Electronic supplementary information for the manuscript:

Direct arylation of $C_{60}Cl_6$ and $C_{70}Cl_8$ with carboxylic acids: a synthetic avenue to water-soluble fullerene derivatives with promising antiviral activity

Table of contents

Experimental procedures

Water-soluble fullerene C_{60} derivatives were synthesized according to the following procedure: Chlorofullerene $C_{60}Cl_6$ (200 mg, 0.214 mmol) was dissolved in nitrobenzene (for phenyl substituted acids, 50 mL) or 1,2-dichlorobenzene (for thienyl substituted acids, 50 mL) under inert conditions. Then aromatic acid (5.36 mmol, 25 eq) and $SnCl_4$ (5 μL) were added. Reaction mixture was heated to 70-80^{o}C and stirred at this temperature for ~1 hour. Then reaction mixture was allowed to cool down, precipitated with 100 mL of hexane and centrifuged. Residue was washed with hexane, dried, washed with acetonitrile 5 times. Obtained orange powder was dissolved in aqueous K_2CO_3 (73.8 mg, 0.535 mmol, 2.5 eq), filtered through syringe filter (average pore size 0.45 μm), precipitated with HCl or acetic acid (1-2 mL), washed with water 10 times and dried in vacuo.

Procedure for the synthesis of the C_{70} derivatives was similar, except 40 equivalents of the aromatic acid was used. Fullerene derivatives **1-11** were obtained with ~95% yields.

Potassium salts of the fullerene derivatives **1-11** were obtained by the dissolution of the compounds **1-11** (1 eq) in aqueous K_2CO_3 (2.5 eq for C_{60} derivatives and 4 eq for C_{70} derivatives) and filtration through the PES syringe filter followed by freeze-drying. Compounds **1-11** in form of potassium salts were obtained with quantitative yields.

Amides of aromatic acids ((3-phenylpropanoyl)glycine (**12a**), 6-(3-phenylpropanamido)hexanoic acid (**12b**), (3-(thiophen-2-yl)propanoyl)glycine (**12c**), 6-(2-(thiophen-2-yl)acetamido)hexanoic acid (**12d**)) were synthesized using standard method reported previously (*Chem. Commun*., 2016, **52**, 7043-7046).

Selected data

Compounds **1** and **2** have been fully characterized earlier (*Org. Biomol. Chem*., 2019, **17**, 7155-7160).

3 (Yield 95%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$, δ, ppm): 1.23 – 1.44 (m, 15H), 3.61 – 3.80 (m, 5H), 6.95 – 7.66 (m, 25H), 7.67 – 8.11 (m, 10H).
^{13}C NMR (151 MHz, $(CD_3)_2SO$, δ, ppm): 18.72 ($\underline{C}H_3$), 44.48 ($\underline{C}$H), 44.56 ($\underline{C}$H), 57.85 ($\underline{C}_{sp3\ fullerene\ cage}$), 60.56 ($\underline{C}_{sp3\ fullerene\ cage}$), 63.18 ($\underline{C}_{sp3\ fullerene\ cage}$), 76.24 ($\underline{C}_{sp3\ fullerene\ cage}$), 115.61, 115.62, 115.77, 118.59, 124.35, 124.39, 124.50, 124.53, 125.99, 126.08, 126.10, 126.11, 126.19, 128.61, 128.70, 128.71, 128.75, 129.90, 129.92, 129.95, 129.98, 130.00, 131.04, 131.06, 135.30, 135.37, 136.32, 137.71, 137.73, 142.63, 142.88, 143.17, 143.68, 143.88, 144.47, 144.93, 145.28, 145.32,

146.22, 146.59, 147.14, 148.51, 148.63, 148.67, 150.25, 151.07, 154.15, 156.59, 158.45, 158.56, 160.19, 175.24 ($\underline{C}$O), 175.29 ($\underline{C}$O).
^{19}F NMR (376 MHz, $(CD_3)_2SO$, δ, ppm): -118.20 (s, 2F) -117.90 (s, 3F).
ESI MS: m/z= 1936 ($[M\text{-}Cl]^-$).

4 (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 2.31 – 2.38 (m, 2H), 2.42 – 2.53 (m, 8H), 2.69 – 2.75 (m, 2H), 2.77 – 2.94 (m, 8H), 3.66 – 3.84 (m, 10H), 6.96 – 7.14 (m, 4H), 7.19 – 7.31 (m, 8H), 7.51 – 7.59 (m, 4H), 7.74 – 7.82 (m, 4H), 8.09 – 8.13 (m, 1H), 8.14 – 8.29 (m, 4H), 12.43 (br.s, 5H).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 30.38 ($\underline{C}H_2$), 30.76 ($\underline{C}H_2$), 31.01 ($\underline{C}H_2$), 36.60 ($\underline{C}H_2$), 36.89 ($\underline{C}H_2$), 37.00 ($\underline{C}H_2$), 41.07 ($\underline{C}H_2$), 57.85 ($\underline{C}_{sp3\ fullerene\ cage}$), 60.59 ($\underline{C}_{sp3\ fullerene\ cage}$), 63.15 ($\underline{C}_{sp3\ fullerene\ cage}$), 76.54 ($\underline{C}_{sp3\ fullerene\ cage}$-Cl), 128.33, 128.48, 128.66, 128.74, 129.40, 129.48, 130.06, 134.51, 135.97, 141.28, 141.60, 141.84, 141.99, 142.79, 143.58, 143.61, 143.65, 143.76, 143.89, 144.07, 144.22, 144.39, 144.44, 145.49, 146.97, 147.03, 147.18, 147.22, 147.33, 147.79, 148.08, 148.17, 148.40, 148.53, 148.57, 148.61, 148.72, 150.66, 151.39, 156.82, 156.86, 171.79 ($\underline{C}$O), 171.85 ($\underline{C}$O), 172.16 ($\underline{C}$O), 172.24 ($\underline{C}$O), 174.15 ($\underline{C}$O), 174.20 ($\underline{C}$O).
ESI MS: m/z= 1751 ($[M\text{-}Cl]^-$).

5 (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 1.16 – 1.28 (m, 10H), 1.29 – 1.38 (m, 10H), 1.41 – 1.51 (m, 10H), 2.09 – 2.22 (m, 10H), 2.23 – 2.46 (m, 10H), 2.68 – 2.90 (m, 10H), 2.92 – 3.08 (m, 10H), 6.61 – 7.66 (m, 20H), 7.69 – 7.99 (m, 5H), 11.92 (br.s, 5H).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 24.66 ($\underline{C}H_2$), 26.39 ($\underline{C}H_2$), 29.32 ($\underline{C}H_2$), 31.18 ($\underline{C}H_2$), 34.05 ($\underline{C}H_2$), 37.04 ($\underline{C}H_2$), 37.28($\underline{C}H_2$), 38.77($\underline{C}H_2$), 38.81($\underline{C}H_2$), 57.88 ($\underline{C}_{sp3\ fullerene\ cage}$), 60.59 ($\underline{C}_{sp3\ fullerene\ cage}$), 63.15 ($\underline{C}_{sp3\ fullerene\ cage}$), 74.87 ($\underline{C}_{sp3\ fullerene\ cage}$-Cl), 127.88, 128.02, 128.28, 128.44, 128.68, 129.02, 129.37, 129.45, 129.49, 129.75, 129.97, 132.09, 135.98, 141.83, 142.04, 142.79, 142.89, 143.04, 143.17, 143.39, 143.57, 143.64, 143.77, 143.87, 144.08, 144.24, 144.38, 144.43, 145.50, 146.97, 147.02, 147.09, 147.19, 147.25, 147.35, 147.75, 147.83, 148.02, 148.11, 148.19, 148.32, 148.42, 148.56, 148.74, 150.65, 151.40, 156.87, 171.41 ($\underline{C}$O), 171.46 ($\underline{C}$O), 171.50 ($\underline{C}$O), 174.87 ($\underline{C}$O). ESI MS: m/z= 1016 ($[M\text{-}Cl]^{2-}$).

6 (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 3.75 (s, 2H), 3.83 (s, 4H), 3.86 (s, 4H), 6.78 (d, 1H, J = 3.5 Hz), 6.83 (d, 1H, J = 3.6 Hz), 6.91 (d, 2H, J = 3.5 Hz), 6.95 (d, 2H, J = 3.6 Hz), 7.09 (d, 2H, J = 3.5 Hz), 7.37 (d, 2H, J = 3.5 Hz).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 35.37 ($\underline{C}H_2$), 35.54 ($\underline{C}H_2$), 35.62 ($\underline{C}H_2$), 54.07 ($\underline{C}_{sp3\ fullerene\ cage}$), 56.28 ($\underline{C}_{sp3\ fullerene\ cage}$), 59.64 ($\underline{C}_{sp3\ fullerene\ cage}$), 75.47 ($\underline{C}_{sp3\ fullerene\ cage}$-Cl), 126.87, 127.04, 127.45, 127.47, 127.60, 129.88, 137.90, 138.11, 138.45, 139.14, 140.21, 142.26, 142.54, 142.82, 143.10, 143.37, 143.89, 144.29, 144.30, 144.33, 144.61, 144.66, 144.93, 145.60, 146.00, 147.12, 147.26, 147.83, 148.18, 148.23, 148.32, 148.62, 148.69, 148.72, 149.95, 150.58, 153.32, 155.70, 171.79 ($\underline{C}$O), 171.92 ($\underline{C}$O), 171.93 ($\underline{C}$O).
ESI MS: m/z= 1425 ($[M\text{-}Cl]^-$).

7 (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 3.69 – 3.84 (m, 20H), 6.75 – 6.82 (m, 2H), 6.89 (d, 2H, J = 3.2 Hz), 6.92 (d, 2H, J = 3.3 Hz), 7.07 (d, 2H, J = 3.4 Hz), 7.35 (d, 2H, J = 3.5 Hz), 8.33 – 8.52 (m, 5H), 12.58 (br.s, 5H).

^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 36.59 (<u>C</u>H_2), 36.81 (<u>C</u>H_2), 36.87 (<u>C</u>H_2), 41.28 (<u>C</u>H_2), 54.12 ($\underline{C}_{sp3\ fullerene\ cage}$), 56.35 ($\underline{C}_{sp3\ fullerene\ cage}$), 59.65 ($\underline{C}_{sp3\ fullerene\ cage}$), 75.49 ($\underline{C}_{sp3\ fullerene\ cage}$-Cl), 126.44, 126.72, 126.78, 127.48, 129.85, 138.89, 139.01, 139.25, 139.60, 139.94, 142.29, 142.54, 142.81, 143.00, 143.06, 143.38, 143.93, 144.30, 144.58, 144.63, 144.93, 145.41, 146.05, 147.12, 147.26, 147.83, 148.16, 148.22, 148.31, 148.60, 148.71, 149.98, 150.60, 153.36, 155.76, 169.49 (<u>C</u>O), 169.67 (<u>C</u>O), 171.57 (<u>C</u>O), 171.61 (<u>C</u>O).
ESI MS: m/z= 1711 ($[M-Cl]^-$).

8 (Yield 95%)**.** ^1H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 1.19 – 1.29 (m, 10H), 1.33 – 1.42 (m, 10H), 1.43 – 1.51 (m, 10H), 2.12 – 2.20 (m, 10H), 2.94 – 3.08 (m, 10H), 3.55 (s, 2H), 3.62 (s, 4H), 3.65 (s, 4H), 6.73 – 6.78 (m, 1H), 6.81 – 6.83 (m, 1H), 6.84 – 6.86 (m, 2H), 6.87 – 6.90 (m, 2H), 7.05 – 7.08 (m, 2H), 7.32 – 7.36 (m, 2H), 7.98 – 8.18 (m, 5H), 12.18 (br.s, 5H).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 24.66 (<u>C</u>H_2), 26.43 (<u>C</u>H_2), 29.19 (<u>C</u>H_2), 34.05 (<u>C</u>H_2), 37.10 (<u>C</u>H_2), 37.27 (<u>C</u>H_2), 37.36 (<u>C</u>H_2), 39.07 (<u>C</u>H_2), 54.12 ($\underline{C}_{sp3\ fullerene\ cage}$), 56.34 ($\underline{C}_{sp3\ fullerene\ cage}$), 59.65 ($\underline{C}_{sp3\ fullerene\ cage}$), 75.51 ($\underline{C}_{sp3\ fullerene\ cage}$-Cl), 126.12, 126.37, 126.54, 126.71, 127.03, 127.41, 129.80, 137.89, 138.09, 138.44, 138.73, 138.79, 139.24, 139.55, 139.76, 140.25, 142.29, 142.36, 142.53, 142.56, 142.81, 143.00, 143.07, 143.38, 143.91, 144.21, 144.31, 144.59, 144.64, 144.93, 145.28, 146.04, 147.12, 147.26, 147.83, 148.17, 148.22, 148.31, 148.60, 148.70, 149.98, 150.60, 153.34, 155.75, 169.03 (<u>C</u>O), 174.88 (<u>C</u>O).
ESI MS: m/z= 1991 ($[M-Cl]^-$).

9 (Yield 95%)**.** ^1H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 3.67 (s, 4H), 3.70 (s, 4H), 3.73 (s, 4H), 3.82 (s, 4H), 6.56 (d, 2H, *J* = 3.4 Hz), 6.60 (d, 2H, *J* = 3.5 Hz), 6.66 (d, 2H, *J* = 3.5 Hz), 6.69 (d, 2H, *J* = 3.4 Hz), 6.72 (d, 2H, *J* = 3.5 Hz), 6.78 (d, 2H, *J* = 3.4 Hz), 6.83 (d, 2H, *J* = 3.5 Hz), 6.99 (d, 2H, *J* = 3.5 Hz).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 35.46 (<u>C</u>H_2), 35.57 (<u>C</u>H_2), 56.05 ($\underline{C}_{sp3\ fullerene\ cage}$), 56.76 ($\underline{C}_{sp3\ fullerene\ cage}$), 57.11 ($\underline{C}_{sp3\ fullerene\ cage}$), 57.59 ($\underline{C}_{sp3\ fullerene\ cage}$), 108.13, 126.48, 126.58, 126.72, 126.85, 126.97, 126.99, 127.35, 131.57, 132.79, 134.03, 134.56, 135.65, 137.15, 137.50, 137.69, 140.56, 140.84, 141.19, 142.12, 142.14, 142.30, 142.34, 145.31, 145.44, 145.48, 145.77, 146.28, 146.52, 148.06, 148.42, 149.45, 149.50, 150.39, 150.51, 151.74, 152.16, 152.55, 153.05, 153.29, 153.79, 154.48, 155.01, 161.06, 171.88 (<u>C</u>OOH), 171.91 (<u>C</u>OOH), 171.92 (<u>C</u>OOH), 172.03 (<u>C</u>OOH).
ESI MS: m/z= 1969 ($[M]^-$).

10 (Yield 95%)**.** ^1H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 2.43 – 2.62 (m, 16H), 2.84 – 3.06 (m, 16H), 6.53 (d, 2H, *J* = 3.5 Hz), 6.57 (d, 2H, *J* = 3.4 Hz), 6.61 (d, 2H, *J* = 3.6 Hz), 6.64 (d, 2H, *J* = 3.5 Hz), 6.66 (d, 2H, *J* = 3.5 Hz), 6.73 (d, 2H, *J* = 3.4 Hz), 6.77 (d, 2H, *J* = 3.5 Hz), 6.94 (d, 2H, *J* = 3.5 Hz).
^{13}C NMR (126 MHz, $(CD_3)_2SO$, δ, ppm): 25.26 (<u>C</u>H_2), 25.29 (<u>C</u>H_2), 25.41(<u>C</u>H_2), 35.85(<u>C</u>H_2), 35.89 (<u>C</u>H_2), 35.91(<u>C</u>H_2), 56.70 ($\underline{C}_{sp3\ fullerene\ cage}$), 57.07 ($\underline{C}_{sp3\ fullerene\ cage}$), 57.33 ($\underline{C}_{sp3\ fullerene\ cage}$), 57.53 ($\underline{C}_{sp3\ fullerene\ cage}$), 124.99, 125.02, 125.13, 125.45, 126.73, 126.76, 126.89, 127.12, 131.57, 132.67, 133.97, 134.52, 135.63, 139.26, 139.47, 139.89, 141.02, 141.99, 142.20, 142.27, 144.37, 144.39, 144.72, 144.89, 145.24, 145.38, 145.78, 146.12, 146.41, 147.99, 148.46, 149.42, 149.51,

150.36, 150.47, 151.76, 152.14, 152.50, 152.58, 153.01, 153.25, 153.74, 154.44, 154.98, 155.05, 160.97, 173.72 (COOH), 173.80 (COOH).
ESI MS: m/z= 1041 ($[M]^{2-}$).

11 (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 1.65 – 1.90 (m, 16H), 2.12 – 2.32 (m, 16H), 2.60 – 2.82 (m, 16H), 6.48 (d, 2H, *J* = 3.5 Hz), 6.52 (d, 2H, *J* = 3.4 Hz), 6.59 (d, 2H, *J* = 3.2 Hz), 6.68 – 6.71 (m, 4H), 6.73 (d, 2H, *J* = 3.5 Hz), 6.82 (d, 2H, *J* = 3.4 Hz), 7.02 (d, 2H, *J* = 3.5 Hz).
^{13}C NMR (126 MHz, DMSO, δ, ppm): 26.92 (CH_2), 27.00 (CH_2), 29.08 (CH_2), 29.21 (CH_2), 33.04 (CH_2), 33.07 (CH_2), 33.17 (CH_2), 56.72 ($C_{sp3\ fullerene\ cage}$), 57.11 ($C_{sp3\ fullerene\ cage}$), 57.36 ($C_{sp3\ fullerene\ cage}$), 57.56 ($C_{sp3\ fullerene\ cage}$), 124.78, 124.81, 124.88, 125.27, 126.78, 126.81, 126.96, 127.13, 131.62, 132.82, 134.13, 134.52, 135.63, 139.22, 139.48, 139.91, 140.90, 142.00, 142.14, 142.25, 145.18, 145.22, 145.29, 145.40, 145.50, 145.67, 145.73, 146.23, 146.57, 148.03, 148.48, 149.40, 149.45, 150.34, 150.47, 151.83, 152.10, 152.15, 152.55, 152.95, 153.11, 153.72, 154.30, 155.02, 161.05, 174.55 (COOH), 174.57 (COOH), 174.59 (COOH).
ESI MS: m/z= 1097 ($[M]^{2-}$).

12a (Yield 95%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$, δ, ppm): 2.44 (t, 2H, *J* = 7.9 Hz), 2.82 (t, 2H, *J* = 7.9 Hz), 3.75 (d, 2H, *J* = 5.8 Hz), 7.16 – 7.29 (m, 5H), 8.19 (s, 1H).
12b (Yield 95%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$, δ, ppm): 2.32 – 2.38 (m, 4H), 2.79 (t, 2H, *J* = 7.8 Hz), 3.20 – 3.26 (m, 2H), 7.15 – 7.29 (m, 5H), 7.91 (s, 1H).
12c (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 3.69 (d, 2H, *J* = 0.4 Hz), 3.76 (d, 2H, *J* = 5.9 Hz), 6.89 – 6.99 (m, 2H), 7.34 (dd, 1H, *J* = 5.1, 1.3 Hz), 8.39 (t, 1H, J = 5.7 Hz).
12d (Yield 95%). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, δ, ppm): 2.41 (t, 2H, *J* = 6.8 Hz), 3.24 – 3.31 (m, 2H), 3.64 (s, 2H), 6.89 – 6.92 (m, 1H), 6.96 (dd, 1H, *J* = 5.1, 3.4 Hz), 7.37 (dd, 2H, *J* = 5.1, 1.3 Hz), 8.19 (t, 1H, *J* = 5.1 Hz).

