Electronic Supplementary Information

γ-Cyclodextrin duplex connected with two disulfide bonds: synthesis, structure and inclusion complexes

Sergey Volkov,[a] Lukáš Kumprecht,[a] Miloš Buděšínský,[a] Michal Dušek,[b] Martin Lepšík[a] and Tomáš Kraus[a*]

[a]*Institute of Organic Chemistry and Biochemistry AS CR, v.v.i., Flemingovo nám. 2, 166 10 Prague 6, Czech Republic. E-mail: kraus@uochb.cas.cz*

[b]*Institute of Physics, AS CR, v.v.i., Na Slovance 2, 182 21 Praha 8, Czech Republic.*

CONTENTS **page**

1. Experimental procedures

1.1. General Experimental Methods

High resolution mass spectra were measured using nanoESI ionization on LTQ Orbitrap XL (Thermo Fisher Scientific). Optical rotations were recorded on AUTOPOL IV (Rudolph Research Analytical). Preparative reversed–phase chromatography (RP) was carried out on medium pressure columns containing C-18 modified silica (Phenomenex Luna, 15 μm). HPLC preparative separations were carried out on Waters 2525 system equipped with ELS detector Waters 2420 and columns Supelcosil LC-SI 250 x 21.2 mm (silica 5 μm). Thin-layer (TLC) and reversed-phase thin-layer chromatography (RPTLC) were performed with precoated Silica Gel 60F and RP-18 F plates (E. Merck) respectively, which were developed by spraying with an aqueous solution of phosphomolybdic acid containing 5% of H_2SO_4 and heating. All chemicals used were commercially available. γ-Cyclodextrin was dried at 100 °C for 10 hours under vacuum prior use. Satisfactory elemental analysis could not be obtained for hydrophilic compounds **4** – **7** unless variable numbers of water molecules were taken into account. Thus, calculations based on weights of these compounds (molarity, yield, optical rotation) are related to the hydrated molecules.

The analyses of cavity volumes in cyclodextrin duplexes were performed using SURFNET[1] implemented in UCSF CHIMERA,[2] version 1.8. A probe sphere possessing a minimal radius of 1.4 A (the "grid interval" parameter) and a maximal of 8.5 A (the "distance cutoff" parameter) was used to fill the cavities in cyclodextrin duplexes. The volumes and sizes of the cavities were output by CHIMERA.

1.2. Procedures for preparation of compounds 1-7

Per(2,3,6-*O*-benzyl)- γ-cyclodextrin (1)

Dried γ-cyclodextrin (9.7 g; 7.48 mmol) was dissolved in dry DMSO (80 mL). The solution was added slowly to the stirred dispersion of sodium hydride (14.4 g, 600 mmol; oil-free) in DMSO (150 mL) under argon atmosphere. Benzyl chloride (41.1 mL; 357 mmol) was added dropwise to the reaction mixture within 60 min, the temperature of the reaction mixture being carefully monitored and kept below 40°C by external water cooling bath. The reaction mixture was allowed to react for 24 hours. After this reaction

time, TLC (silica, toluene-acetone 95:5) showed only one major spot of the product. The excess of hydride was decomposed by addition of methanol under cooling. The reaction mixture was diluted with water (220 mL) and the product was extracted to diethyl ether (4x300 mL). Organic layers were combined, washed with brine, dried with sodium sulfate and evaporated. Crude product was purified by column chromatography (silica, gradient elution from toluene to toluene-acetone 97:3) to give colorless amorphous material after extensive drying under vacuum (22.8 g; 88%). R_f 0.50 (TLC, silica, toluene-acetone 98:2); ^{1}H-NMR, see Table S1; ^{13}C-NMR, see Table S2; HR-MS (MALDI) calcd. for $C_{216}H_{224}O_{40}$ [M + Na] $^+$ m/z 3457.54940, found m/z 3457.54924. Anal. Calcd for $C_{216}H_{224}O_{40}$: C, 74.98; H, 6.53; found C, 74.71; H, 6.45.

2^{I}, 2^{II}, 2^{III}, 2^{IV}, 2^{V}, 2^{VI}, 2^{VII}, 2^{VIII} 3^{I}, 3^{II}, 3^{III}, 3^{IV}, 3^{V}, 3^{VI}, 3^{VII}, 3^{VIII}, 6^{II}, 6^{III}, 6^{IV}, 6^{VI}, 6^{VII}, 6^{VIII} -Docosa-*O*-benzyl- γ-cyclodextrin (2a).

2^{I}, 2^{II}, 2^{III}, 2^{IV}, 2^{V}, 2^{VI}, 2^{VII}, 2^{VIII} 3^{I}, 3^{II}, 3^{III}, 3^{IV}, 3^{V}, 3^{VI}, 3^{VII}, 3^{VIII}, 6^{II}, 6^{III}, 6^{V}, 6^{VI}, 6^{VII}, 6^{VIII} -Docosa-*O*-benzyl- γ-cyclodextrin (2b).

Diisobutylaluminum hydride (1.5 M solution in toluene, 273 mL, 410 mmol) was introduced into a calibrated Schlenk flask. The flask was placed into a water-ice cooling bath at ~ 0 ºC and the solution of DIBAL-H was concentrated to a half of the volume under reduced pressure. Concentrated solution was added to a dried perbenzylated γ-cyclodextrin **1** (6 g, 1.73 mmol) at t ~ 0 ºC. The reaction mixture was allowed to react for 41h at ~ 0 ºC under argon atmosphere. Then it was diluted with toluene (500 ml) and was poured onto ice (aprox. 1 kg). After the decomposition of the excess of hydride, 1 M aqueous solution of HCl (400 mL) was added slowly with stirring. The crude product was partitioned between water and toluene by means of a multiple extraction. The combined organic layers were washed with aqueous sodium carbonate and dried with magnesium sulfate. The evaporation of the solvent gave a crude mixture (4. 628 g) that was subjected to column chromatography (silica, isocratic elution with dichloromethane-acetone 98:2) which allowed isolation of diol **2a** (0.776 g, 14%, white amorphous material) and a mixture of **2a** and **2b** (1.838 g, 32%, white amorphous material). Further separation of the mixture of diols by preparative HPLC (silica, eluent – dichlormethane-acetone 98:2) allowed another crop of compound **2a** (580 mg, 10%, white amorphous material) and

compound **2b** (567 mg, 10 %). Analytical data for compound **2a**: R_f 0.52 (TLC, silica, dichloromethane-acetone 98:2); $[\alpha]^{25}_D$ = +37 (c = 0.24 in $CHCl_3$); ^{1}H-NMR, see Table S1; ^{13}C-NMR, see Table S2; HR MS (ESI) calcd. for $C_{202}H_{212}O_{40}$ $[M + 2Na]^{2+}$ m/z 1661.71697, found 1661.71625. Anal. Calcd for $C_{202}H_{212}O_{40}$: C, 73.97; H, 6.52, found C, 73.81; H, 6.52. R_f 0.48, dichlormethane-acetone 98:2; $[\alpha]^{25}_D$ = +40 (c = 0.25 in $CHCl_3$); ^{1}H-NMR, see Table S1; ^{13}C-NMR, see Table S2; HR-MS (ESI) calc. for $C_{202}H_{212}O_{40}$ $[M + 2Na]^{2+}$ m/z 1661.71697, found 1661.71691. Anal. Calcd for $C_{202}H_{212}O_{40}$: C, 73.97; H, 6.52, found C, 73.72; H, 6.74.

2^{I}, 2^{II}, 2^{III}, 2^{IV}, 2^{V}, 2^{VI}, 2^{VII}, 2^{VIII} 3^{I}, 3^{II}, 3^{III}, 3^{IV}, 3^{V}, 3^{VI}, 3^{VII}, 3^{VIII}, 6^{II}, 6^{III}, 6^{IV}, 6^{VI}, 6^{VII}, 6^{VIII} -Docosa-*O*-benzyl-6^{I},6^{V}-dibromo-6^{I},6^{V}-dideoxy-γ-cyclodextrin (3).

Triphenylphosphane (358 mg, 1.36 mmol), tetrabromomethane (454 mg, 1.37 mmol) and diol **2a** (560 mg, 0.17 mmol) were dissolved subsequently in dry DMF (2 mL) in a Schlenk flask under argon atmosphere. The reaction mixture was heated to 65° C for 4 hours and then the reaction was quenched by addition of MeOH (0.5 mL). Solvents were evaporated in vacuum and product **3** was isolated from the resulting solid by column chromatography (50 g silica, gradient elution toluene to toluene – acetone 99:1) as amorphous solid (440 mg, 76 %). R_f 0.34 (TLC, silica, toluene-acetone 97:3); $[\alpha]^{25}_D$ = +32 (c = 0.284 in $CHCl_3$); ^{1}H-NMR, see Table S1; ^{13}C-NMR, see Table S2; HR-MS (ESI) calcd for $C_{202}H_{210}Br_2O_{38}$ $[M + 2Na]^{2+}$ m/z 1723.63256, found 1723.62909. Anal. Calcd for $C_{202}H_{210}Br_2O_{38}$: C, 71.24; H, 6.22; Br, 4.69, found C, 71.05; H, 6.47; Br, 4.90.

6^{I},6^{V}-Dibromo-6^{I},6^{V}-dideoxy-γ-cyclodextrin (4)

Compound **3** (338 mg, 0.099 mmol) was dissolved in degassed mixture of DMF-ethanol (1:1, 12 mL). Then palladium on charcoal (10% w/w, 98 mg) was added and the reaction mixture was placed into an autoclave equipped with a magnetic stirring bar. The autoclave was flushed with argon and filled with hydrogen to a pressure of 40 bar at room temperature. The mixture was allowed to react for 6 hours under stirring. Then the excessive hydrogen was released and the catalyst was removed from the reaction mixture by filtration through cellite. The solvents were evaporated under reduced pressure. The residue was dissolved in water (20 mL) and charged onto C-18 reversed-phase column.

Gradient elution from methanol-water 5:95 to 4:6 allowed isolation of the product **4** as as amorphous colorless solid (106 mg, 71 % calcd for pentahydrate). R_f 0.5 (reversed-phase, methanol-water 1:1); $[\alpha]^{25}_D$ = +145 (c = 0.36 in DMSO); ^{1}H NMR, see Table S1; ^{13}C NMR, see Table S2; HR-MS (ESI) calcd for $C_{48}H_{78}Br_2O_{38}$ [M + Na]$^+$ *m/z* 1443.24300, found 1443.24333. Anal. Calcd for $C_{48}H_{78}Br_2O_{38}\cdot 5H_2O$: C, 38.10; H, 5.86; Br 10.56, found C, 37.93; H, 5.51, Br 10.66.

6^I,6^V-Bis(acetylthio)-6^I,6^V-dideoxy-γ-cyclodextrin (5).

The compound **4** (170 mg, 0.112 mmol, calcd for pentahydrate) was dissolved in dry degassed DMF (1.7 mL) in a Schlenk flask equipped with a magnetic stirring bar under argon atmosphere. The solution of potassium thioacetate (28 mg, 0.247 mmol) in dry degassed DMF (0.4 mL) was added dropwise at room temperature. The mixture was allowed to react for 12 hours. Then DMF was evaporated under reduced pressure at 40 ºC and the crude product was dissolved in water (50 mL). Subsequent purification by column chromatography (C-18 reversed-phase, gradient elution from methanol-water 5:95 to 4:6) afforded white material (126 mg, 73 %, calcd for heptahydrate). R_f 0.46 (reversed-phase, methanol-water 4:6); ^{1}H NMR, see Table S1; ^{13}C NMR, see Table S2; HR-MS (ESI) calcd for $C_{52}H_{84}O_{40}S_2$ [M + Na]$^+$ *m/z* 1435.38725, found 1435.38745. Anal. Calcd for $C_{52}H_{84}O_{40}S_2\cdot 7H_2O$: C, 40.57; H, 6.42, S 4.17, found: C, 40.53; H, 6.42, S 3.97.

6^I,6 ^{V}-Disulfanyl-6^I,6^V-dideoxy-γ-cyclodextrin (6).

