Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

5'- Vs 3'-end sugar conformational control in shaping up dinucleotides

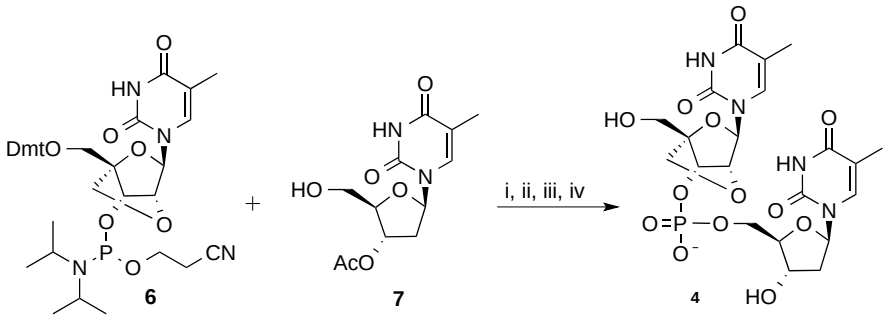
Jouda Jakhlal, Stéphanie Coantic-Castex, Clément Denhez, Christian Pertermann, Agathe Martinez, Dominique Harakat, Dominique Guillaume, Pascale Clivio*

Table of contents

General Remarks. Solvents and chemicals used for the reactions were purchased from commercial suppliers and used without further purification unless otherwise stated. 2′-*O*-4′-*C*-methylene-5′-*O*-dimethoxytrityl-5-methyluridine 3′-*N*,*N*′-diisopropyl(cyanoethyl)phosphoramidite (**6**) and 2′-*O*-4′-*C*-methylene-5′-*O*-dimethoxytrityl-5-methyluridine (**9**) were prepared from 2′-*O*-4′-*C*-methylene-5-methyluridine ($\mathbf{T_{LN}}$) (purchased from Exiqon) using published procedures.[1] 3'-*O*-Acetylthymidine (**7**) was prepared from thymidine (**T**) (purchased from Aldrich) using standard procedures.[2] 5′-*O*-dimethoxytritylthymidine 3′-*N*,*N*′-diisopropyl(cyanoethyl)phosphoramidite (**8**) was from Eurogentec. Phosphoramidites and 3'-acetates were dried overnight at room temperature in a desiccator over P_2O_5 prior to use. Acetonitrile and dichloromethane were dried by distillation from calcium hydride. Chromatography was performed on silica gel 60, particle size 35-70 μm, unless otherwise stated. ^{1}H NMR and ^{13}C NMR spectra were recorded on 300 or 500 or 600 MHz spectrometers. Observed chemical shift (δ) values are given in ppm and coupling constants (J) in Hz. For the determination of the sugar conformer populations, spectra were analyzed using PERCH NMR software (v 2014.1). Following abbreviations are used: s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublet)..., Tp and pT represents the 5'-end and the 3'-end nucleosides, respectively, of **3**-**5**. Prochiral H2',H2" proton was attributed from its NOE with the H6 proton of the base. Prochiral H5',H5" and H6',H6" protons have not been assigned and the most deshielded one has been arbitrarily labeled H5' and H6'. ^{1}H NMR chemical shifts were calibrated using residual solvent signals at the following values: CD_3OD δ_H 3.31 and δ_C 49.15, D_2O δ_H 4.80. ^{13}C NMR spectra recorded in D_2O were calibrated from dioxane (δ_C 67.8 ppm). ^{31}P NMR and ^{19}F NMR chemical shifts were reported from an external capillary standard of 85% phosphoric acid (δ_P 0.00 ppm) and $CFCl_3$ (δ_F -77.00 ppm), respectively. High Resolution Mass Spectra (HRMS) were recorded on a Q-Tof Micromass spectrometer. HPLC purifications were performed on a Sunfire C18 (5 μm, 10 x 250 mm) column using a 67 min, 4 mL/min gradient of 0-20% CH_3CN in 0.05 M aqueous ammonium acetate. The detection was set at 260 nm.

Synthetic procedures and characterization data for all new compounds.

$T_{LN}pT_{LS}$ (3). Synthetic details will be reported elsewhere. ^{1}H NMR (D_2O, 600 MHz): δ 7.60 (1H, d, $J_{6,Me}$ = 1.2 Hz, H_6 Tp), 7.57 (1H, d, $J_{6,Me}$ = 1.2 Hz, H_6 pT), 6.49 (1H, d, $J_{1',2'}$ = 7.2 Hz, $H_{1'}$ pT), 5.62 (1H, s, $H_{1'}$ Tp), 5.26 (1H, d, $J_{3',2'}$ = 4.6 Hz, $H_{3'}$ pT), 4.94 (1H, d, $J_{6',6''}$ = 8.5 Hz, $H_{6'}$ pT), 4.67 (1H, br s, $H_{2'}$ Tp), 4.64 (1H, d, $J_{6'',6'}$ = 8.5 Hz, $H_{6''}$ pT), 4.43 (1H, d, $J_{3',P}$ = 5.4 Hz, $H_{3'}$ Tp), 4.26 (1H, dd, $J_{5',P}$ = 4.0 Hz, $J_{5',5''}$ = 11.4 Hz, $H_{5'}$ pT), 4.25 (1H, dd, $J_{2',3'}$ = 4.6 Hz, $J_{2',1'}$ = 7.2 Hz, $H_{2'}$ pT), 4.16 (1H, dd, $J_{5'',P}$ = 4.5 Hz, $J_{5''-5'}$ = 11.4 Hz, $H_{5''}$ pT), 4.06 (1H, d, $J_{6',6''}$ = 8.5 Hz, $H_{6'}$ Tp), 4.03 (2H, m, $H_{5'}H_{5''}$ Tp), 3.96 (1H, d, $J_{6'',6'}$ = 8.5 Hz, $H_{6''}$ Tp), 1.87 (3H, br s, CH_3 Tp), 1.82 (3H, br s, CH_3 pT). ^{13}C NMR (D_2O, 150.9 MHz): δ 167.6 (C_4 Tp), 167.3 (C_4 pT), 153.2 (C_2 pT), 152.3 (C_2 Tp), 137.8 (C_6 pT), 137.1 (C_6 Tp), 113.8 (C_5 Tp), 112.0 (C_5 pT), 89.7 (d, J_{C-P} = 8.7 Hz, $C_{4'}$ pT), 89.0 ($C_{1'}$ pT), 88.1 ($C_{1'}$ Tp), 86.8 ($C_{3'}$ pT), 86.0 (d, J_{C-P} = 10.2 Hz, $C_{4'}$ Tp), 79.4 ($C_{6'}$ pT), 79.0 ($C_{2'}$ Tp), 75.6 ($C_{2'}$ pT), 73.4 (d, J_{C-P} = 5.2 Hz, $C_{3'}$ Tp), 72.9 ($C_{6'}$ Tp), 65.7 (d, J_{C-P} = 4.9 Hz, $C_{5'}$ pT), 57.3 ($C_{5'}$ Tp), 13.0 (CH_3 Tp), 12.6 (CH_3 pT). ^{31}P NMR (D_2O, 202.5 MHz): δ -1.8 ppm. HRMS ($(M+Na)^+$, MeOH): calc. for $C_{22}H_{27}N_4O_{14}PNa$ 625.1159, found 625.1162.