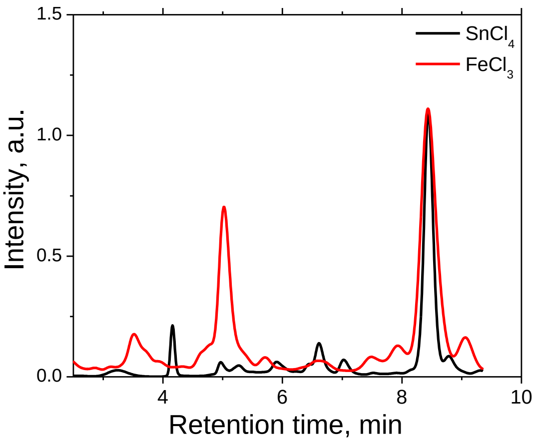

Fig. S1. HPLC profiles of the $C_{60}Cl_6$ reaction with hydrocinnamic acid methyl ester catalyzed by $FeCl_3$ and $SnCl_4$ (C18 Cosmosil column, elution with toluene/acetonitrile mixtures 80%/20% v/v, 30oC, flow rate 1 mL/min)

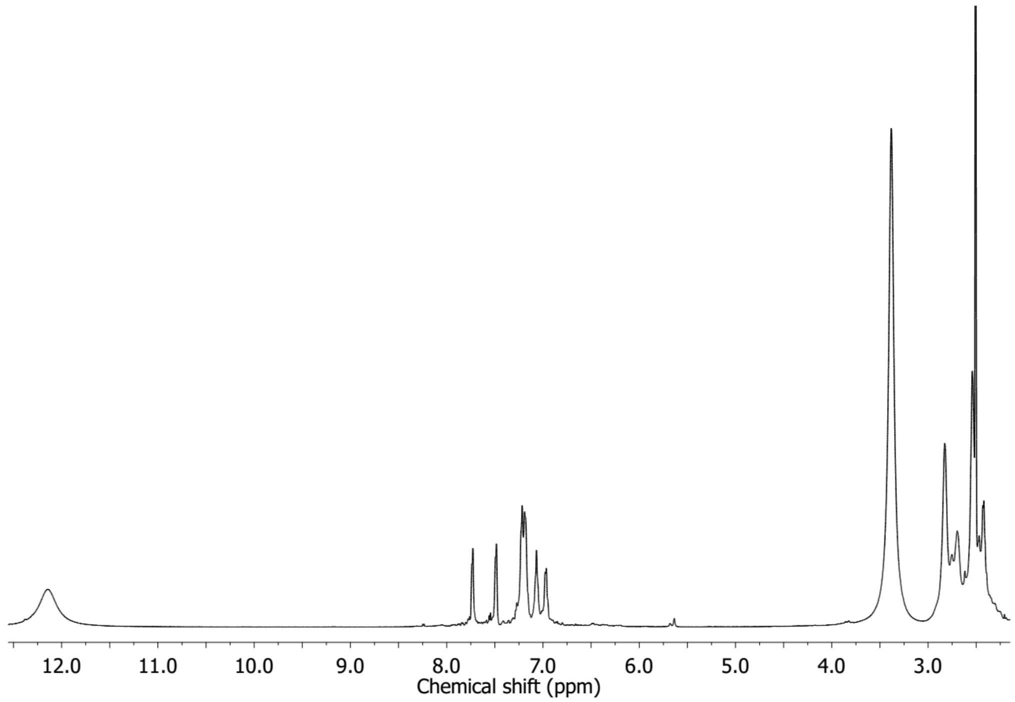

Fig. S2. ^{1}H NMR spectrum of compound **1**

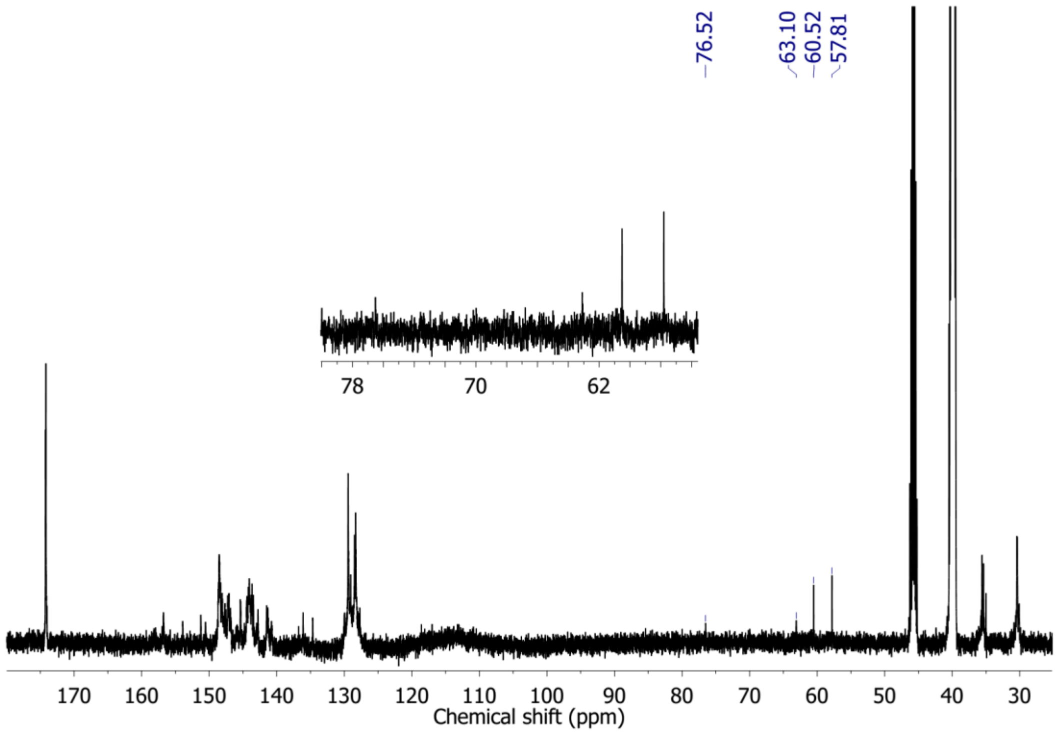

Fig. S3. ^{13}C NMR spectrum of compound **1**

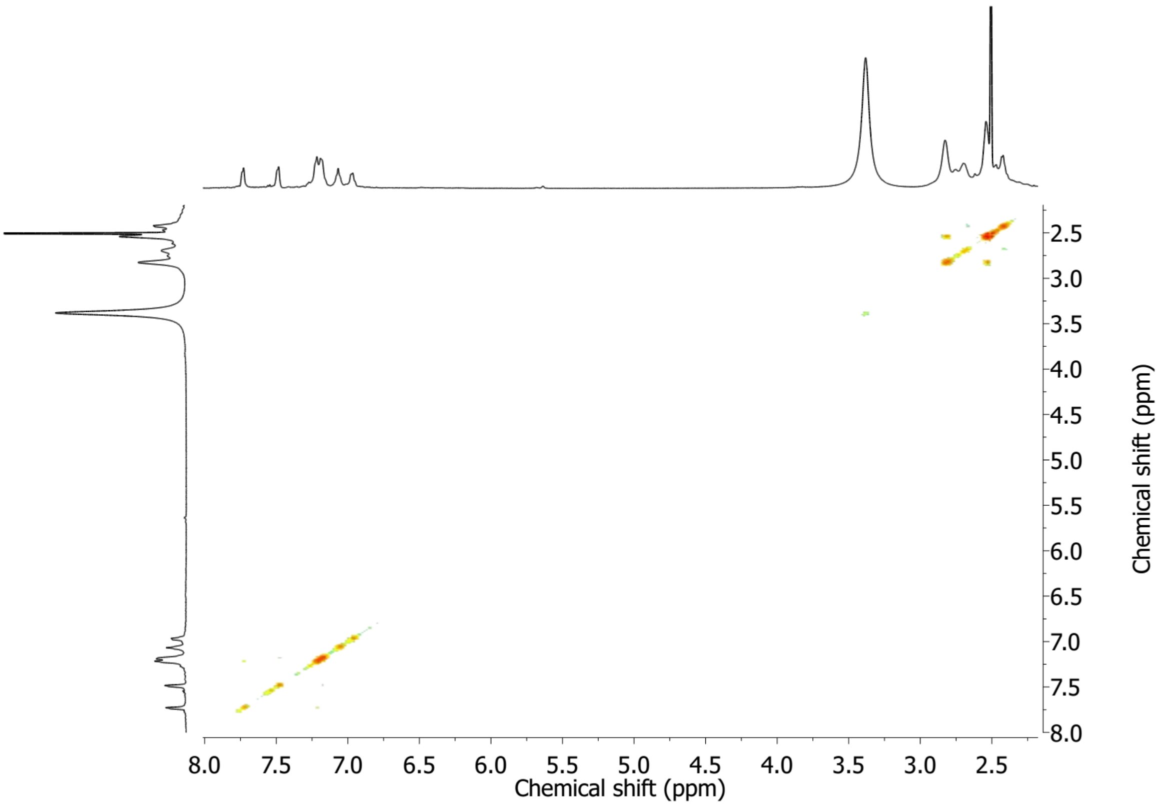

Fig. S4. ^{1}H-^{1}H COSY NMR spectrum of compound **1**

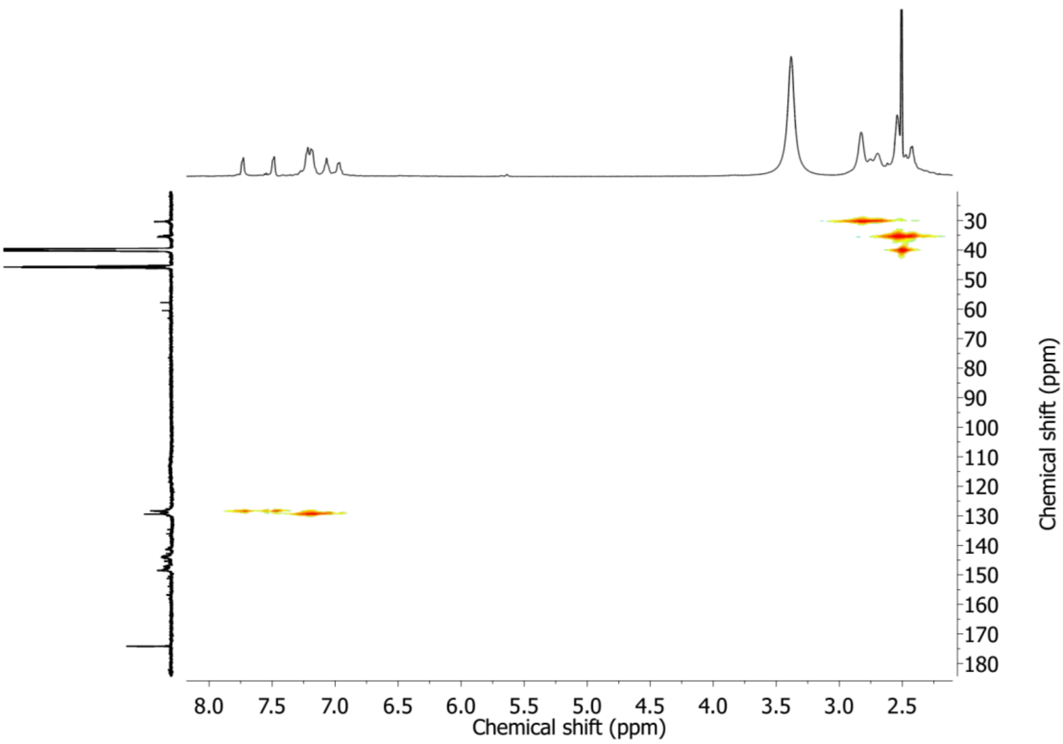

Fig. S5. ^{1}H-^{13}C HSQC NMR spectrum of compound **1**

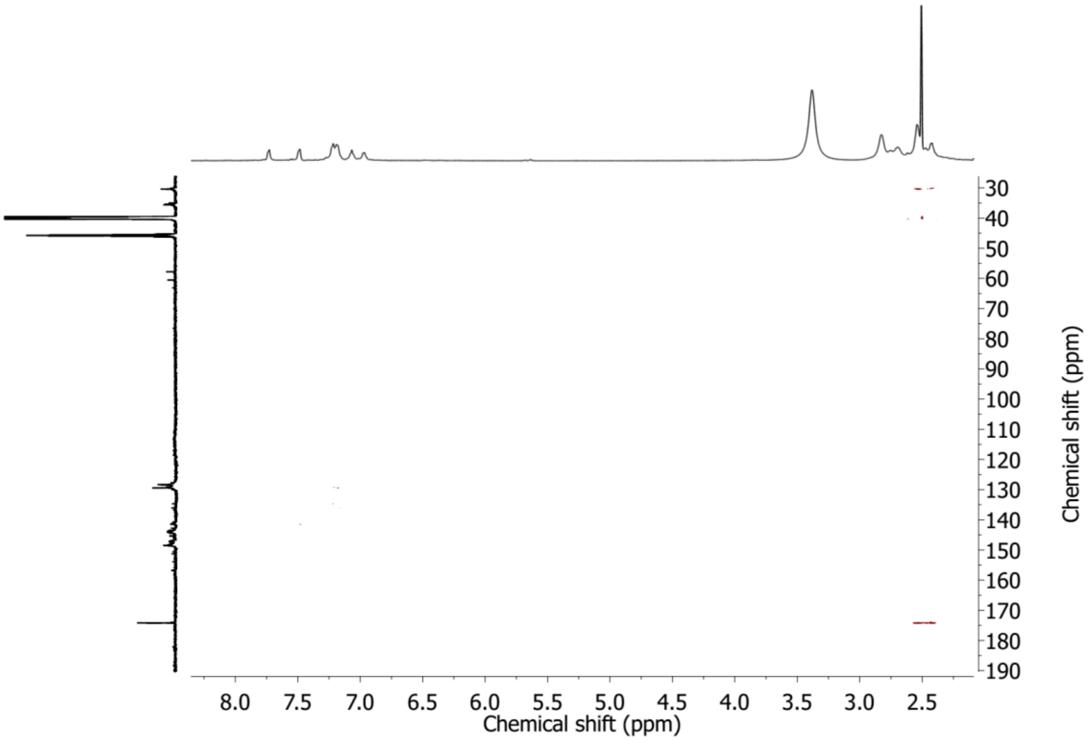

Fig. S6. ^{1}H-^{13}C HMBC NMR spectrum of compound **1**

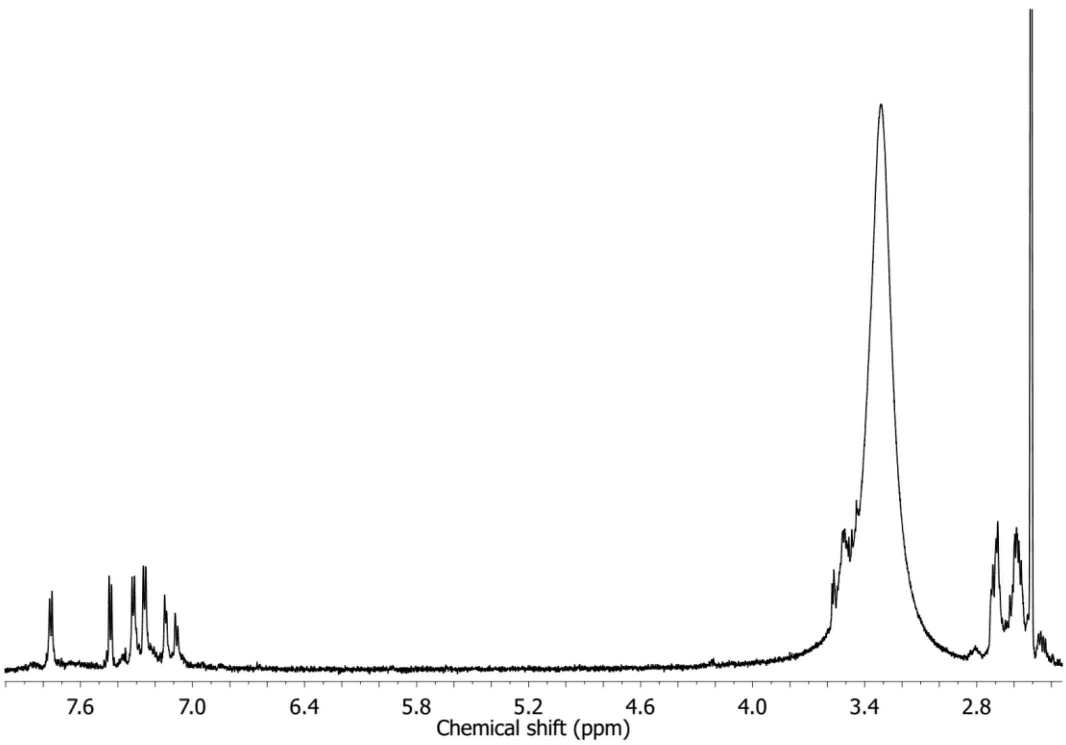

Fig. S7. ^{1}H NMR spectrum of compound **2**

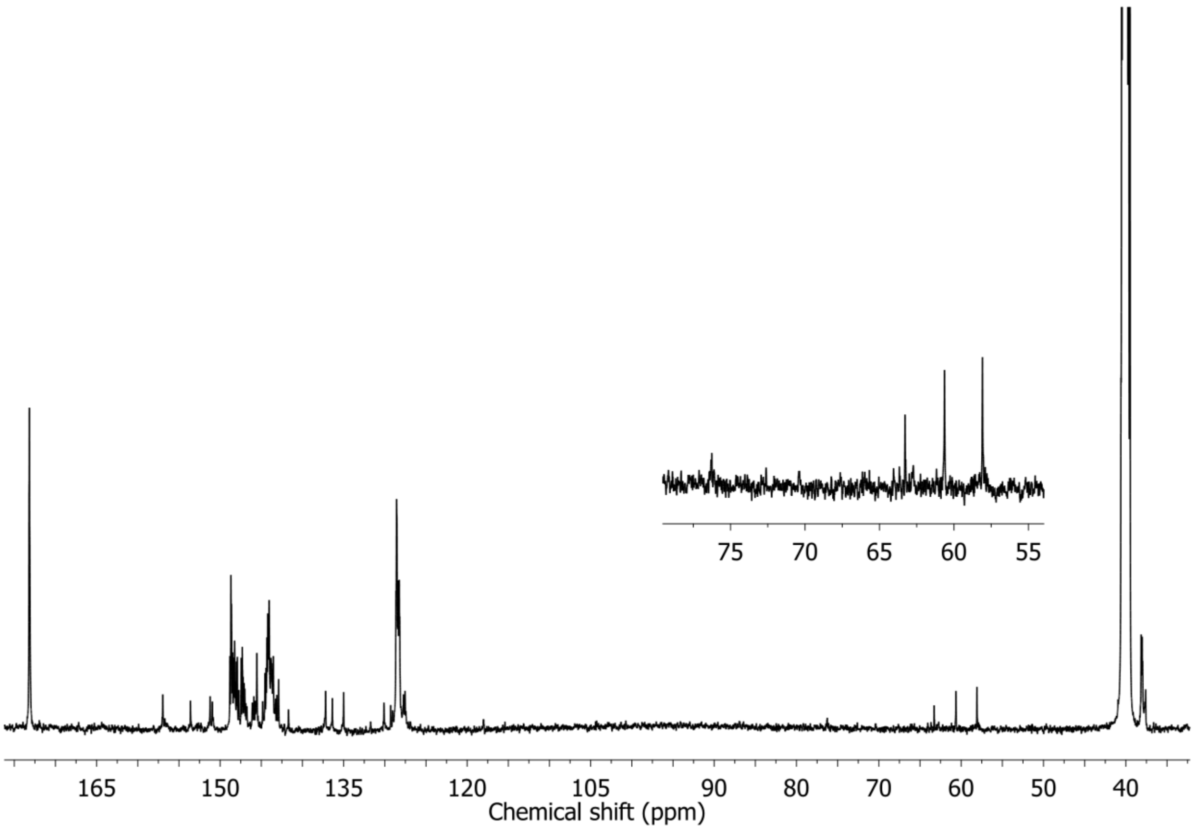

Fig. S8. ^{13}C NMR spectrum of compound **2**

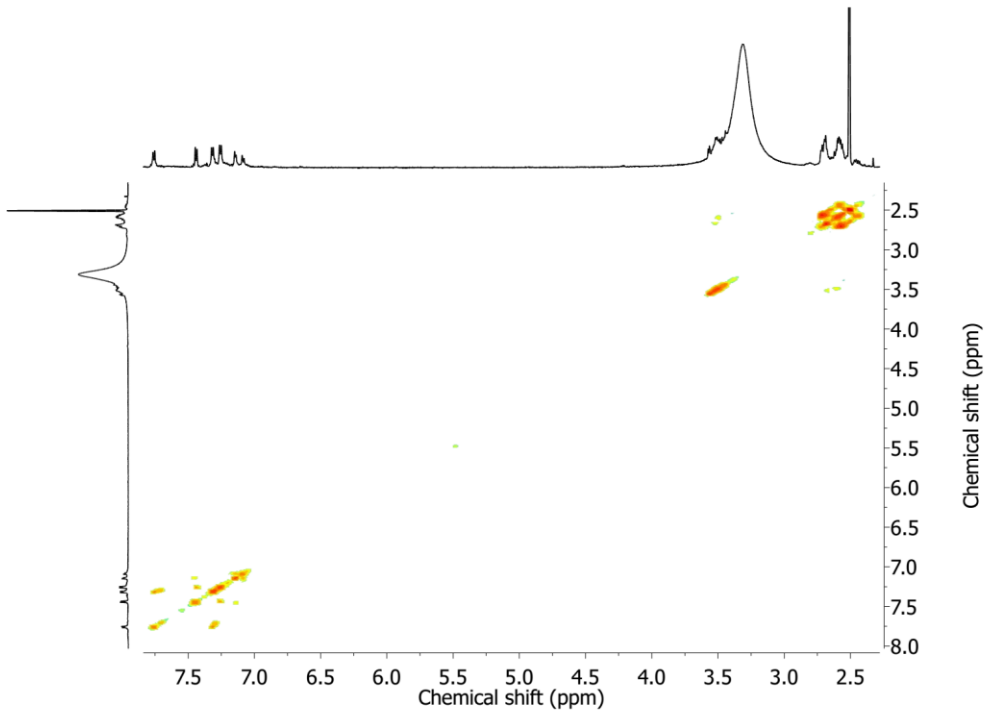

Fig. S9. ^{1}H-^{1}H COSY NMR spectrum of compound **2**

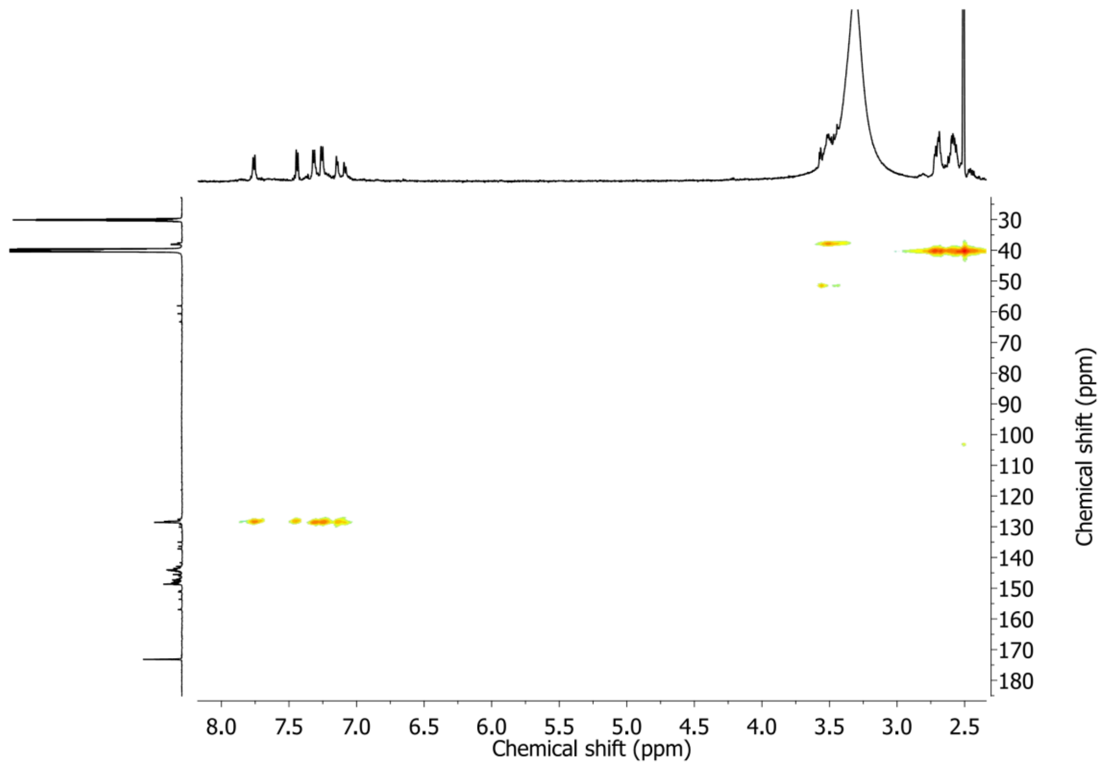

Fig. S10. ^{1}H-^{13}C HSQC NMR spectrum of compound **2**