The solution of dithioacetate **5a** (30 mg, 0.019 mmol, calcd for heptahydrate) in water (2 ml) was degassed by vaccuum - argon cycle. Then the degassed solution of 1M sodium hydroxid (0.41 mL) was added under argon atmosphere. The mixture was allowed to react under stirring for 5h. Then 1M HCl was added dropwise untill pH 6 of the solution was reached. Subsequent purification by column chromatography (C-18 reversed-phase, gradient elution from methanol-water 5:95 to 4:6) afforded white material (19 mg, 66%, calcd for octahydrate). R_f 0.34 (reversed-phase, methanol-water 4:6); ^{1}H NMR, see Table S1; ^{13}C NMR, see Table S2; HR-MS (ESI) calcd for $C_{48}H_{80}O_{38}S_2$ [M + Na]$^+$ *m/z*

1351.36612, found 1351.36607. Anal. Calcd for $C_{48}H_{80}O_{38}S_2 \cdot 8H_2O$: C, 39.13; H, 6.57, S 4.35, found: C, 39.23; H, 6.26, S 4.35.

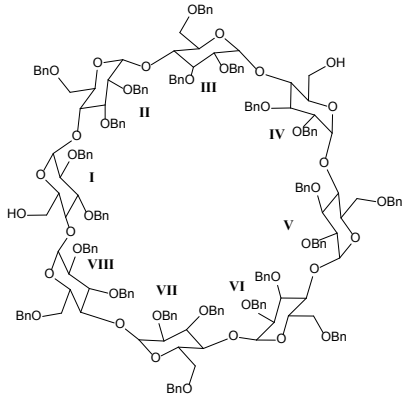
Duplex 7.

The solution of dithioacetate **5** (30 mg, 0.019 mmol, calcd for heptahydrate) in water (2 mL) was degassed by vaccuum - argon cycle and the degassed solution of 1M sodium hydroxid (0.41 mL) was added under argon atmosphere. The reaction mixture was degassed again and stirred for 5h at room temperature under argon atmosphere. Then water (24.3 mL) was added to the mixture and the pH was adjusted to 9.0 by introduction of carbon dioxide. The mixture was allowed to react under vigorous stirring in open bottle for 72 hours. Then the reaction mixture was neutralized with 1M HCl to pH~9. White precipitate was isolated by centrifugation, washed with water and dried *in vacuo* at room temperature (23 mg, 82% calcd for dodecahydrate). ^{1}H NMR, see Table S1; ^{13}C NMR, see Table S2; HR-MS (MALDI) calcd for $C_{96}H_{156}O_{76}S_4$ $[M + Na]^+$ m/z 2675.7117, found 2675.7163 Anal. calc. for $C_{96}H_{156}O_{76}S_4 \cdot 12H_2O$: C, 40.17; H, 6.32, S 4.47, found C, 39.86; H, 5.97, S 4.53.

1.3. ^{1}H and ^{13}C NMR data

The NMR spectra were measured on Bruker AVANCE-600 instrument (^{1}H at 600.13 MHz and ^{13}C at 150.9 MHz) with a cryoprobe in $CDCl_3$, DMSO or D_2O at 27 °C. Spectra in $CDCl_3$ were referenced to TMS (^{1}H) or solvent peak (^{13}C, using $\delta_C(CDCl_3)$ = 77.0 ppm). Spectra in DMSO were referenced to solvent peak (using δ_H(DMSO) = 2.50 ppm; δ_C(DMSO) = 39.7 ppm). Spectra in D_2O were referenced to dioxane (added as internal standard) using δ_H(dioxane) = 3.75 ppm and δ_C(dioxane) = 69.3 ppm. Partial structural assignment of proton and carbon signals was achieved combining 1D-^{1}H and ^{13}C-spectra with homonuclear 2D-H,H-COSY, 2D-H,H-TOCSY, 2D-H,H-ROESY and heteronuclear 2D-H,C-HSQC and 2D-H,C-HMBC spectra; detailed procedure was described elsewhere.[3]

The ^{1}H and ^{13}C NMR data are summarized in Tables S1 and S2. Complete sequential structural assignment of proton and carbon signals has been achieved only for γ-cyclodextrin derivatives **2a, 2b** and **3**. The sequential positions of their glucose units in Table 1 and 2 are labeled with Roman numerals I to VIII (see Figure 1) given in brackets following the chemical shift number. For debenzylated compounds **4**-**7** only partial structural assignment was possible due to the poor separation or overlap of many signals; numbers of protons or carbon atoms are given in arabic numeral notation in brackets. All observed coupling constants *J*(H,H) of ring protons of glucose units (not given in Tables 1 and 3) showed the values characteristic for $^{4}C_1$ conformation (*J*(1,2) = 3.5 – 4.0 Hz, *J*(2,3) ~ *J*(3,4) ~ *J*(4,5) = 8.5 – 10.0 Hz).



2b

Figure S1. Numbering of glucose residues used in this work exemplified with compound **2b**.

Table S1. ^{1}H NMR data for compounds **1-7**.

Compound (solvent)	**H-1**	**H-2**	**H-3**	**H-4**	**H-5**	**H-6a + H-6b**
1[a] (CDCl3)	5.18 (8)	3.45 (8)	3.99 (8)	3.92 (8)	3.85 (8)	3.90 (8) + 3.46 (8)
2a[b] ($CDCl_3$)	5.321 (IV+VIII) 5.043 (II+VI) 4.863 (III+VII) 4.857 (I+V)	3.454 (IV+VIII) 3.386 (II+VI) 3.360 (III+VII) 3.282 (I+V)	3.944 (I+V) 3.883 (III+VII) 3.881 (II+VI) 3.864 (IV+VIII)	3.807 (III+VII) 3.804 (II+VI) 3.737 (I+V) 3.704 (IV+VIII)	~3.83 (I+II+III+IV +V+VI+VII+VIII)	3.88(2) + 3.44(2) 3.85(2) + 3.47(2) 3.80(2) + 3.61(2) 3.745(2) + 3.50(2) (I+V)
2b[c] ($CDCl_3$)	5.365 (VII) 5.313 (VIII) 5.286 (V) 5.145 (II) 5.025 (I) 5.009(III+VI) 4.951 (IV)	3.524 (VIII) 3.520 (VII) 3.466 (V) 3.455 (II) 3.438 (VI) 3.434 (III) 3.363 (IV) 3.360 (I)	3.989 (II) 3.963 (VIII) 3.961 (IV) 3.942 (VII) 3.937 (I) 3.934 (V) 3.882 (III+VI)	3.948 (VI) 3.944 (II) 3.861 (III) 3.843 (V) 3.831 (IV) 3.805 (VIII) 3.786 (VII) 3.770 (I)	3.978 (VII) 3.950 (VIII) 3.845 (V) 3.828 (II) 3.806 (I) 3.840 (III+VI) 3.806 (IV)	3.942 + 3.464 (V) 3.922 + 3.540 (VI) 3.920 + 3.542 (III) 3.917 + 3.491 (II) 3.894 + 3.673 (VII) 3.876 + 3.688 (VIII) 3.837 + 3.586 (IV) 3.808 + 3.586 (I)
3[d] ($CDCl_3$)	5.176 (III+VII) 5.154 (II+VI) 5.114 (IV+VIII) 5.049 (I+V)	3.477 (IV+VIII) 3.461 (III+VII) 3.456 (II+VI) 3.366 (I+V)	3.981 (III+VII) 3.978 (I+V) 3.968 (II+VI) 3.966 (IV+VIII)	3.952 (IV+VIII) 3.938 (II+VI) 3.919 (III+VII) 3.514 (I+V)	4.019 (I+V) 3.877 (IV+VIII) 3.834 (II+VI) 3.821 (III+VII)	4.17+3.52 (IV+VIII) 3.97+3.60 (II+VI) 3.92+3.52 (III+VII) 3.76+3.51 (I+V)
4 (D_2O)	5.197(2) 5.114(2) 5.112(2) 5.105(2)	3.681(2) 3.662(2) 3.658(2) 3.656(2)	3.956(2) 3.931(2) 3.926(2) 3.923(2)	3.620(2) 3.593(2) 3.587(2) 3.566(2)	4.058(2) 3.90-3.88(6)	3.90 – 3.85(12) 3.824(2)+3.786(2)
5[e] (D_2O)	5.193(2) 5.130(2) 5.109(2) 5.066(2)	3.677(2) 3.660(2) 3.653(2) 3.637(2)	3.951(2) 3.942)2) 3.917(4)	3.640(2) 3.605(2) 3.567(2) 3.459(2)	4.005(2) 3.882(4) 3.828(2)	3.92 – 3.85(12) 3.724(2)+2.892(2)
6[f] (DMSO)	4.918(2) 4.913(2) 4.881(2) 4.870(2)	3.353(2) 3.338(2) 3.310(2) 3.298(2)	3.605(2) 3.604(2) 3.599(2) 3.573(2)	3.403(2) 3.387(2) 3.349(2) 3.300(2)	3.677(2) 3.61-3.52(6)	3.714(2)+3.645(2) 3.66(2)+3.57(2) 3.65 – 3.62(4) 2.948(2)+2.724(2) (I+V)
7 (DMSO +CD_3COOD) T = 320 K	4.949(4) 4.903(4) 4.892(4) 4.851(4)	3.386(4) 3.345(4) 3.336(4) 3.298(4)	3.49-3.64(16)	3.178(4) 3.335(4) 3.432(4) 3.486(4)	3.49-3.64(16)	3.728(4)+3.637(4) 3.841(4)+3.718(4) 3.865(4)+3.697(4) 3.317(4)+3.012(4)

[a] **24x Bn: 24x C_6H_5:** 7.24 – 7.00 m; 24**x CH_2**: 5.11 (d, *J*=11.0, 8H) and 4.745 (d, *J*=11.0, 8H), 4.485 (d, *J*=12.4, 8H) and 4.44 (d, *J*=12.4, 8H), 4.35 (d, *J*=12.2, 8H) and 4.31 (d, *J*=12.2, 8H);

[b] **22x Bn: 22x C_6H_5:** 7.19 – 7.00 m; 22**x CH_2**: 5.080 (d, *J*=10.9, 2H), 5.066 (d, *J*=10.8, 2H), 4.984 (d, *J*=11.0, 2H), 4.684 (d, *J*=11.4, 2H), 4.640 (d, *J*=10.8, 4H), 4.616 (d, *J*=10.9, 2H), 4.564 (d, *J*=11.0, 2H), 4.545 (d, *J*=12.2, 2H), 4.442 (d, *J*=12.2, 2H), 4.438 (d, *J*=12.2, 2H), 4.386 (d, *J*=12.2, 2H), 4.382 (d, *J*=12.5, 2H), 4.375 (d, *J*=12.4, 2H), 4.350 (d, *J*=12.4, 4H), 4.343 (d, *J*=12.4, 2H), 4.339 (d, *J*=12.5, 2H), 4.329 (d, *J*=12.4, 4H), 4.322 (d, *J*=12.5, 2H), 4.285 (d, *J*=12.2, 2H);

[c] **22x Bn: 22x C_6H_5:** 7.28-7.06 m; 22**x CH_2**: 5.052(2), 4.972, 4.94-4.90(5), 4.64-4.25(36H);

[d] **22x Bn: 22x C_6H_5:** 7.28 – 7.07 m; 22x **CH_2**: 5.214 (d, *J*=10.8, 2H), 5.203 (d, *J*=10.8, 2H), 5.098 (d, *J*=10.9, 2H), 4.886 (d, *J*=10.9, 2H), 4.778 (d, *J*=10.9, 2H), 4.772 (d, *J*=10.9, 2H), 4.722 (d, *J*=10.8, 2H), 4.716 (d, *J*=11.0, 2H), 4.514

(d, *J*=12.3, 2H), 4.507 (d, *J*=12.5, 2H), 4.498 (d, *J*=12.5, 2H), 4.486 (d, *J*=12.6, 2H), 4.466 (d, *J*=11.8, 2H), 4.462 (d, *J*=12.1, 2H), 4.451 (d, *J*=12.3, 2H), 4.443 (d, *J*=12.6, 2H), 4.421 (d, *J*=12.4, 2H), 4.406 (s, 4H), 4.396 (d, *J*=12.0, 2H), 4.379 (d, *J*=11.9, 2H), 4.372 (d, *J*=12.0, 2H).

