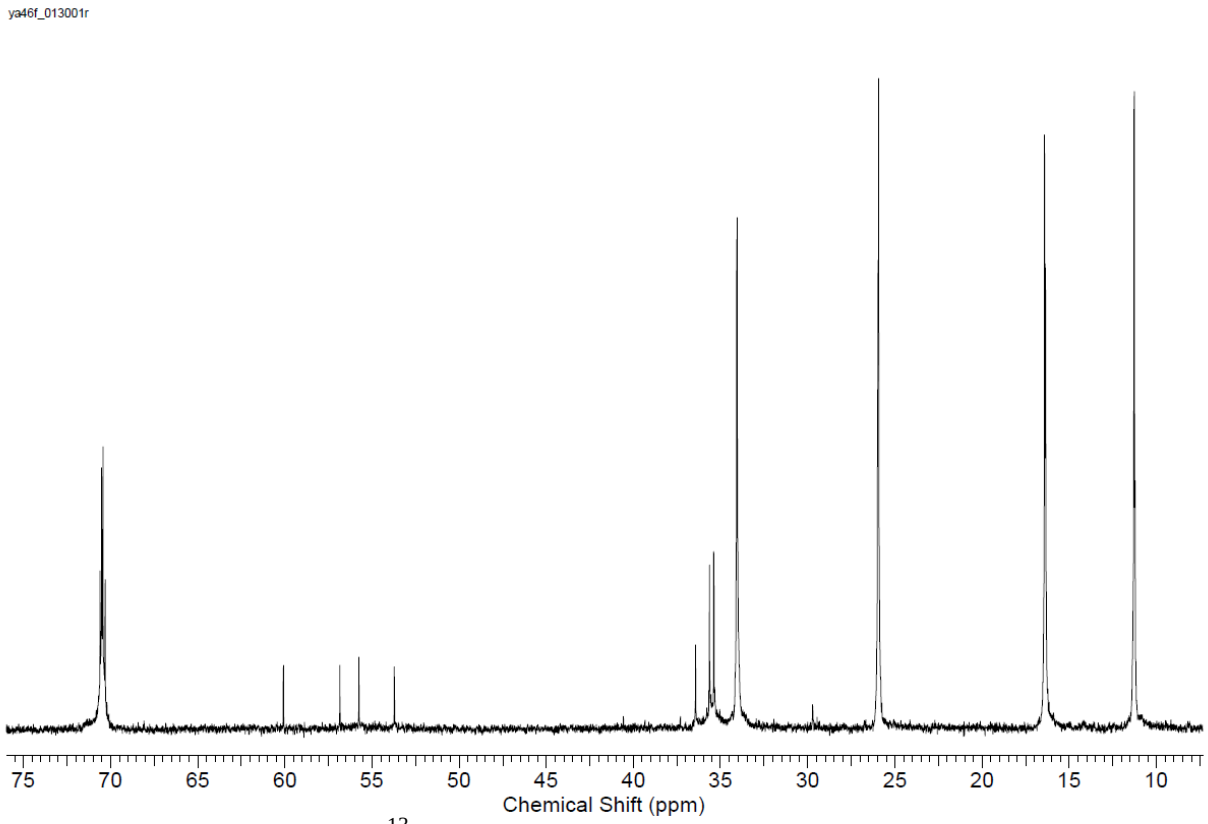


Fig. S15. High-field part of the ^{13}C NMR spectrum of compound **1d**

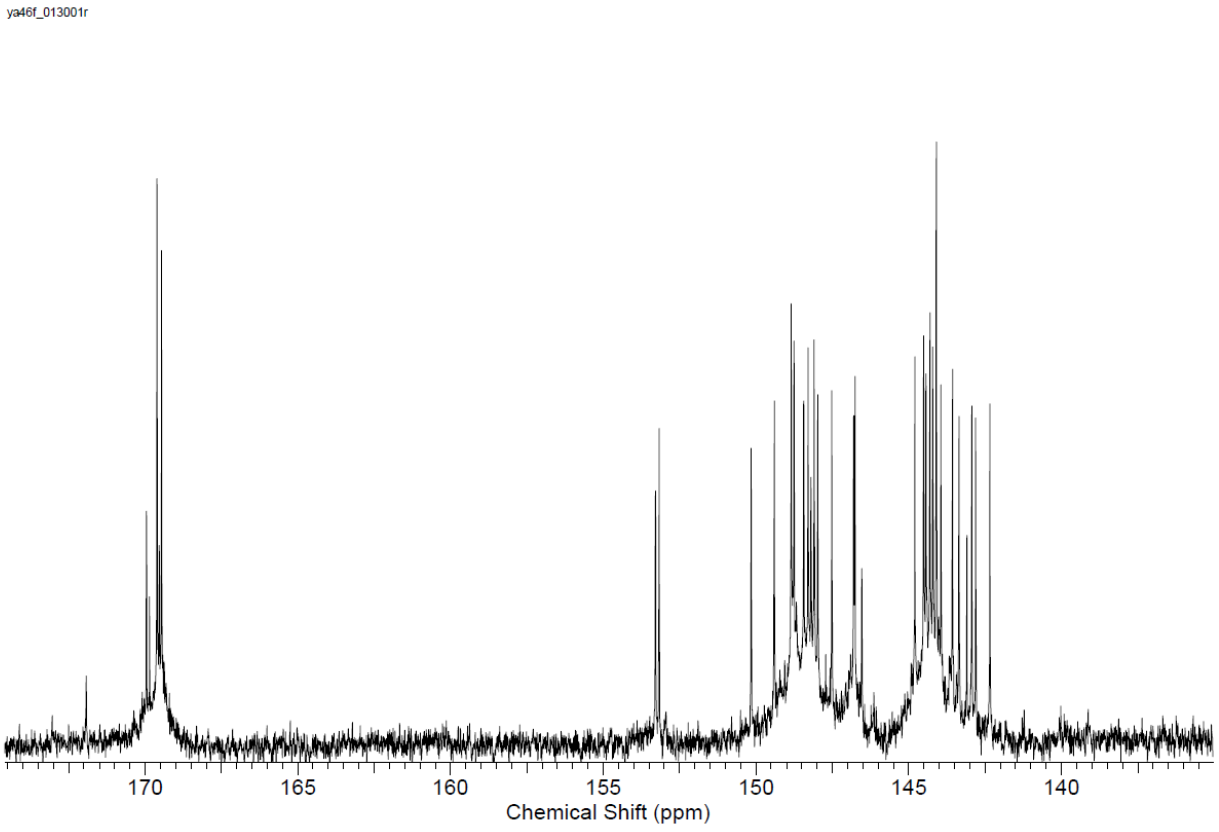


Fig. S16. Low-field part of the ^{13}C NMR spectrum of compound **1d**

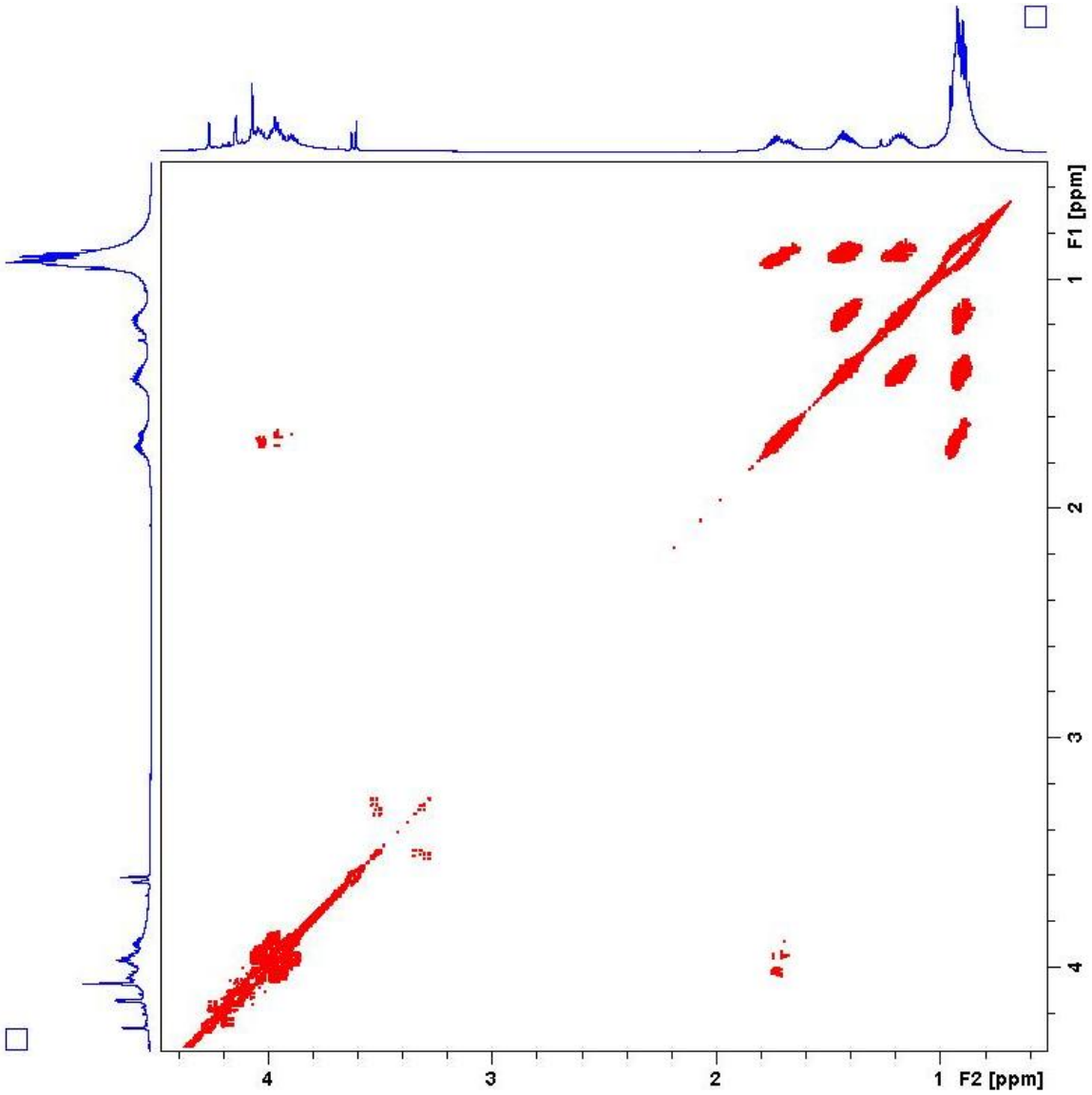


Fig. S17. H-H COSY NMR spectrum of compound **1d**

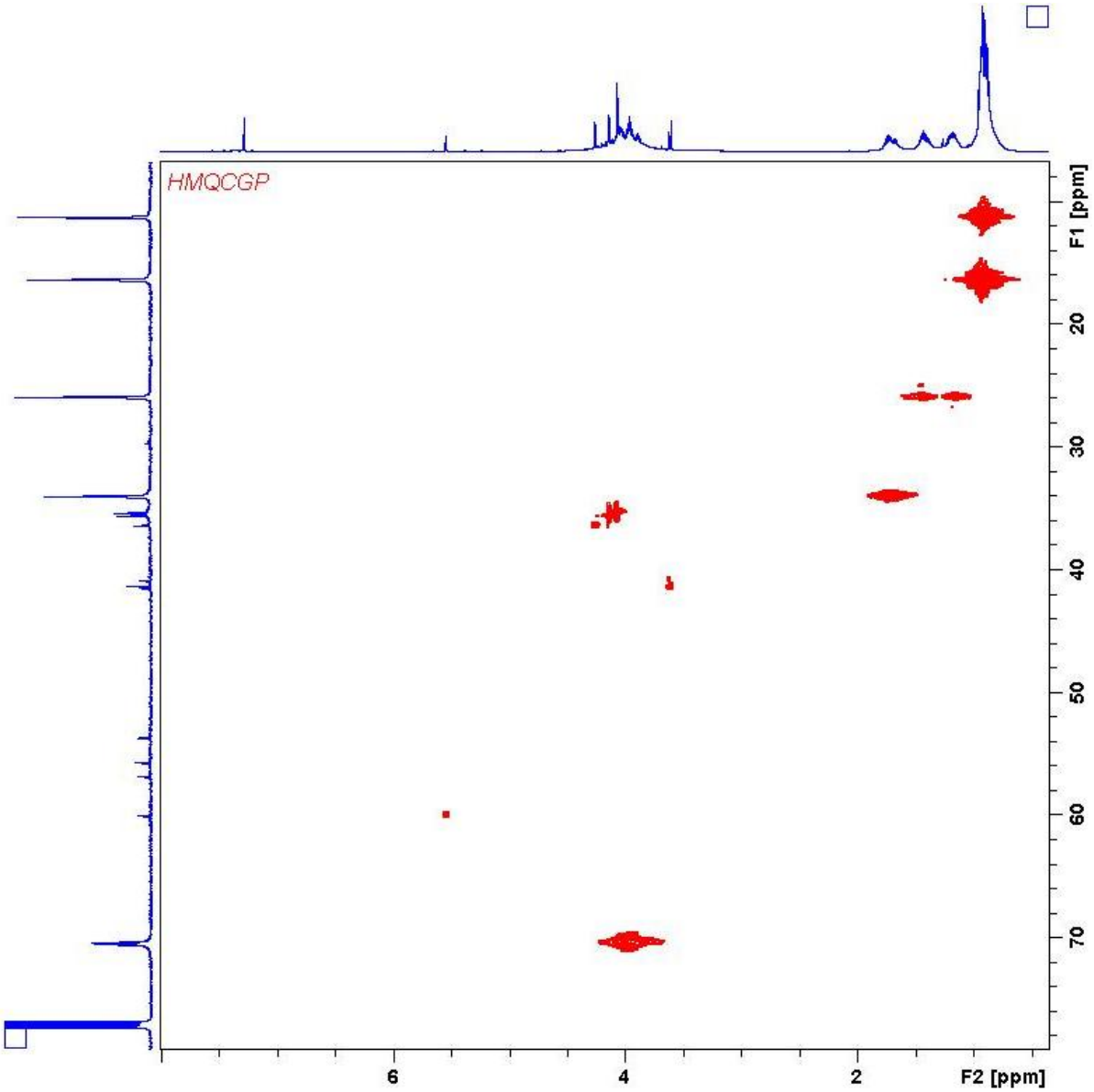