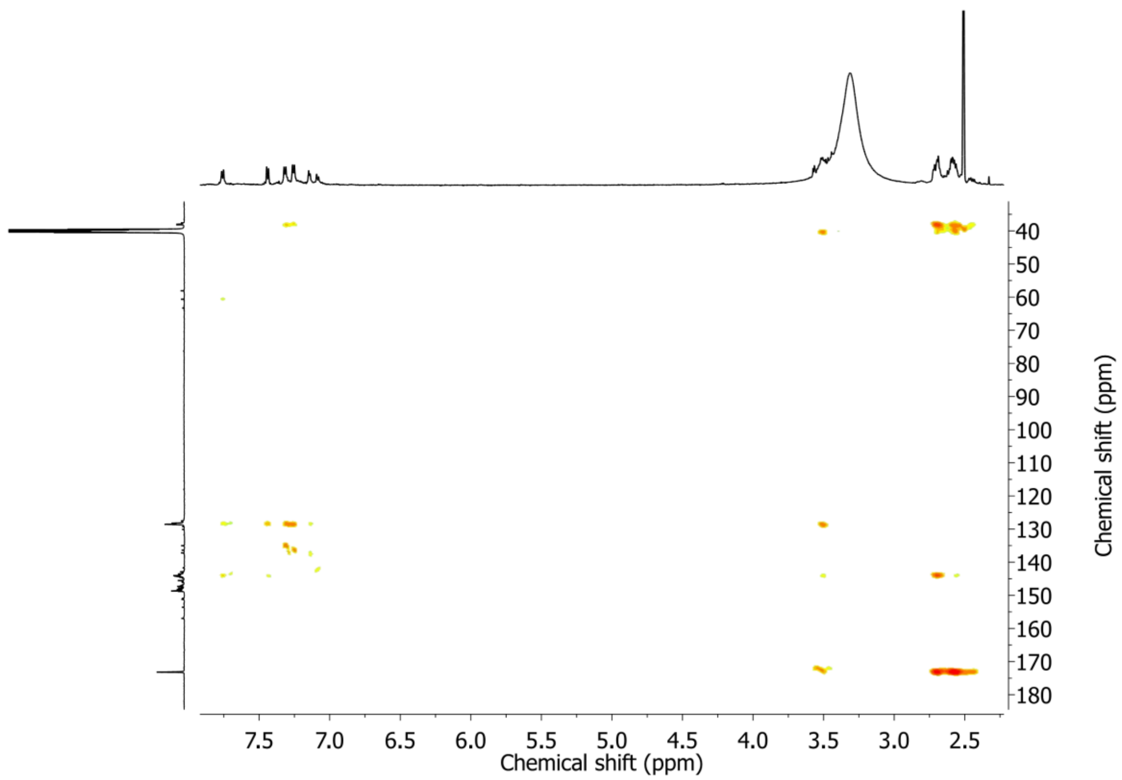

Fig. S11. ^{1}H-^{13}C HMBC NMR spectrum of compound **2**

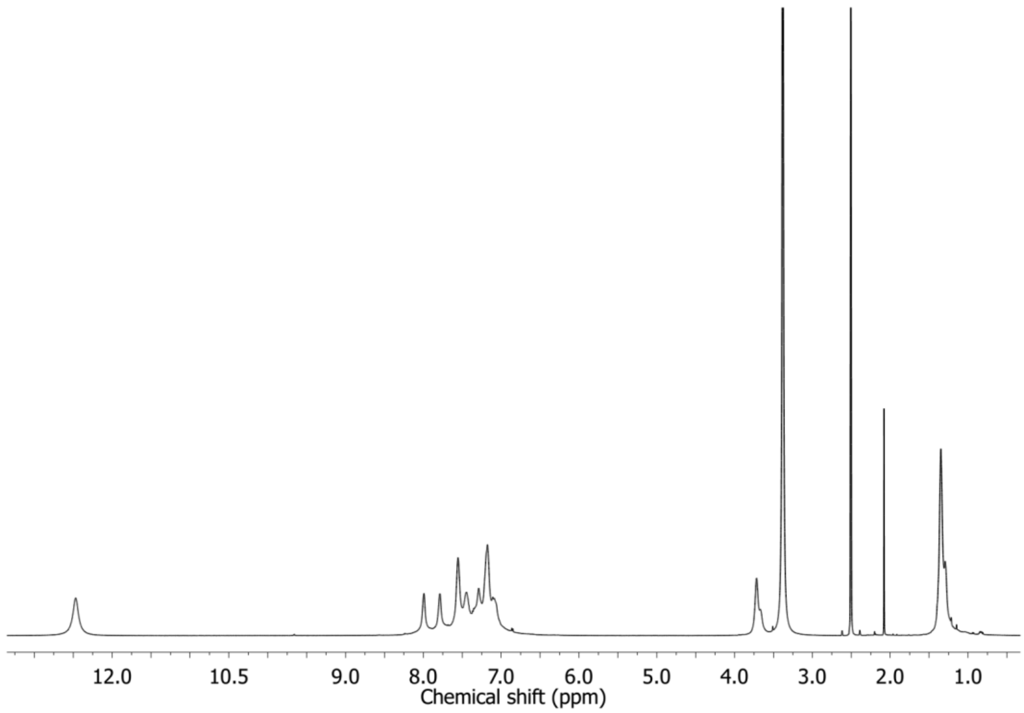

Fig. S12. ^{1}H NMR spectrum of compound **3**

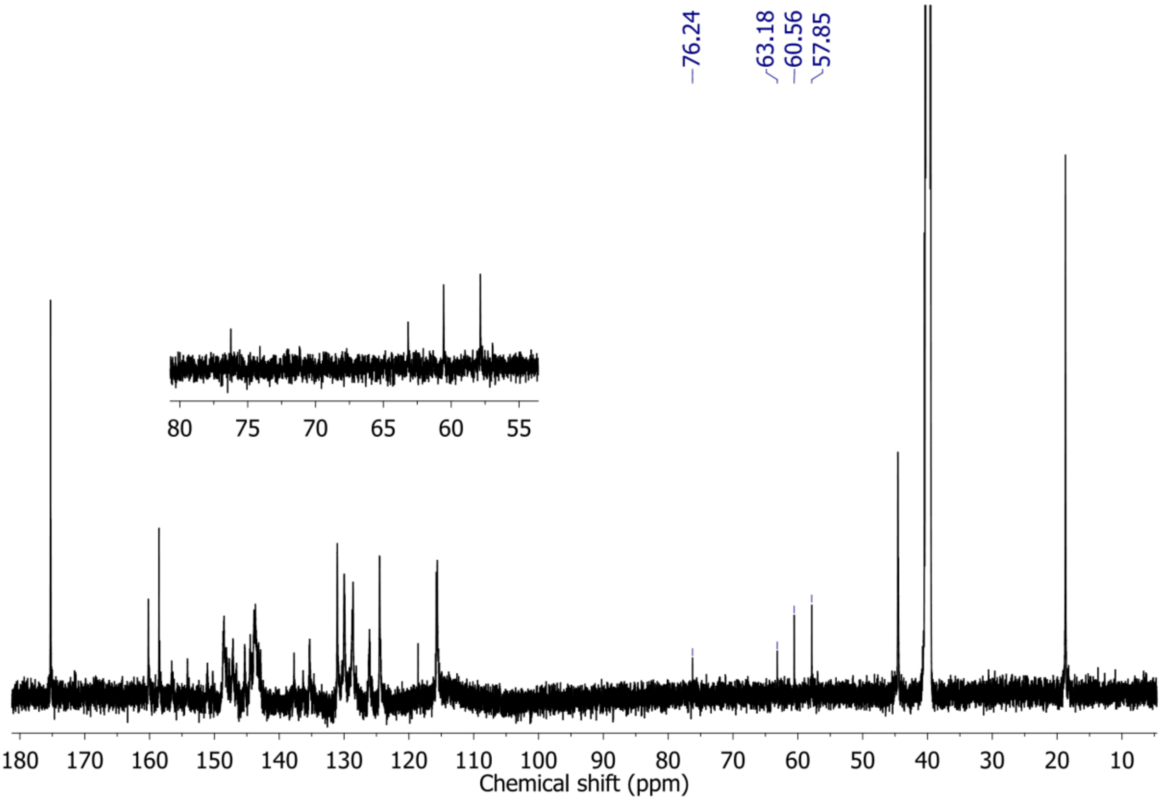

Fig. S13. ^{13}C NMR spectrum of compound **3**

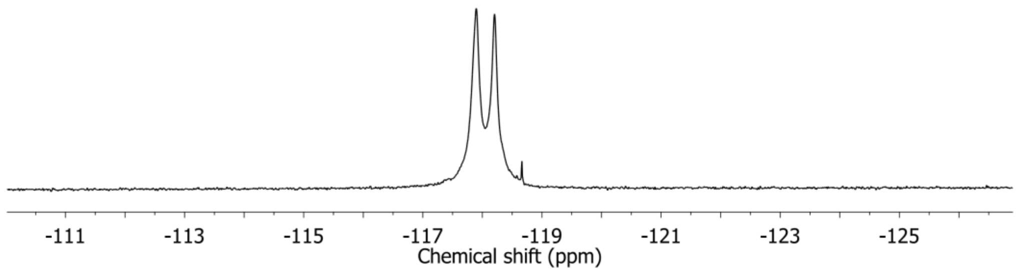

Fig. S14. ^{19}F NMR spectrum of compound **3**

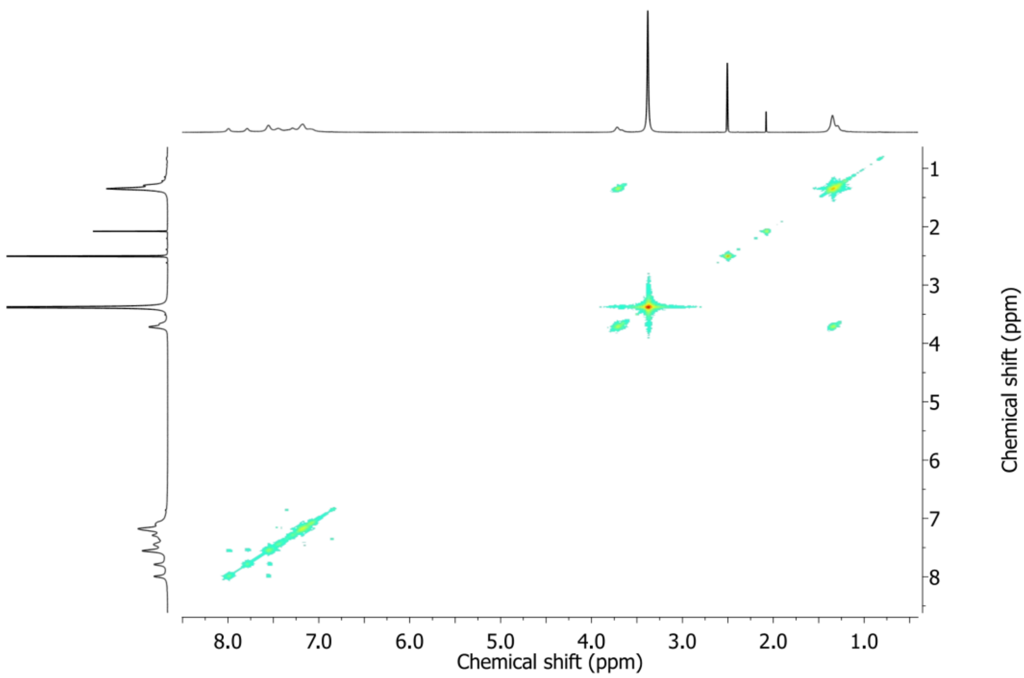

Fig. S15. ^{1}H-^{1}H COSY NMR spectrum of compound **3**

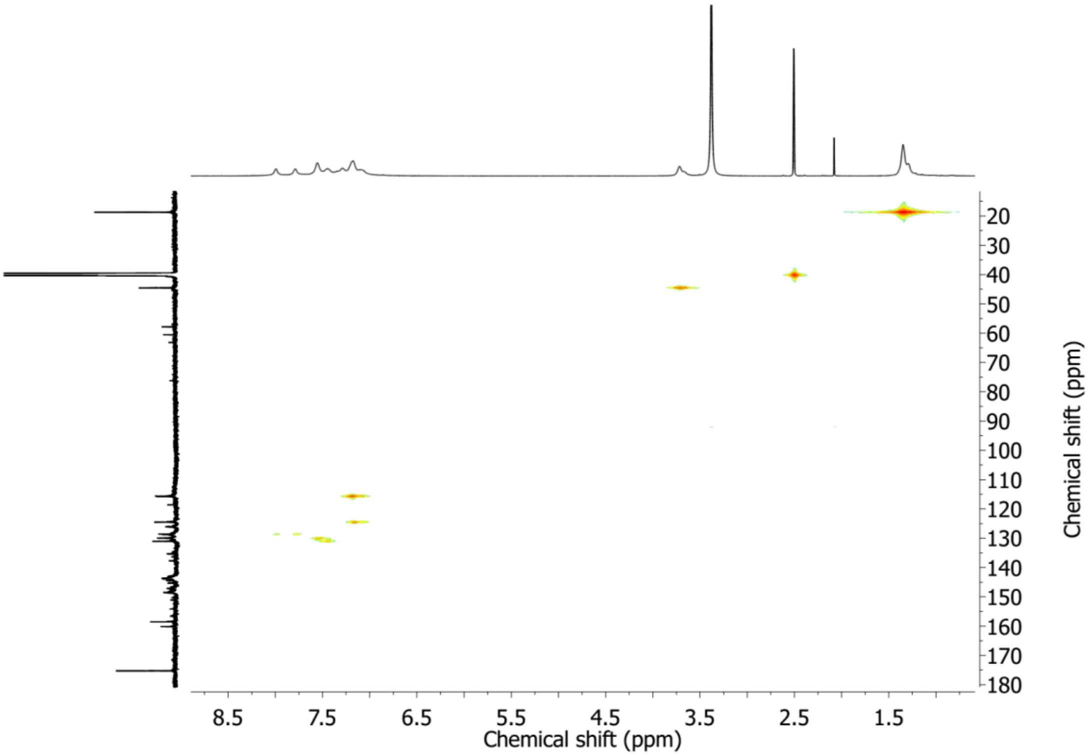

Fig. S16. ^{1}H-^{13}C HSQC NMR spectrum of compound **3**

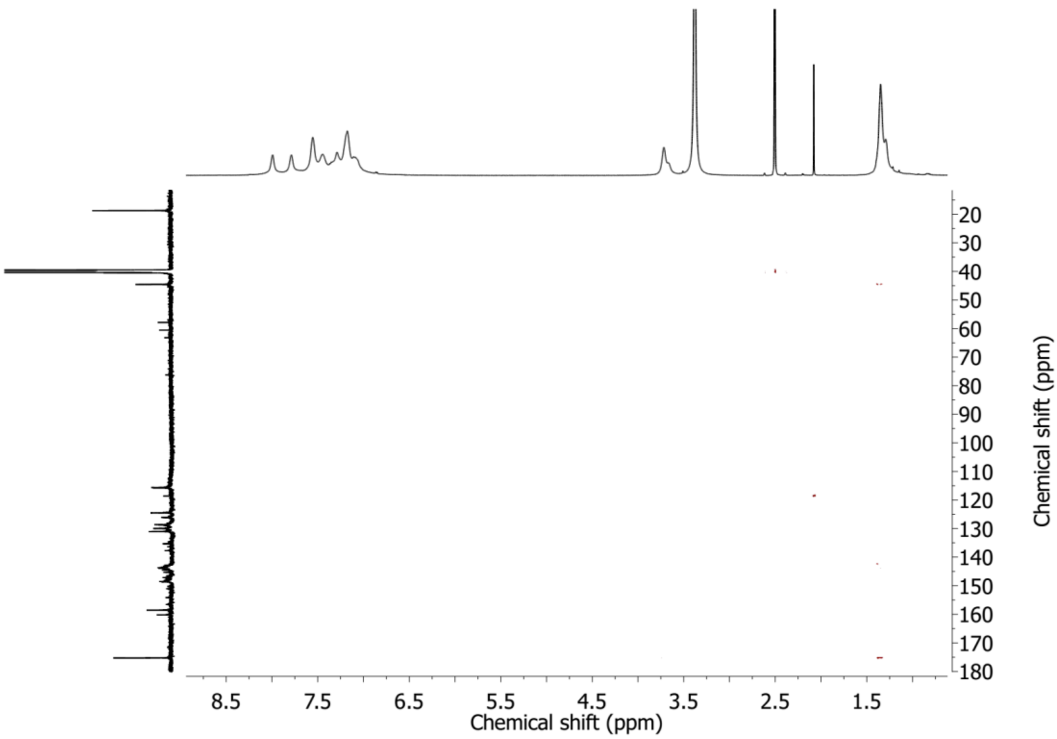

Fig. S17. ^{1}H-^{13}C HMBC NMR spectrum of compound **3**

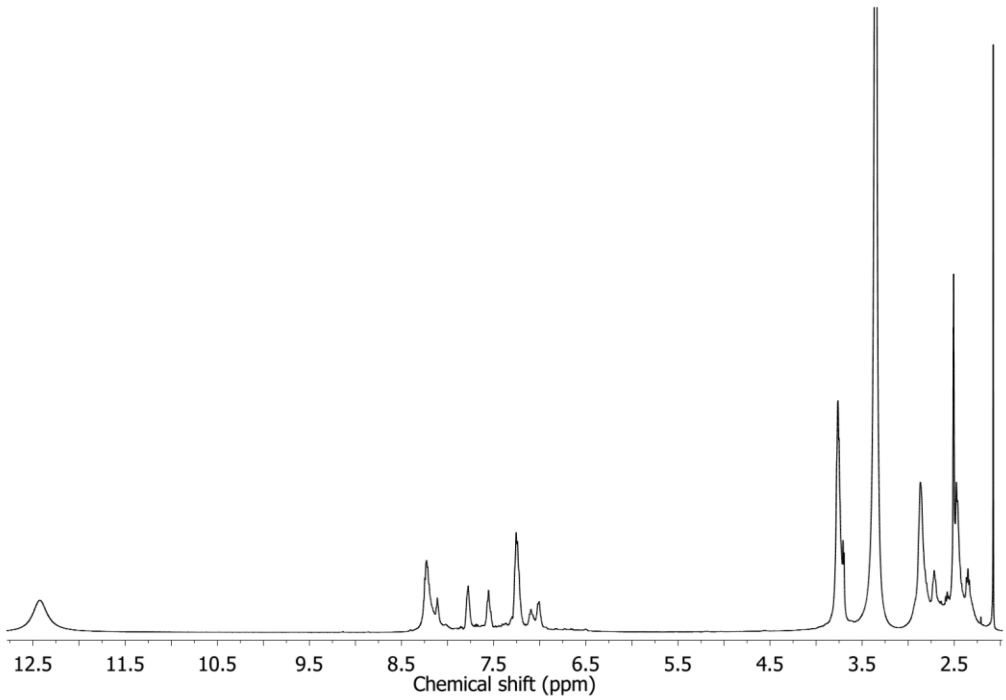

Fig. S18. ^{1}H NMR spectrum of compound **4**

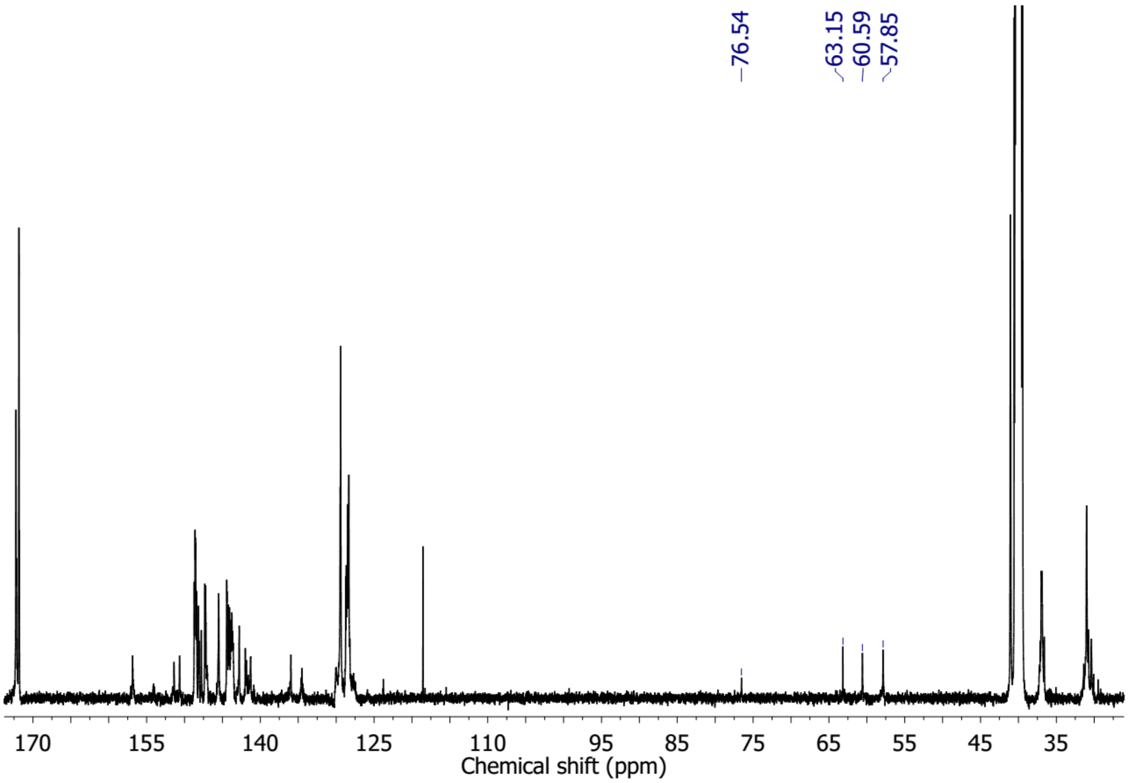

Fig. S19. ^{13}C NMR spectrum of compound **4**

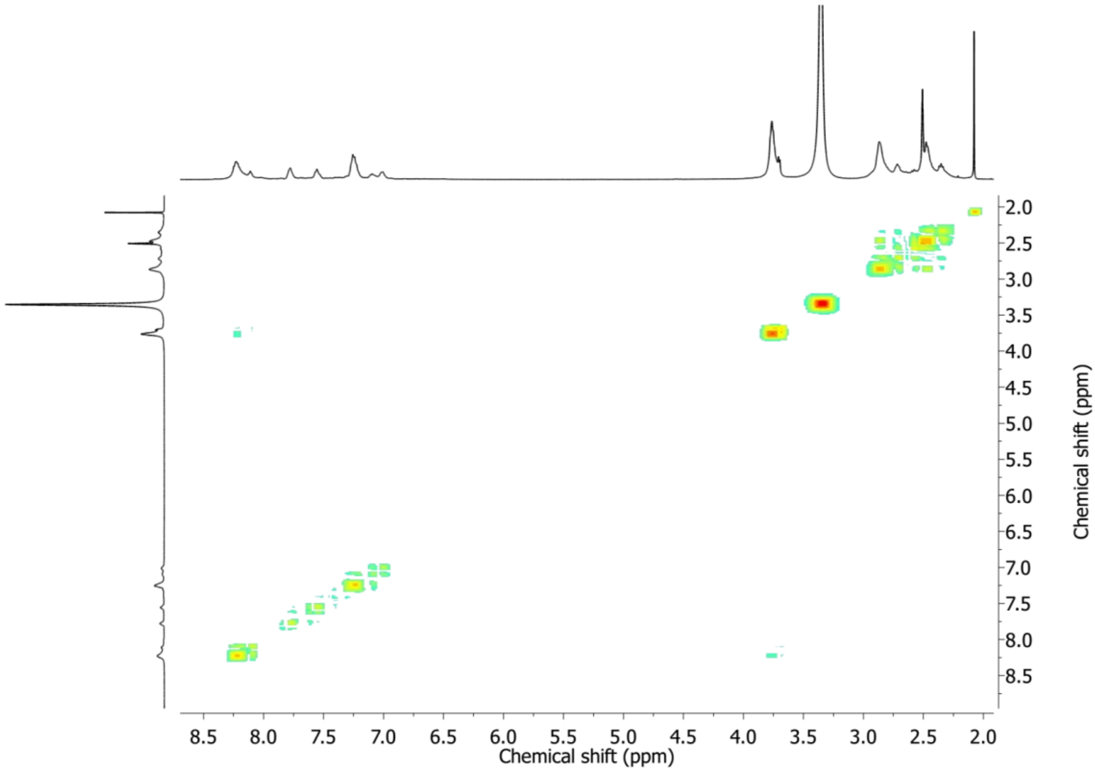

Fig. S20. ^{1}H-^{1}H COSY NMR spectrum of compound **4**

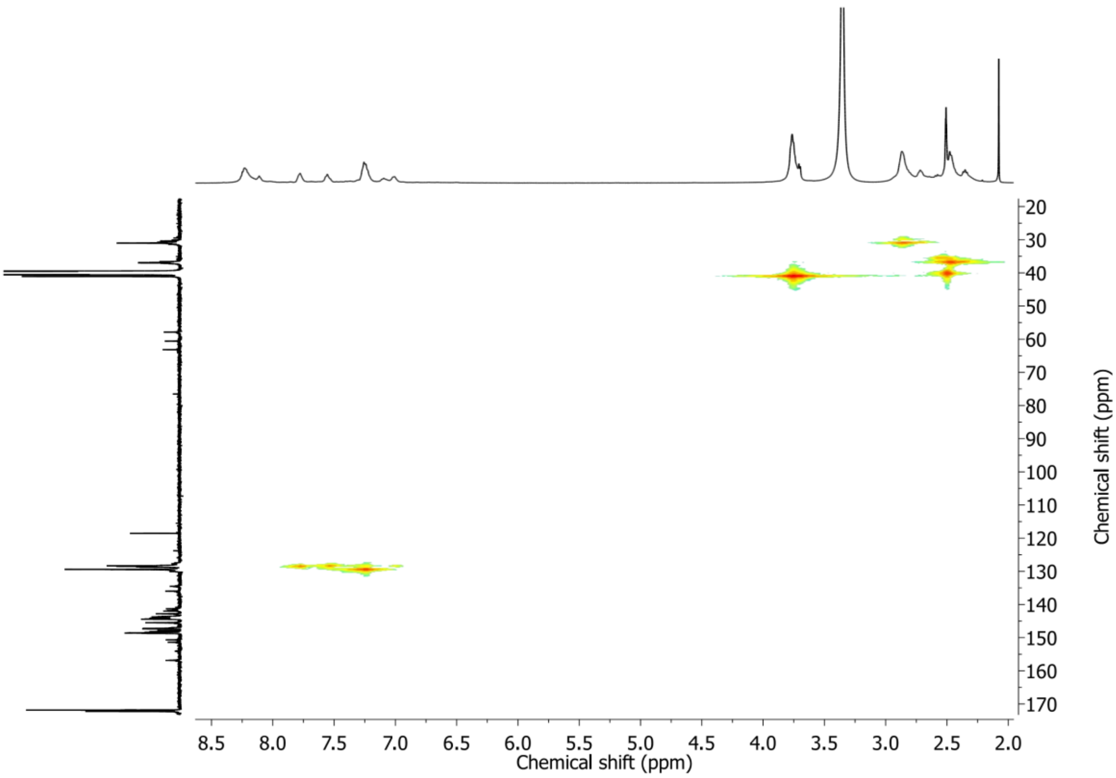

Fig. S21. ^{1}H-^{13}C HSQC NMR spectrum of compound **4**