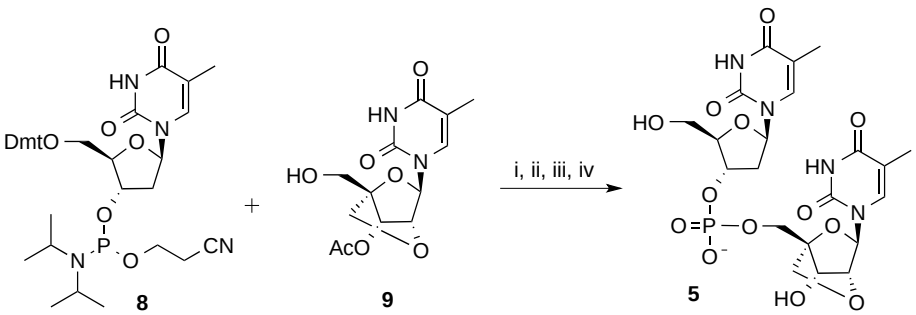


Conditions: (i) 5-Ethylthiotetrazole, CH_3CN; (ii) I_2, THF/H_2O/2,6-lutidine; (iii) conc. aqueous NH_4OH; (iv) 80% aqueous AcOH.

$T_{LN}pT$ (4). To an anhydrous acetonitrile solution (3.7 mL) of phosphoramidite **6**[1] (218 mg, 0.28 mmol) and 3'-*O*-acetylthymidine **7**[2] (96 mg, 0.34 mmol) under argon was added 5-ethylthiotetrazole (121 mg, 0.93 mmol). The mixture was stirred for 35 min at room temperature. A 0.2 M iodine solution (107 mg in 2.1 mL THF/H_2O/2,6-lutidine (2/1/1)) was then added. After 55 min of stirring at room temperature, a saturated aqueous solution of $Na_2S_2O_3$ was added until discoloration. The solution was diluted with CH_2Cl_2 (9 mL) and washed with water (9 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated. The residue was dissolved in conc. aqueous NH_4OH (3.6 mL) and stirred at room temperature overnight. The solution was concentrated and the residue

dissolved in 80% aqueous acetic acid (2.9 mL). The resulting solution was stirred at room temperature for 4 h and concentrated under reduced pressure. Water (10 mL) and CH_2Cl_2 (10 mL) were added and the aqueous phase was separated, washed with CH_2Cl_2 (2 x 10 mL), concentrated in vacuo then purified by HPLC to give **4** after lyophilization (22 mg, 0.038 mmol, 14% based on **7**). ^{1}H NMR (D_2O, 600 MHz): δ 7.81 (1H, d, $J_{6,Me}$ = 1.2 Hz, H_6 pT), 7.63 (1H, d, $J_{6,Me}$ = 1.2 Hz, H_6 Tp), 6.24 (1H, t, $J_{1',2''} = J_{1',2'}$ = 6.0 Hz, $H_{1'}$ pT), 5.60 (1H, s, $H_{1'}$ Tp), 4.74 (1H, s, $H_{2'}$ Tp), 4.49 (1H, q, $J_{3',4'} = J_{3',2'} = J_{3',2''}$ = 6.0 Hz, $H_{3'}$ pT), 4.40 (1H, d, $J_{3',P}$ = 6.1 Hz, $H_{3'}$ Tp), 4.19 (1H, ddd, $J_{5',4'}$ = 2.6 Hz, $J_{5',P}$ = 3.9 Hz, $J_{5',5''}$ = 11.5 Hz, $H_{5'}$ pT), 4.11 (1H, m, $H_{4'}$ pT), 4.07 (1H, d, $J_{6',6''}$ = 8.4 Hz, $H_{6'}$ Tp), 4.06 (1H, m, $H_{5''}$ pT), 4.03 (2H, m, $H_{5'}H_{5''}$ Tp), 3.97 (1H, d, $J_{6'',6'}$ = 8.4 Hz, $H_{6''}$ Tp), 2.39 (1H, td, $J_{2'',1'} = J_{2'',3'}$ = 6.0 Hz, $J_{2'',2'}$ = 13.7 Hz, $H_{2''}$ pT), 2.28 (1H, td, $J_{2',3'} = J_{2',1'}$ = 6.0 Hz, $J_{2',2''}$ = 13.7 Hz, $H_{2'}$ pT), 1.85 (3H, br s, CH_3 Tp), 1.82 (3H, br s, CH_3 pT). ^{13}C NMR (D_2O, 150.9 MHz): δ 166.0 (C_4 Tp), 166.0 (C_4 pT), 151.3 (C_2 Tp), 150.7 (C_2 pT), 136.5 (C_6 pT), 135.6 (C_6 Tp), 111.2 (C_5 Tp), 110.6 (C_5 pT), 88.6 (d, J_{C-P} = 8.3 Hz, $C_{4'}$ Tp), 86.9 ($C_{1'}$ Tp), 85.1 ($C_{1'}$ pT), 84.7 (d, J_{C-P} = 9.1 Hz, $C_{4'}$ pT), 77.8 ($C_{2'}$ Tp), 71.9 (d, J_{C-P} = 5.0 Hz, $C_{3'}$ Tp), 71.6 ($C_{6'}$ Tp), 69.1 ($C_{3'}$ pT), 63.9 (d, J_{C-P} = 5.1 Hz, $C_{5'}$ pT), 56.0 ($C_{5'}$ Tp), 39.2 ($C_{2'}$ pT), 11.7 (CH_3 Tp), 11.3 (CH_3 pT). ^{31}P NMR (D_2O, 202.5 MHz): δ -1.5 ppm. HRMS ($(M+Na)^+$, MeOH): calc. for $C_{21}H_{27}N_4O_{13}PNa$ 597.1210, found 597.1205.

2′-*O*-4′-*C*-methylene-2′-*O*-acetyl-5-methyluridine (9). 2′-*O*-4′-*C*-methylene-5′-*O*-dimethoxytrityl-5-methyluridine[1] (180 mg, 0.31 mmol) was dissolved in anhydrous pyridine (250 µL). Acetic anhydride (150 µL, 0.57 mmol) was added to the solution. The mixture was stirred at room temperature for 2 h and diluted with CH_2Cl_2 (3 mL) then washed with brine (2 x 3 mL). The organic layer was dried over anhydrous $MgSO_4$ and concentrated in vacuo. The residue was dissolved with anhydrous CH_2Cl_2 (900 µL). Trifluoroacetic acid (97 µL, 1.10 mmol) was added. The mixture was stirred at room temperature for 1 h then MeOH was dropped into until discoloration. The solution was concentrated and purified by silica gel chromatography using a gradient of methanol in dichloromethane (0 to 2%) to afford **9** (90 mg, 0.29 mmol) in 93% yield. ^{1}H NMR (CD_3OD, 300 MHz): δ 7.75 (1H, q, J = 1.2 Hz), 5.63 (1H, s), 4.87 (1H, s), 4.55 (1H, s), 3.90 (4H, m), 2.11 (3H, s), 1.90 (3H, d, J = 1.2 Hz). ^{13}C NMR (CD_3OD, 75 MHz): δ 171.8, 166.4, 151.9, 136.3, 111.1, 89.6, 88.4, 79.1, 72.9, 72.2, 57.5, 20.5, 12.7. HRMS ($(M+Na)^+$, MeOH): calc. for $C_{13}H_{16}N_2O_7Na$ 335.0855, found 335.0852.



Conditions: (i) 5-Ethylthiotetrazole, CH_3CN; (ii) I_2, THF/H_2O/2,6-lutidine; (iii) conc. aqueous NH_4OH; (iv) 80% aqueous AcOH.