Fig. S18. H-C HMQC NMR spectrum of compound **1d**

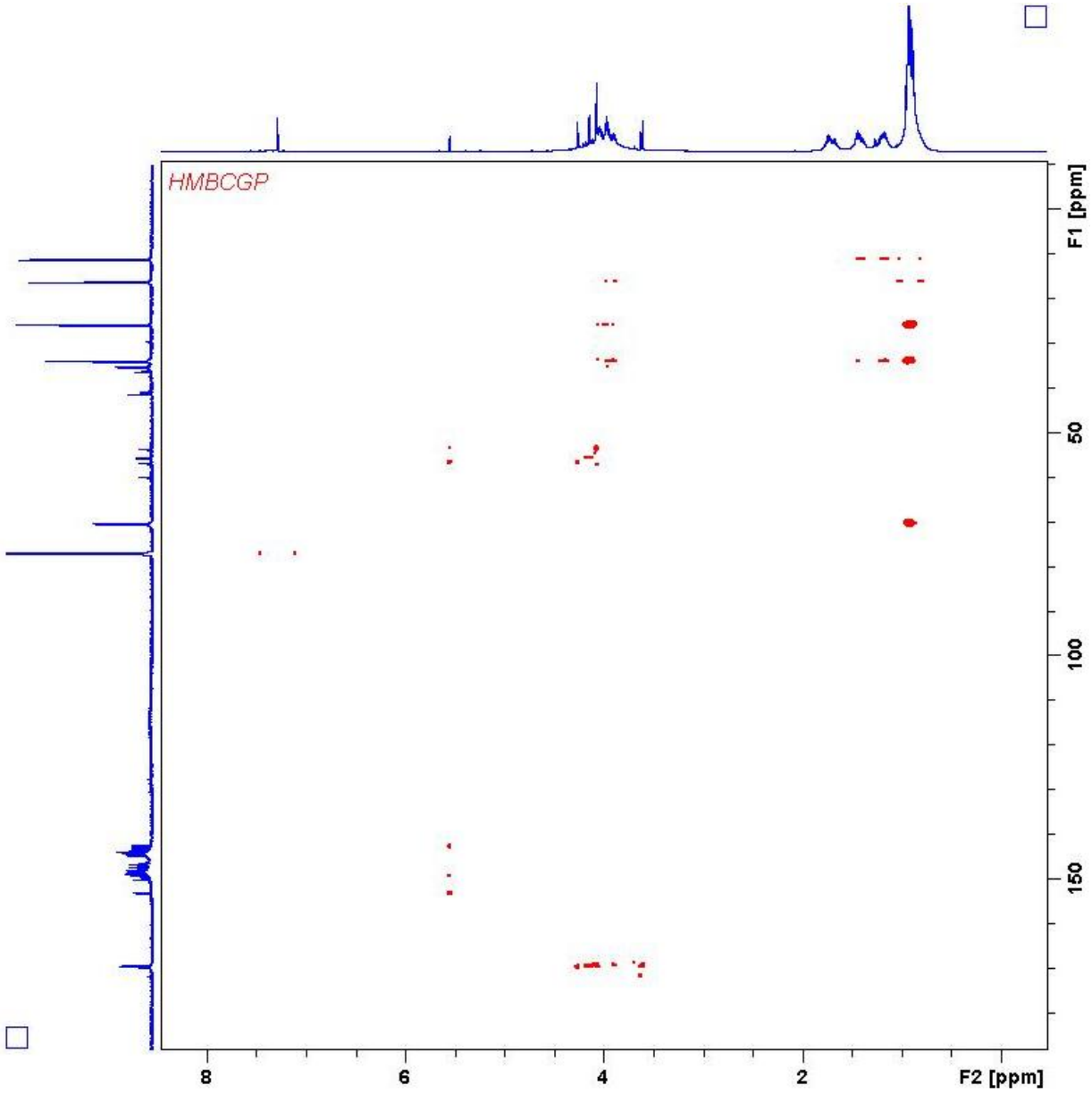


Fig. S19. H-C HMBC NMR spectrum of compound **1d**

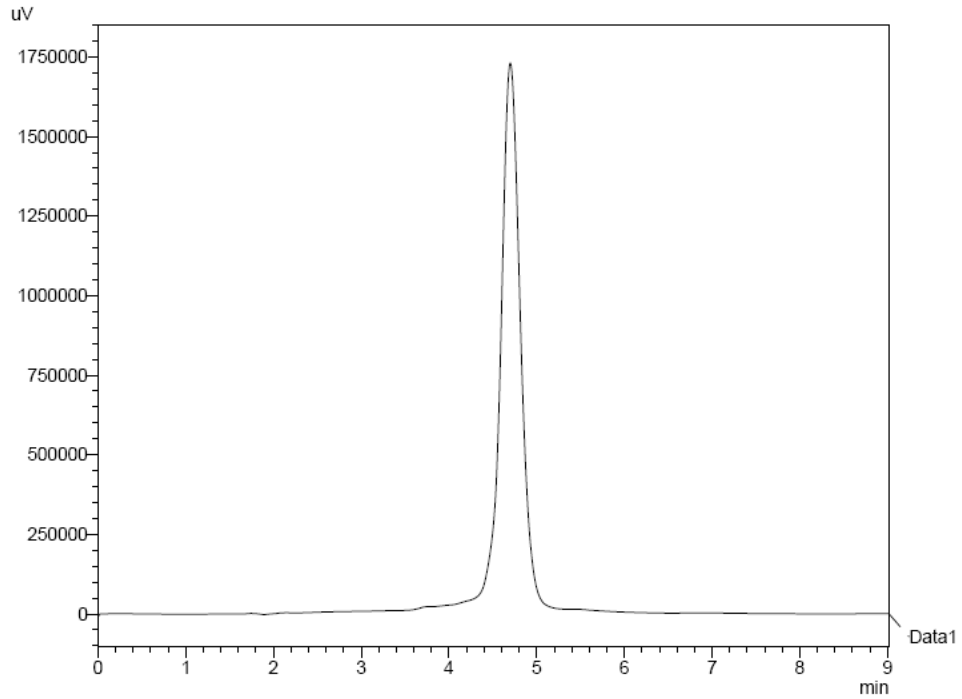


Fig. S20. HPLC profile of compound **1e** (Phenomenex Luna 5u C18(2) column, 150 x 4.6 mm, acetonitrile/toluene 70/30 v/v, flow rate 1 mL min^{-1}).