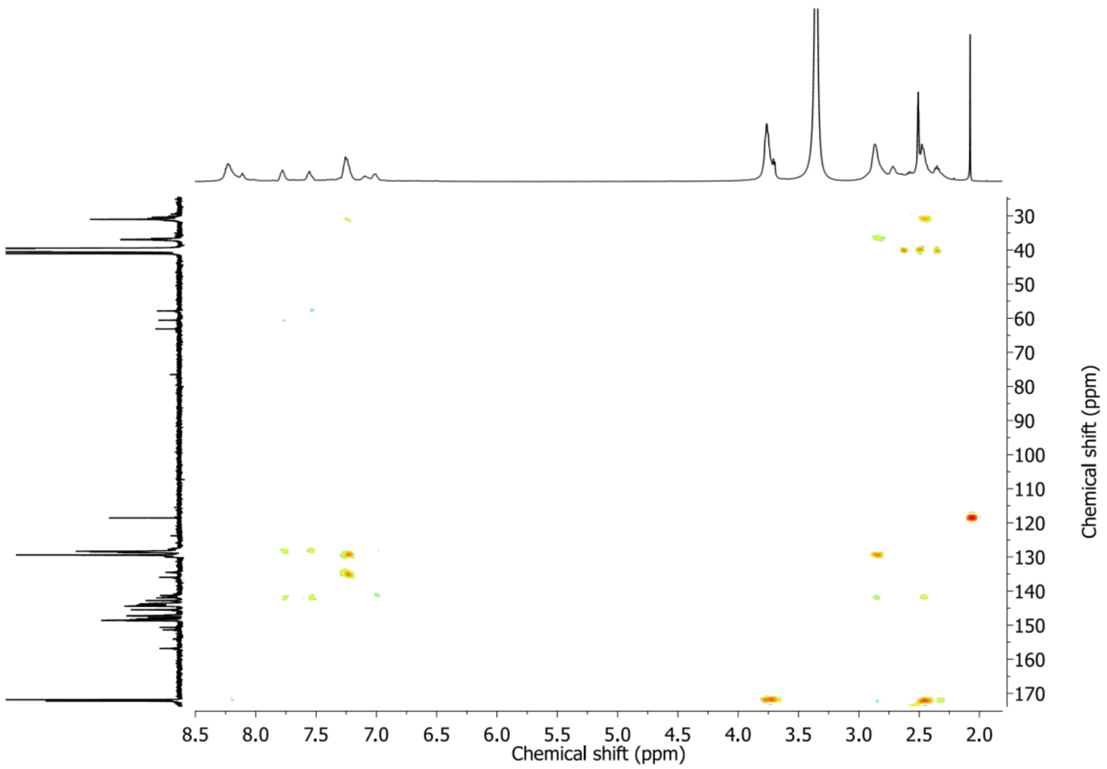

Fig. S22. ^{1}H-^{13}C HMBC NMR spectrum of compound **4**

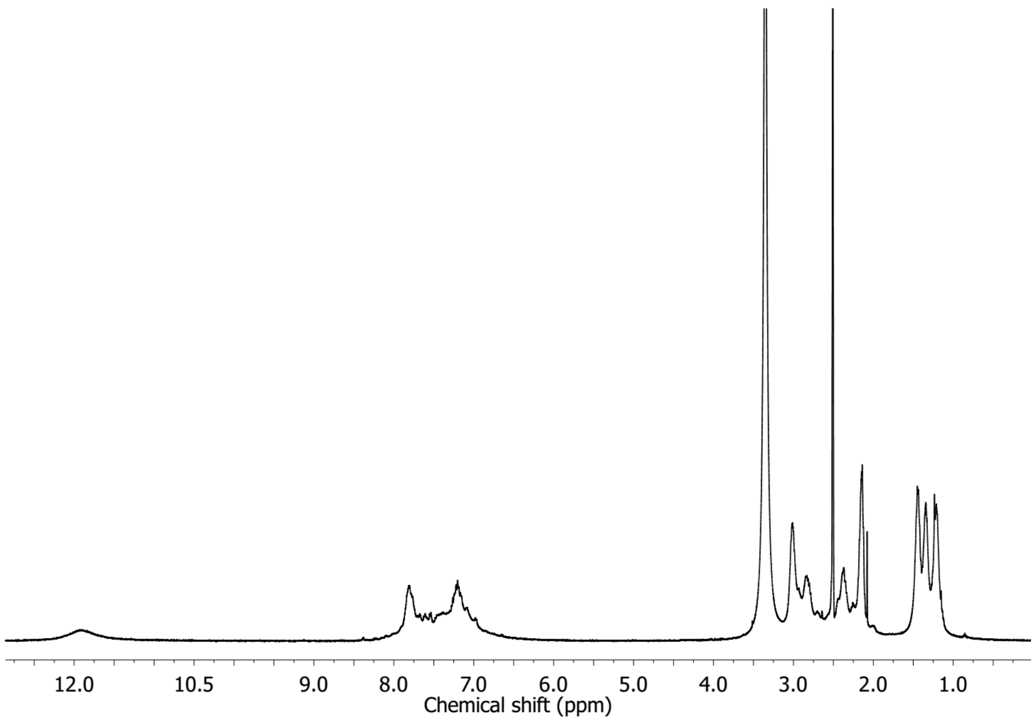

Fig. S23. ^{1}H NMR spectrum of compound **5**

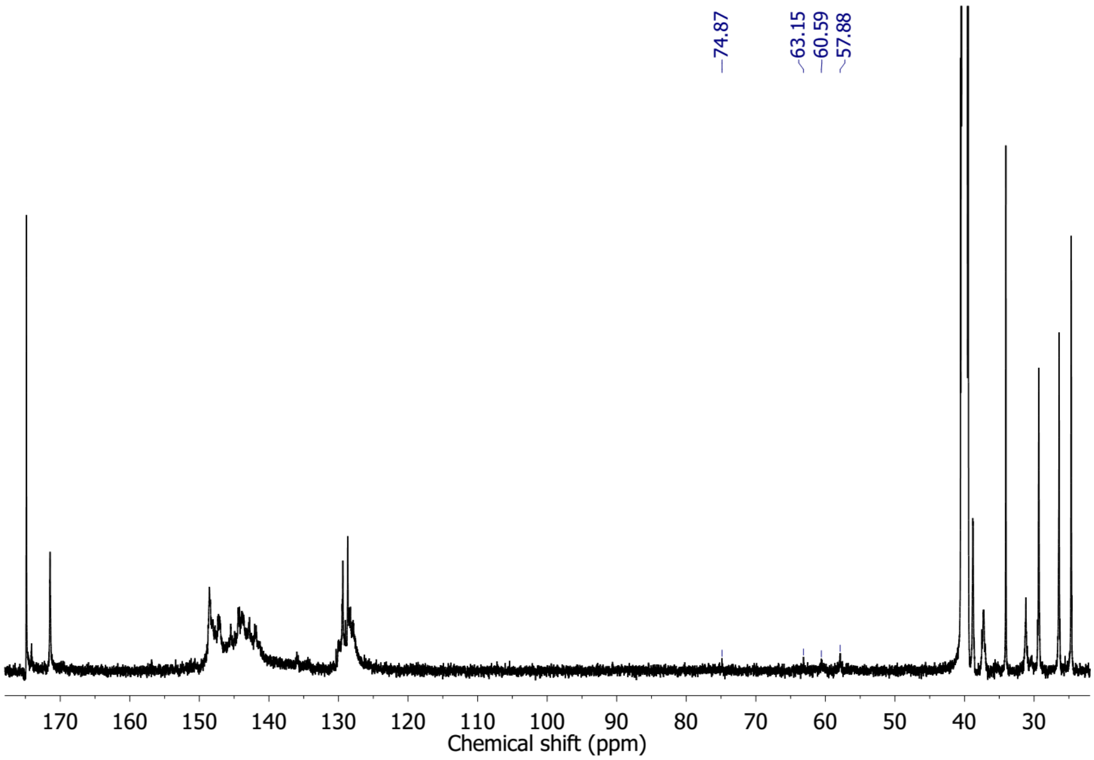

Fig. S24. ^{13}C HMBC NMR spectrum of compound **5**

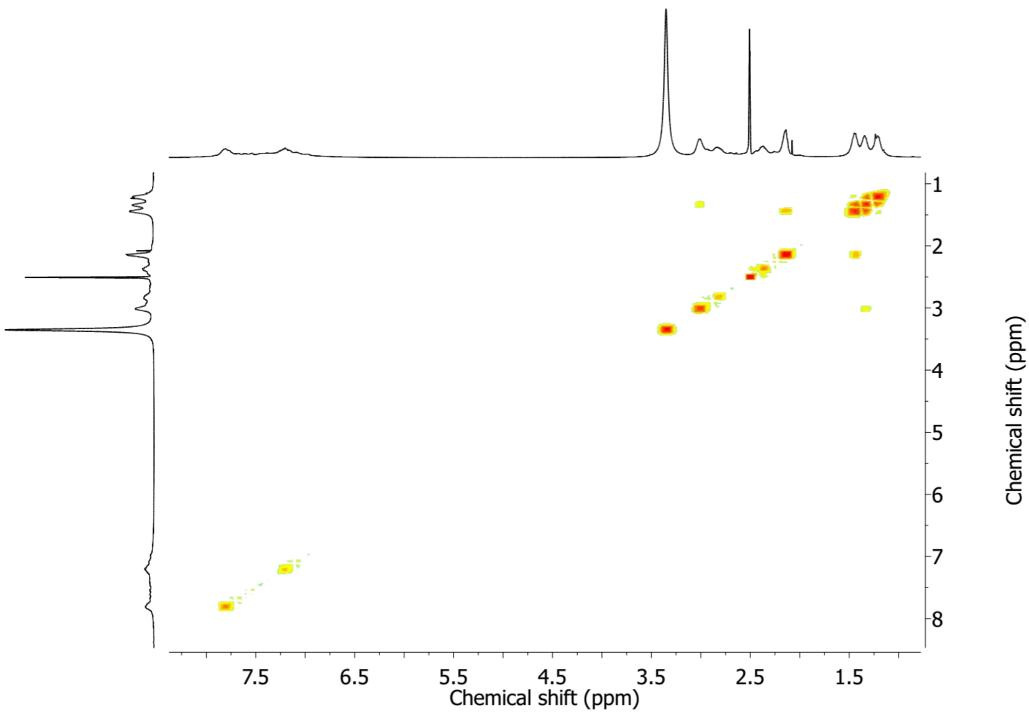

Fig. S25. ^{1}H-^{1}H COSY NMR spectrum of compound **5**

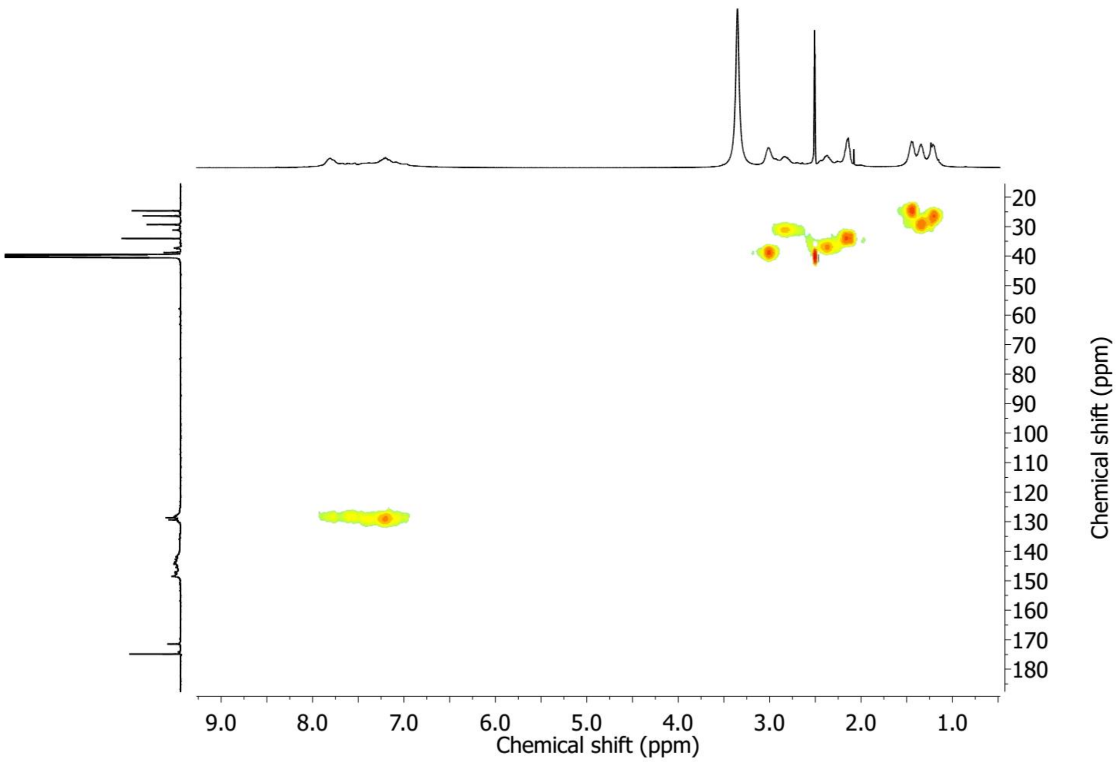

Fig. S26. ^{1}H-^{13}C HSQC NMR spectrum of compound **5**

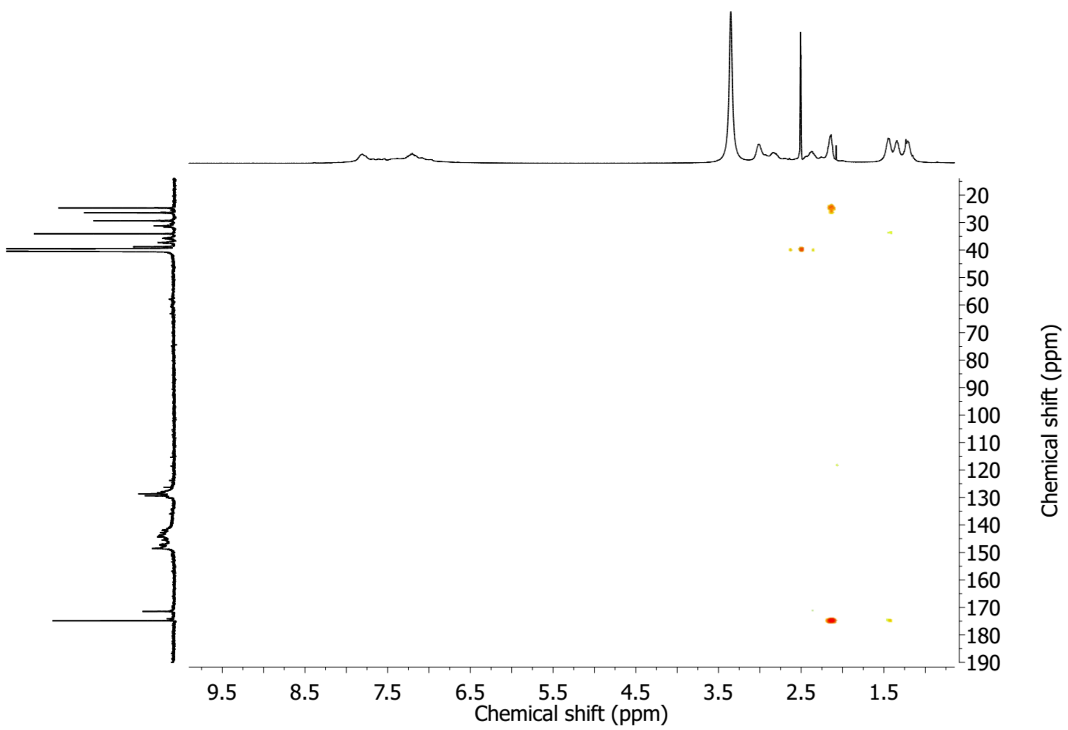

Fig. S27. ^{1}H-^{13}C HMBC NMR spectrum of compound **5**

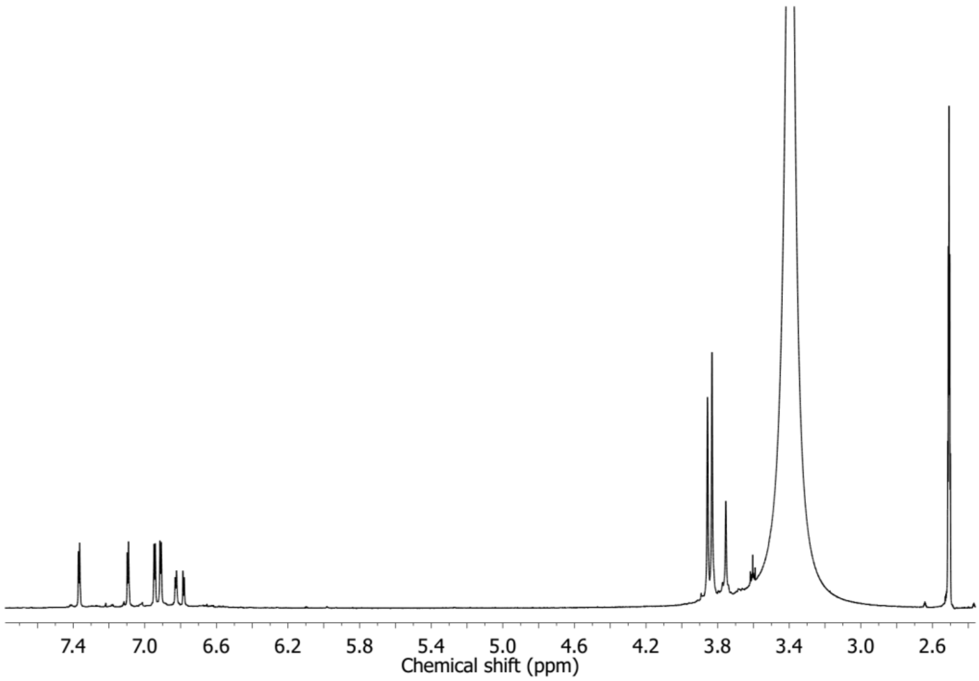

Fig. S28. ¹H NMR spectrum of compound **6**

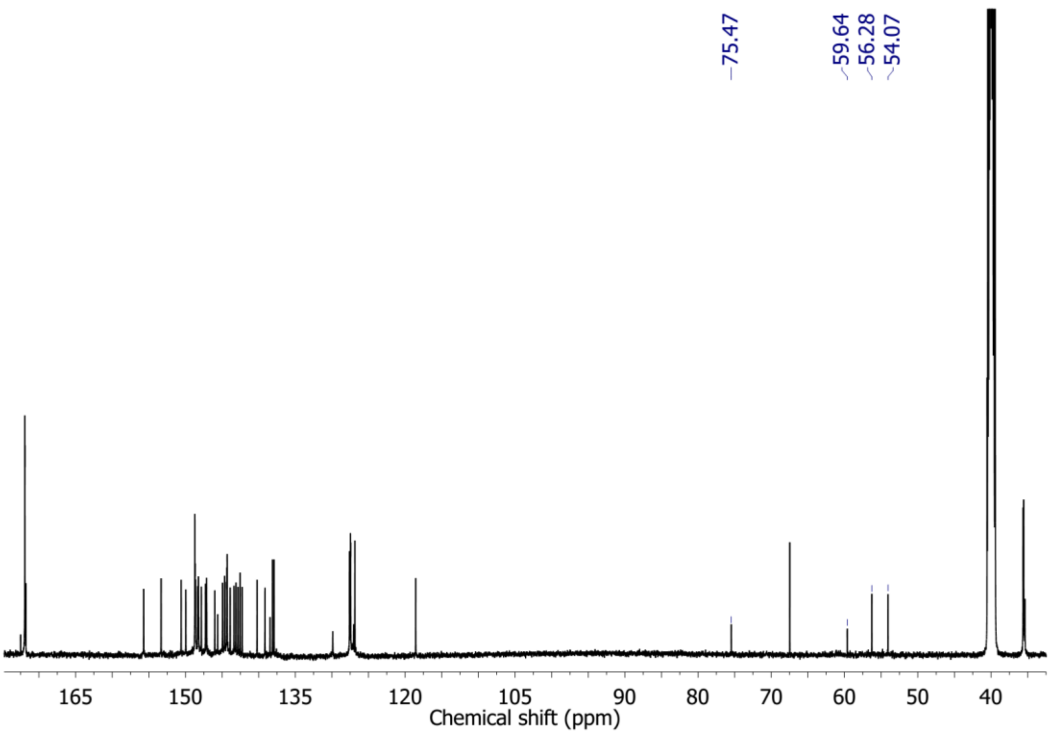

Fig. S29. ¹³C NMR spectrum of compound **6**

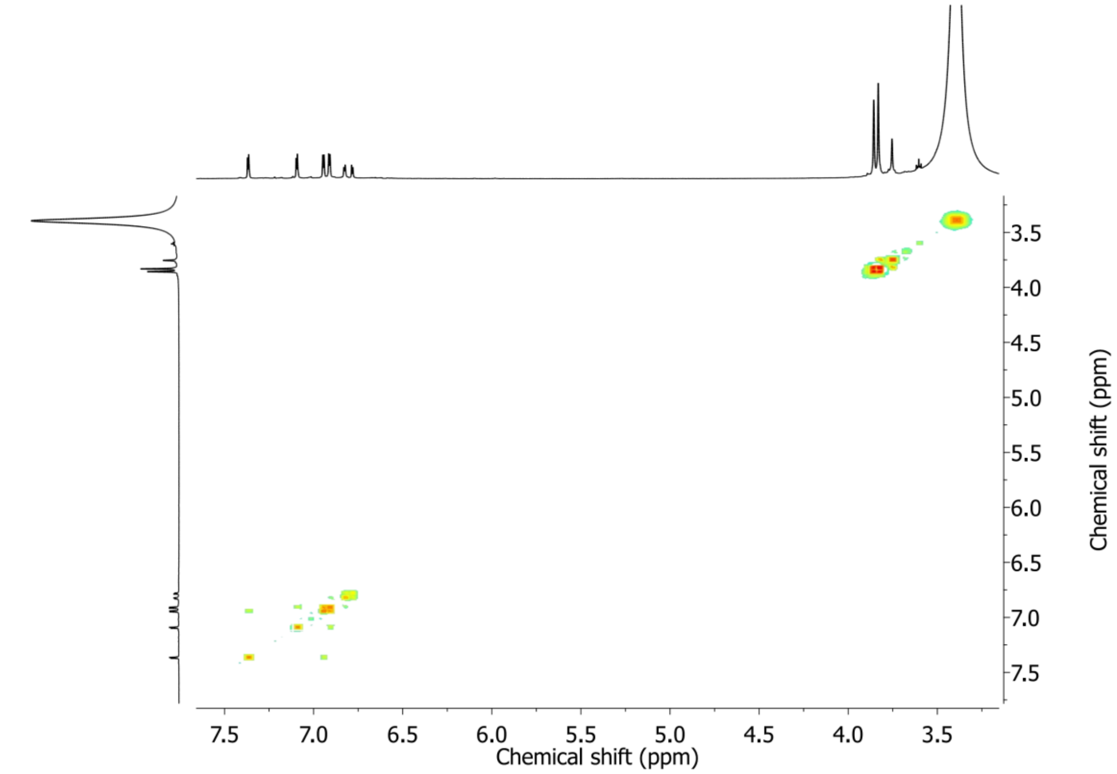

Fig. S30. ^{1}H-^{1}H COSY NMR spectrum of compound **6**

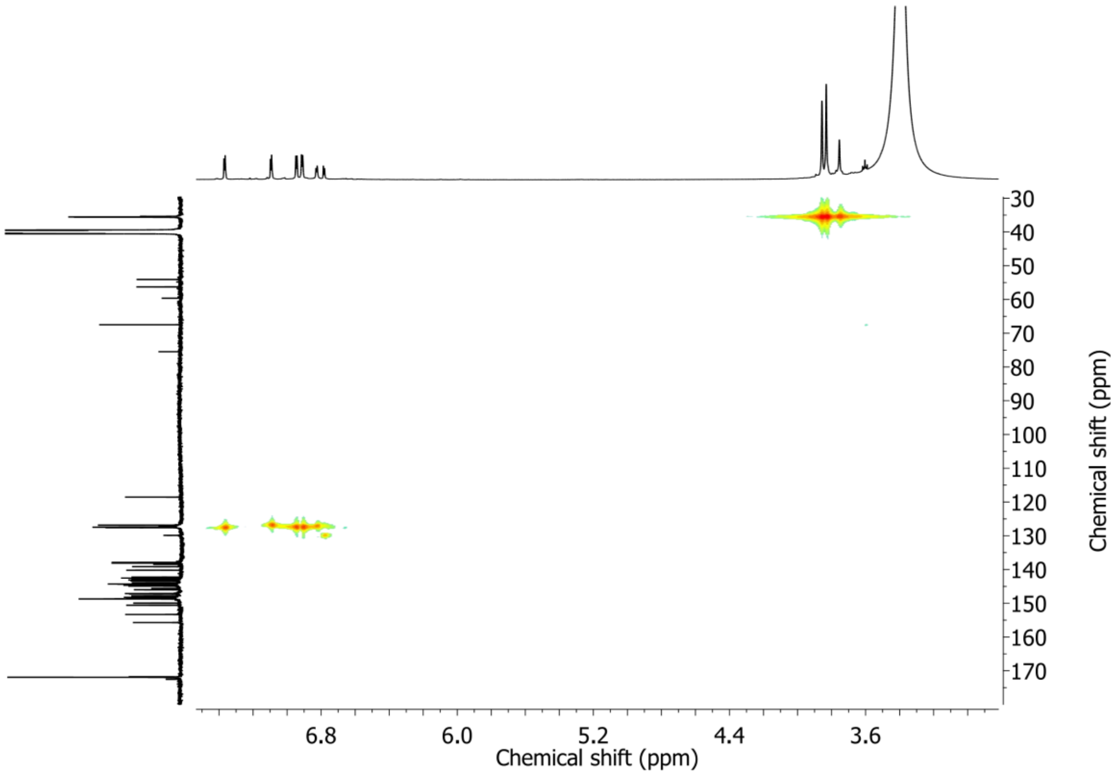

Fig. S31. ^{1}H-^{13}C HSQC NMR spectrum of compound **6**

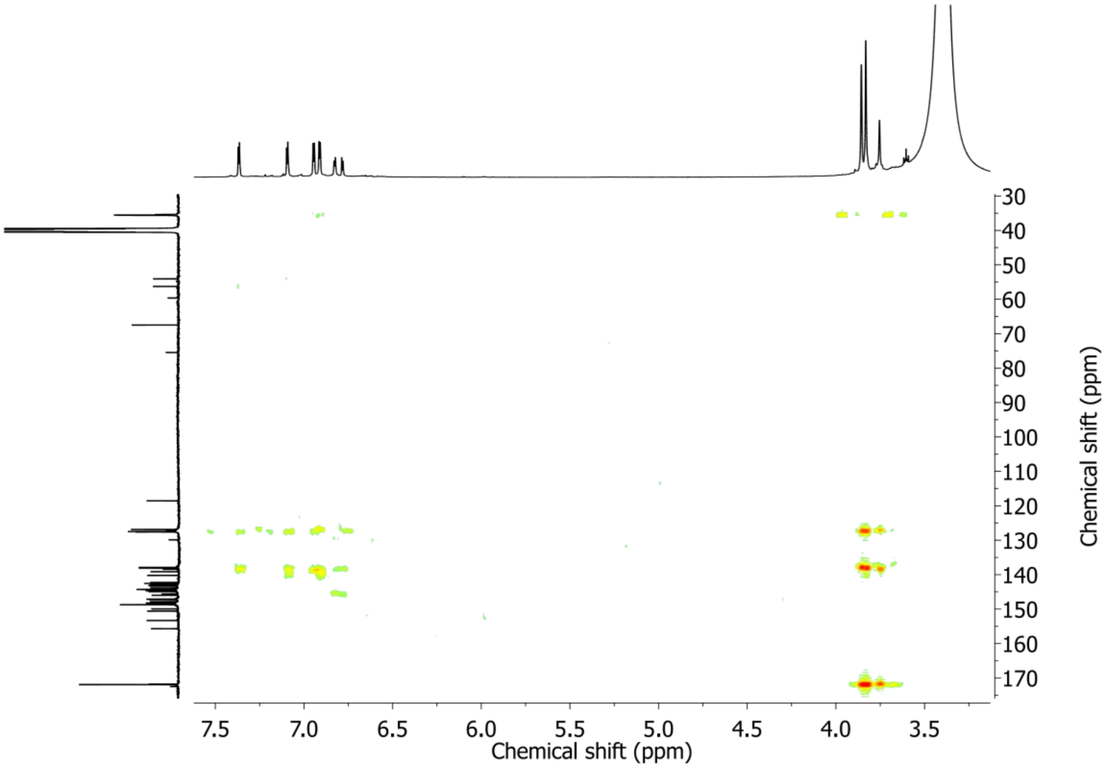

Fig. S32. ^{1}H-^{13}C HMBC NMR spectrum of compound **6**

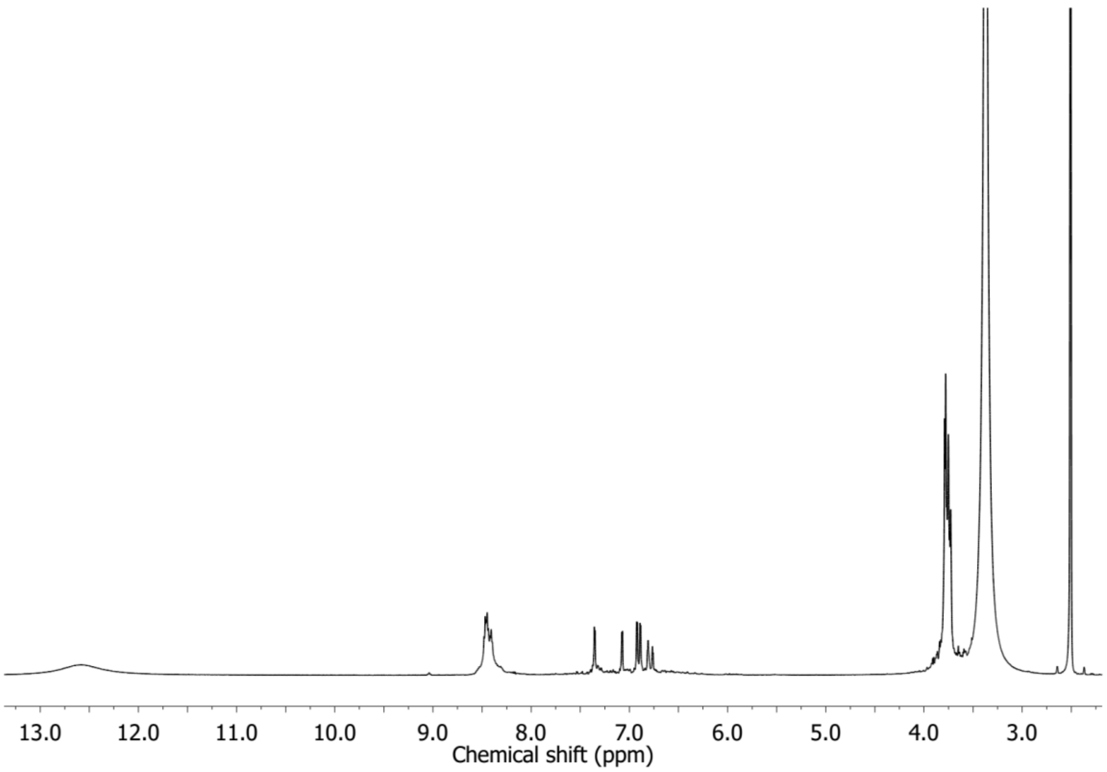