TpT_{LN} (5). Phosphoramidite **8** (153 mg, 0.21 mmol) and alcohol **9** (77 mg, 0.25 mmol) were dissolved in anhydrous CH_3CN (2.2 mL) under argon. 5-Ethylthiotetrazole (88 mg, 0.68 mmol) was added to the solution. The mixture was stirred for 30 min at room temperature. A 0.2 M iodine solution (78 mg in 1.54 mL THF/H_2O/2,6-lutidine (2/1/1)) was then added. After 50 min of stirring at room temperature, a saturated aqueous solution of $Na_2S_2O_3$ was added until discoloration. The solution was diluted with CH_2Cl_2 (7 mL) and washed with water (6 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated. The residue was dissolved in conc. aqueous NH_4OH (2.7 mL) and stirred at room temperature overnight. The solution was concentrated and the residue dissolved in 80% aqueous acetic acid (2.7 mL). The resulting solution was stirred at room temperature for 4 h and concentrated. Water (10 mL) and CH_2Cl_2 (10 mL) were added, and aqueous phase separated, extracted with CH_2Cl_2 (2 x 10 mL), concentrated and purified by HPLC. The interest fractions were lyophilized to give **5** (24 mg, 0.042 mmol, 20% based on **8**). ^{1}H NMR (D_2O, 600 MHz): δ 7.70 (1H, br d, $J_{6,Me}$ = 1.2 Hz, H_6 Tp), 7.65 (1H, br d, $J_{6,Me}$ = 1.2 Hz, H_6 pT), 6.20 (1H, t, $J_{1',2''} = J_{1',2'}$ = 6.4 Hz, $H_{1'}$ Tp), 5.67 (1H, br s, $H_{1'}$ pT), 4.77 (1H, m, $H_{3'}$ Tp), 4.46 (1H, br s, $H_{2'}$ pT), 4.3 (1H, br s, $H_{3'}$ pT), 4.28 (2H, m, $H_{5'}H_{5''}$ pT), 4.20 (1H, q, $J_{4',5'} = J_{4',3'} = J_{4',5''}$ = 4.3 Hz, $H_{4'}$ Tp), 4.05 (1H, d, $J_{6',6''}$ = 8.5 Hz, $H_{6'}$ pT), 3.96 (1H, d, $J_{6'',6'}$ = 8.5 Hz, $H_{6''}$ pT), 3.84 (1H, m, $H_{5'}H_{5''}$ Tp), 2.55 (2H, m, $H_{2'}H_{2''}$ Tp), 1.88 (3H, d, $J_{Me,6}$ = 1.2 Hz, CH_3 pT), 1.85 (3H, d, $J_{Me,6}$ = 1.2 Hz, CH_3 Tp). ^{13}C NMR (D_2O, 150.9 MHz): δ 167.6 (C_4 pT), 167.5 (C_4 Tp), 152.7 (C_2 Tp), 152.2 (C_2 pT), 138.5 (C_6 Tp), 137.2 (C_6 pT), 112.4 (C_5 Tp), 111.9 (C_5 pT), 88.7 (d, J_{C-P} = 8.9 Hz, $C_{4'}$ pT), 87.0 ($C_{1'}$ pT), 86.7 (d, J_{C-P} = 6.8 Hz, $C_{4'}$ Tp), 86.3 ($C_{1'}$ Tp), 80.3 ($C_{2'}$ pT), 75.5 (d, J_{C-P} = 5.3 Hz, $C_{3'}$ Tp), 72.3 ($C_{6'}$ pT), 70.3 ($C_{3'}$ pT), 61.9 ($C_{5'}$ Tp), 61.4 (d, J_{C-P} = 4.8 Hz, $C_{5'}$ pT), 38.9 (d, J_{C-P} = 2.2 Hz, $C_{2'}$ Tp), 13.1 (CH_3 pT), 12.8 (CH_3 Tp). ^{31}P NMR (D_2O, 202.5 MHz): δ -0.9 ppm. HRMS ($(M+Na)^+$, MeOH): calc. for $C_{21}H_{27}N_4O_{13}PNa$ 597.1210, found 597.1214.

NMR spectra.

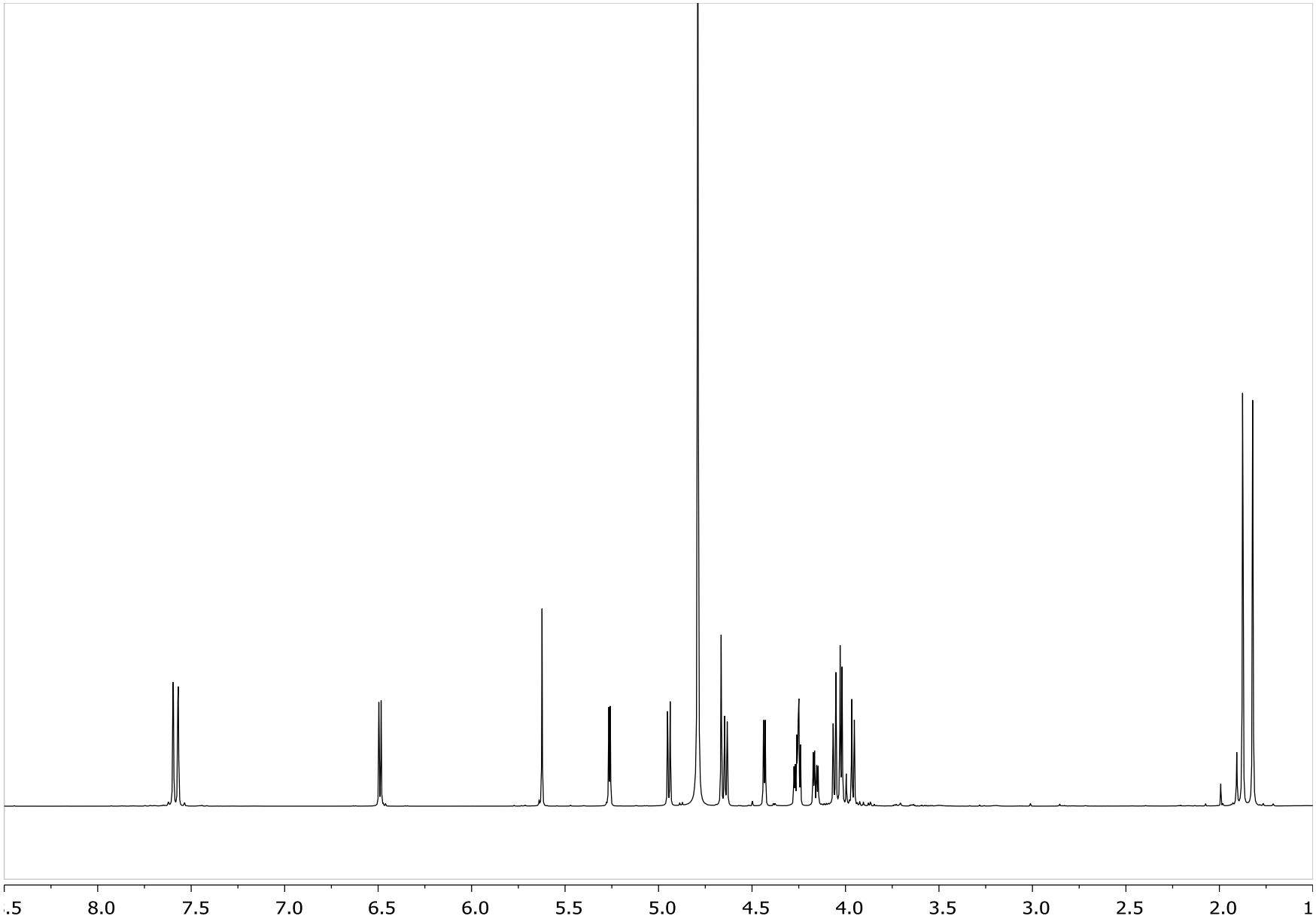


Figure S7: ^{1}H NMR spectrum of **3** (600 MHz, D_2O).