[e] **2x SCOCH$_3$**: 2.400 s;

[f] **8x 2-OH**: 5.872 (d, *J*=6.3), 5.797 (d, *J*=6.9), 5.752 (d, *J*=7.5), 5.709 (d, *J*=7.4); **8x 3-OH**: 5.885 (d, *J*=2.9), 5.819 (d, *J*=2.8), 5.784 (d, *J*=2.6), 5.674 (d, *J*=2.5); **6x 6-OH**: 4.574 (t, *J*=5.5), 4.561 (t, *J*=5.6), 4.533 (t, *J*=5.7); **SH**: 2.119 (t, *J*=8.0).

Table S2. ^{13}C NMR data for compounds **1**-**7**.

Compound (solvent)	**C-1**	**C-2**	**C-3**	**C-4**	**C-5**	**C-6**
1 [a]	98.46 (8)	78.79 (8)	80.99 (8)	78.36 (8)	71.38 (8)	69.00 (8)
($CDCl_3$)						
2a [b]	99.19 (III+VII)	79.19 (I+V)	81.01 (I+V)	80.04 (IV+VIII)	71.78(2)	69.29(2)
($CDCl_3$)	98.71 (I+V)	78.86 (III+VII)	80.94 (III+VII)	79.69 (II+VI)	71.71(2)	68.98(2)
	98.64 (II+VI)	78.72 (II+VI)	80.65 (II+VI)	79.28 (III+VII)	71.43(2)	68.61(2)
	97.42 (IV+VIII)	77.44 (IV+VIII)	80.22 (IV+VIII)	74.26 (I+V)	71.38(2)	61.33 (I+V)
2b [c]	98.87 (VI)	79.24 (I)	80.98 (IV)	79.58 (V)	71.86 (IV)	69.39 (VII)
($CDCl_3$)	98.70 (III)	79.18 (IV)	80.96 (II)	79.36 (IV)	71.85 (I)	69.31 (VIII)
	98.64 (II)	79.01 (III)	80.95 (VI)	79.18 (VII)	71.84 (VIII)	68.53 (V)
	98.46 (V)	78.94 (VI)	80.94 (I)	78.94 (III)	71.76 (VII)	68.62 (II)
	98.32 (IV+VII)	78.85 (II)	80.92 (V)	78.74 (II)	71.48 (III)	61.57 (I+IV)
	98.15 (VIII)	78.68 (V)	80.84 (III)	77.76 (VI)	71.44 (VI)	69.19 (VI)
	97.87 (I)	77.65 (VII)	80.34 (VII)	75.90 (I)	71.37 (II+V)	69.13 (III)
		77.46 (VIII)	80.17 (VIII)	75.74 (VIII)		
	98.57 (II+VI)	79.21 (I+V)	81.08	79.67 (II+VI)	71.56 (II+VI)	69.06 (IV+VIII)
3 [d]	98.27 (IV+VIII)	78.66 (II+VI)	80.89	79.50 (I+V)	71.38 (IV+VIII)	68.93 (B+VI)
($CDCl_3$)	98.14 (III+VII)	78.59 (III+VII)	80.82	78.95 (III+VII)	71.36 (III+VII)	68.86 (III+VII)
	97.94 (I+V)	78.40 (IV+VIII)	80.47	77.50 (IV+VIII)	70.87 (I+V)	34.35 (I+V)
	104.38(2)	74.99(2)	75.65 (2)	85.16(2)	74.70(2)	63.13(2)
4	104.32(4)	74.97(2)	75.63(2)	83.37(2)	74.50(2)	63.07(2)
(D_2O)	104.21(2)	74.94(2)	75.55(2)	83.33(2)	74.44(2)	62.97(2)
		74.87(2)	75.19(2)	83.18(2)	73.06(2)	36.08(2)
	104.61(2)	75.04(2)	75.61(4)	86.70(2)	74.50(2)	63.06(2)
5 [e]	104.26(2)	75.02(2)	75.54(2)	83.14(2)	74.46(2)	62.91(2)
(D_2O)	104.11(2)	74.93(2)	75.52(2)	82.91(2)	74.42(2)	62.70(2)
	103.79(2)	74.82(2)		82.83(2)	73.34(2)	33.42(2)
	104.49(2)	73.06(2)		84.18(2)	72.48(2)	60.32(2)
6	102.46(2)	72.96(2)		81.56(2)	72.36(2)	59.97(4)
(DMSO)	102.08(2)	72.91(2)		80.83(2)	72.27(2)	25.82(2)
	101.76(2)	72.88(4)		80.66(2)	71.39(2)	
		72.83(2)				
		72.63(2)				
		72.50(2)				
7	103.07(4)	73.36(4)		86.27(4)	72.23(8)	60.37(4)
(DMSO	102.80(4)	73.26(4)		82.43(4)	71.96(4)	59.73(4)
+CD_3COOD)	102.77(4)	73.02(4)		81.71(4)	70.02(4)	59.33(4)
T = 320 K	100.94(4)	72.95(4)		79.76(4)		41.25(4)
		72.74(4)				
		72.71(4)				
		72.39(4)				
		72.23(4)				

[a] **24x Bn: 24x Ar-C$_{ipso}$:** 139.21(8), 138.21(8), 138.10(8); **24x Ar-C$_{para}$ + 44x Ar-C$_{ortho}$ + 44x Ar-C$_{meta}$**: 128.28(16), 128.12(16), 128.02(16), 127.95(16), 127.52(24), 127.45(8), 127.33(16), 126.88(8); **24x CH$_2$**: 75.50(8), 73.11(8), 72.78(8).

[b] **22x Bn: 22x Ar-C$_{ipso}$:** 139.32, 139.29, 139.22, 139.15, 138.99, 138.97, 138.89(2), 138.33, 138.25(2), 138.21, 138.20, 138.18, 138.14, 138.13(2), 138.09, 137.93, 137.92, 137.89, 137.82; **22x Ar-C$_{para}$ + 44x Ar-C$_{ortho}$ + 44x Ar-C$_{meta}$**: 128.31(4), 128.30(4), 128.24(4), 128.20(4), 128.15(4), 128.14(4), 128.12(4), 128.10(4), 128.08(4), 128.02(4), 127.99(4), 127.98(4), 127.94(10), 127.92(6), 127.76(8), 127.58(4), 127.56(6), 127.53(2), 127.48(6), 127.46(2), 127.44(2), 127.37(4), 127.30(4), 127.02(2), 126.98(4), 126.88(2); **22x CH$_2$**: 75.80(2), 75.65(2), 75.15(2), 74.56(2), 73.27(4), 73.13(2), 73.04(2), 72.91(2), 72.69(2), 72.59(2).

[c] **22x Bn: 22x Ar-C$_{ipso}$:** 139.33, 139.29, 139.22, 139.15, 138.99, 138.97, 138.89(2), 138.33, 138.25(2), 138.21, 138.20, 138.18, 138.14, 138.13(2), 138.09, 137.93, 137.92, 137.89, 137.82; **22x Ar-C$_{para}$ + 44x Ar-C$_{ortho}$ + 44x Ar-C$_{meta}$**: 128.34 – 126.85(110); **22x CH$_2$**: 75.72, 75.54, 75.44, 75.09, 74.92, 74.88, 74.68, 74.56, 73.31, 73.30, 73.29, 73.25, 73.20(2), 73.11, 73.01, 72.97, 72.96(2), 72.70, 72.63, 72.57.

[d] **22x Bn: 22x Ar-C$_{ipso}$:** 139.30(2), 139.26(2), 139.00(2), 138.87(2), 138.32(2), 138.18(2), 138.14(2), 138.08(2), 138.04(2), 138.02(2), 137.96(2); **22x Ar-C$_{para}$ + 44x Ar-C$_{ortho}$ + 44x Ar-C$_{meta}$**: 128.34(4), 128.29(4), 128.28(4), 128.23(4), 128.22(4), 128.14(8), 128.11(4), 128.09(4), 128.04(20), 127.97(8), 127.93(4), 127.66(2), 127.64(4), 127.59(4), 127.52(6), 127.51(6), 127.50(2), 127.44(2), 127.37(4), 127.21(4), 127.06(2), 127.03(2), 126.93(2), 126.91(2); **22x CH$_2$**: 75.96(2), 75.80(2), 75.44(2), 75.29(2), 73.29(2), 73.25(4), 73.14(2), 72.91(2), 72.83(2), 72.67(2).

[e] **2xCH$_3$-CO-S:** 32.74(2), 203.06(2).

1.4. Crystal structure of duplex 7:

1.4.1. Crystallization experiments: Crystals were grown from DMSO/water solutions using hanging drops. Thus, duplex **7** was dissolved in DMSO at concentrations 0.7-3.5 mg/ml and placed on silanized glass cover slides over vials containing water/DMSO in various proportions (Table S3). The top of each vial was sealed with high vacuum grease. Vials were kept at 18°C for about one week. After this period vials in columns B and C contained crystal suitable for X-ray analyses (Figure S2).

Table S3. Solvent composition

Row of vials/ solvents in each vial	**A**	**B**	**C**	**D**	**E**
H_2O, ml	0.3	0.9	1.5	2.1	2.7
DMSO, ml	2.7	2.1	1.5	0.9	0.3

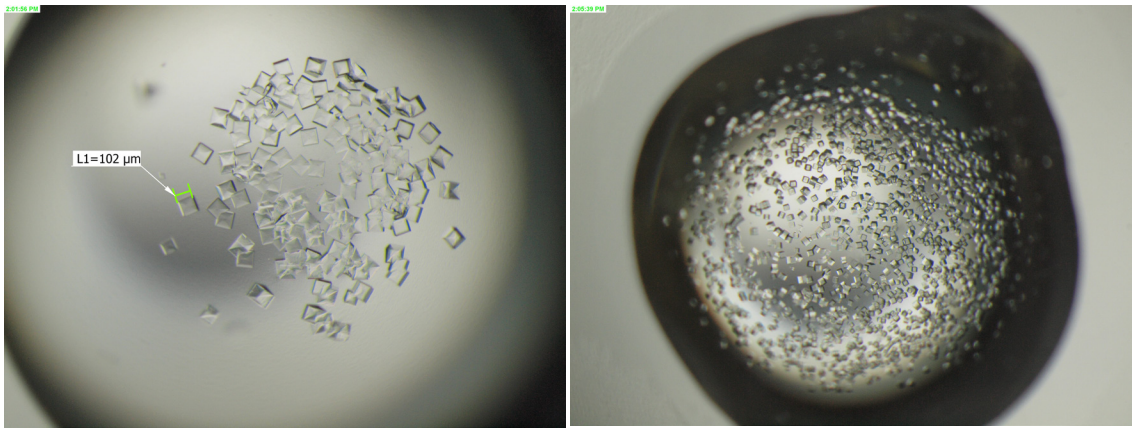


Figure S2. Crystals grown on hanging drop in DMSO/water mixtures.

1.4.2. X-Ray analysis.

Small rectangular single crystal of dimensions 0.1x0.06x0.06 mm was measured at 120K with four-circle kappa diffractometer Gemini of Oxford Diffraction. The radiation CuKα from a classical sealed x-ray tube was collimated and focused by mirrors of the Cu-Ultra collimator, and detected with CCD detector Atlas. The data were processed by Crysalis Pro version 1.171.37.34 of Agilent Technologies, scaling and absorption correction was done by Jana2006, version 16/9/2014.[4] Although the sample was weakly diffracting, the obtained data were of good quality.

Symmetry of the sample was cubic, with non-centrosymmetric space group I23. The structure was partially solved by direct methods using SIR2011;[5] refinement and completing of the structure model was done with Jana2006. The main problem of structure determination was the presence of disordered solvent, which could not be exactly determined. This solvent represents 1/3 of the molecular weight and without correcting for its presence the molecule of the duplex could not be reliably determined. Such correction could be done either using the SQUEEZE technique[6] which removes contribution of the solvent directly from the input diffraction data, or by careful modeling of the solvent's electron density. We used the second more laborious approach because it does not affect input data and provides full information about the investigated structure.