Fig. S33. ^{1}H NMR spectrum of compound **7**

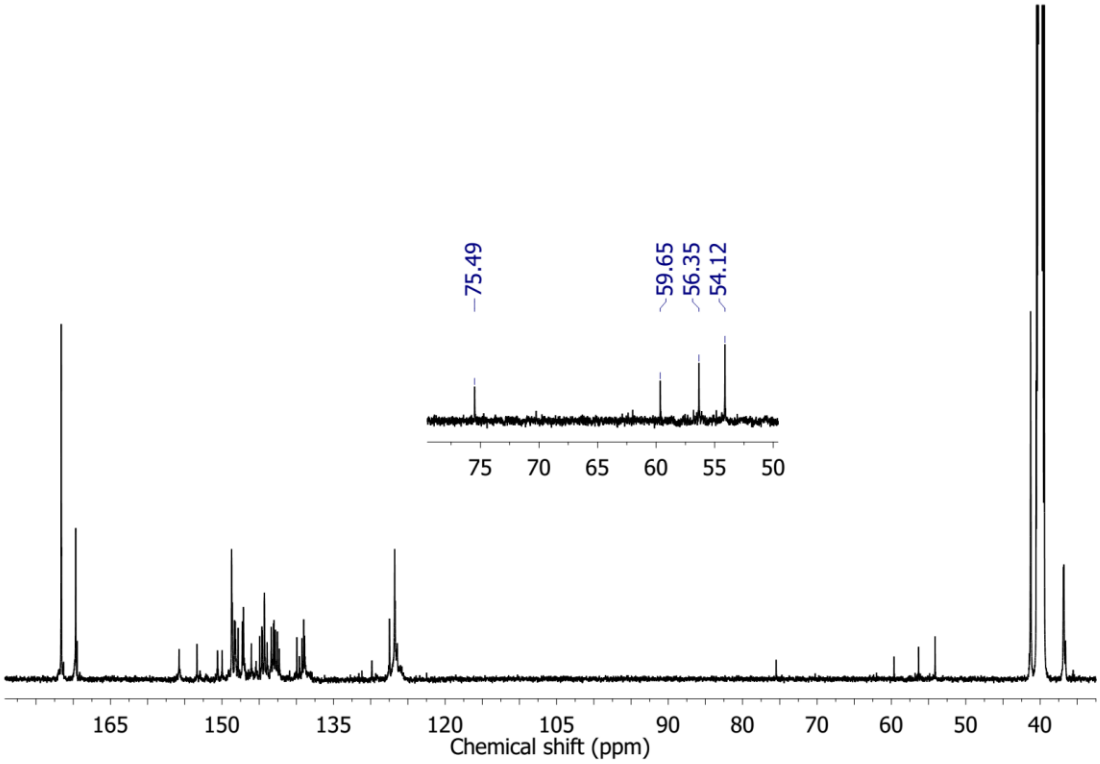

Fig. S34. ^{13}C NMR spectrum of compound **7**

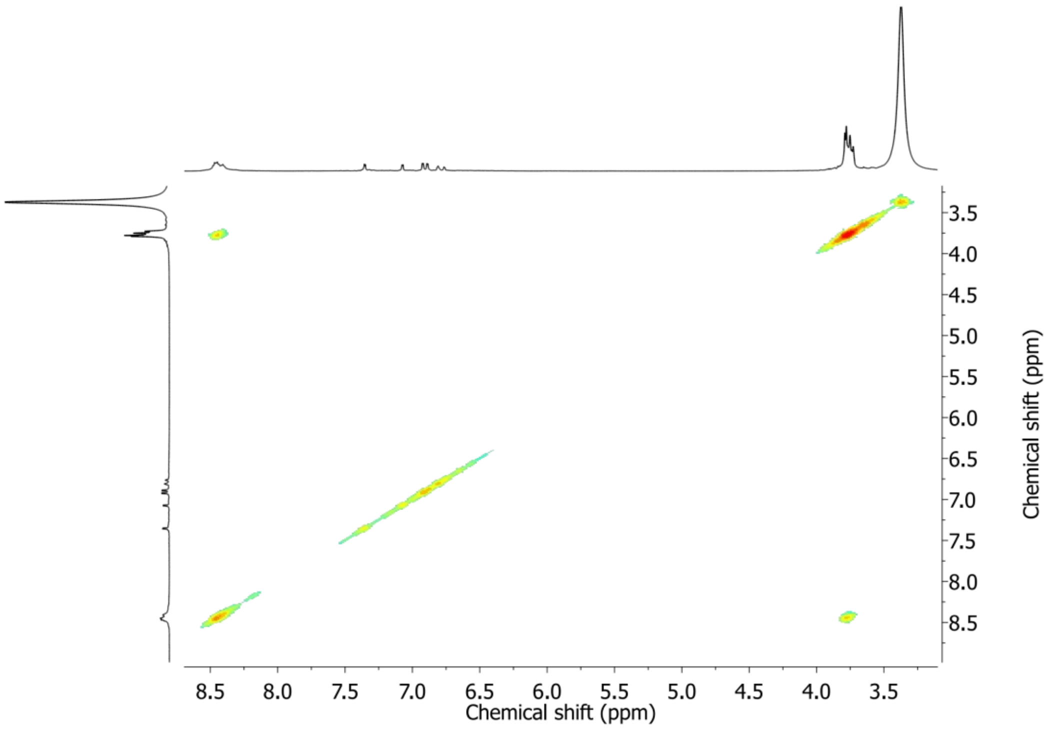

Fig. S35. ^{1}H-^{1}H COSY NMR spectrum of compound **7**

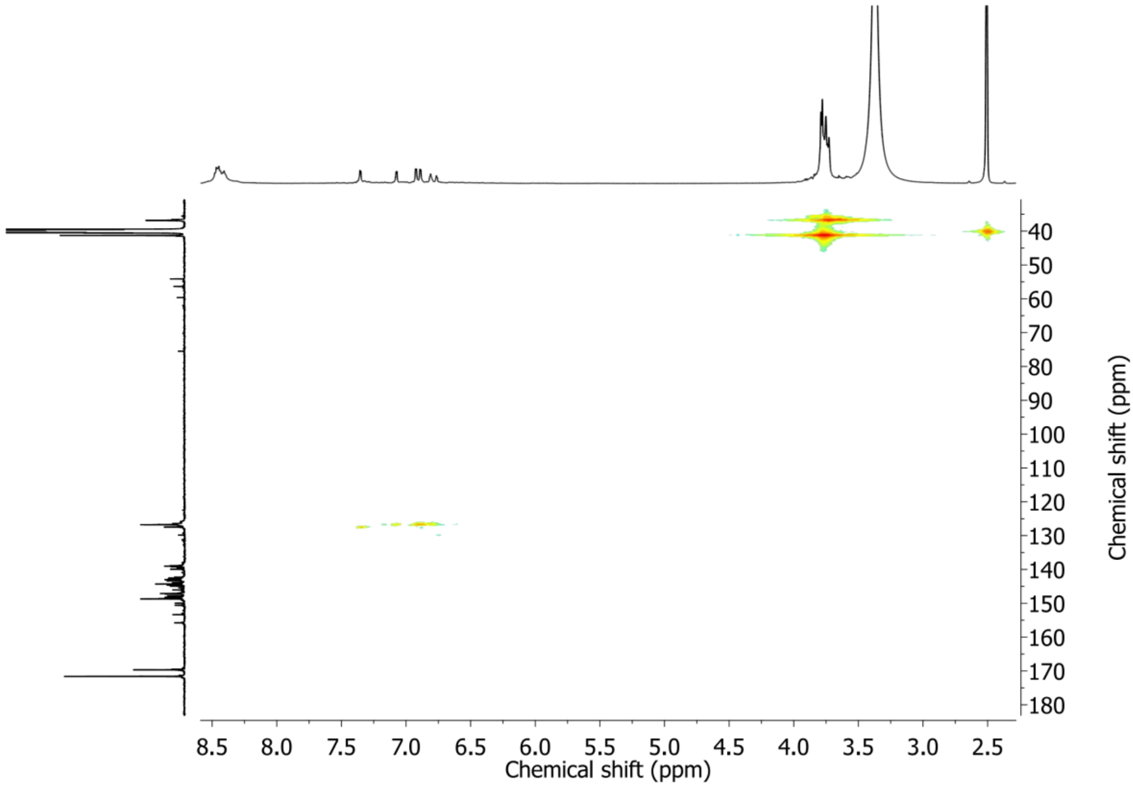

Fig. S36. ^{1}H-^{13}C HSQC NMR spectrum of compound **7**

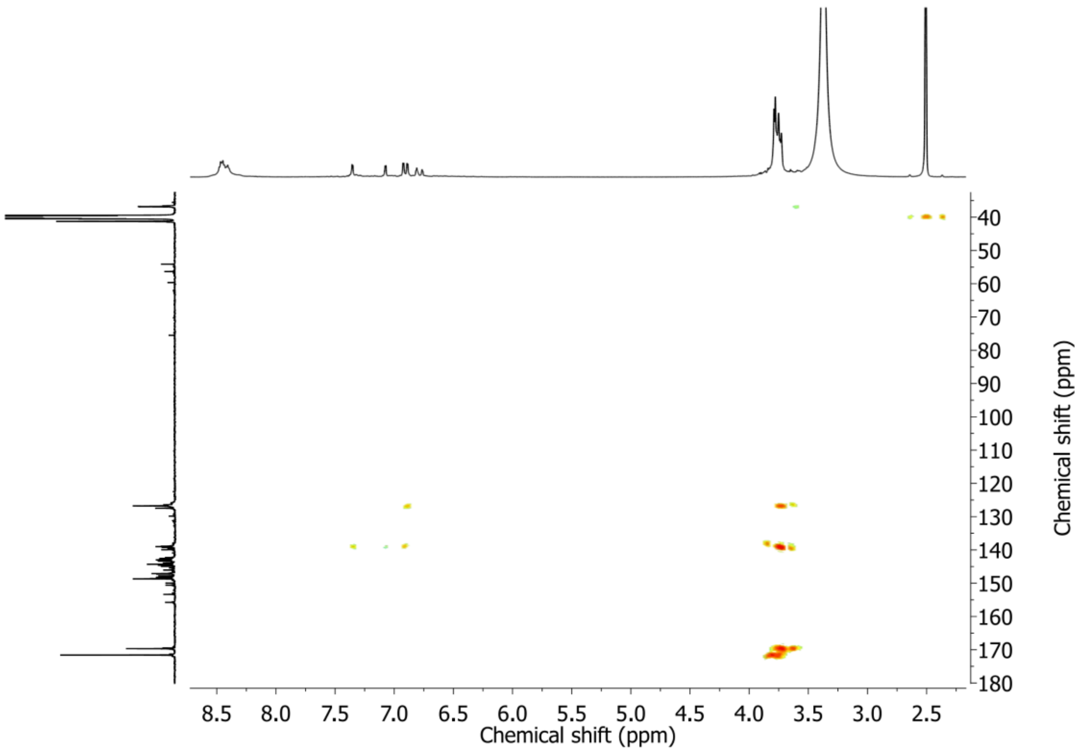

Fig. S37. ^{1}H-^{13}C HMBC NMR spectrum of compound **7**

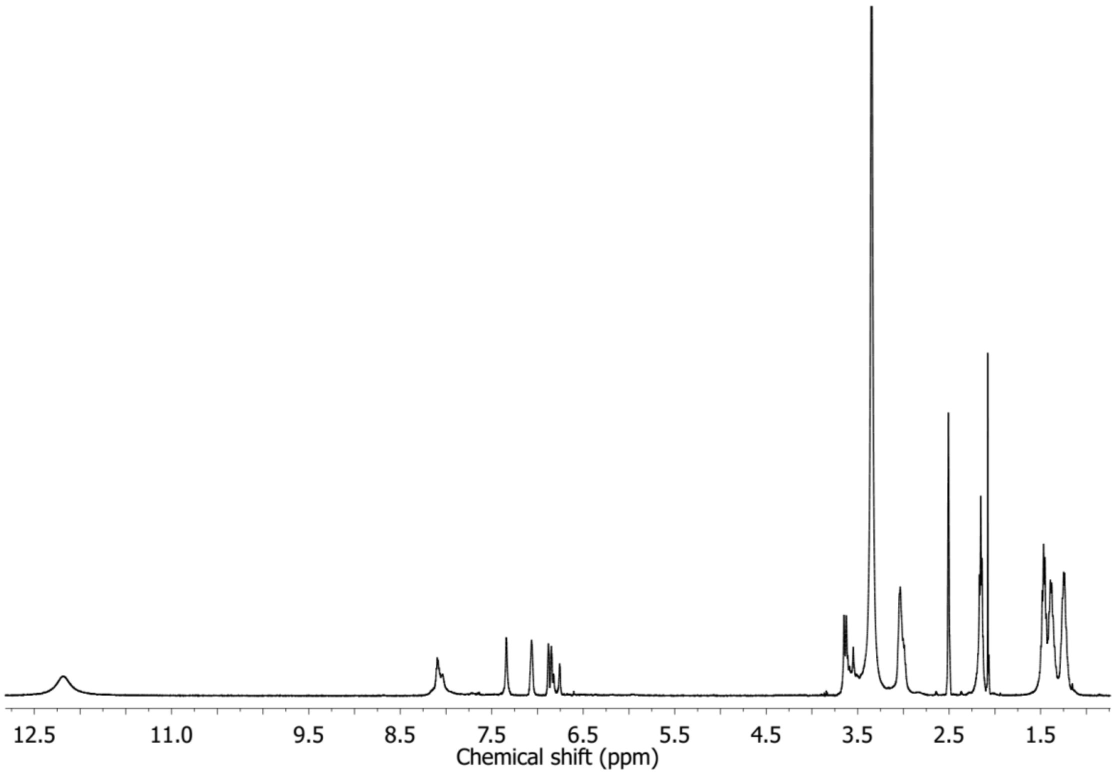

Fig. S38. ^{1}H NMR spectrum of compound **8**

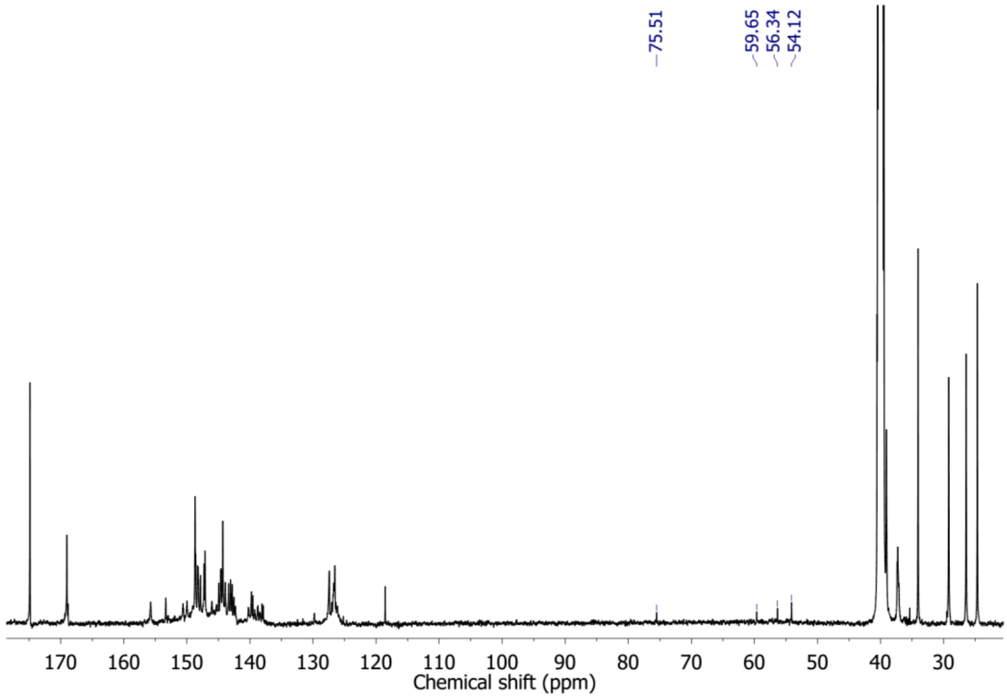

Fig. S39. ^{13}C NMR spectrum of compound **8**

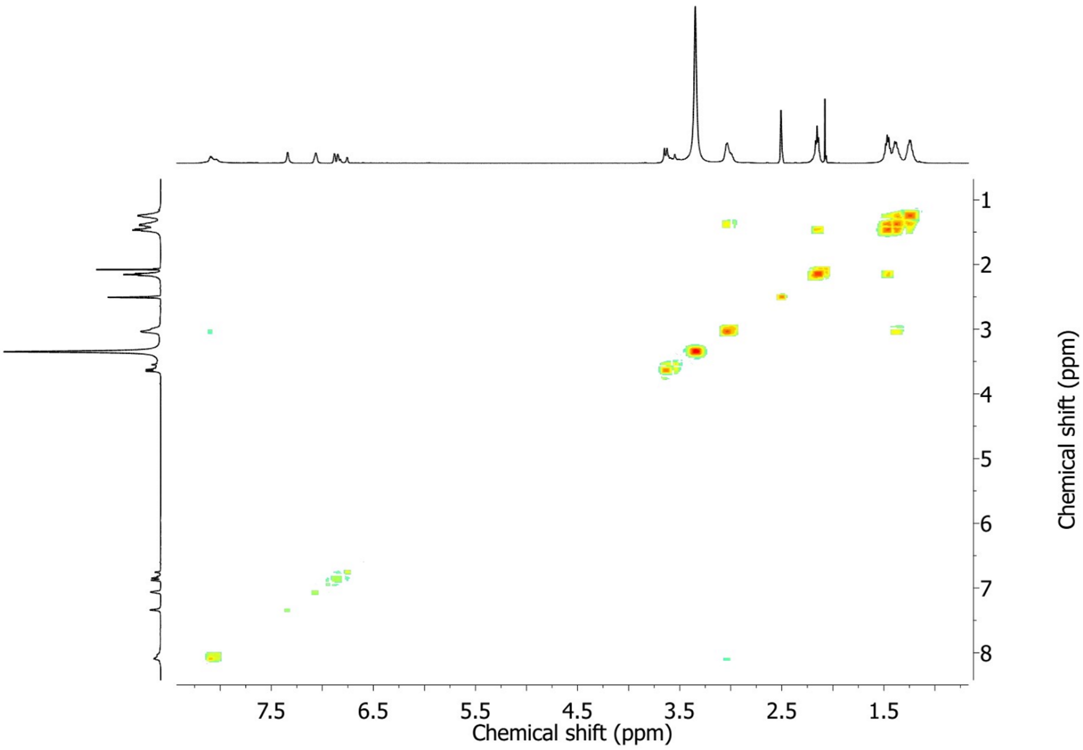

Fig. S40. ^{1}H-^{1}H COSY NMR spectrum of compound **8**

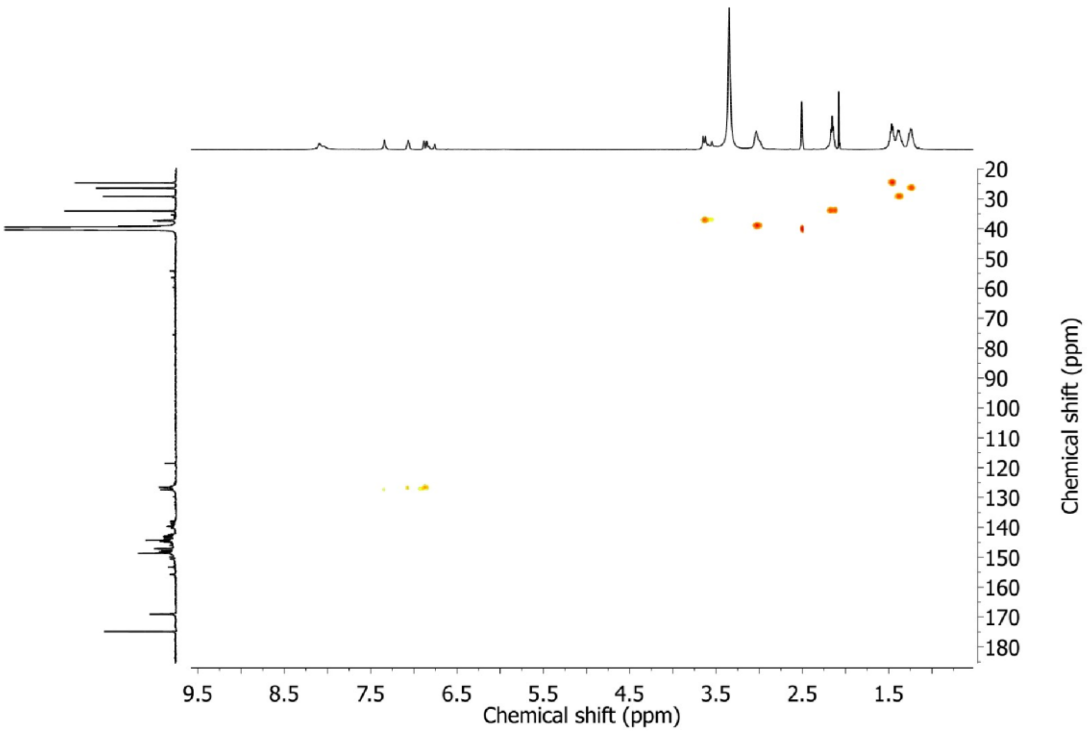

Fig. S41. ^{1}H-^{13}C HSQC NMR spectrum of compound **8**

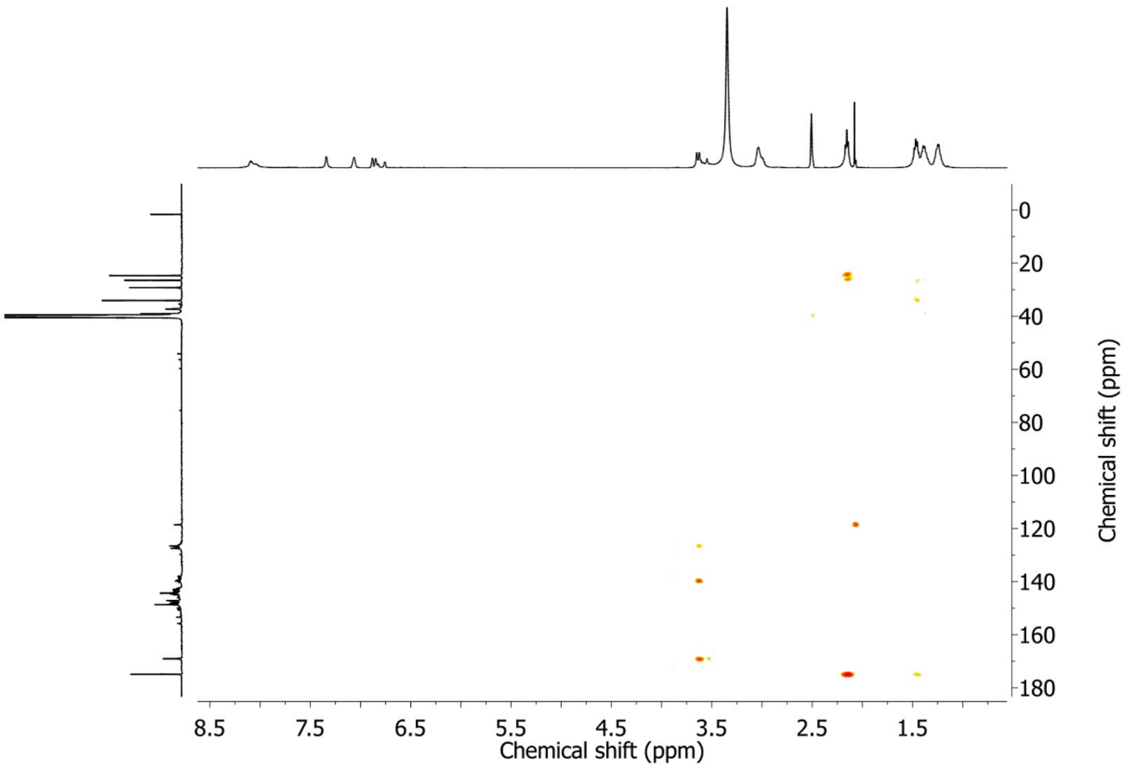

Fig. S42. ^{1}H-^{13}C HMBC NMR spectrum of compound **8**

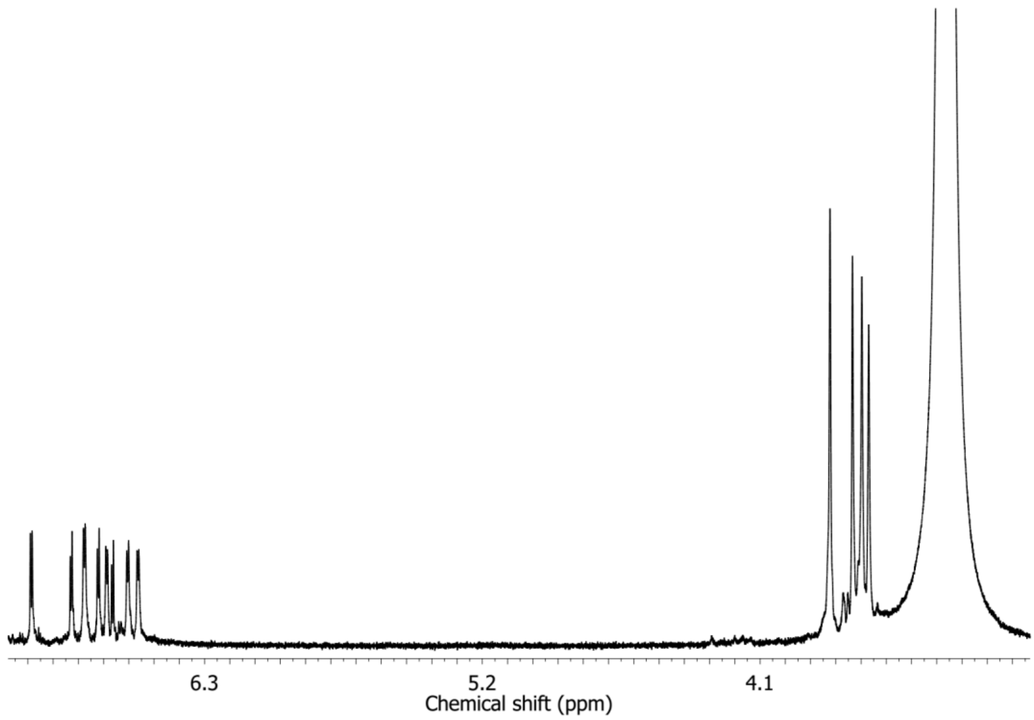

Fig. S43. ^{1}H NMR spectrum of compound **9**

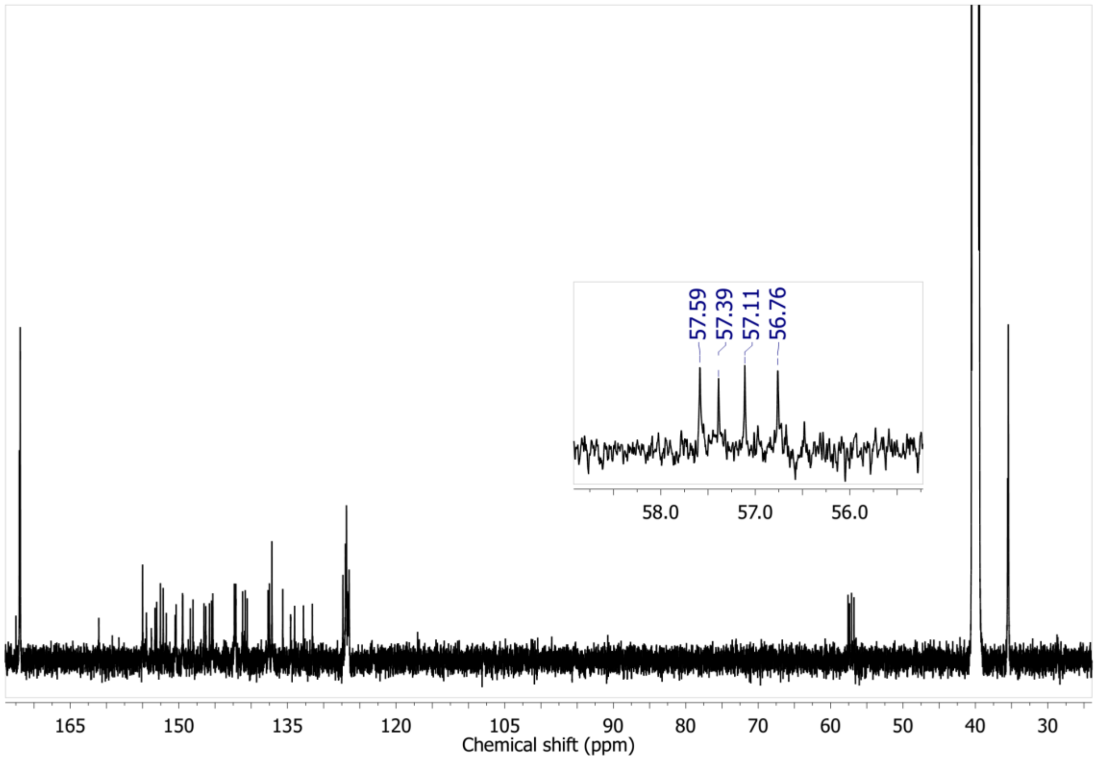

Fig. S44. ^{13}C NMR spectrum of compound **9**