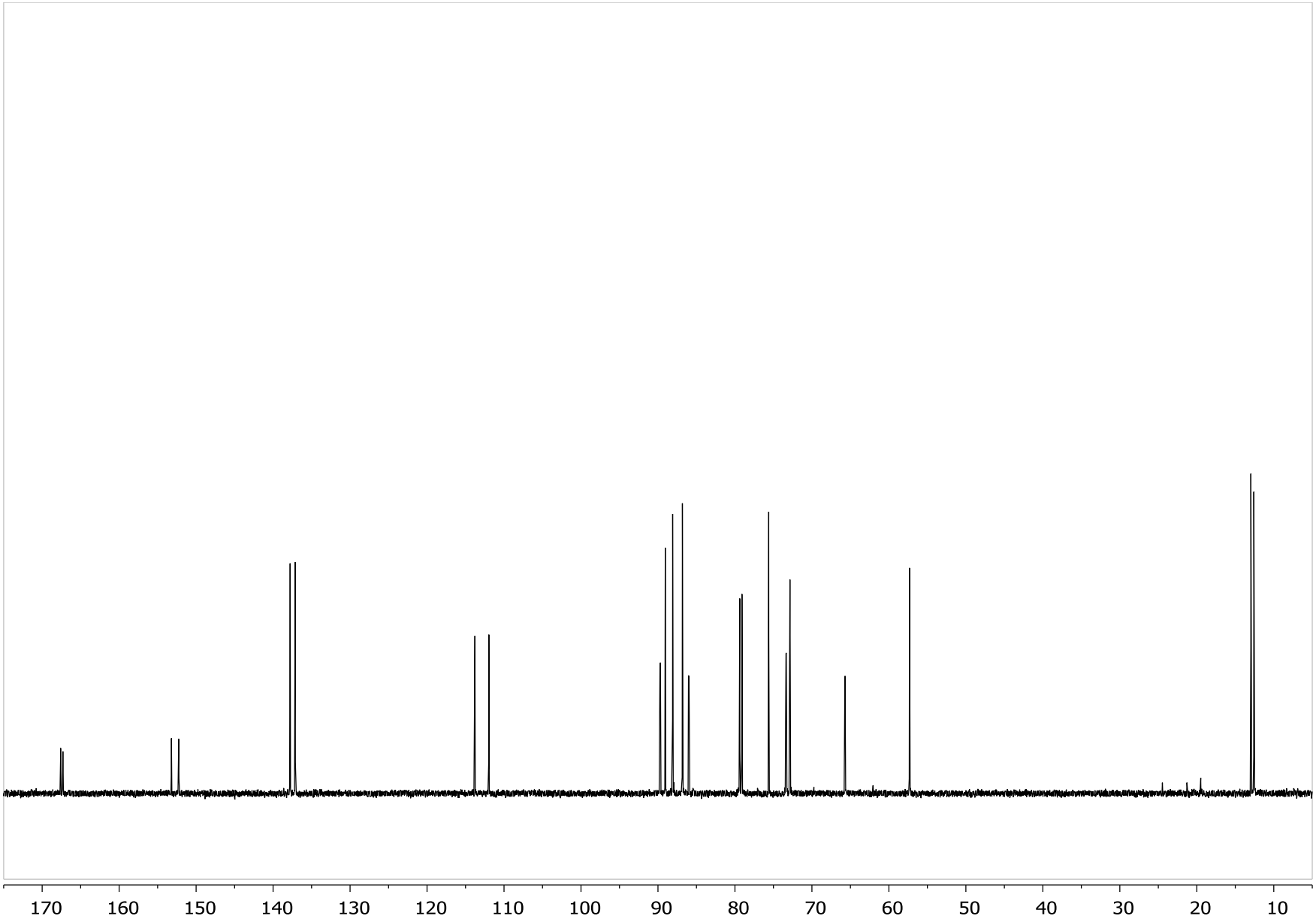


Figure S8: ^{13}C NMR spectrum of **3** (150 MHz, D_2O).

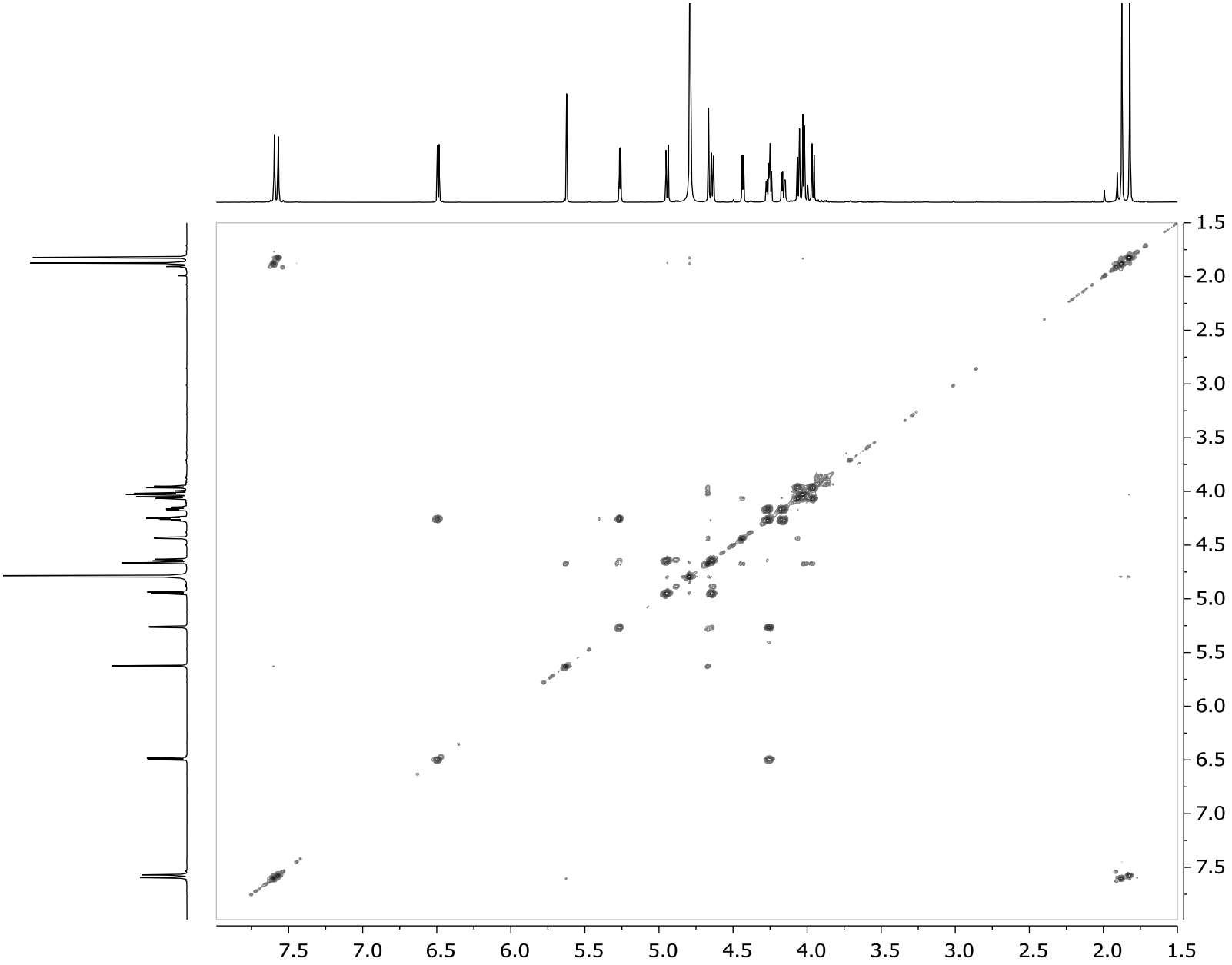


Figure S9: COSY spectrum of **3** (600 MHz, D_2O).