In order to describe electron density of solvent, oxygen atoms were placed to the maxima of the difference Fourier map located outside the duplex. Displacement parameters and occupancy of these "oxygens" were refined. In many cases, anharmonic displacement parameters of such auxiliary atoms help to efficiently describe electron density. In our case, however, neither harmonic nor anharmonic displacement parameters could be used because of extending of such shape during the refinement to the area of the duplex. Finally, we succeeded to describe electron density of the solvent with numerous isotropic auxiliary oxygen atoms with refined occupancy. In the final structure model these atoms are labeled XO* and they have no geometric meaning.

As soon as the solvent was described difference Fourier map revealed details of the duplex, including hints about hydroxyl hydrogen atoms. However, refinement of all individual structure parameters was still not possible. To improve stability of the refinement, we used rigid body approach for glucose units. Only one such unit was

refined and positioned to four places in the unit cell with help of three rotation angles and a translation vector. The rotation angles and the translation vector were refined separately for each of the four symmetry-independent glucose positions. Displacement parameters were refined as TLS tensors separately for each position. Sulfur atoms, the bridging oxygen atoms and the CH_2-OH chains were refined as individual atoms, i.e. without the rigid body approach. For these atoms the distances were restraint by such way that corresponding distances were kept the same, and displacement parameters were refined independently. Hydrogen atoms bonded to carbon were attached geometrically. For hydrogen atoms of hydroxyl O-H distance was constraint to 0.96 Å and the angle C-O-H to 109°.

The final non-weighted R value for observed (>3σ) reflections was 0.1089 and goodness of fit S=3.34. The higher value of S corresponds with the fact that the structure cannot be determined with usual accuracy due to the solvent. It should be noted that Jana2006 does not refine the weighting scheme and calculates the true goodness of fit based on experimental weights. The crystallographic data were deposited into the Cambridge Structural Database under the number CCDC 1026740 and can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

1.5. Isothermal titration calorimetry procedures

All solutions for the titration experiments were prepared using ultra-pure water (Milli-Q Synthesis, total organic carbon content ≤ 5 ppb). Samples were prepared in bottles made of clear Duran borosilicate glass, which had been cleaned with peroxysulfuric and dried thoroughly prior to use. Calculations of the concentrations of commercially available compounds (**1**- **6**) were based on the minimal guaranteed content of compound provided by manufacturers for each compound. Due to hygroscopic nature of duplex **7**, the concentrations in the prepared samples were determined by quantitative NMR measurements (in DMSO + 20 μL of CF_3COOD) of the lyophilized samples using acetanilide (sublimed, 99.99%) as an internal standard. Concentrations of samples are given in Table 4.

Table S4 . Concentrations of stock solutions used for titrations.

entry	burette		cell	
	compd.	conc. (M)	compd.	conc. (M)
1	**7**	8.25×10^{-5}	**8**	4.32×10^{-6}
2	**7**	1.02×10^{-4}	**9**	1.08×10^{-5}
3	**7**	8.63×10^{-5}	**10**	9.02×10^{-6}
4	**7**	1.02×10^{-4}	**11**	1.01×10^{-5}
5	**12**	7.90×10^{-4}	**7**	4.25×10^{-5}
6	**13**	5.75×10^{-4}	**7**	3.15×10^{-5}

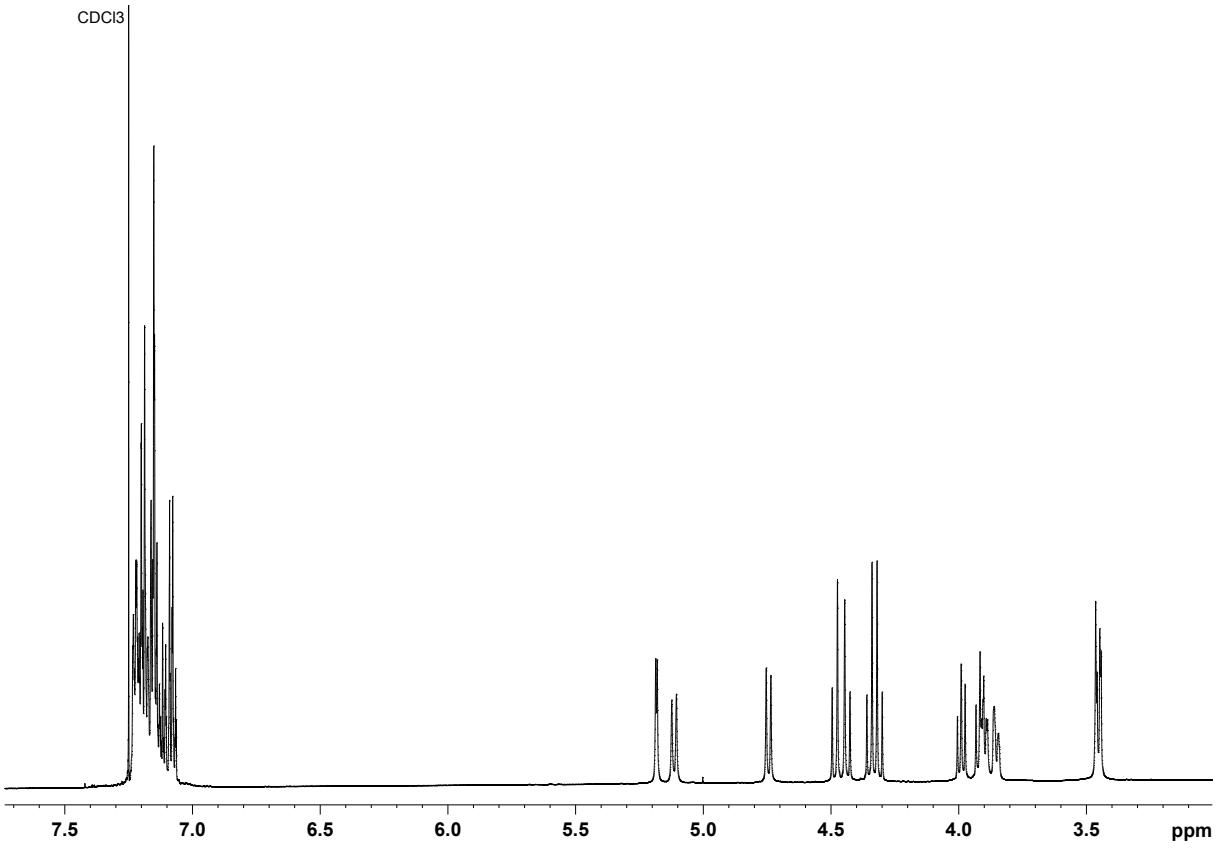
The titrations were carried out with MicroCal VP ITC system, model 2007. In a typical run, clean cell was washed twice with the sample solution and titrated with the solution of the titrant. The added injected volumes varied from 5 to14 μL, except for the first throw-away injection which was 2 μL in all cases. The recorded thermograms were – after manual baseline corrections – analysed using Origine 7 based software supplied by Microcal.

Standard errors estimates for $\Delta H°$ and K were directly produced by Origine 7 as results of fitting procedures; they correspond to square roots of multiples of corresponding diagonal elements of variance-covariance matrix by reported reduced χ^2 values. Standard error of $T\Delta S°$ was computed using equation Eq 1, which was derived[3] from the general formula

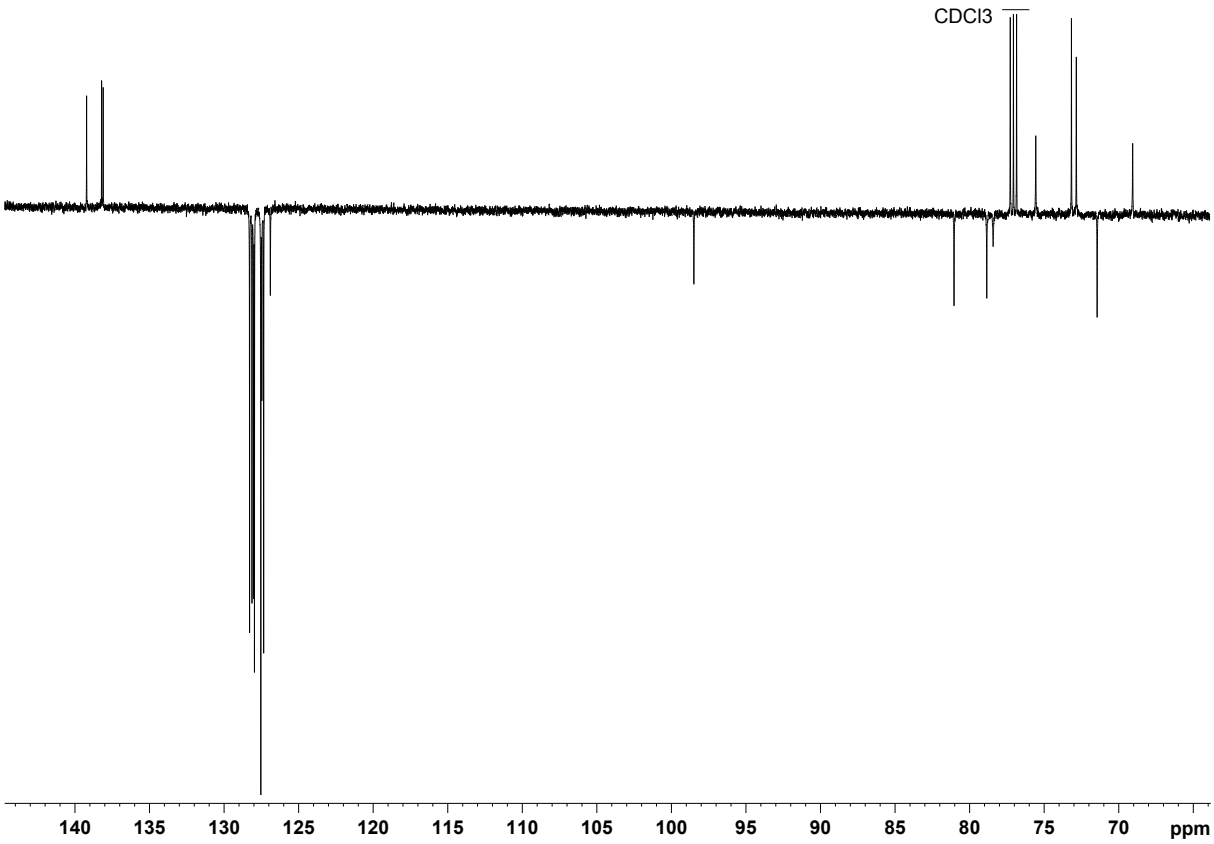
for calculation of error propagation of correlated variables, where $\sigma_{\Delta H^\circ}$ and σ_K stand for the standard errors estimates of ΔH° and K, respectively, and ρ_{K,H° is the correlation coefficient for ΔH° and K calculated from variance-covariance matrix.

$$\sigma_{T\Delta S^\circ} = \sqrt{\sigma^2_{\Delta H^\circ} + \left(\frac{RT}{K}\right)^2 \sigma^2_K + 2\left(\frac{RT}{K}\right)\rho_{K,\Delta H^\circ}\sigma_{\Delta H^\circ}\sigma_K} \qquad \text{(Eq 1)}$$