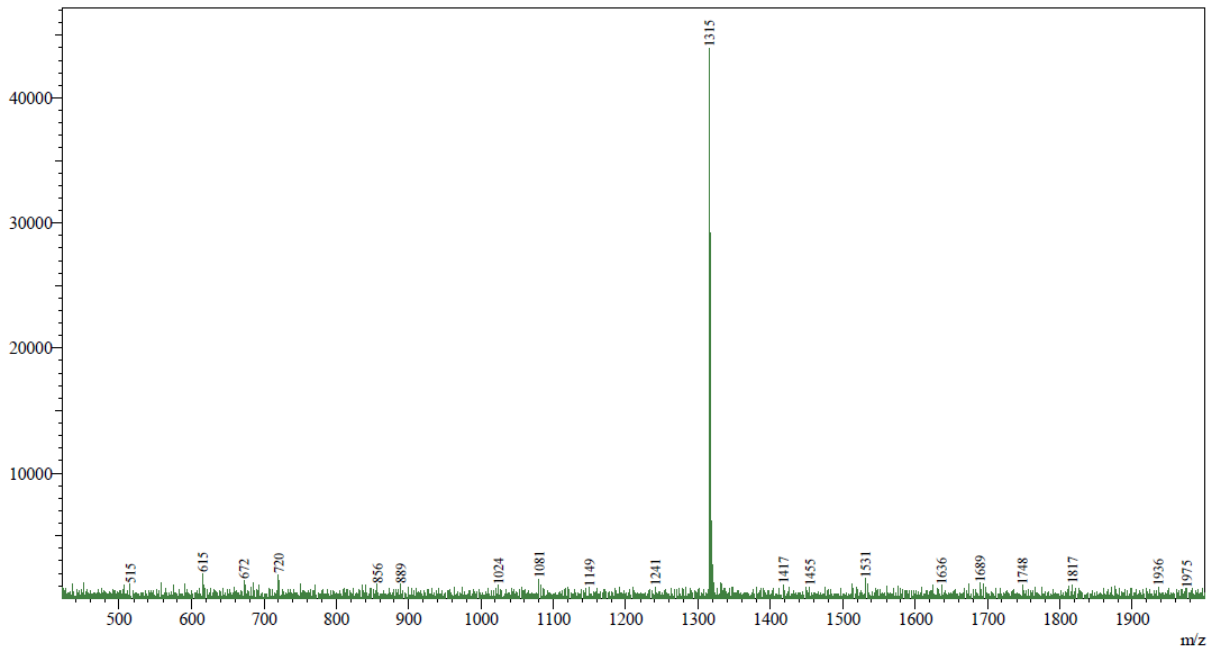


Fig. S21. ESI mass spectrum of compound **1e**

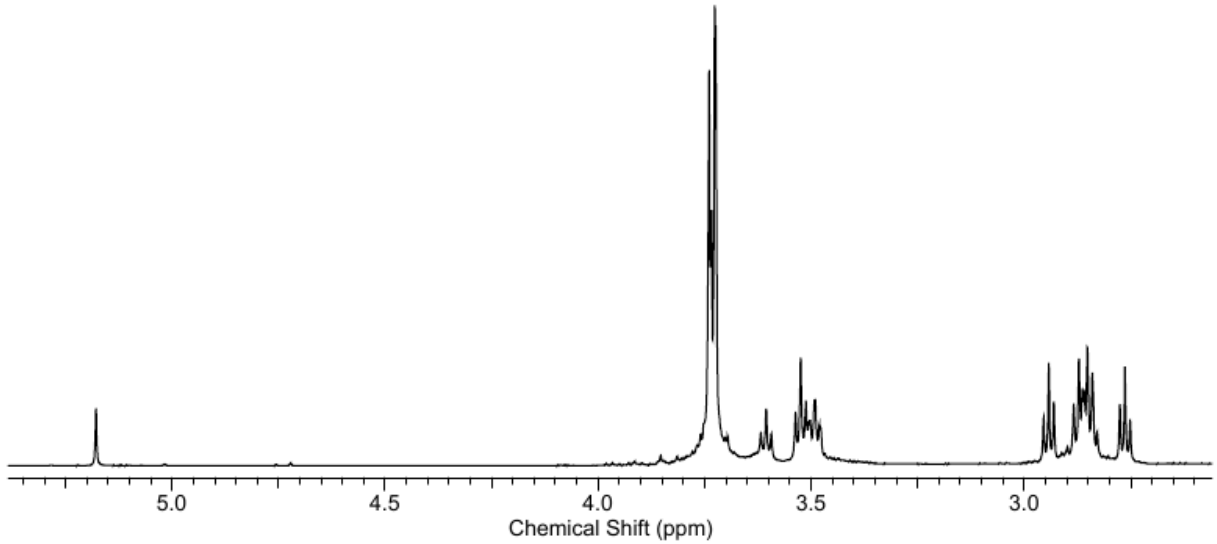


Fig. S22. ^{1}H NMR spectrum of compound **1e**

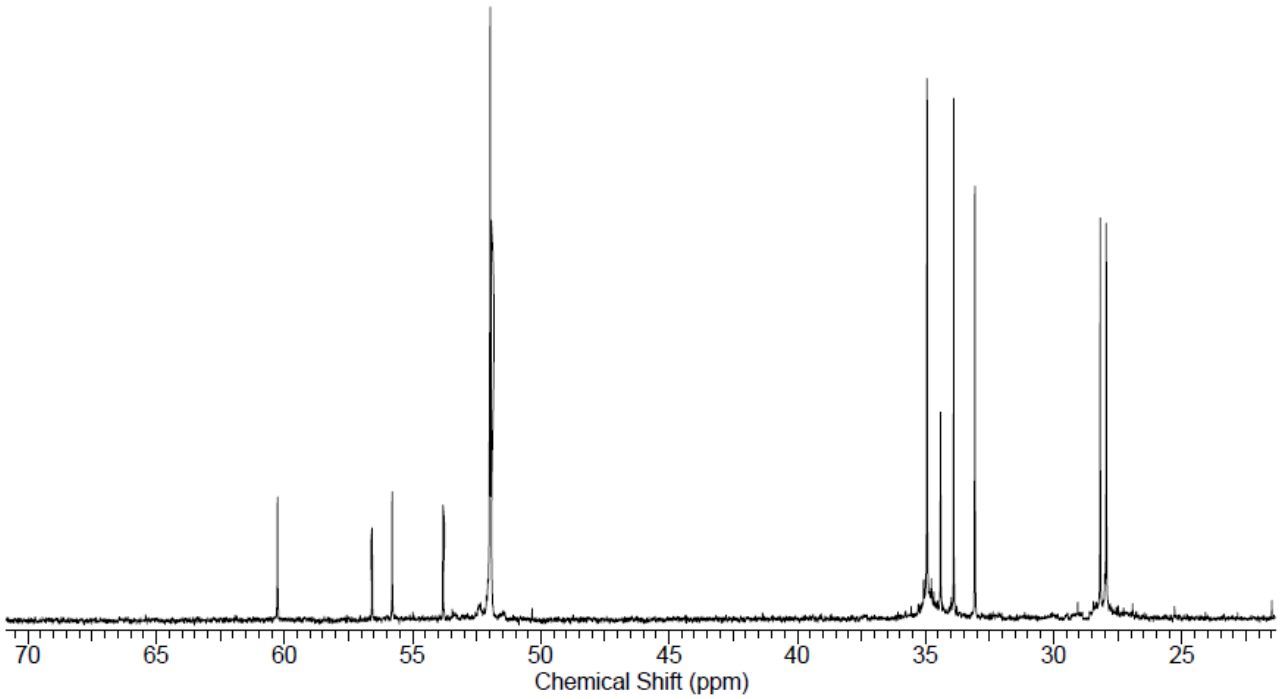


Fig. S23. High-field part of the ^{13}C NMR spectrum of compound **1e**

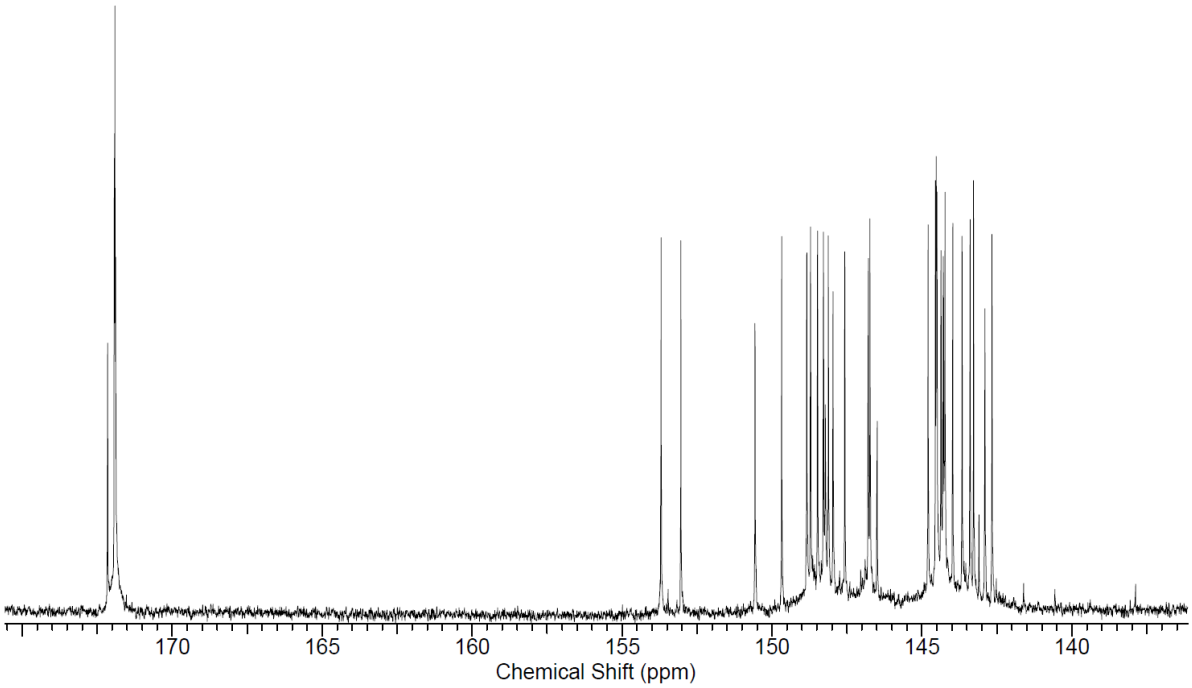


Fig. S24. Low-field part of the ^{13}C NMR spectrum of compound **1e**

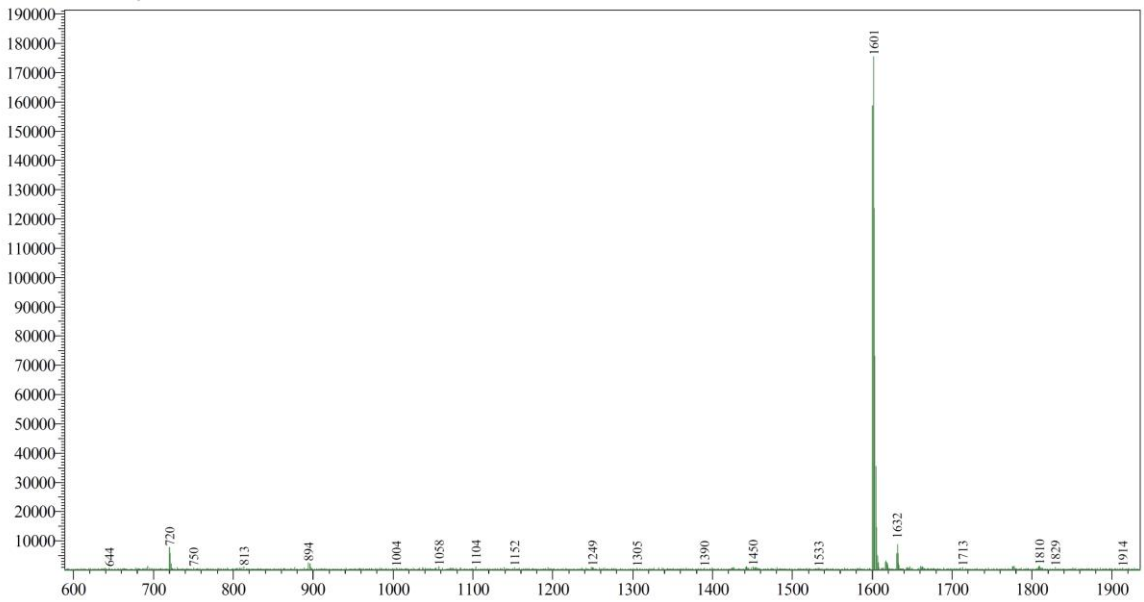