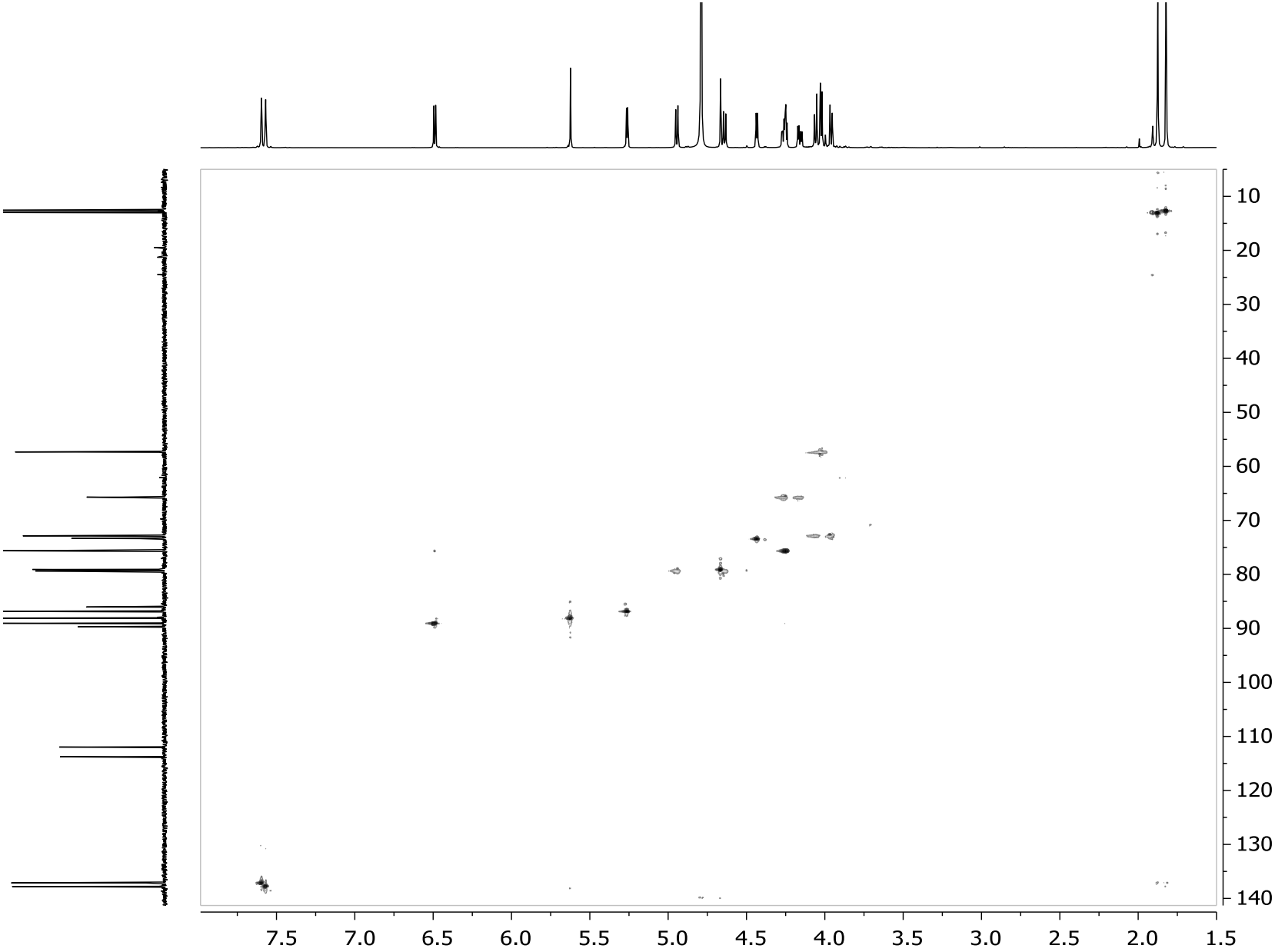


Figure S10: HSQC spectrum of **3** (600 MHz, D_2O).

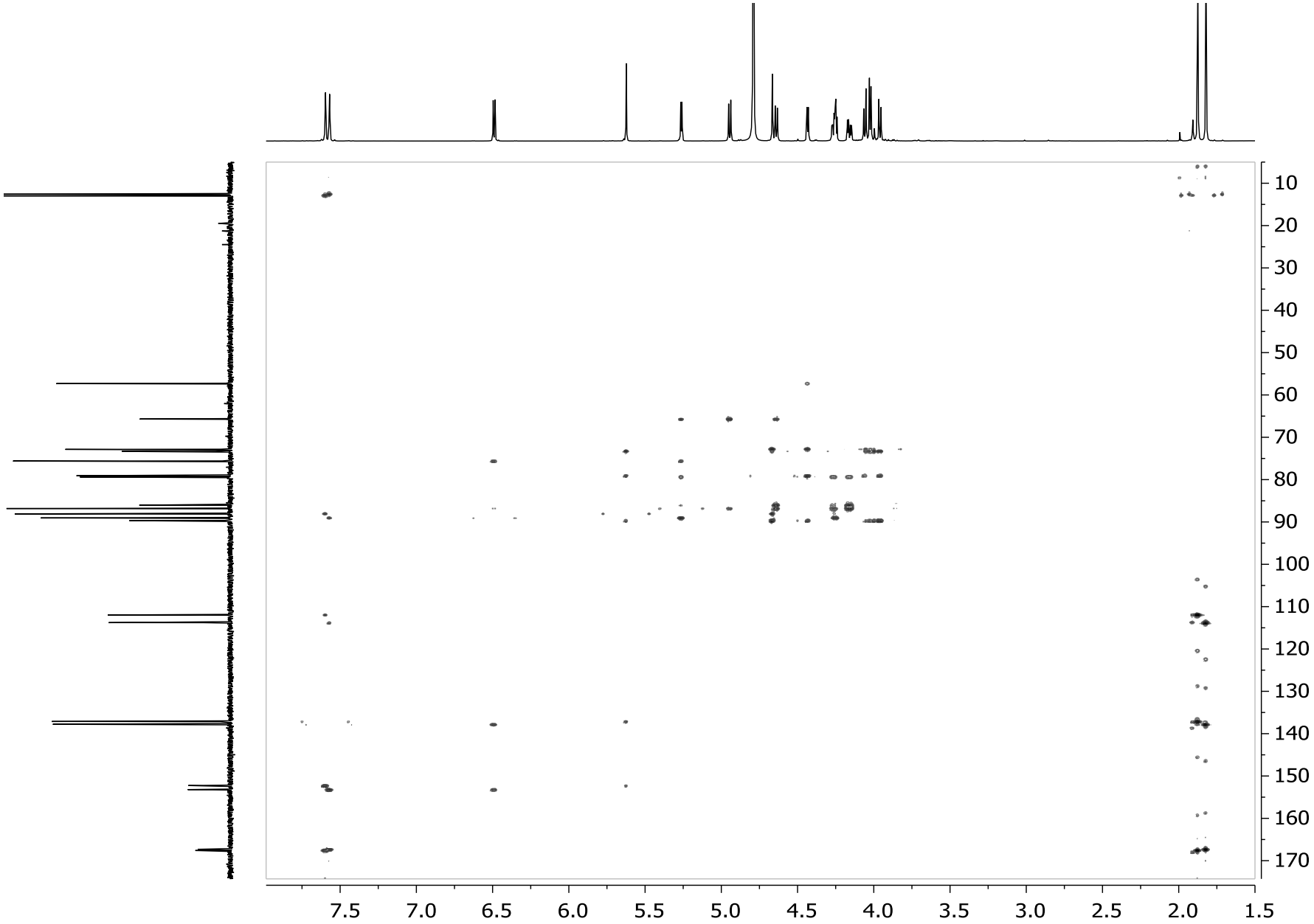


Figure S11: HMBC spectrum of **3** (600 MHz, D_2O).