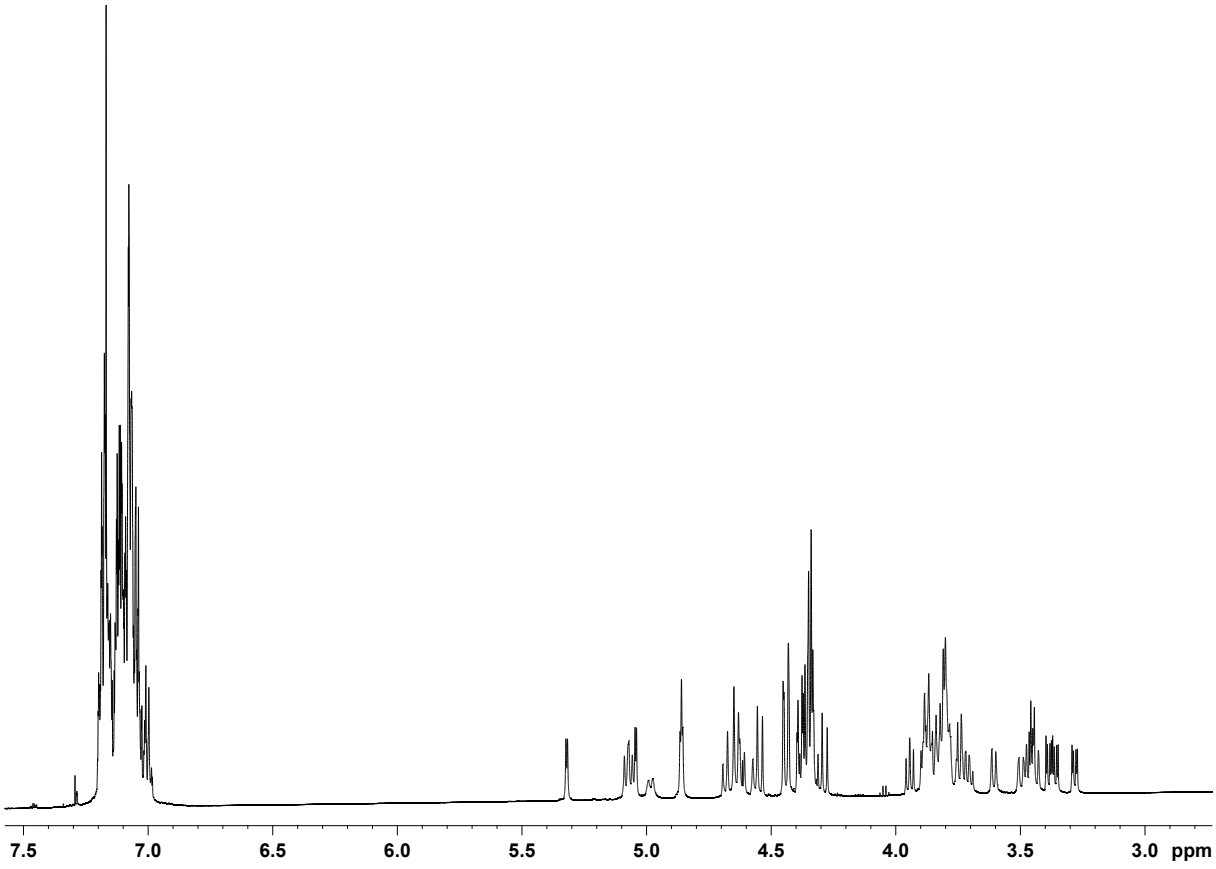
2. Copies of 1H and 13C NMR spectra of compounds 1-7



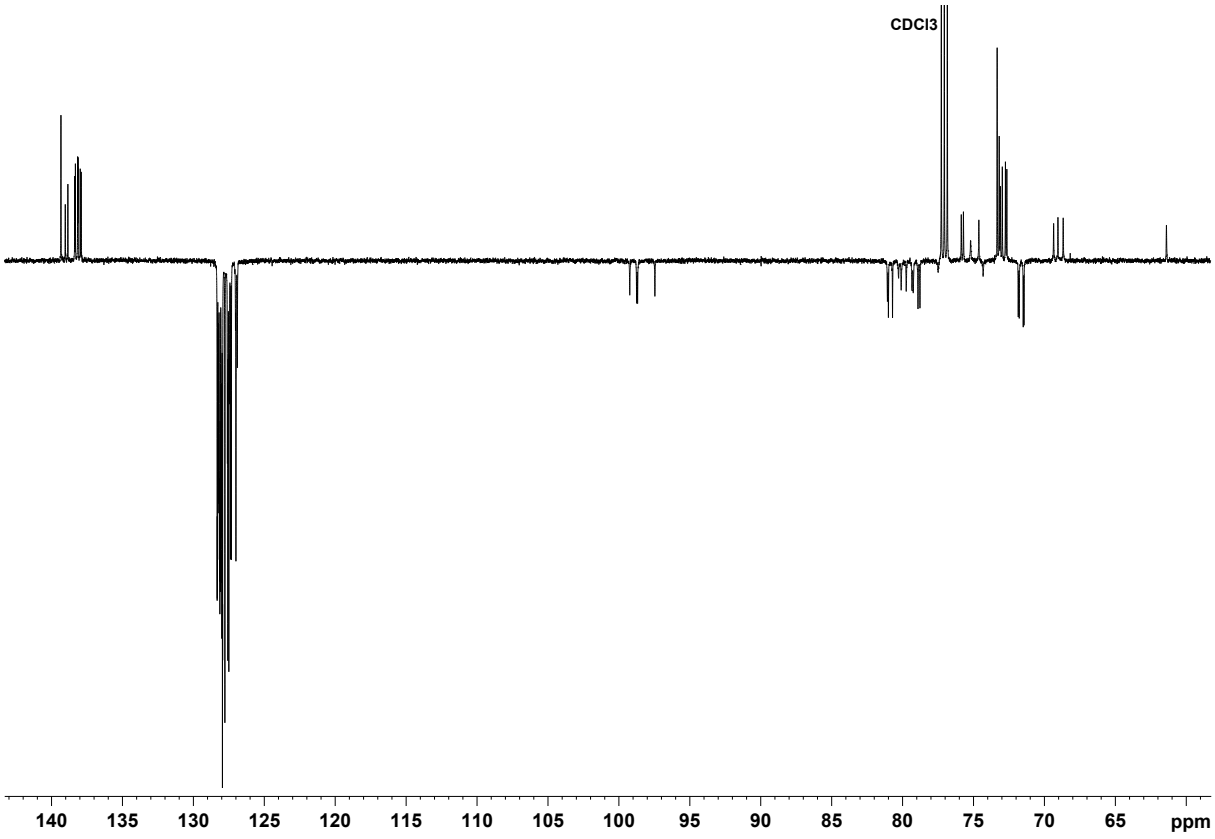
Compound **1**
(600 MHz ^{1}H NMR spectrum in $CDCl_3$; T = 300 K)



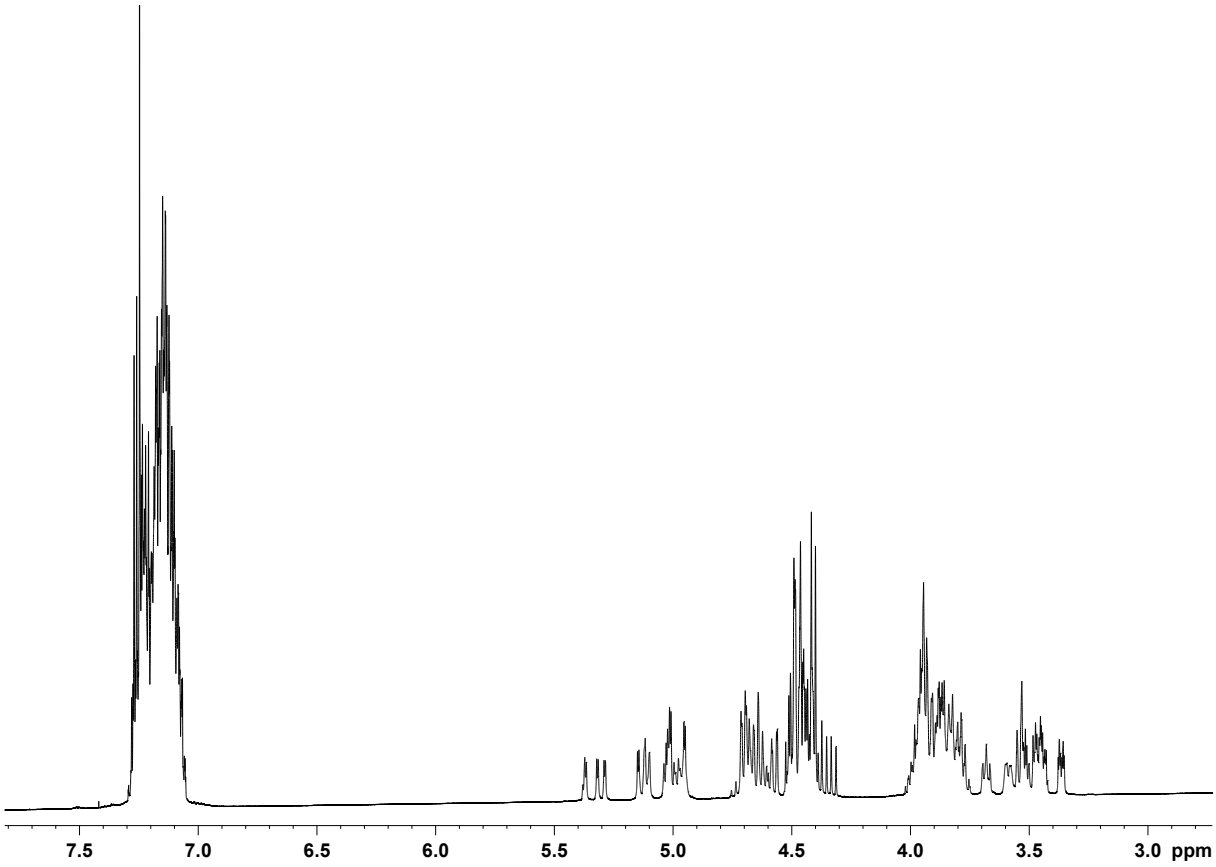
Compound **1**
(150.9 MHz ^{13}C NMR spectrum in $CDCl_3$; T = 300)



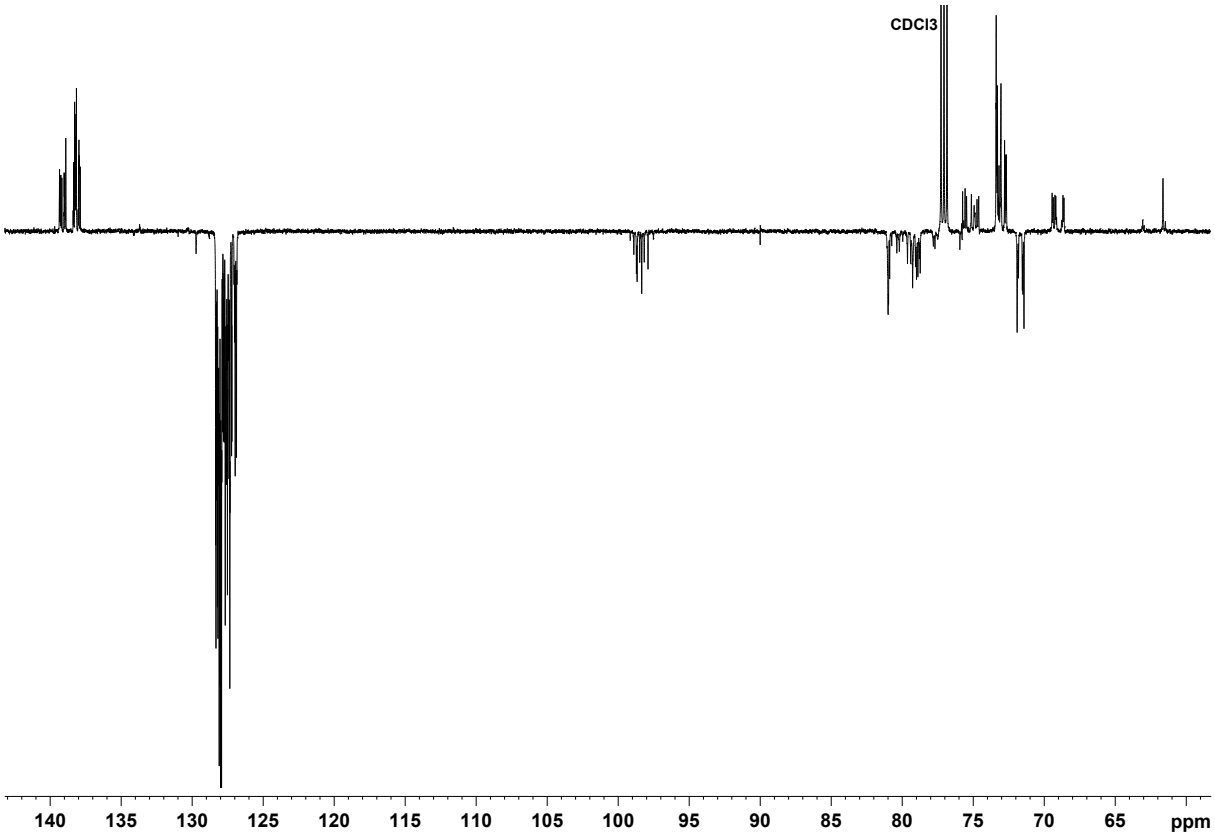
Compound **2a**
(600 MHz ^{1}H NMR spectrum in $CDCl_3$; T = 300 K)