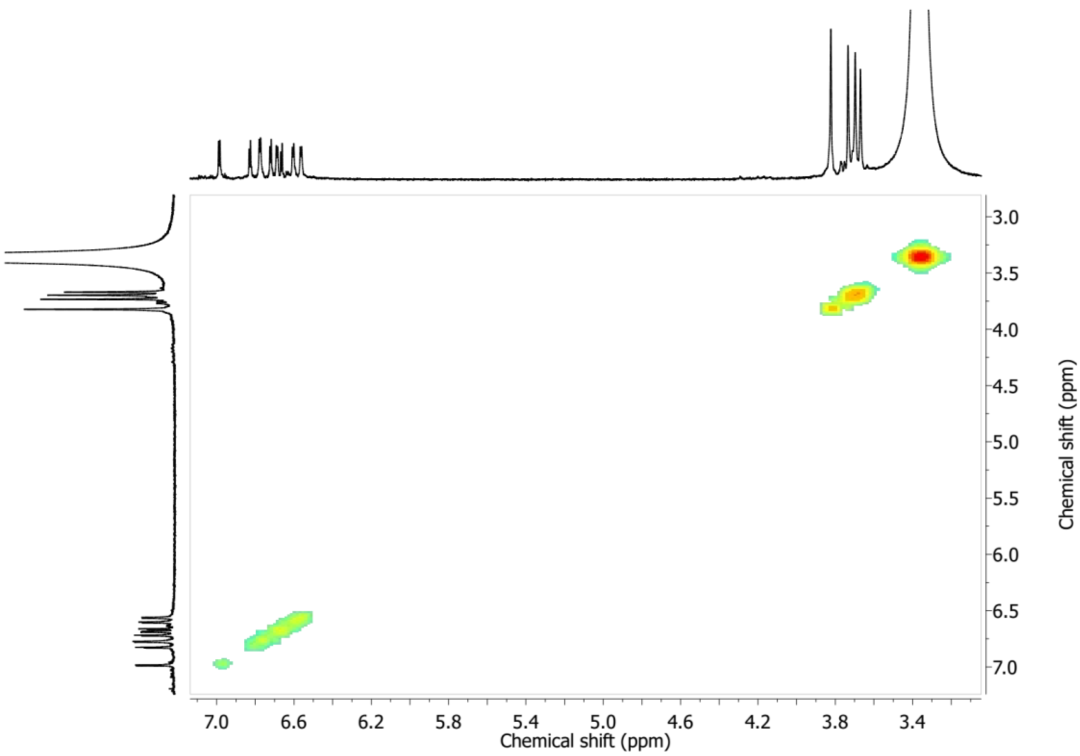

Fig. S45. ^{1}H-^{1}H COSY NMR spectrum of compound **9**

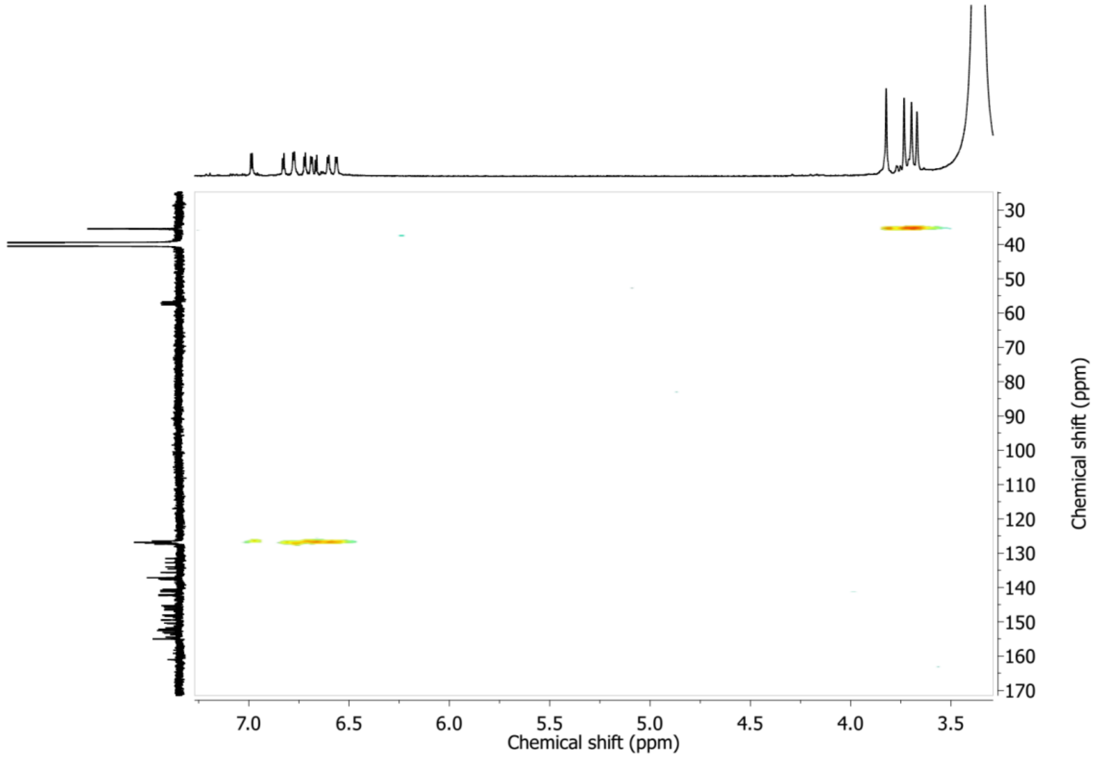

Fig. S46. ^{1}H-^{13}C HSQC NMR spectrum of compound **9**

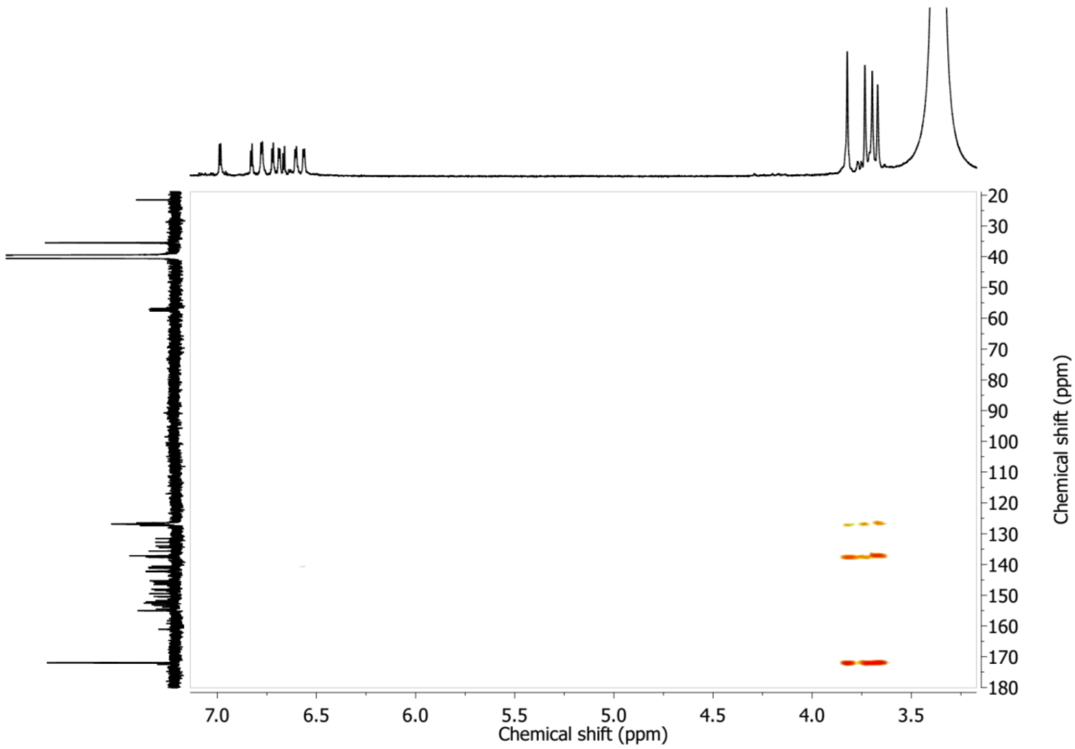

Fig. S47. ^{1}H-^{13}C HMBC NMR spectrum of compound **9**

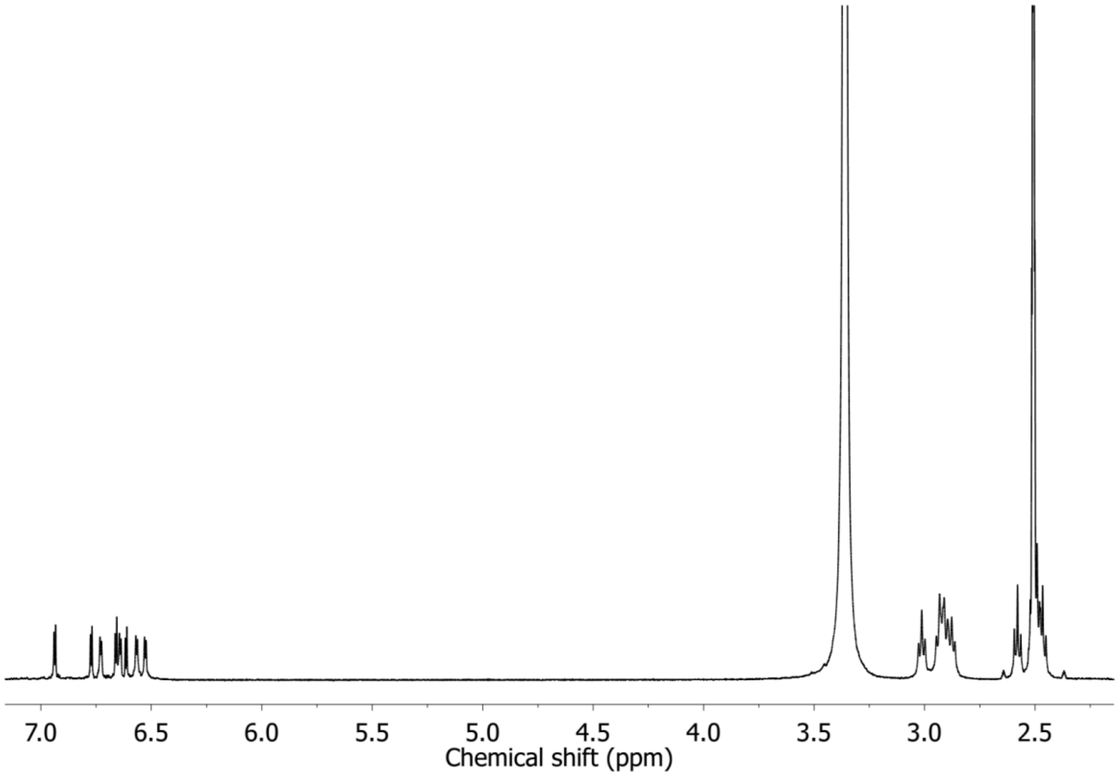

Fig. S48. ^{1}H NMR spectrum of compound **10**

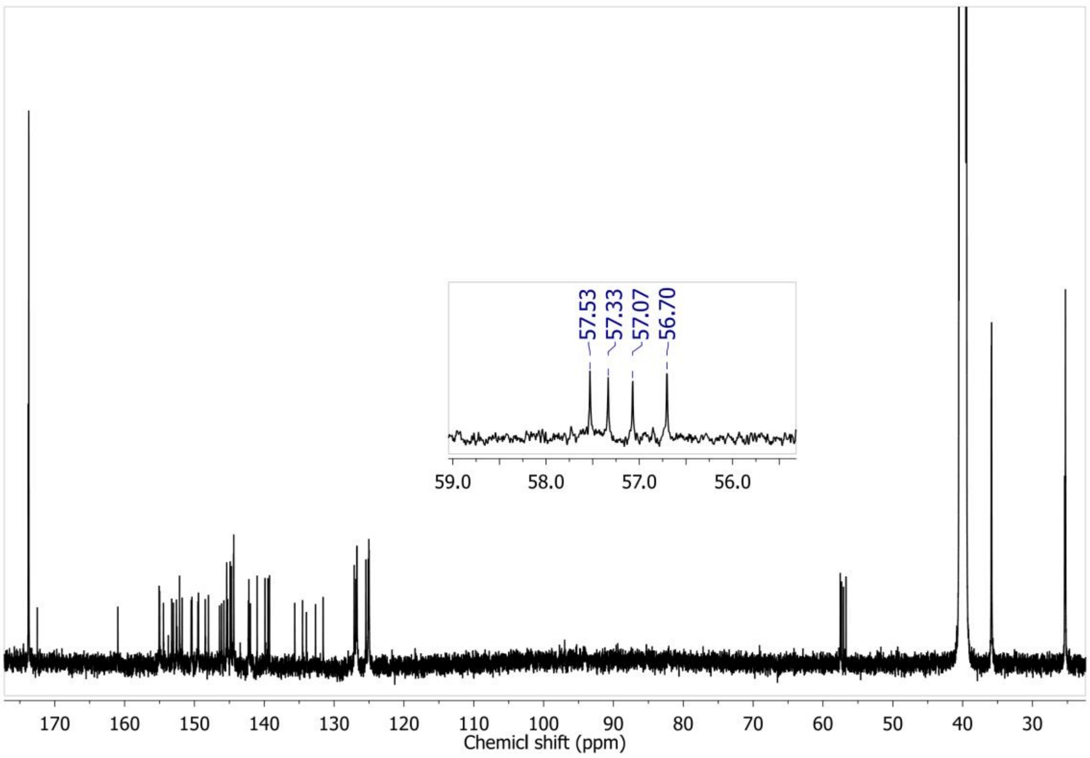

Fig. S49. ^{13}C NMR spectrum of compound **10**

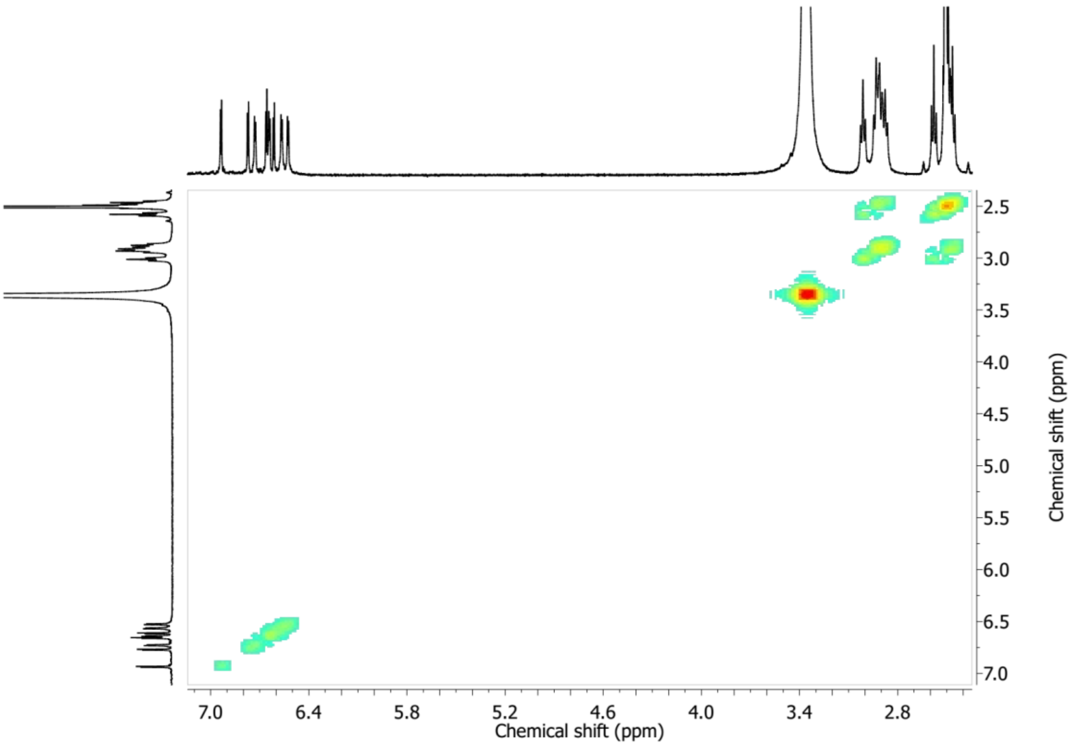

Fig. S50. ^{1}H-^{1}H COSY NMR spectrum of compound **10**

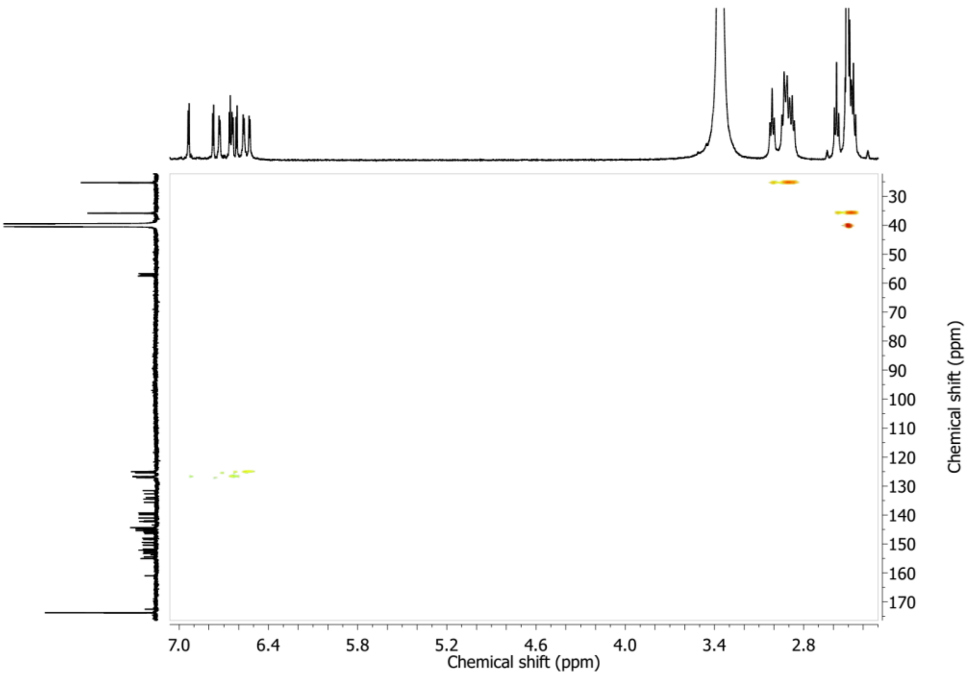

Fig. S51. ^{1}H-^{13}C HSQC NMR spectrum of compound **10**

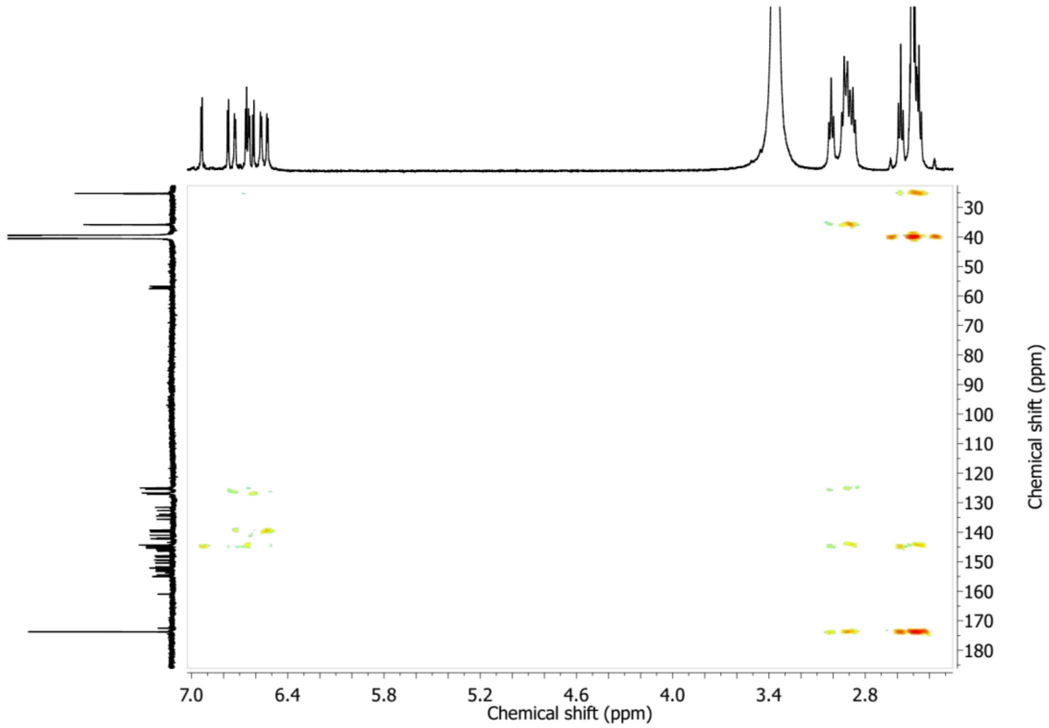

Fig. S52. ^{1}H-^{13}C HMBC NMR spectrum of compound **10**

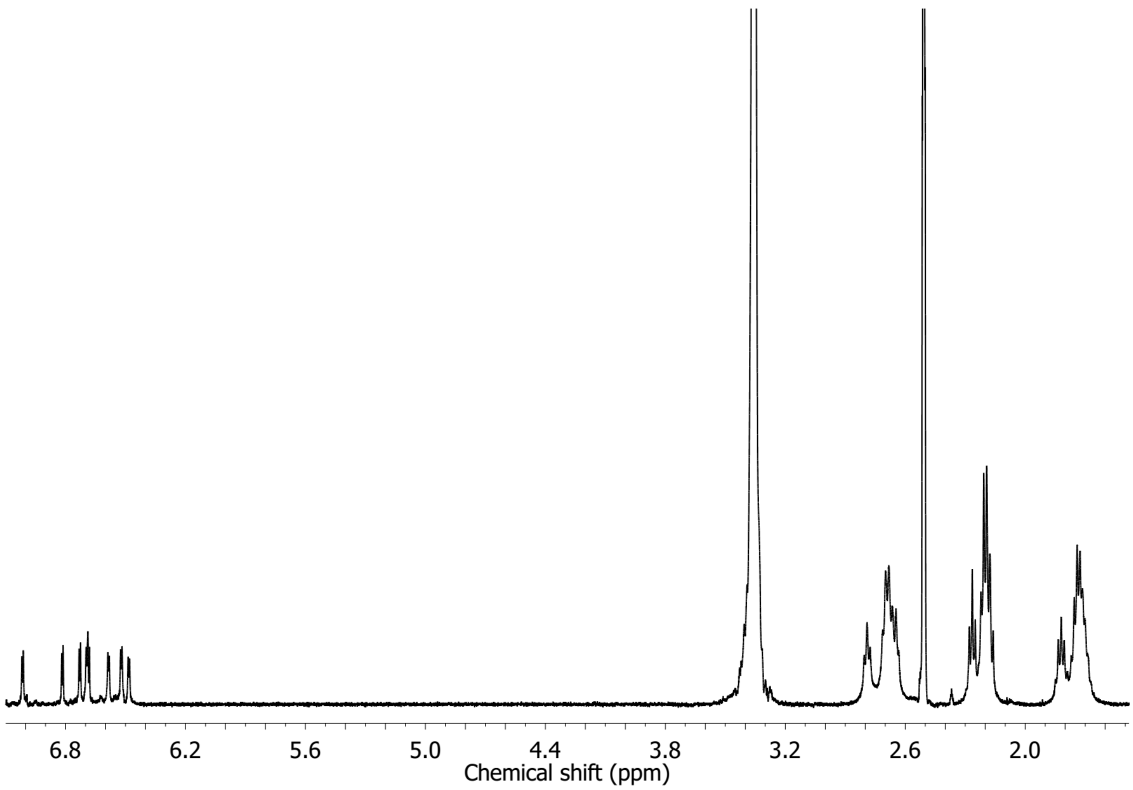

Fig. S53. ^{1}H NMR spectrum of compound **11**

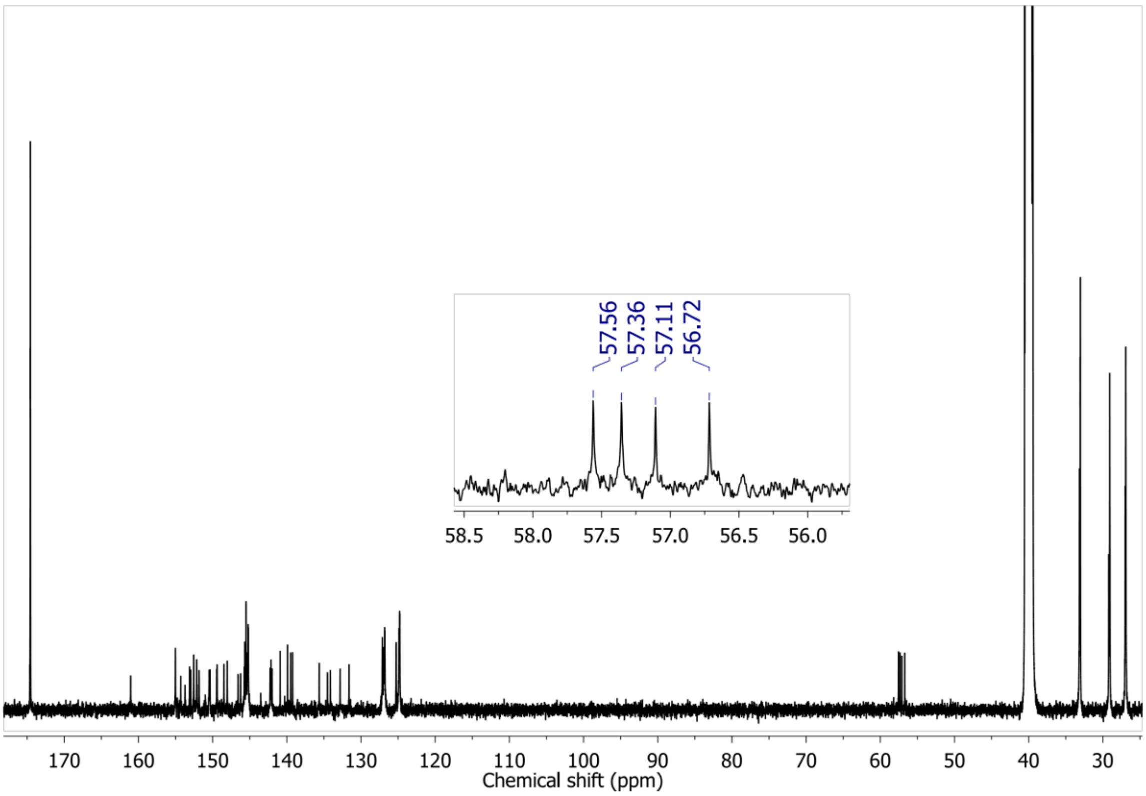

Fig. S54. ^{13}C NMR spectrum of compound **11**

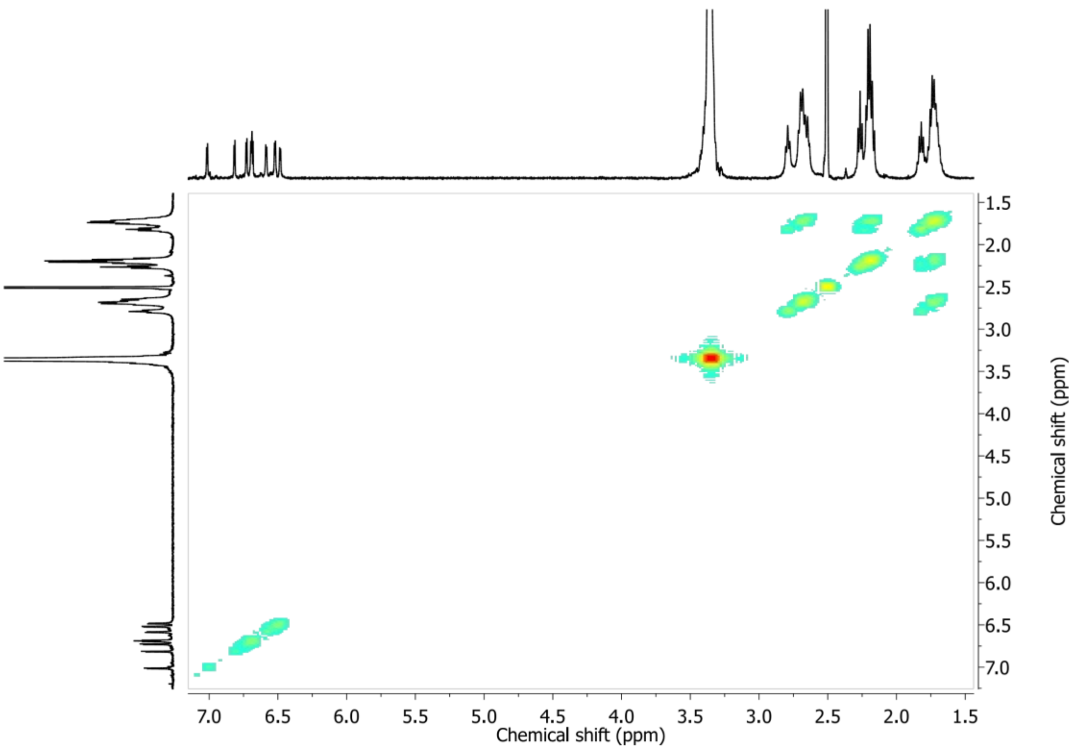

Fig. S55. ^{1}H-^{1}H COSY NMR spectrum of compound **11**

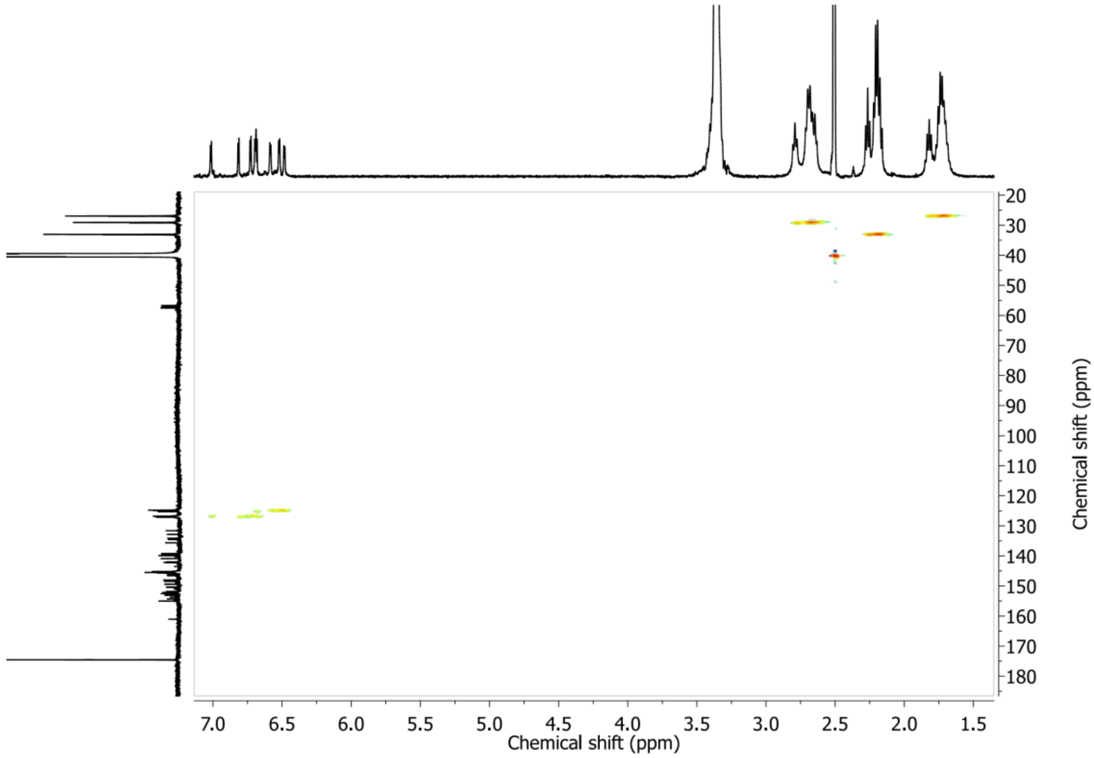