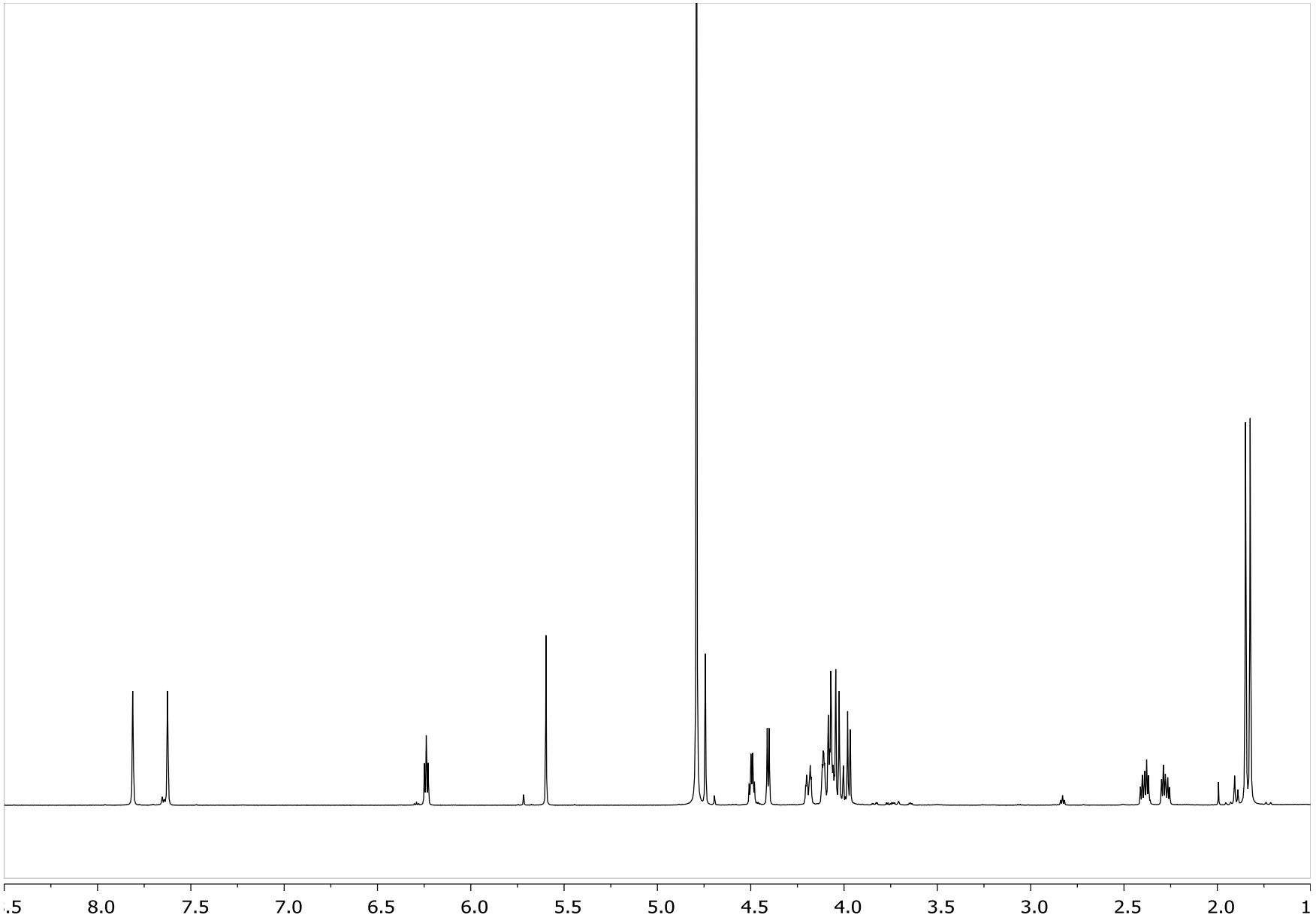


Figure S12: ^{1}H NMR spectrum of **4** (600 MHz, D_2O).

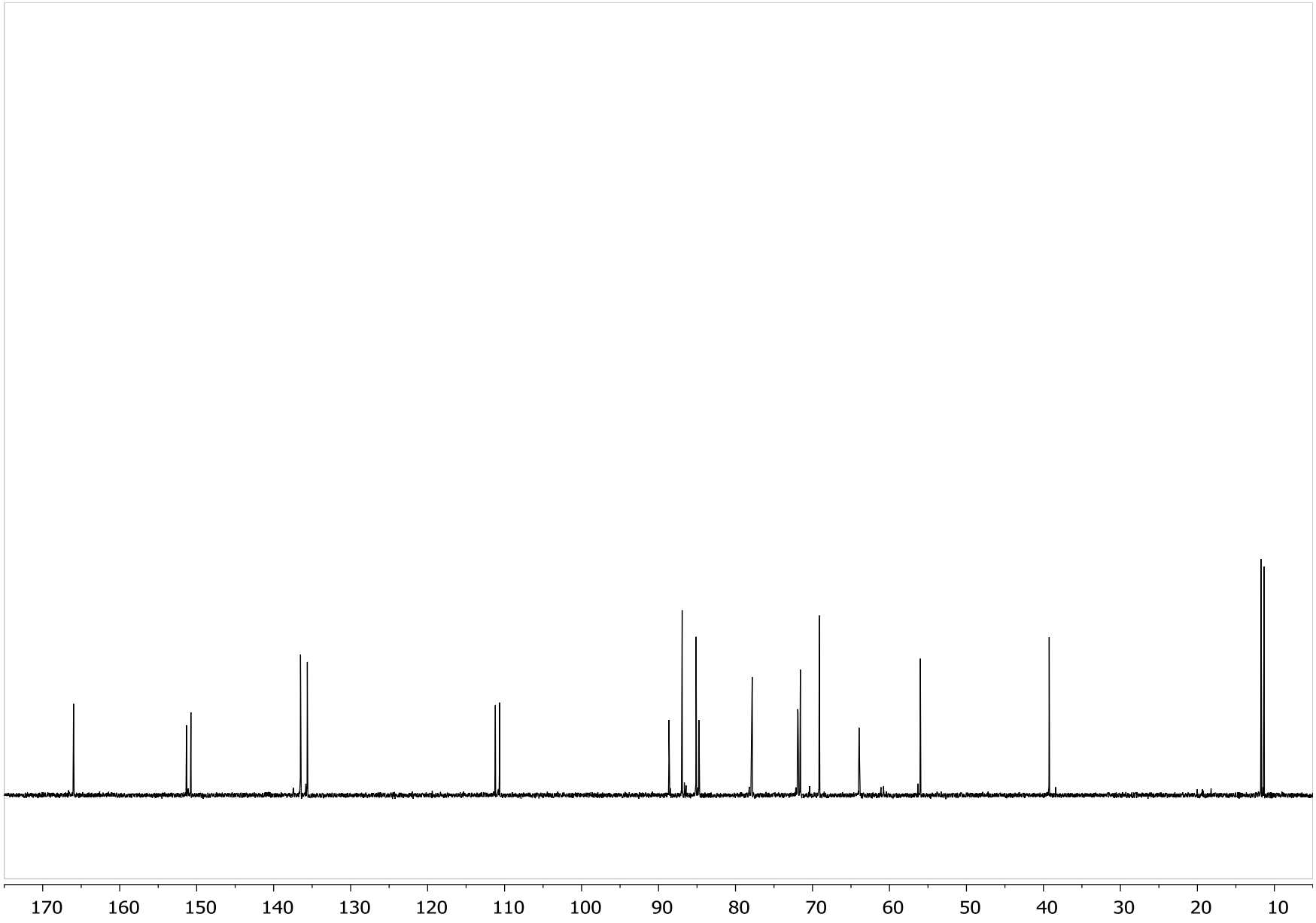


Figure S13: ^{13}C NMR spectrum of **4** (150 MHz, D_2O).