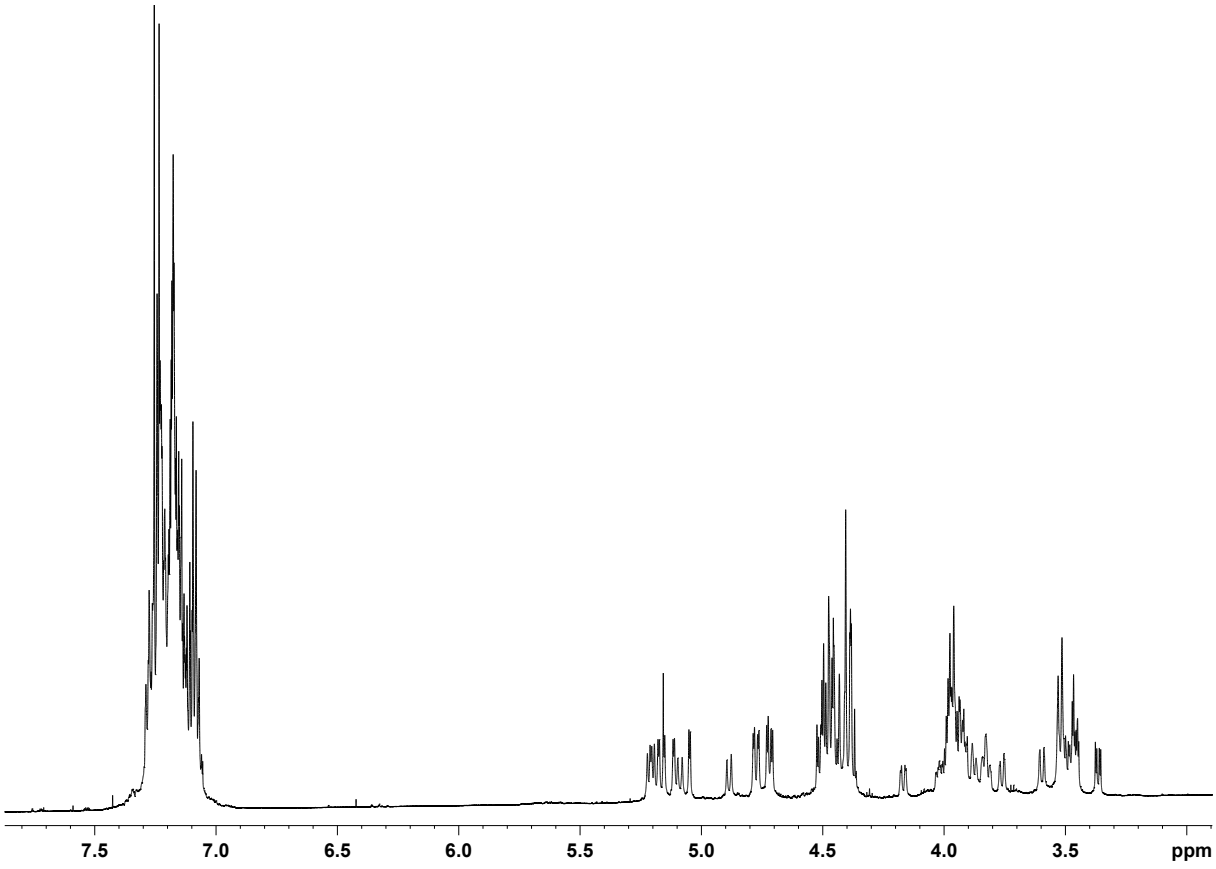
Compound **2a**
(150.9 MHz ^{13}C NMR spectrum in $CDCl_3$; T = 300 K)



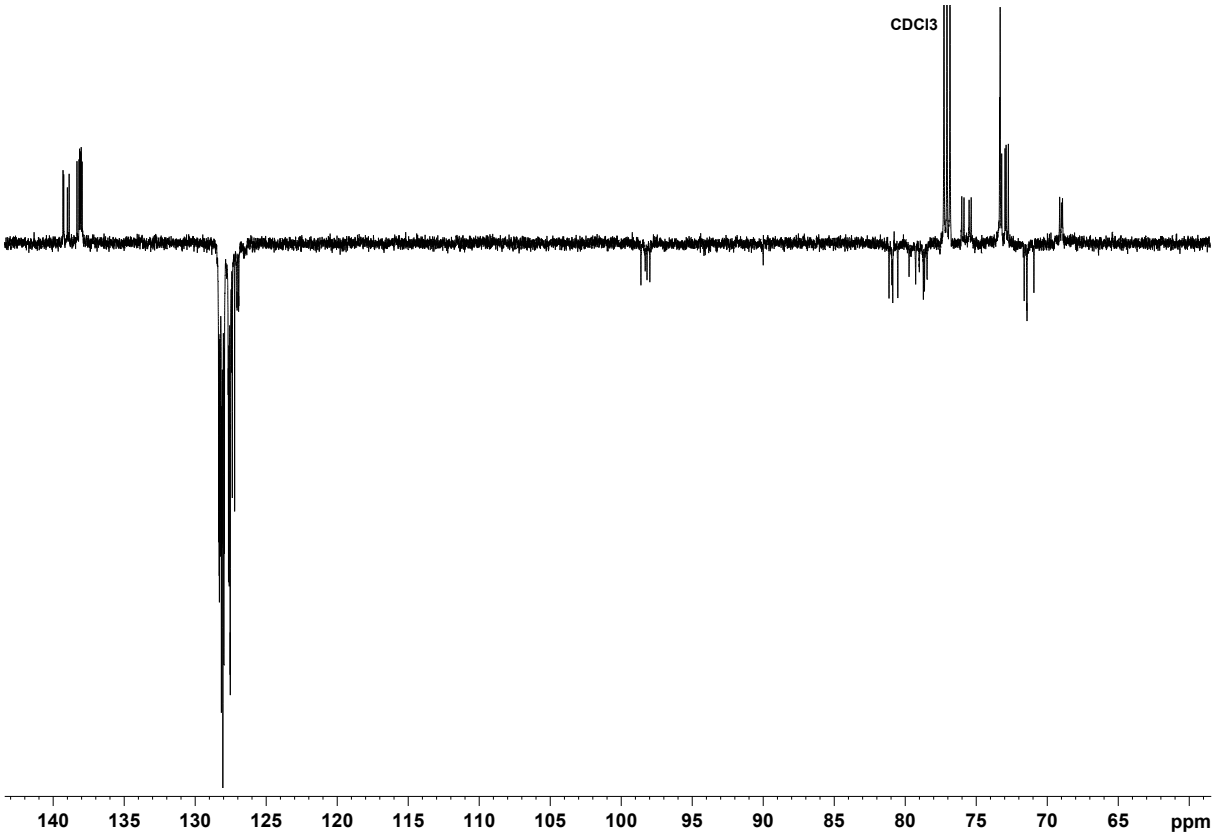
Compound **2b**
(600 MHz ^{1}H NMR spectrum in $CDCl_3$; T = 300 K)



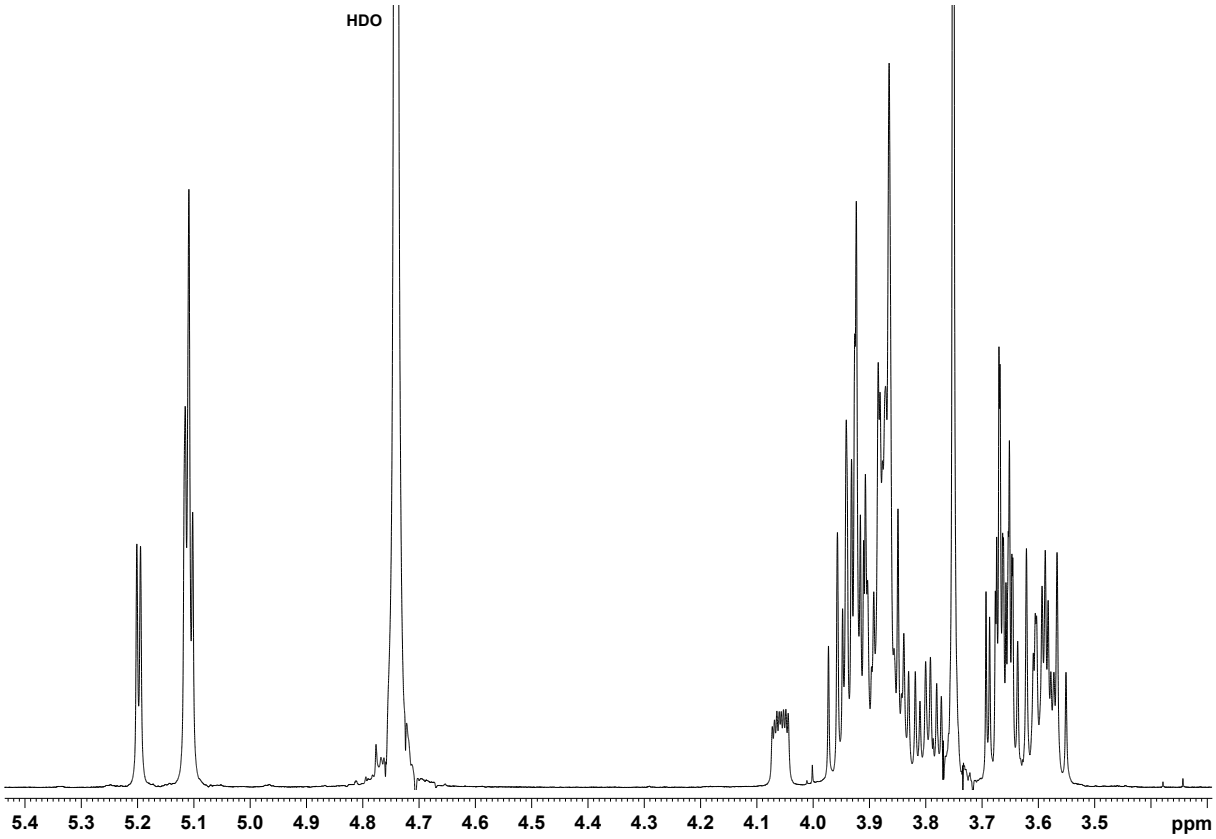
Compound **2b**
(150.9 MHz ^{13}C NMR spectrum in $CDCl_3$; T = 300 K)