Fig. S25. ESI mass spectrum of compound **1f**

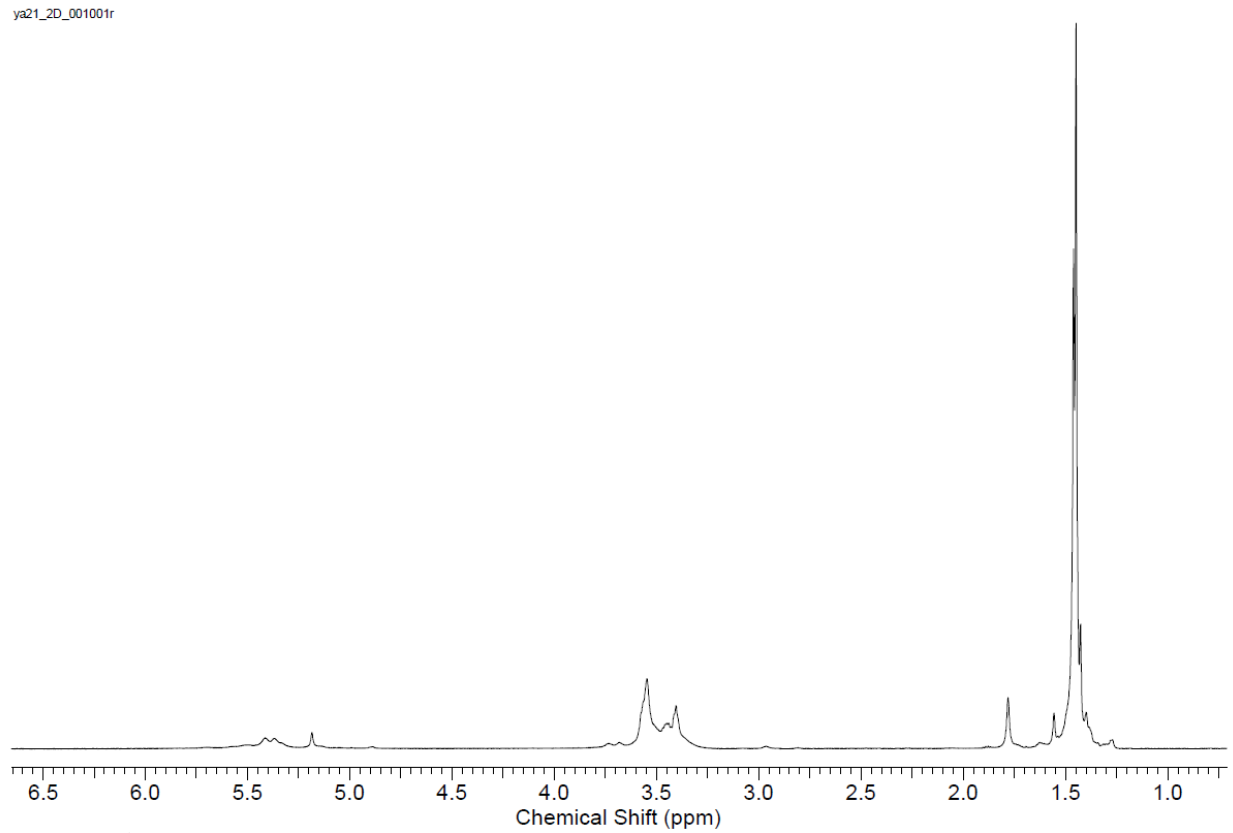


Fig. S26. ^{1}H NMR spectrum of compound **1f**

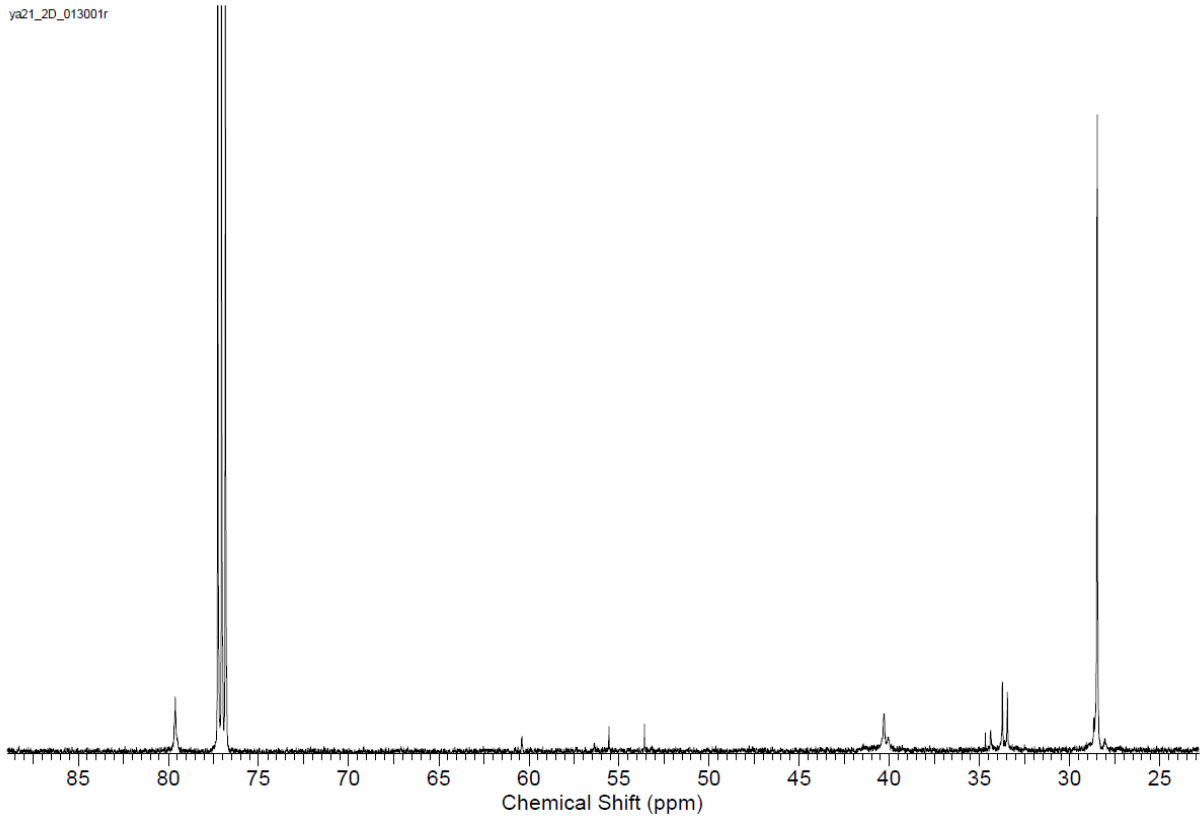


Fig. S27. High-field part of the ^{13}C NMR spectrum of compound **1f**

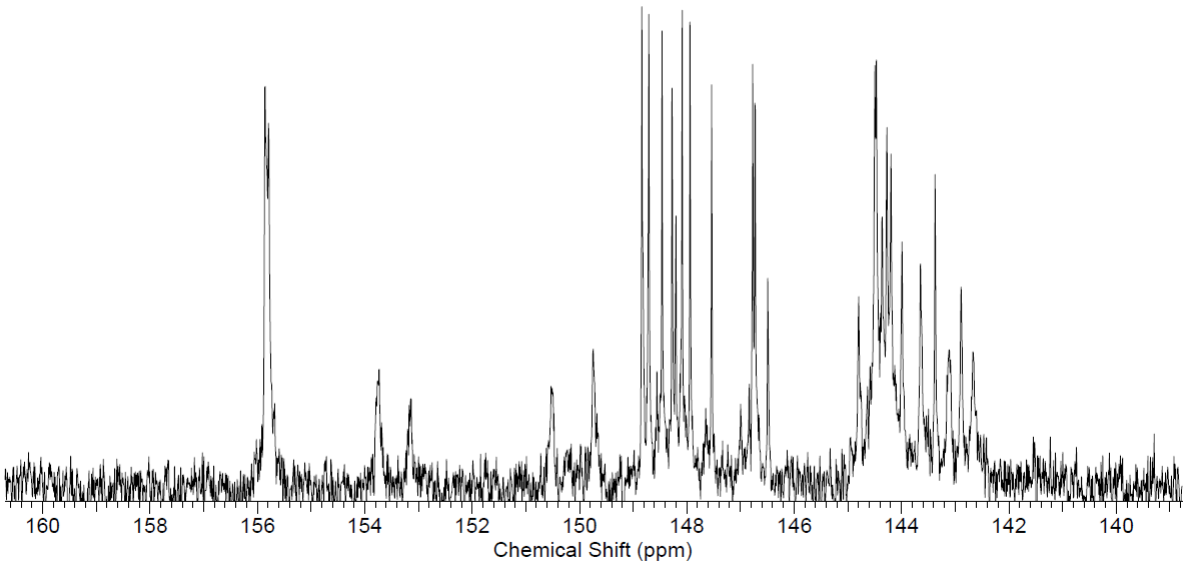


Fig. S28. Low-field part of the ^{13}C NMR spectrum of compound **1f**

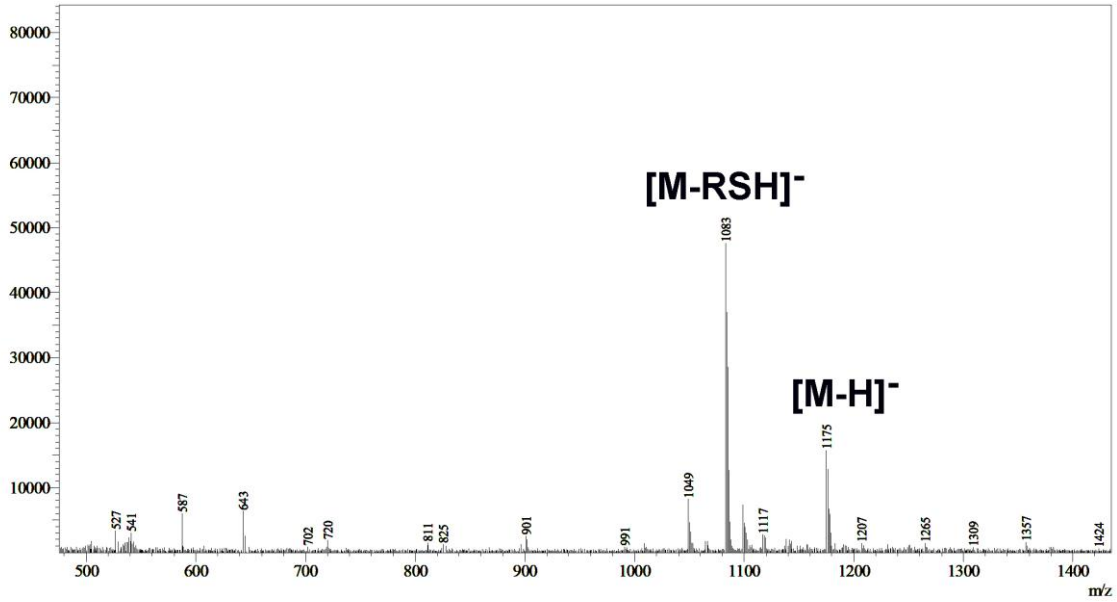