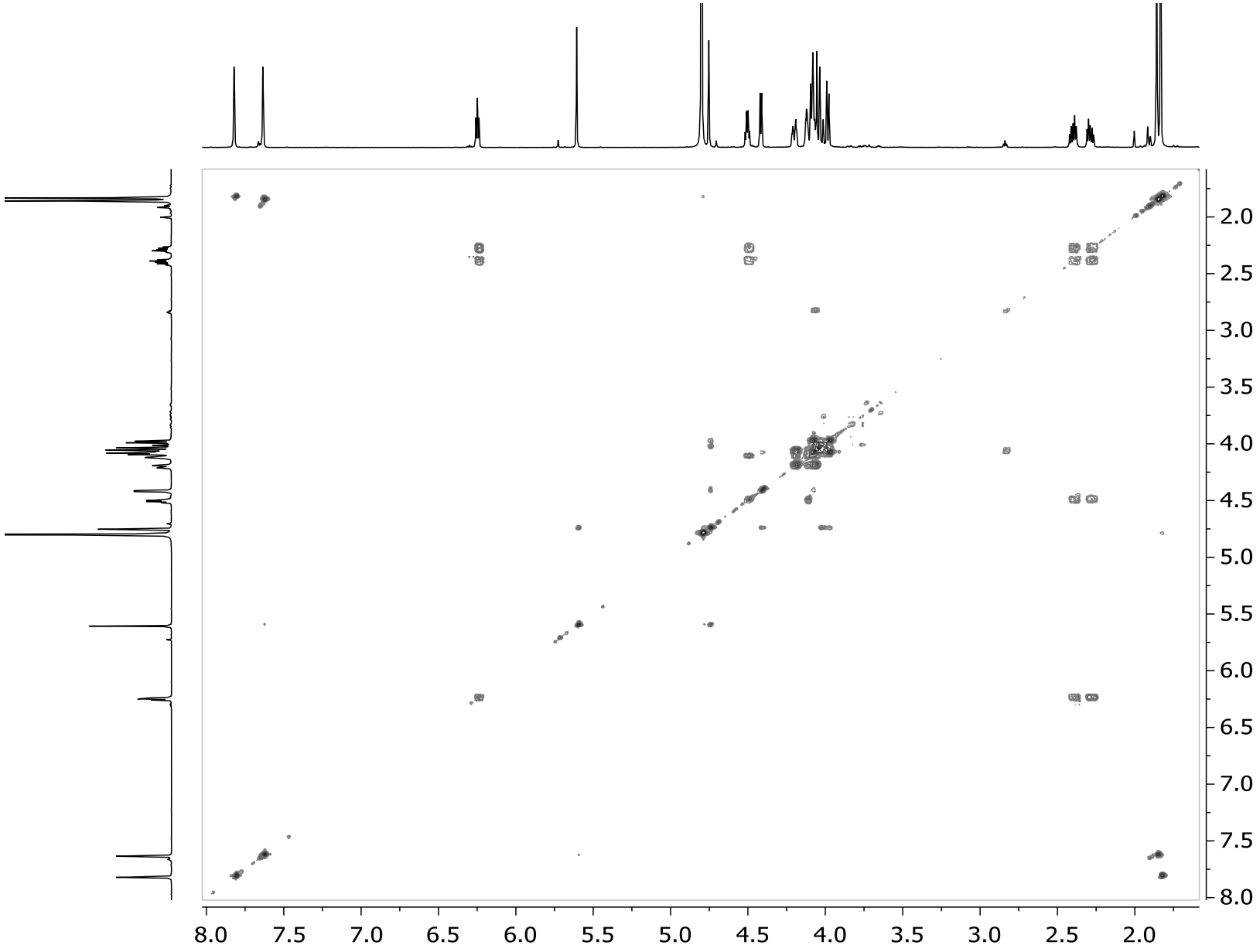


Figure S14: COSY spectrum of **4** (600 MHz, D_2O).

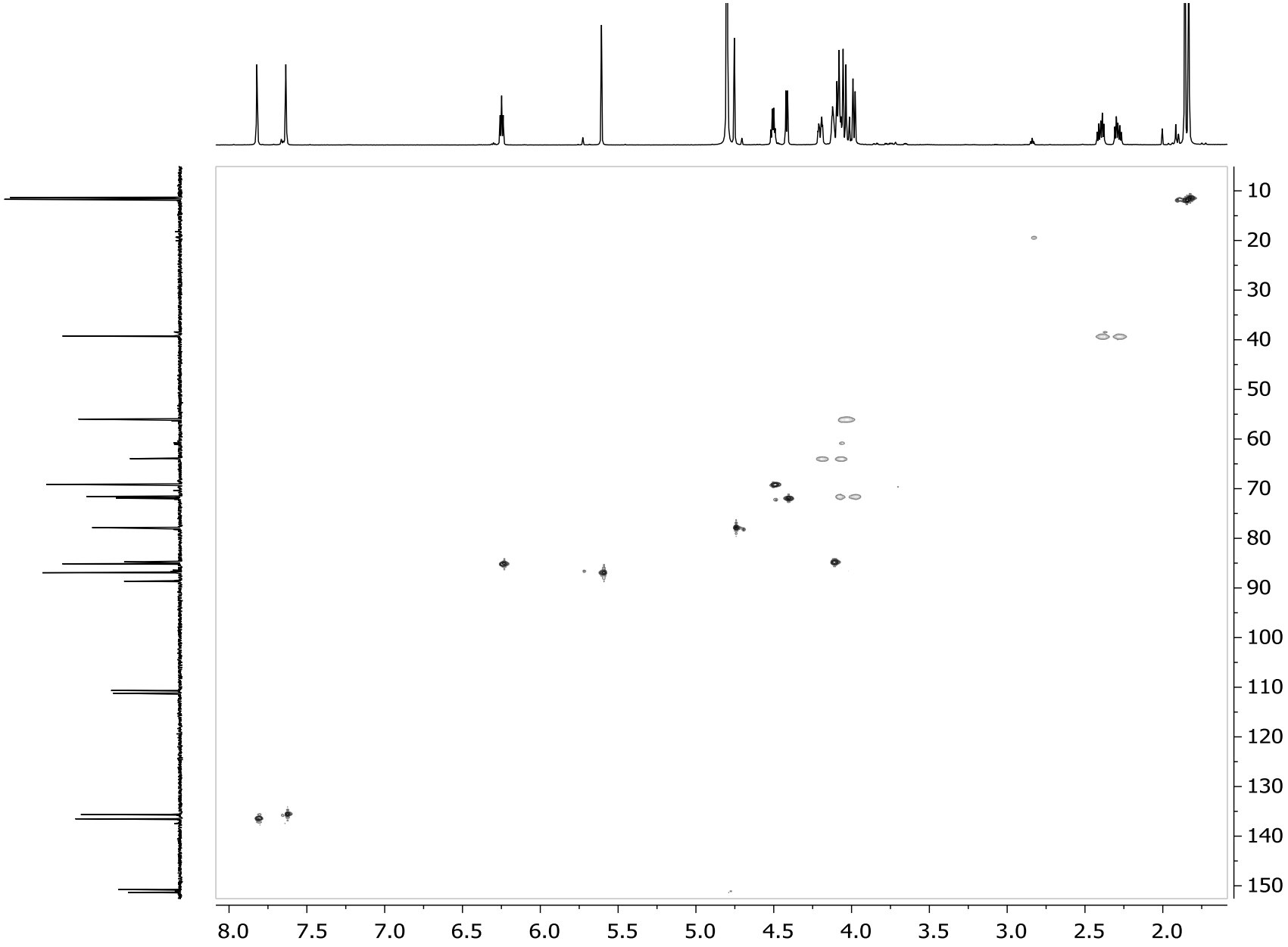


Figure S15: HSQC spectrum of **4** (600 MHz, D_2O).