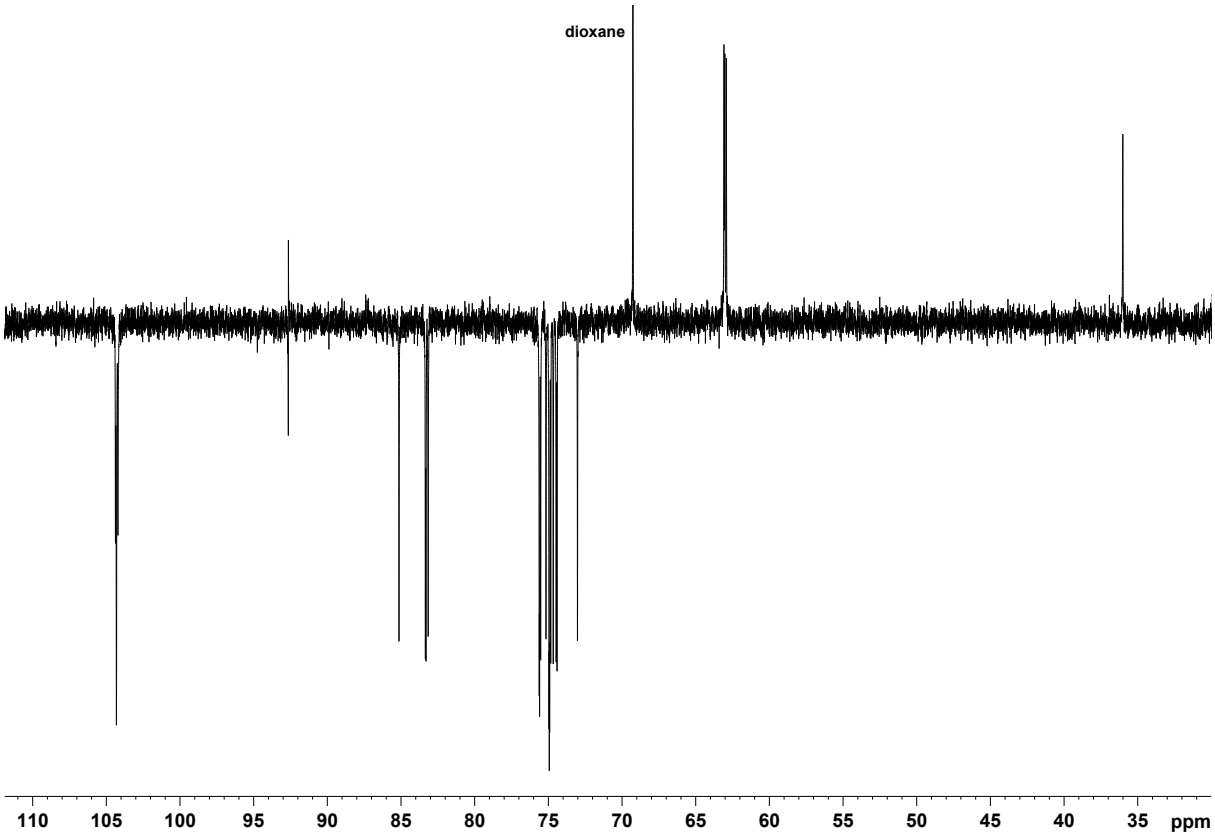
Compound **3**
(600 MHz ^{1}H NMR spectrum in $CDCl_3$; T = 300 K)



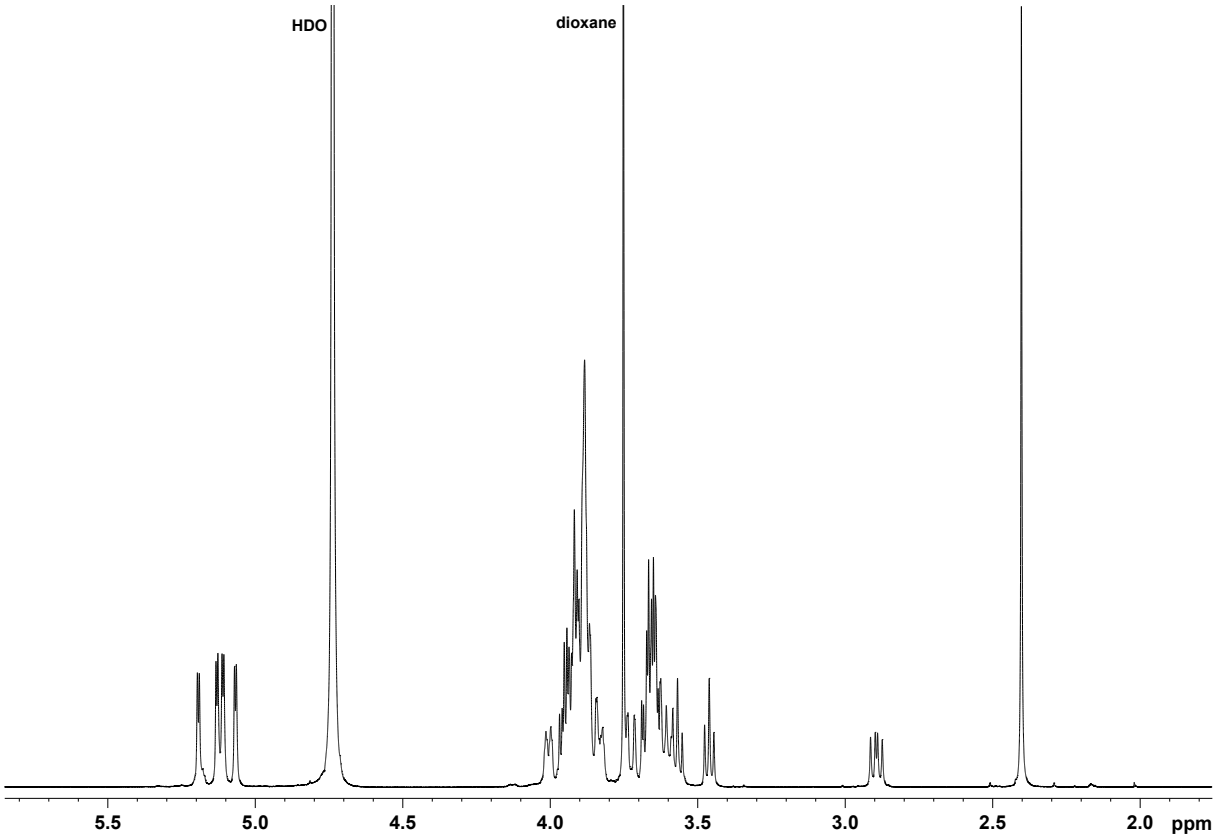
Compound **3**
(150.9 MHz ^{13}C NMR spectrum in $CDCl_3$; T = 300 K)



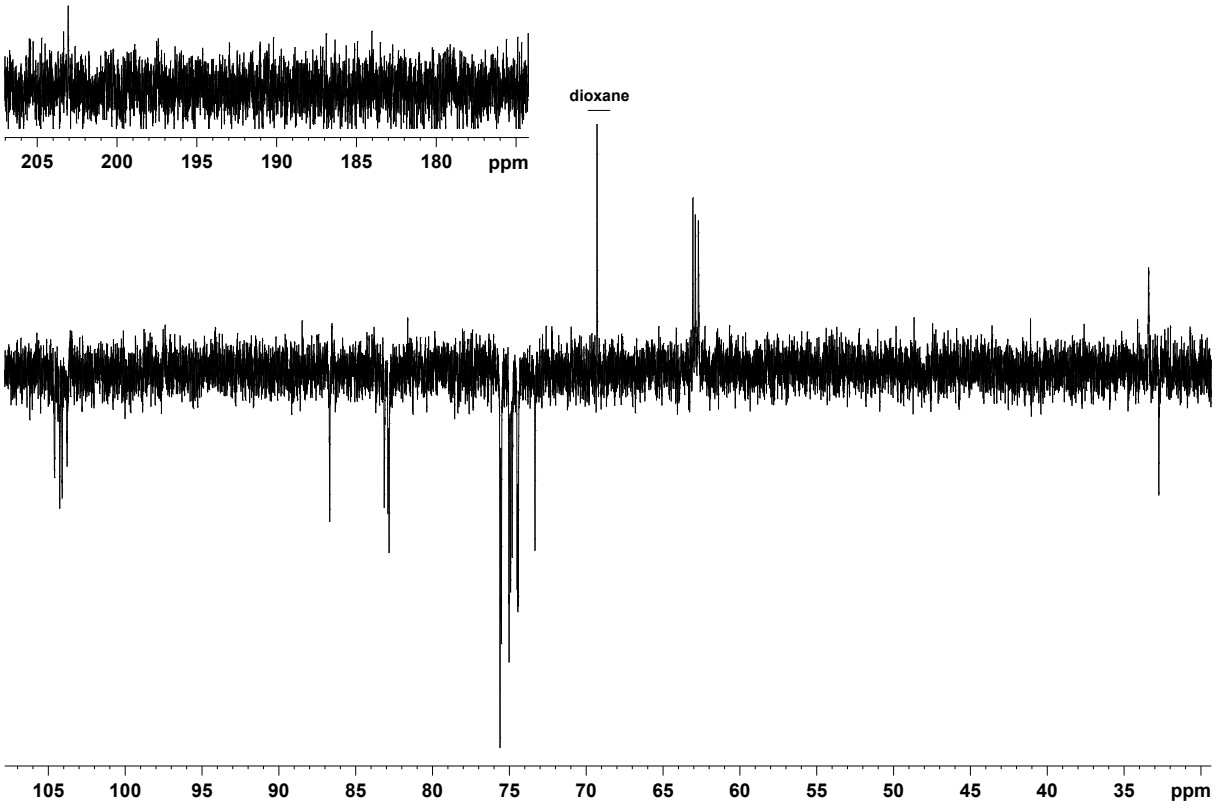
Compound **4**
(600 MHz ^{1}H NMR spectrum in D_2O; T = 300 K)



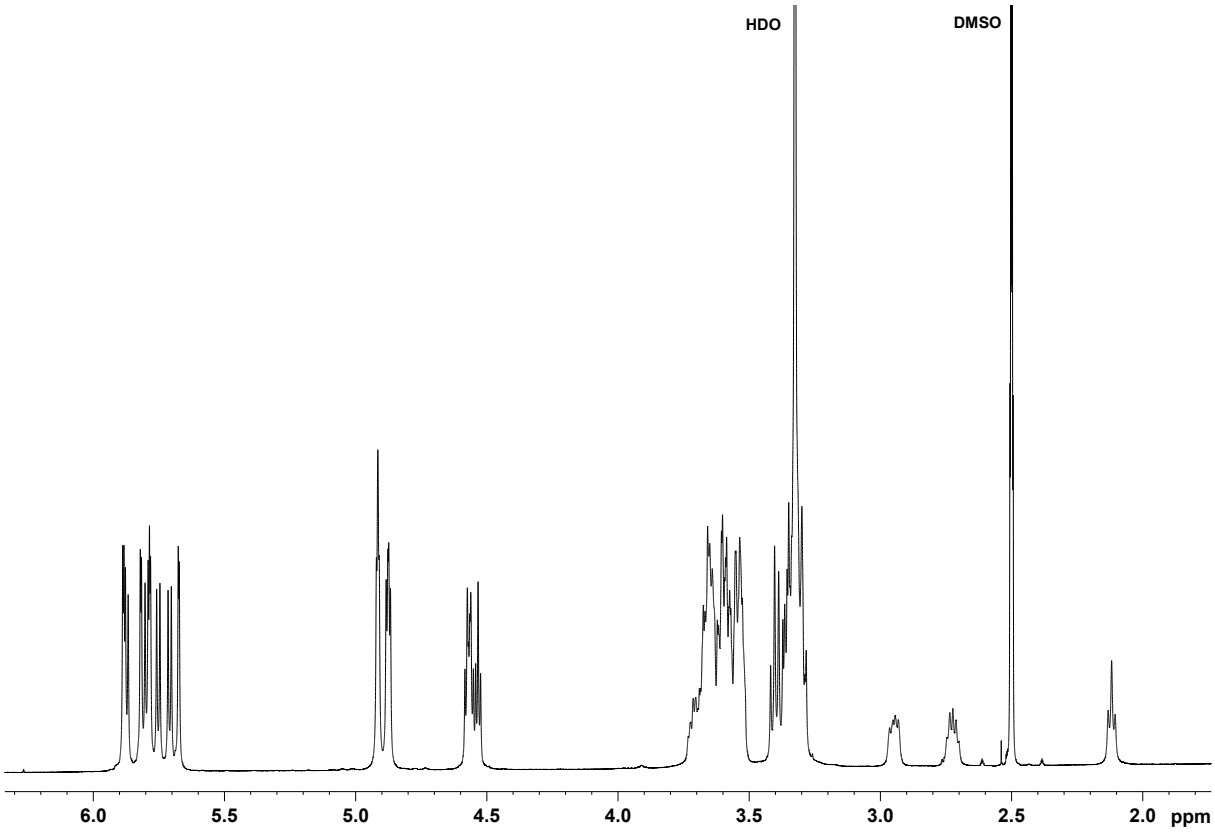
Compound **4**
(150.9 MHz ^{13}C NMR spectrum in D_2O; T = 300 K)