Fig. S29. ESI mass spectrum of compound **1g**

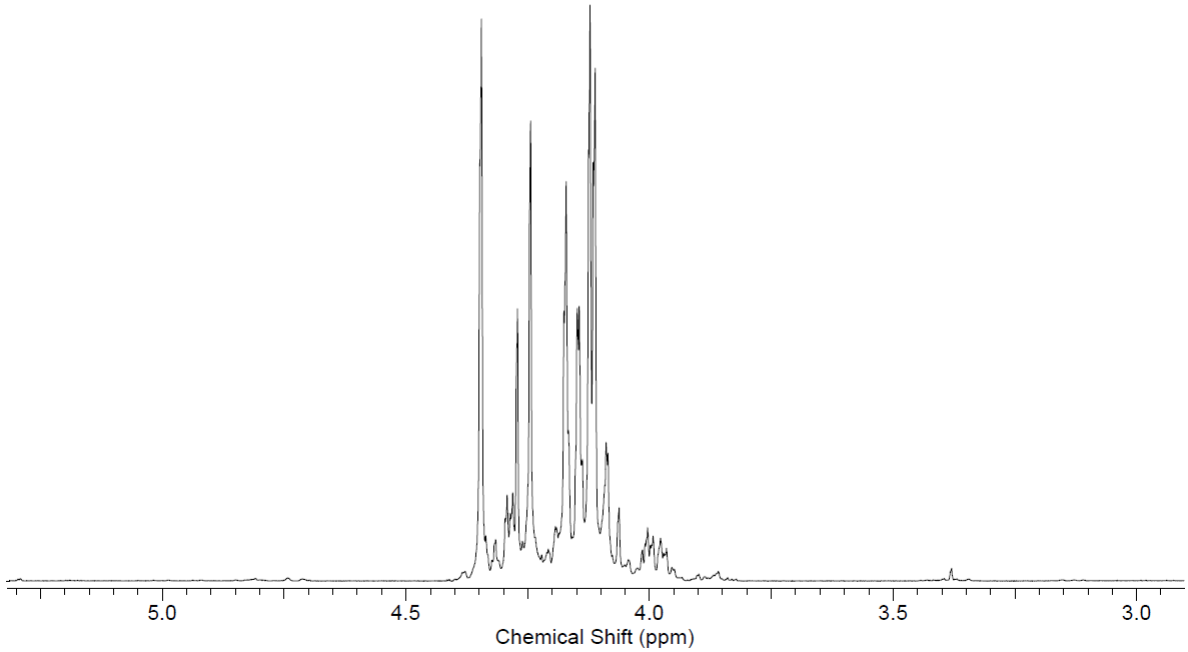


Fig. S30. ^{1}H NMR spectrum of compound **1g**

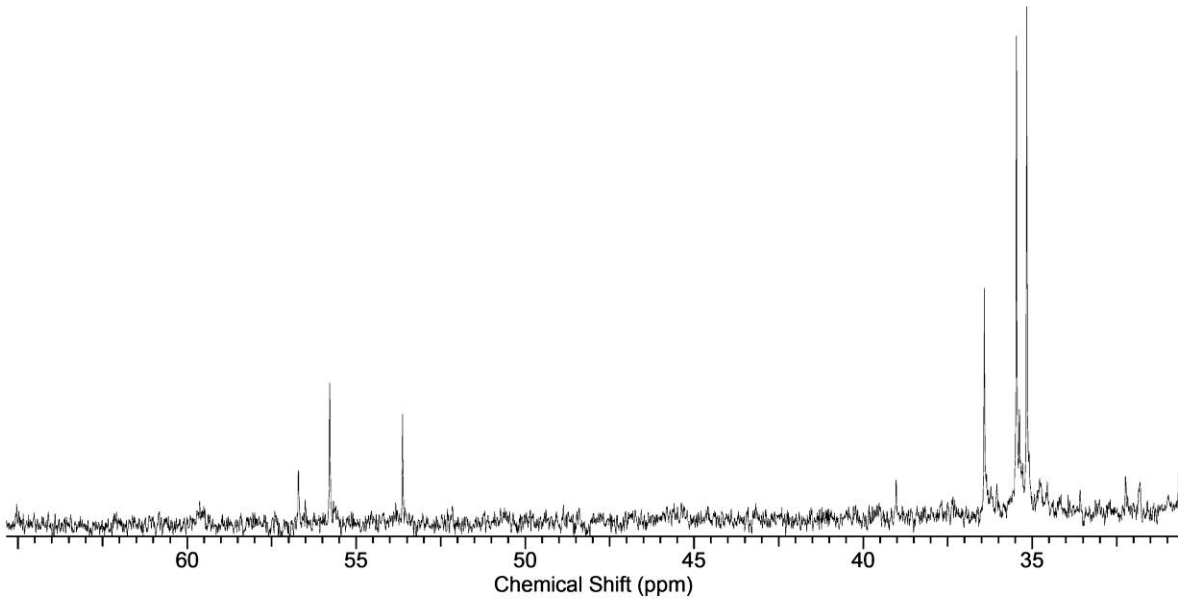


Fig. S31. High-field part of the ^{13}C NMR spectrum of compound **1g**

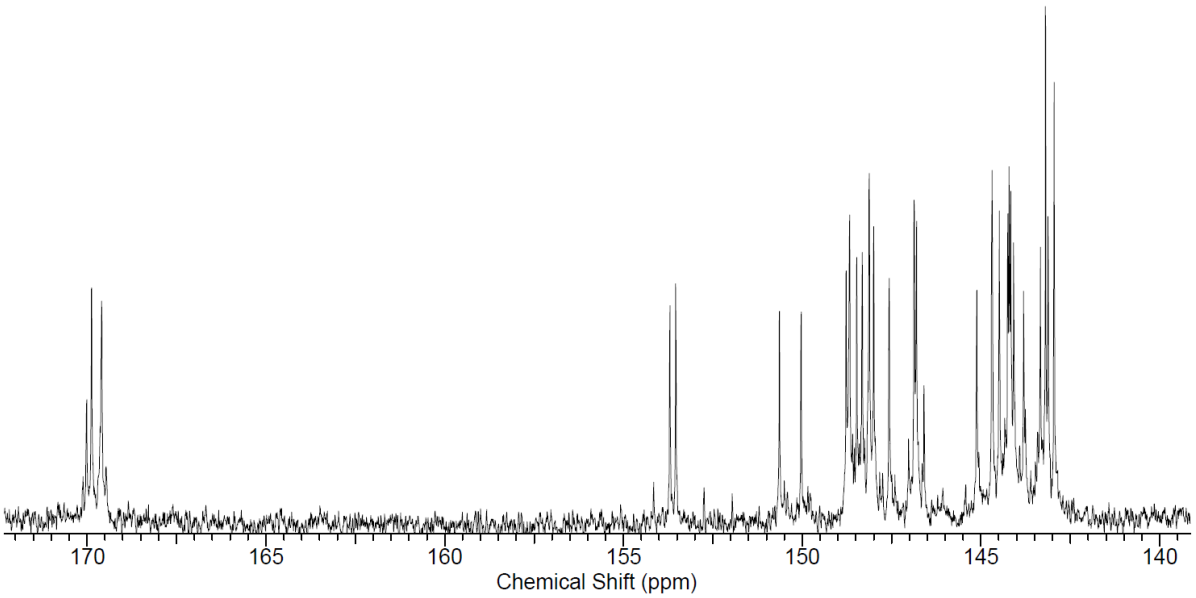


Fig. S32. Low-field part of the ^{13}C NMR spectrum of compound **1g**

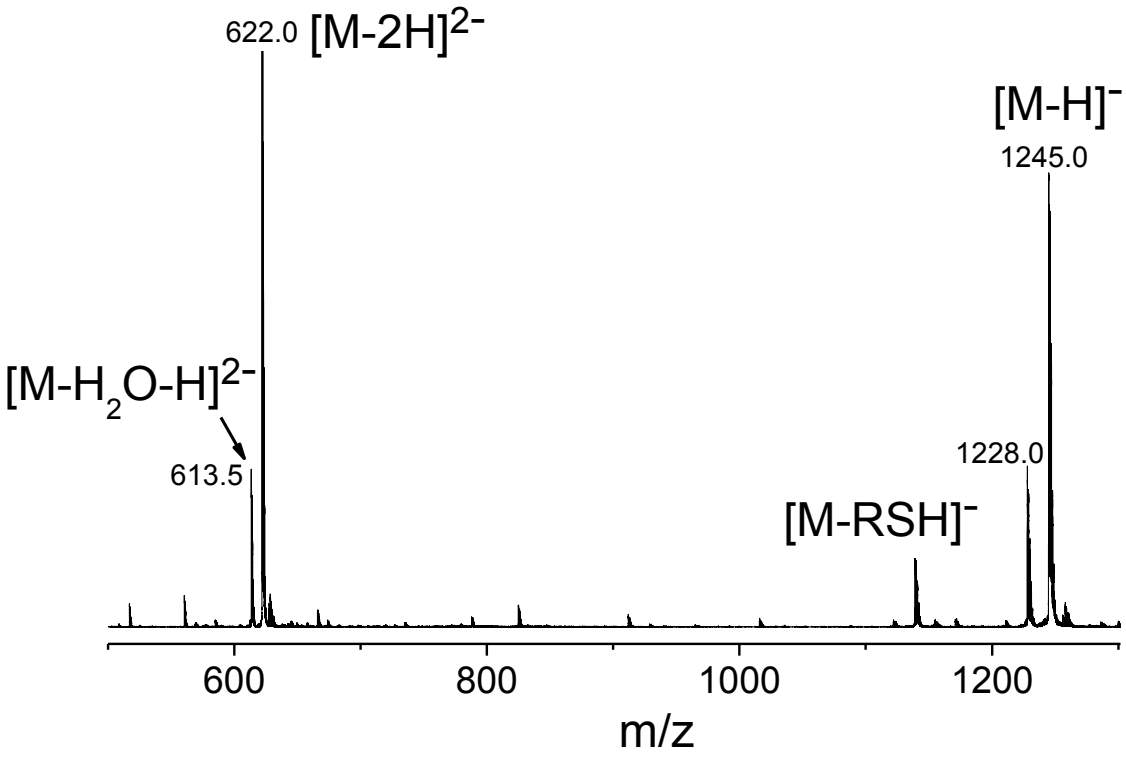


Fig. S33. ESI mass spectrum of compound **1h**