Fig. S56. ^{1}H-^{13}C HSQC NMR spectrum of compound **11**

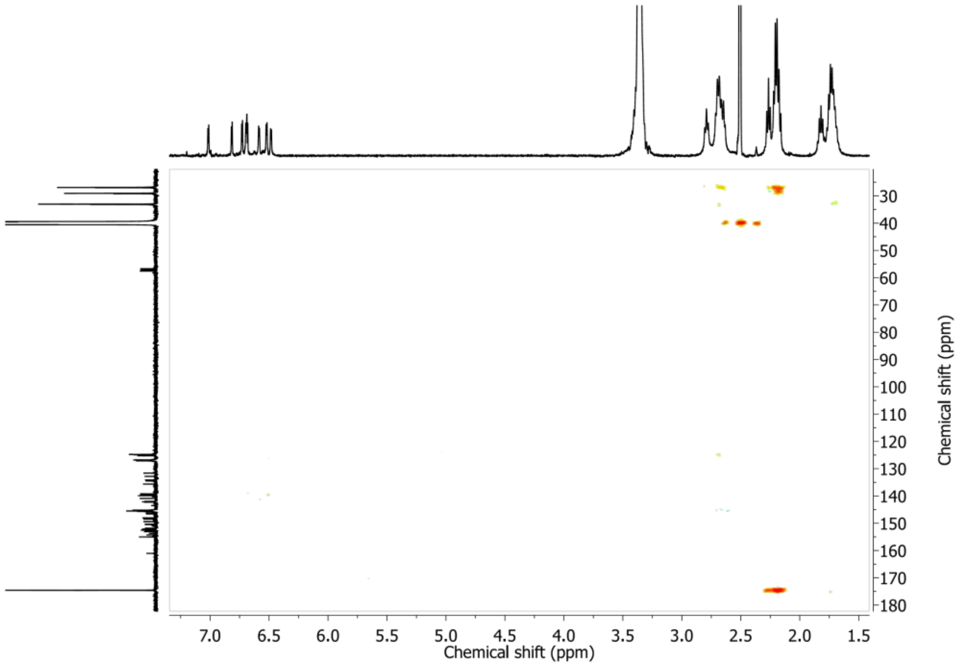

Fig. S57. ^{1}H-^{13}C HMBC NMR spectrum of compound **11**

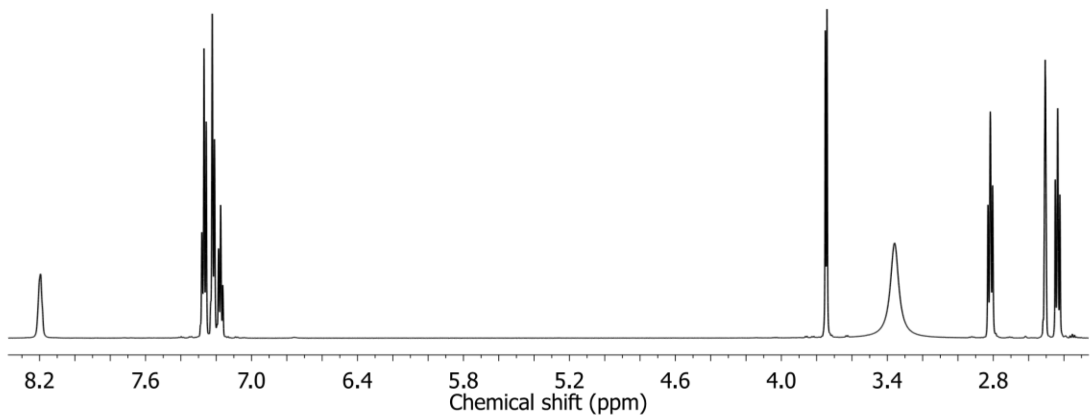

Fig. S58. ^{1}H NMR spectrum of compound **12a**

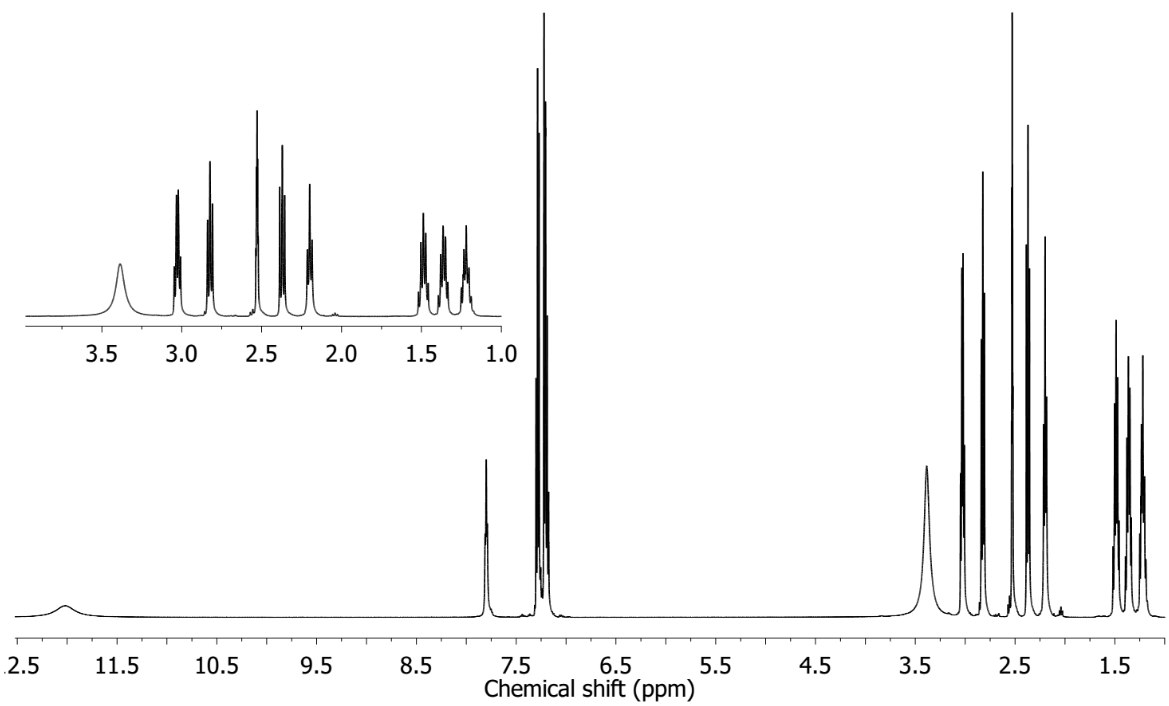

Fig. S59. ^{1}H NMR spectrum of compound **12b**

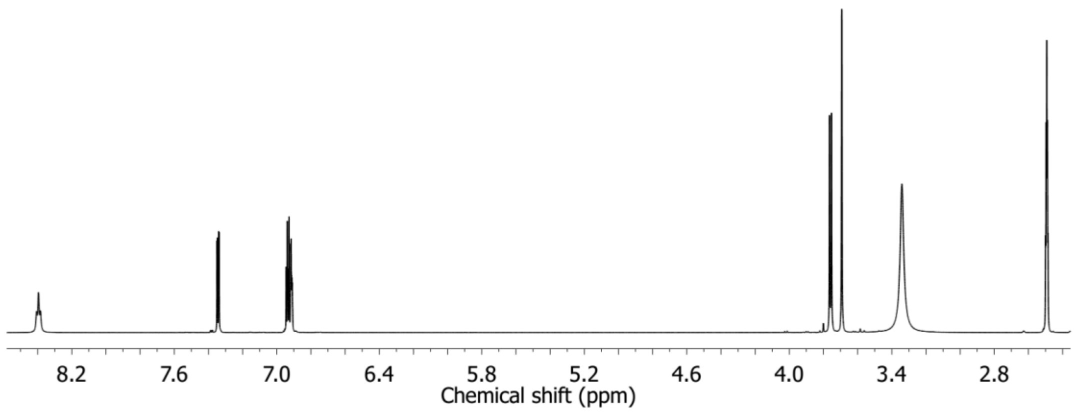

Fig. S60. ^{1}H NMR spectrum of compound **12c**

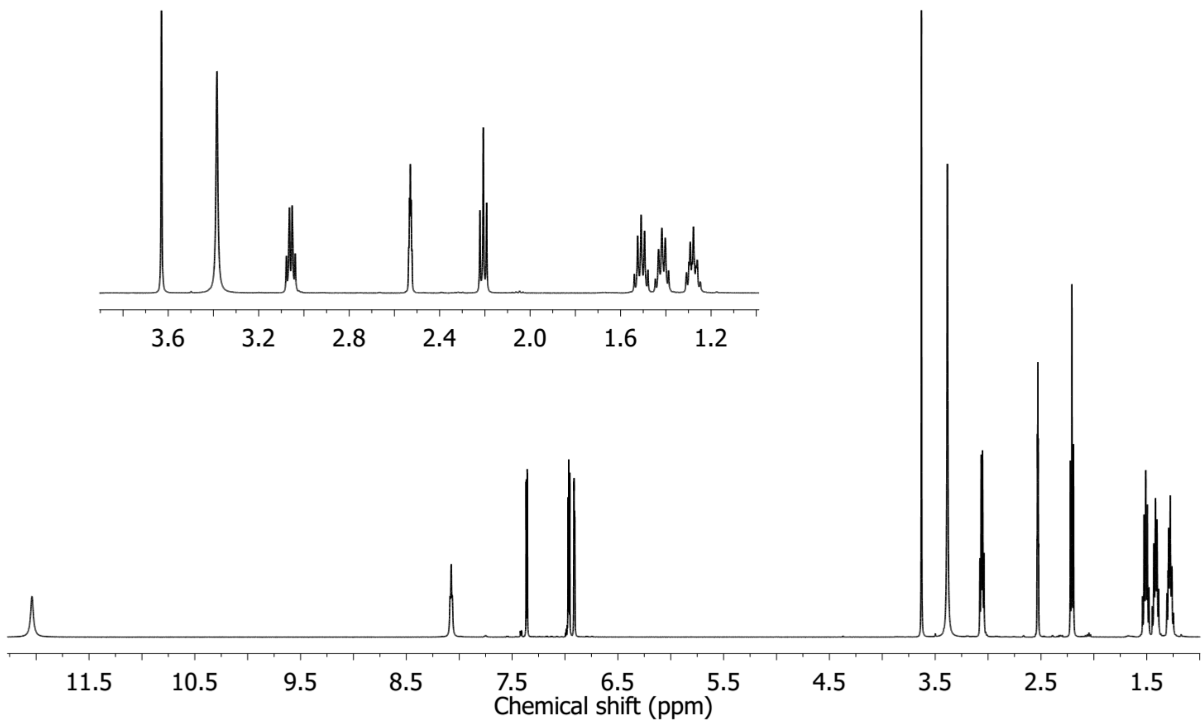

Fig. S61. ^{1}H NMR spectrum of compound **12d**

Biological assays

The antiviral assay in TZM-bl cells has been described in detail by Gordts *et al.* 2015 (S.C. Gordts, G. Ferir, T. D'Huys, M.I. Petrova, S. Lebeer, R. Snoeck, et al., The low cost compound lignosulfonic acid (LA) exhibits broad-spectrum anti-HIV and anti-HSV activity and has potential for microbicidal applications, *PLoS One*, **2015**, 10, 7, e0131219).

Briefly, 1.7×104 TZM-bl cells were pre-incubated with various concentrations of compounds in standard cell culture medium supplemented with diethylaminoethyl-dextran (15 µg/ml; Sigma-Aldrich) at 37 °C for 30 min. Then the cells were infected with HIV-1 NL4-3 or HIV-1 BaL virus or HIV-2 ROD stock. After two days of incubation at 37 °C, steady-lite plus substrate solution (PerkinElmer, Waltham, MA, USA) was added. The luminescent signal, which is proportional to the amount of viral replication, was measured using the SpectraMax L® microplate reader (Molecular Devices) and analyzed using Softmax Pro® software (Molecular Devices).

The anti-HIV assays in MT-4 cells were performed as follows. Compound-treated MT-4 cells (5×104 cells per sample in cell culture medium; 30 min incubation at 37 °C) were infected with HIV-1 NL4-3 or HIV-2 ROD virus stocks. After five days, virus-induced cytopathogenic effect was checked using optical microscopy and cell viability (i.e. measurement of viral replication) was evaluated using the MTS/PES-based CellTiter 96 Aqueous One Solution Cell Proliferation assay (Promega, Fitchburg, WI, USA). Absorbance was recorded using the VersaMax ELISA™ microplate reader (Molecular Devices) and analyzed with the Softmax Pro® software (Molecular Devices).