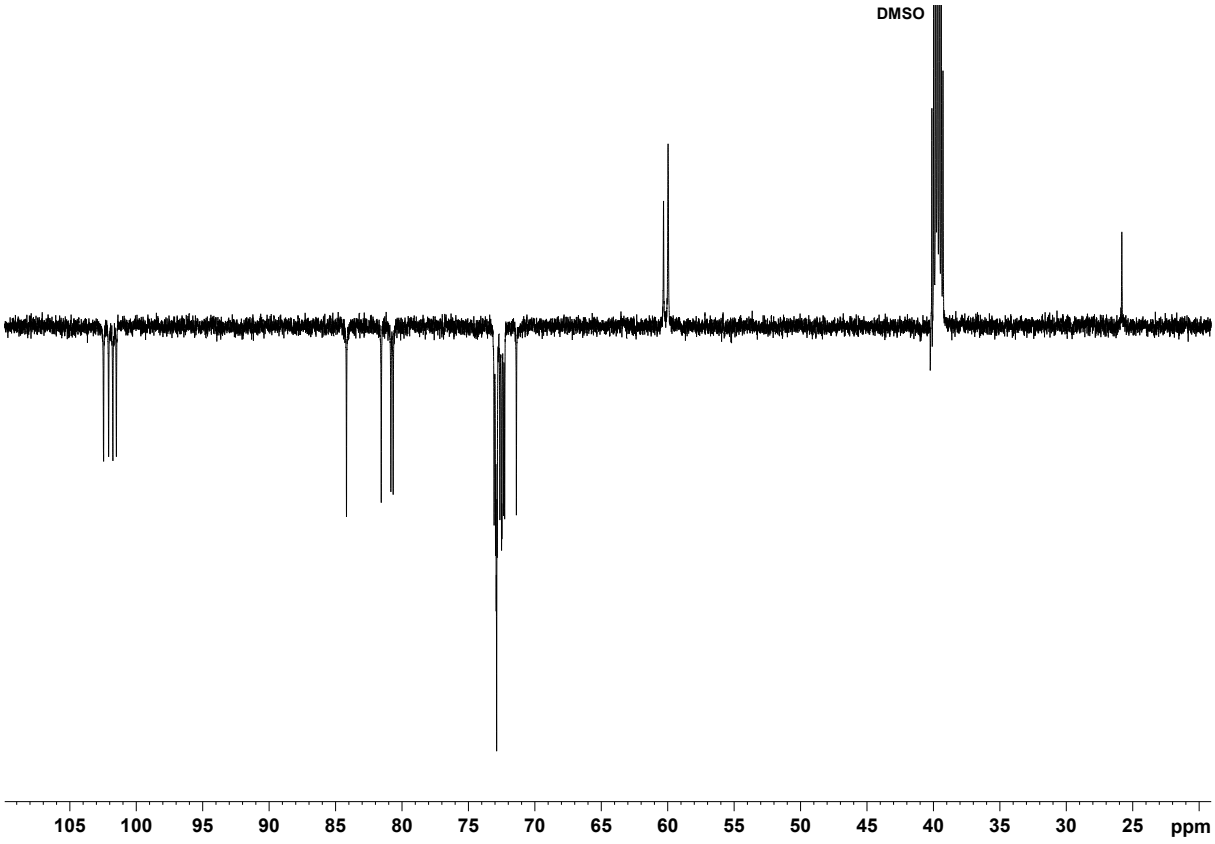
Compound **5**
(600 MHz ^{1}H NMR spectrum in D_2O; T = 300 K)



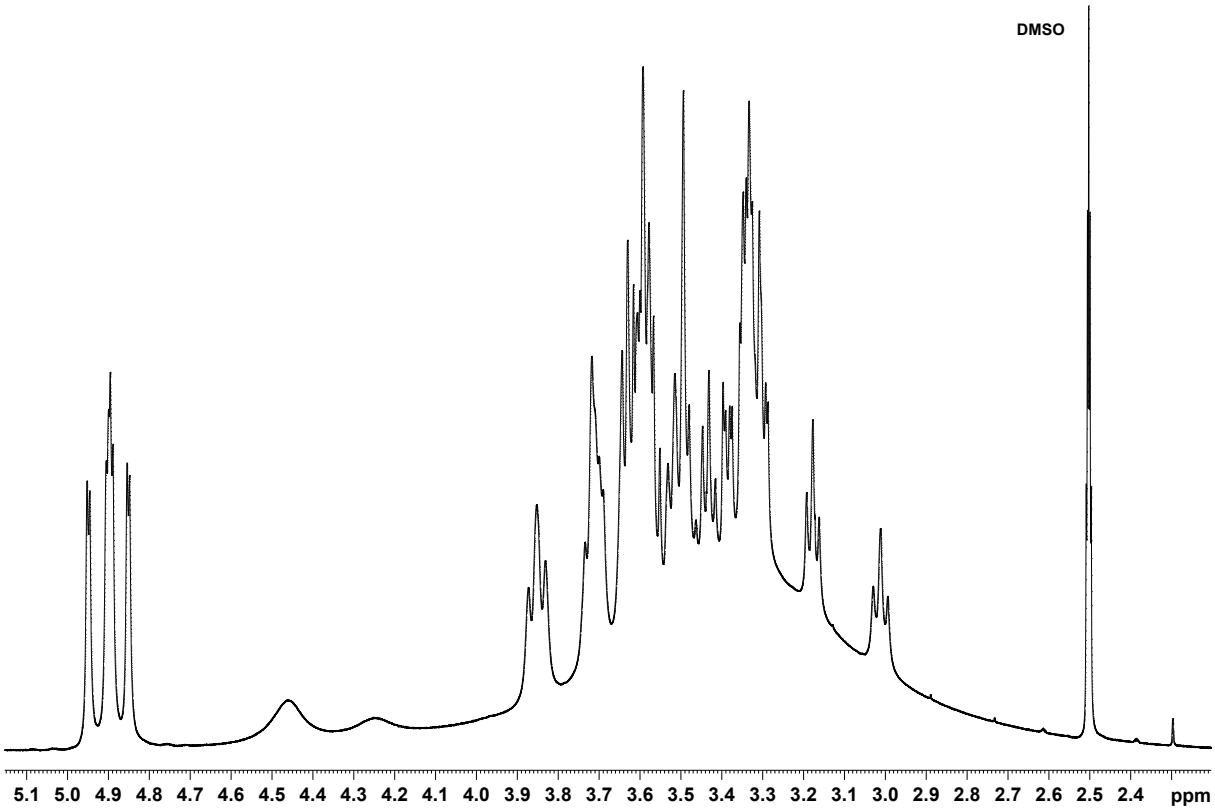
Compound **5**
(150.9 MHz ^{13}C NMR spectrum in D_2O; T = 300 K)



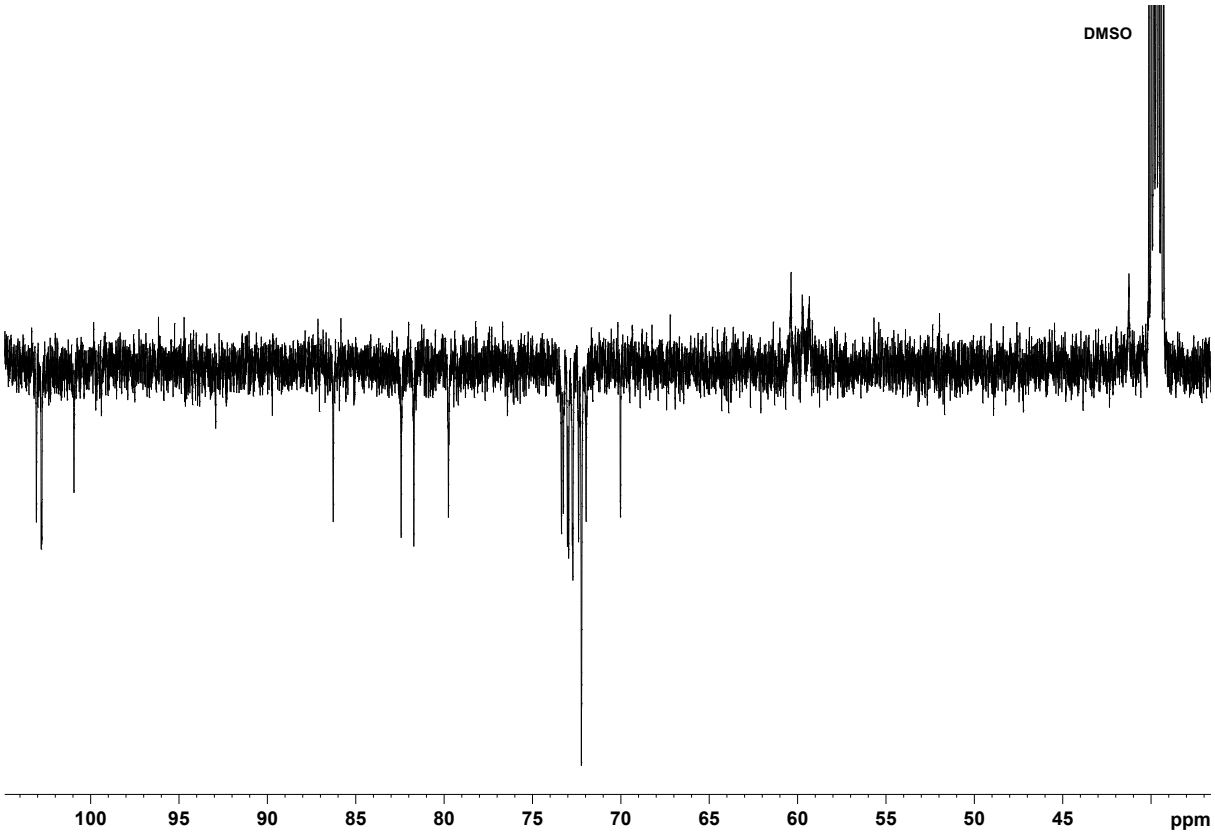
Compound **6**
(600 MHz ^{1}H NMR spectrum in DMSO; T = 300 K)



Compound **6**
(150.9 MHz ^{13}C NMR spectrum in DMSO; T = 300 K)



Compound **7**
(600 MHz ^{1}H NMR spectrum in DMSO; T = 320 K)



Compound **7**
(150.9 MHz ^{13}C NMR spectrum in DMSO; T = 320 K)

References

1. Laskowski, R. A. *J. Mol. Graph.* **1995**, *13*, 323.
2. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G.S.; Greenblatt D.M.; Meng E. C. and Ferrin, T. E., *J. Comput. Chem.* **2004,** 25, 1605–1612.
3. L. Kumprecht, M. Budesinsky, J. Vondrasek, J. Vymetal, J. Cerny, I. Cisarova, J. Brynda, V. Herzig, P. Koutnik, J. Zavada and T. Kraus, *J. Org. Chem.*, 2009, **74**, 1082-1092.
4. Petricek, V., Dusek, M. & Palatinus, L. *Z. Kristallogr.* **2014**, *229*(5), 345-352
5. Burla, M.C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., Mallamo. M., Mazzone, A., Polidori, G., Spagna, R. SIR2011: a new package for crystal structure determination and refinement, *J. Appl. Cryst.* **2012**, *45*, 357-361
6. A.L.Spek (2005) PLATON, A Multipurpose Crystallographic Tool, Utrecht University, Utrecht, The Netherlands