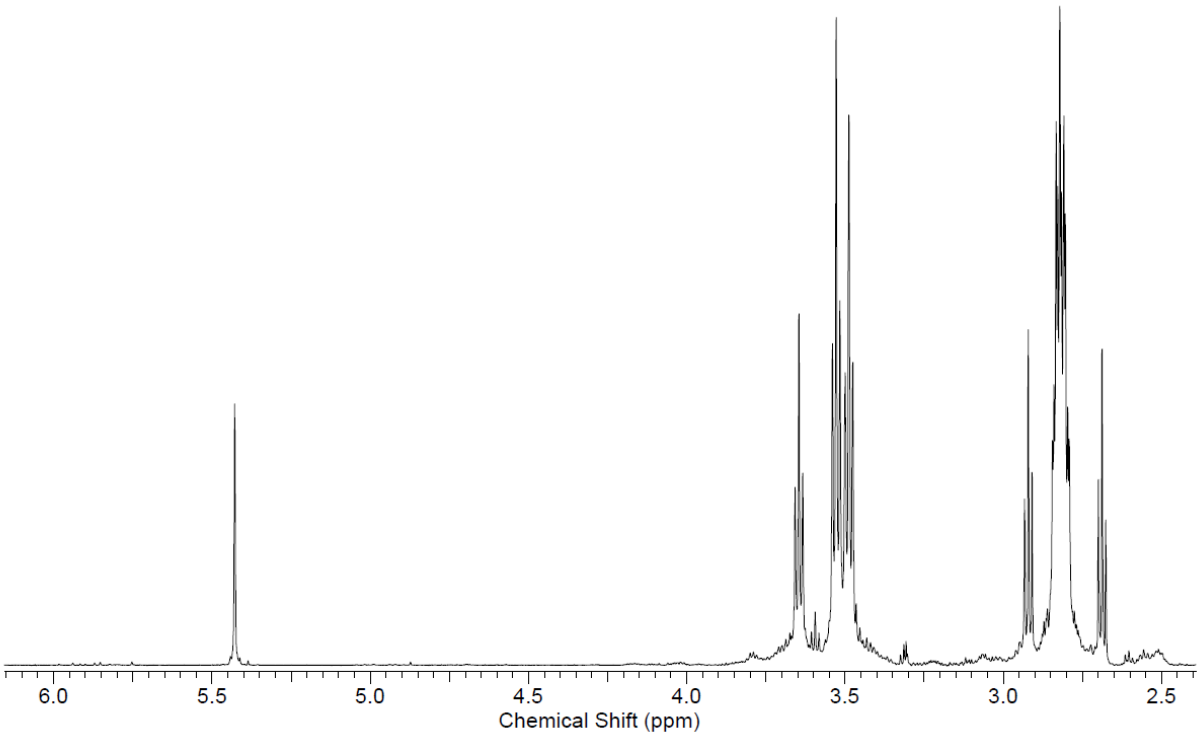


Fig. S34. ^{1}H NMR spectrum of compound **1h**

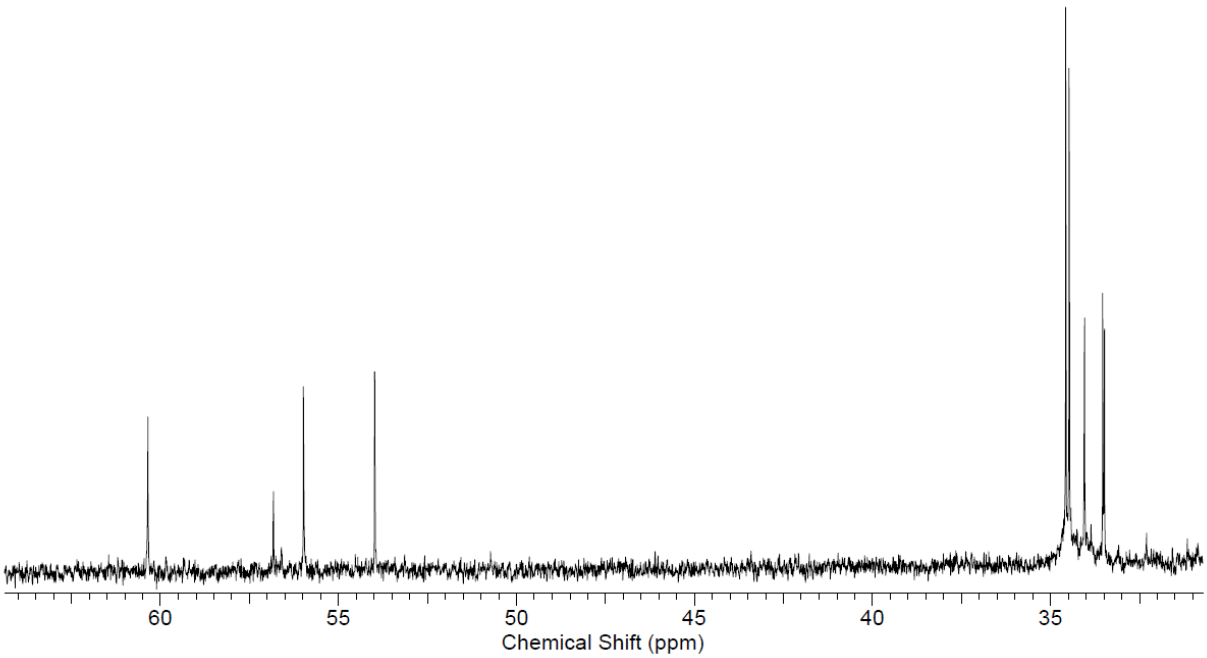


Fig. S35. High-field part of the ^{13}C NMR spectrum of compound **1h**

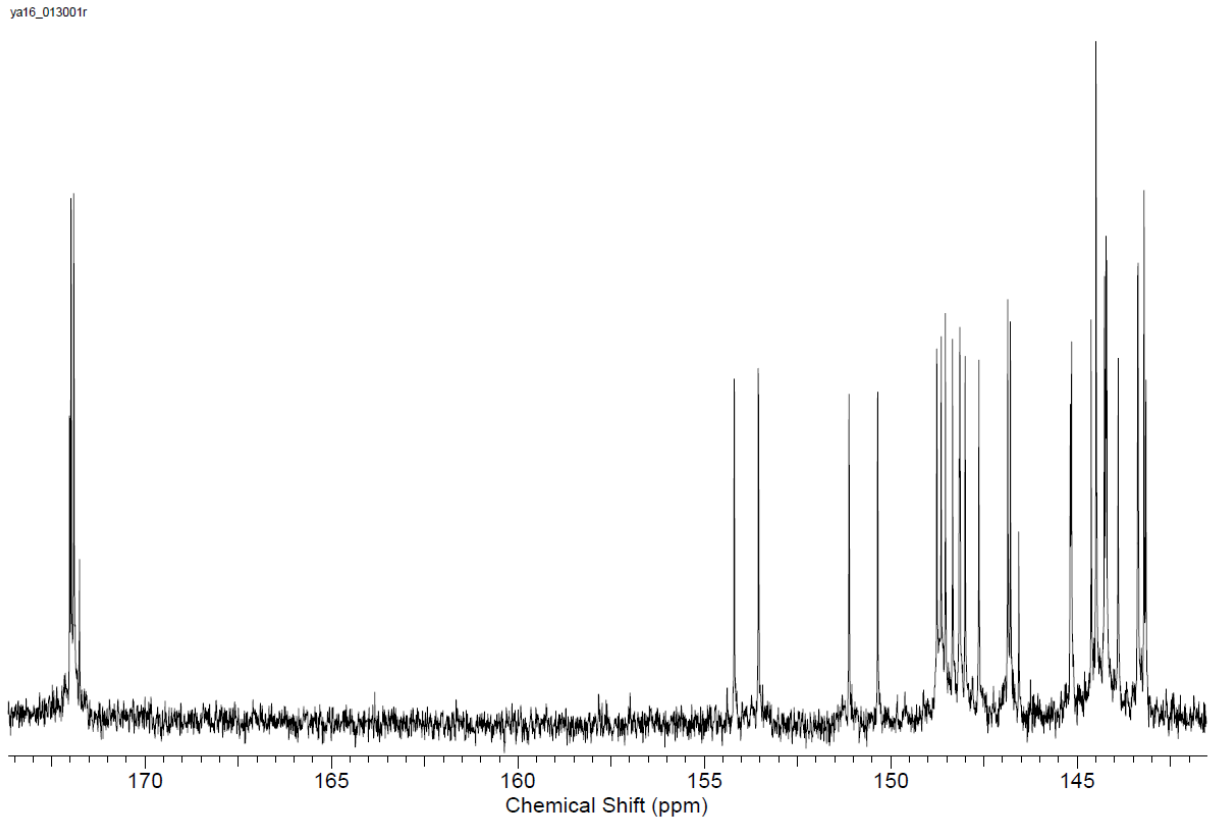


Fig. S36. Low-field part of the ^{13}C NMR spectrum of compound **1h**

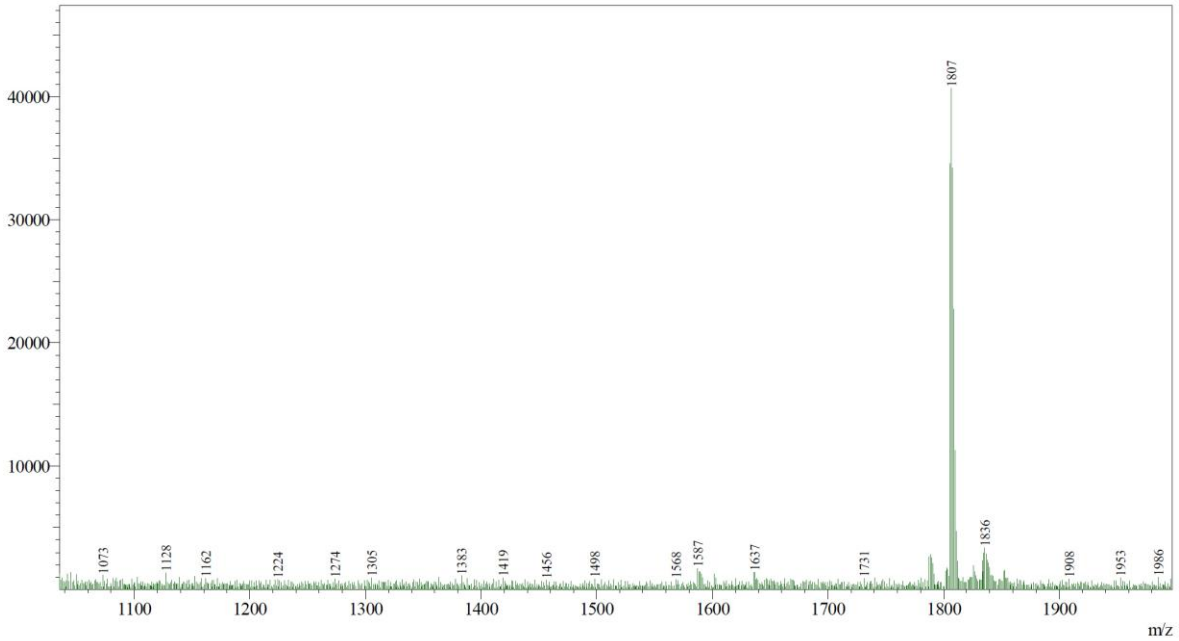


Fig. S37. ESI mass spectrum of compound **1i**

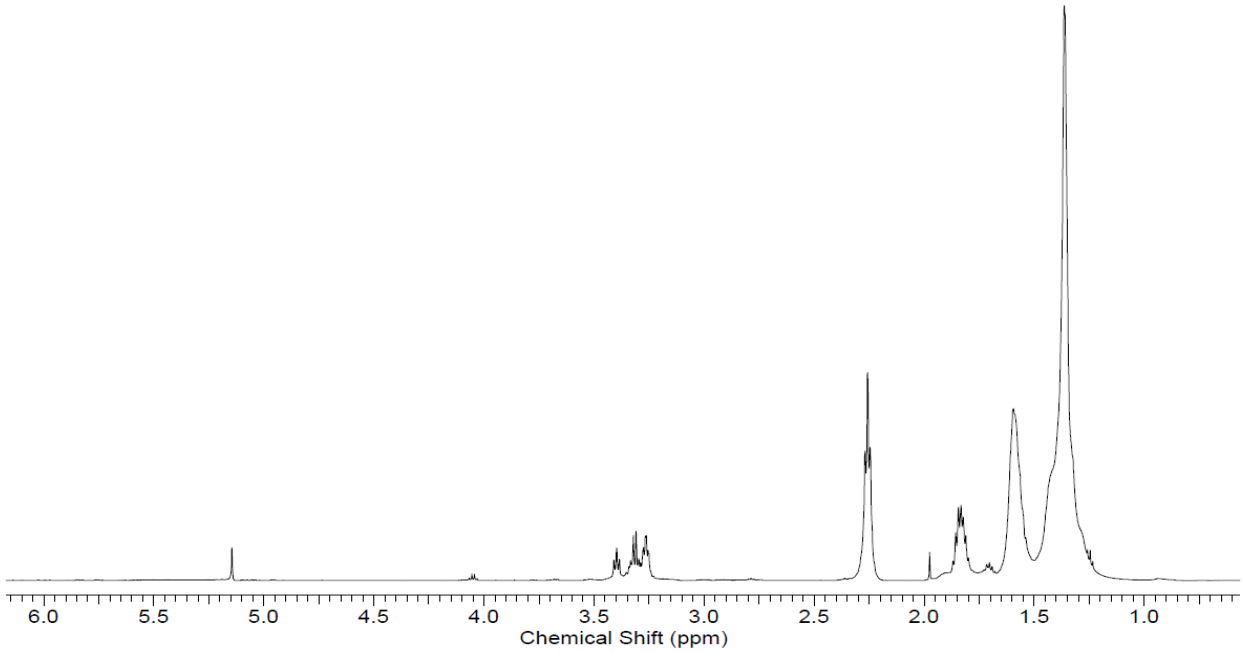


Fig. S38. ^{1}H NMR spectrum of compound **1i**

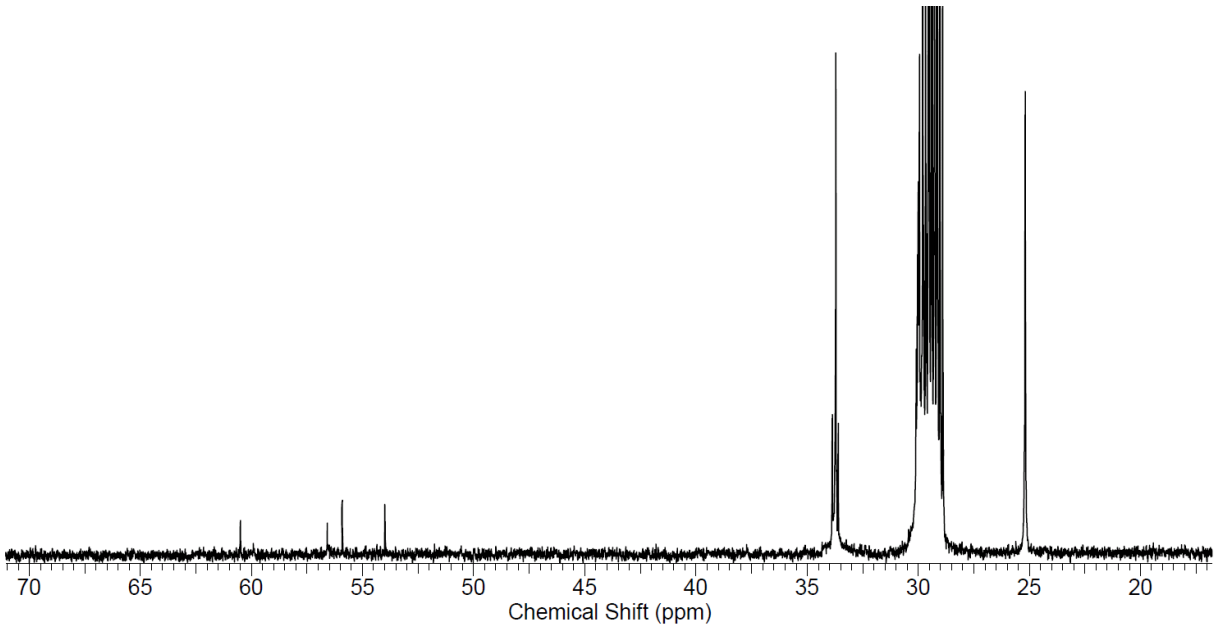


Fig. S39. High-field part of the ^{13}C NMR spectrum of compound **1i**

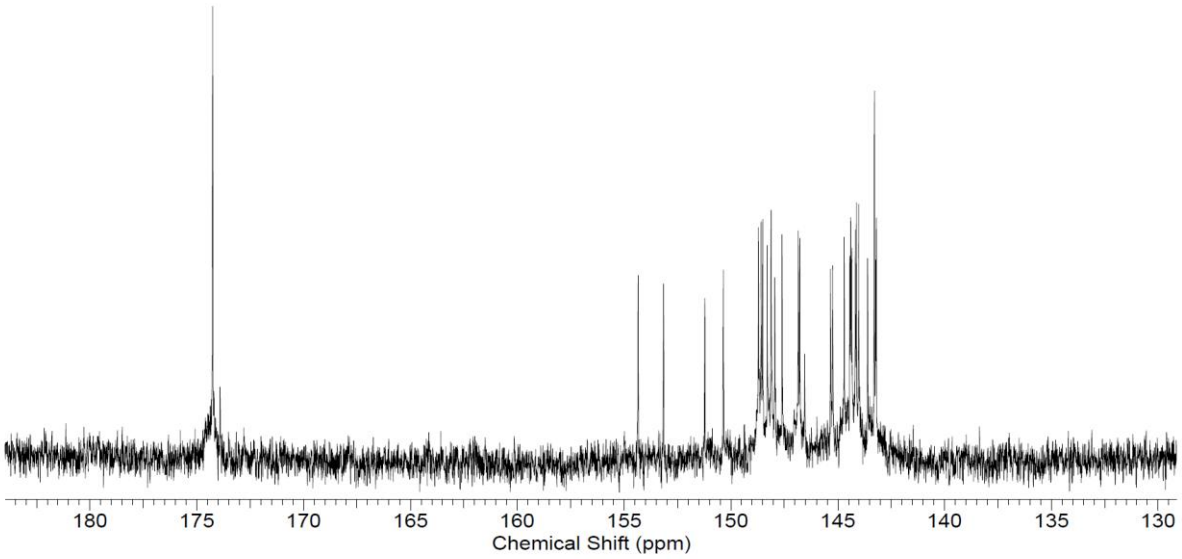


Fig. S40. Low-field part of the ^{13}C NMR spectrum of compound **1i**

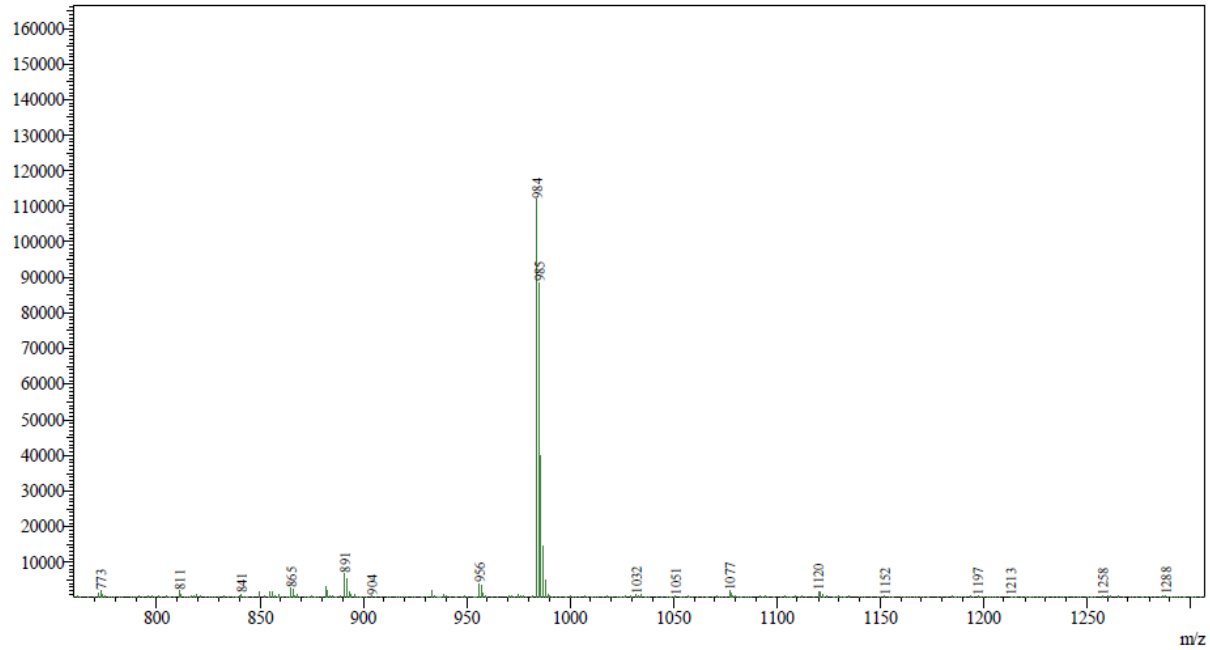


Fig. S41. ESI mass spectrum of compound **2e** (peak with m/z=984 corresponds to the $[M+CN]^-$ anion)

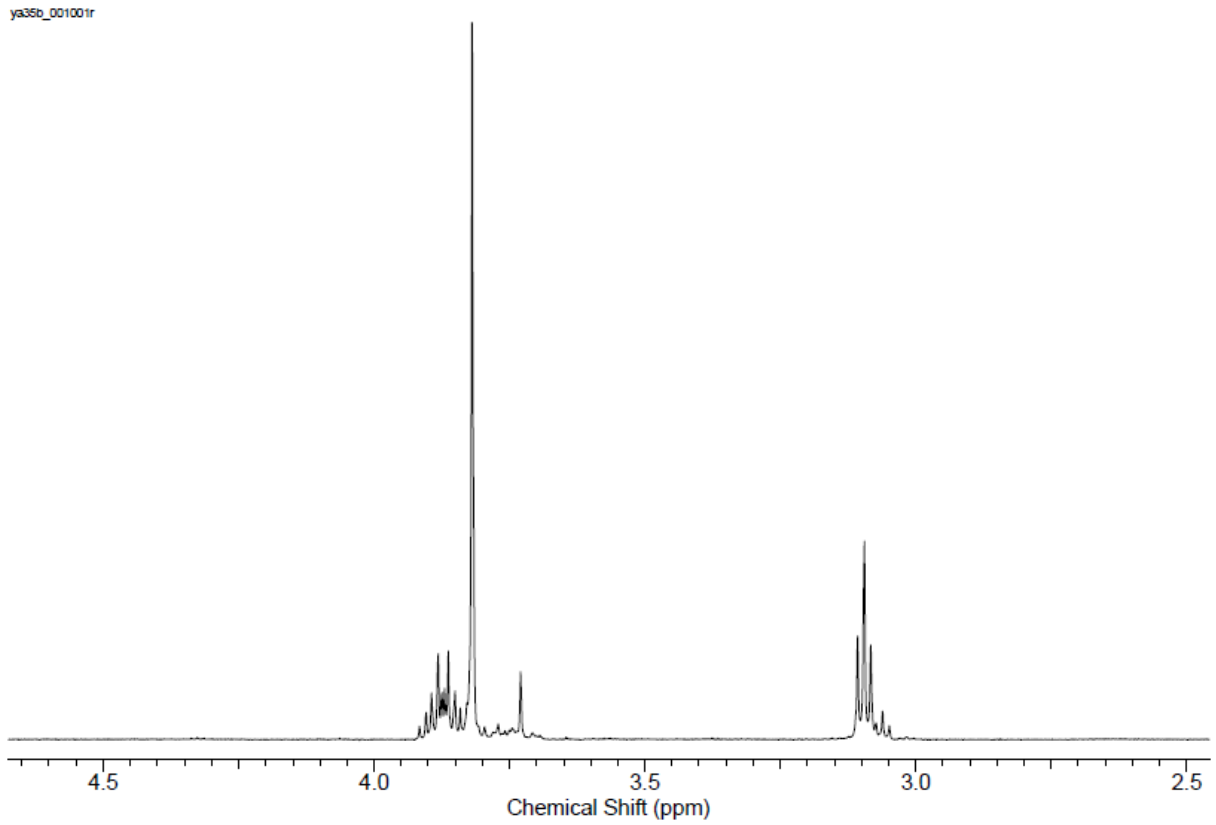


Fig. S42. ^{1}H NMR spectrum of compound **2e**

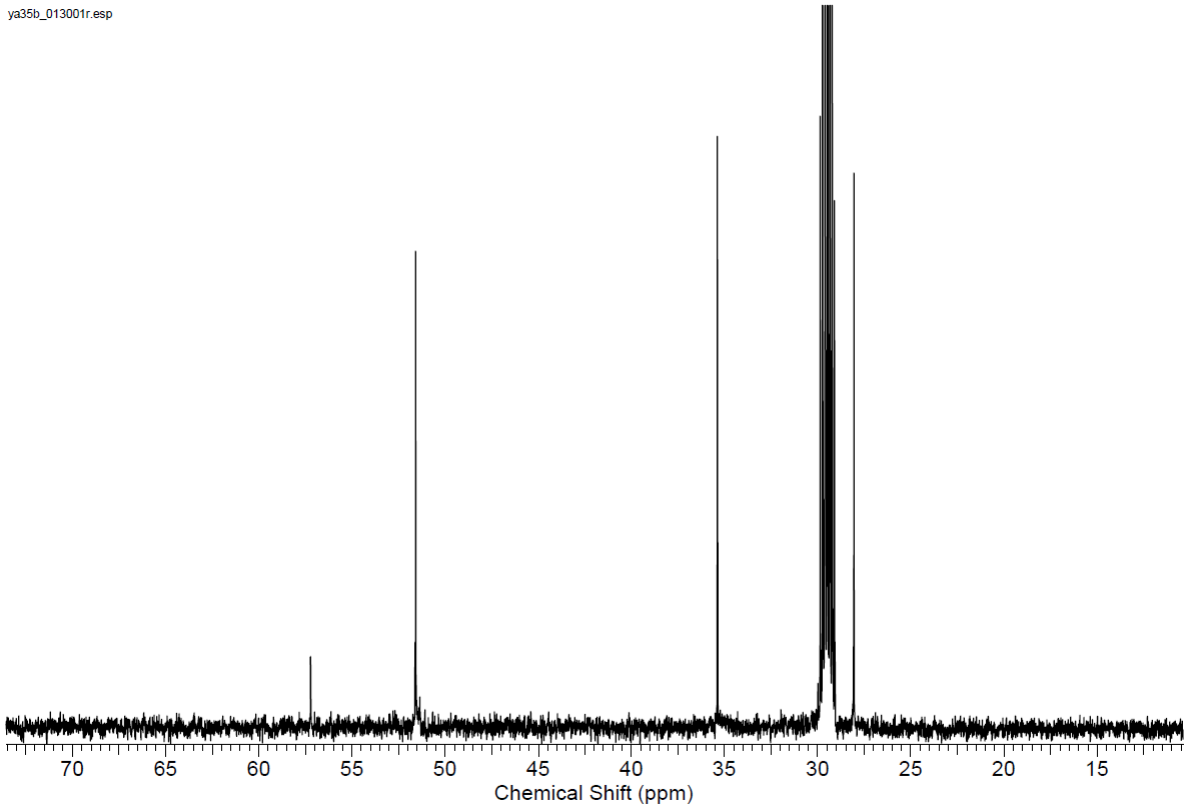


Fig. S43. High-field part of the ^{13}C NMR spectrum of compound **2e**

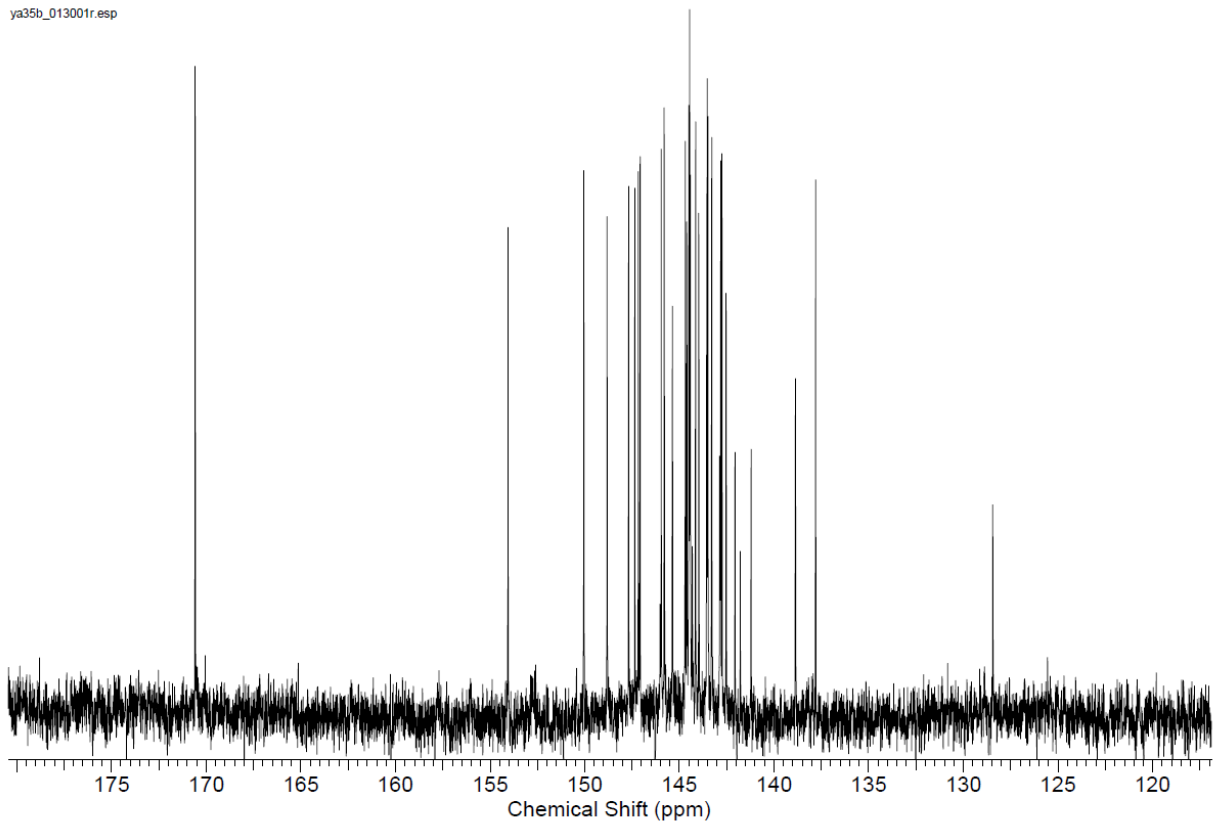


Fig. S44. Low-field part of the ^{13}C NMR spectrum of compound **2e**

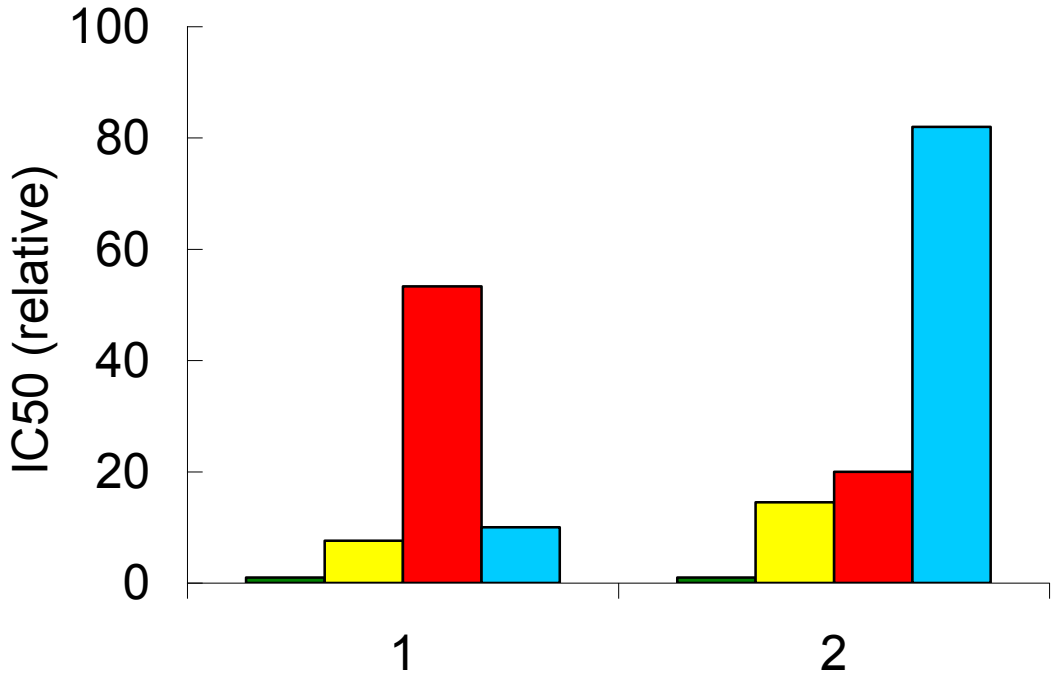


Fig. S45 Relative IC_{50} values characterizing inhibitory activities of **1i-K** (1) and **1h-K** (2) with respect to protein tyrosine phosphatases PTP1B (green), TC-PTP (yellow), PTPβ (red) и LAR-PTP (blue)

The activities of **1i-K** and **1h-K** were determined in 0.05 M Bis-Tris buffer, pH 7.2 (PTP1B, TC-PTP, LAR-PTP) and pH 7.0 (PTP-β). *p*-Nitrophenylphosphate was used as a substrate (2 mM for PTP1B, TC-PTP, LAR-PTP and 1 mM for PTP-β)