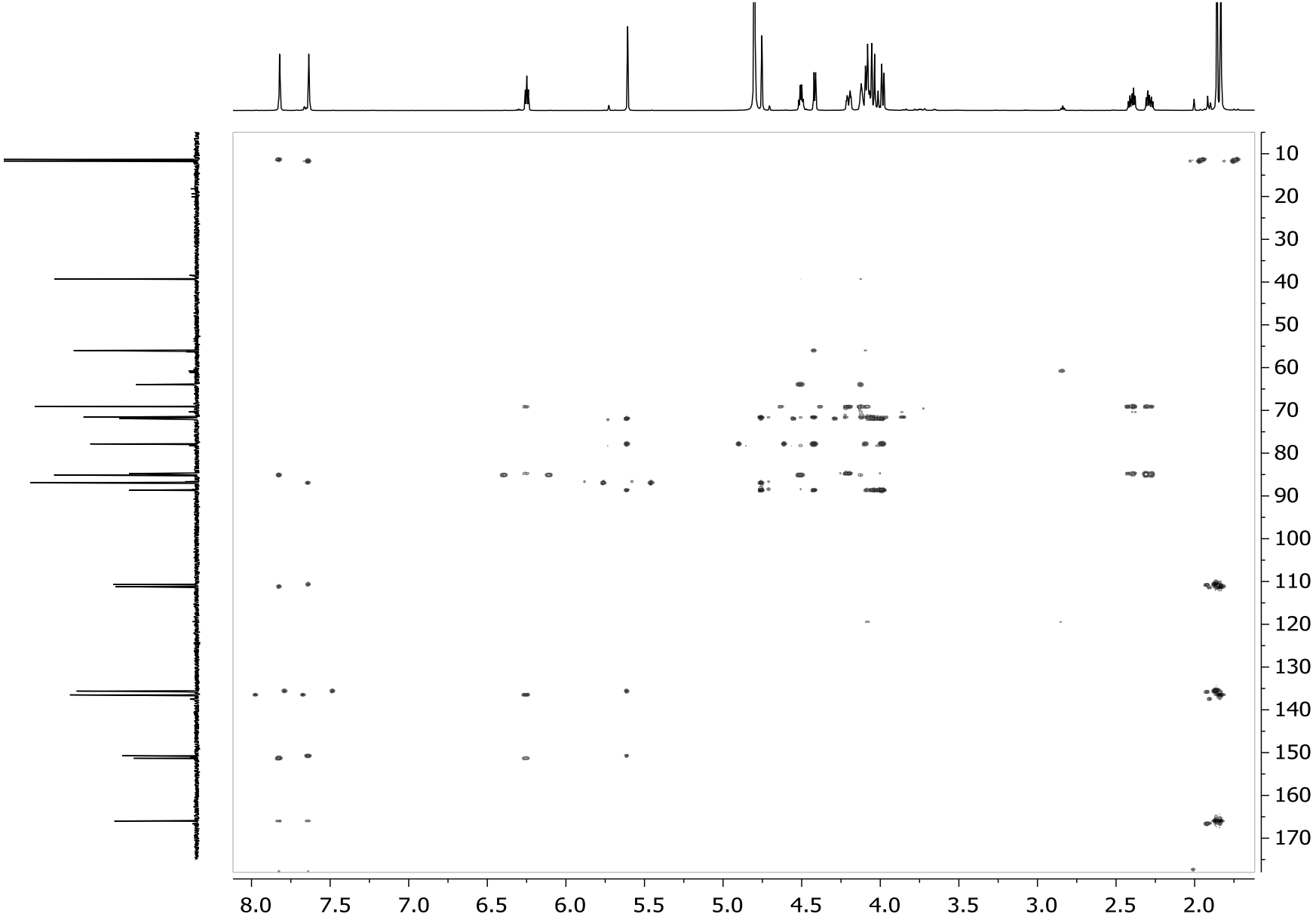


Figure S16: HMBC spectrum of **4** (600 MHz, D_2O)

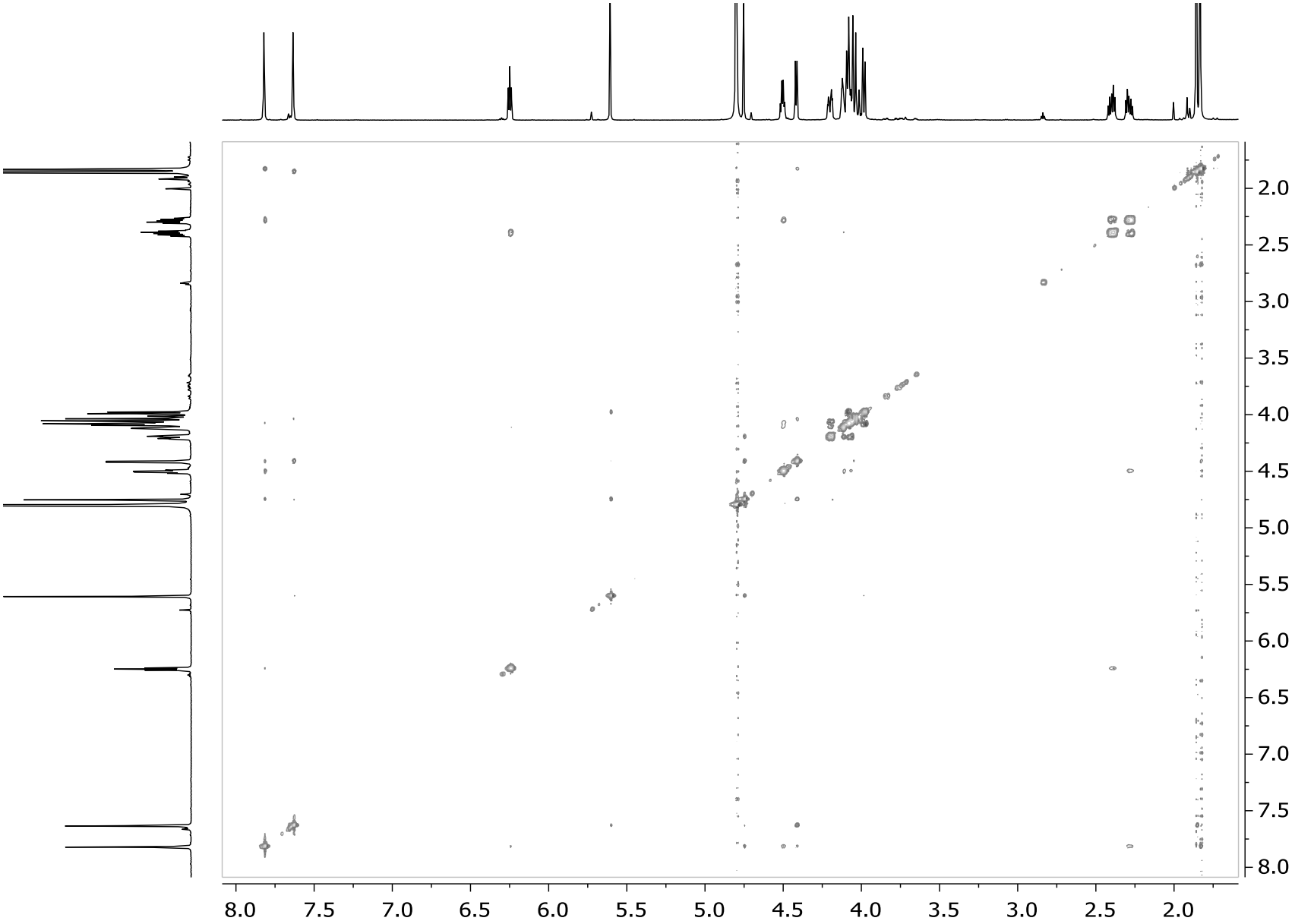


Figure S17: NOESY spectrum of **4** (600 MHz, D_2O, mixing time 800 ms).

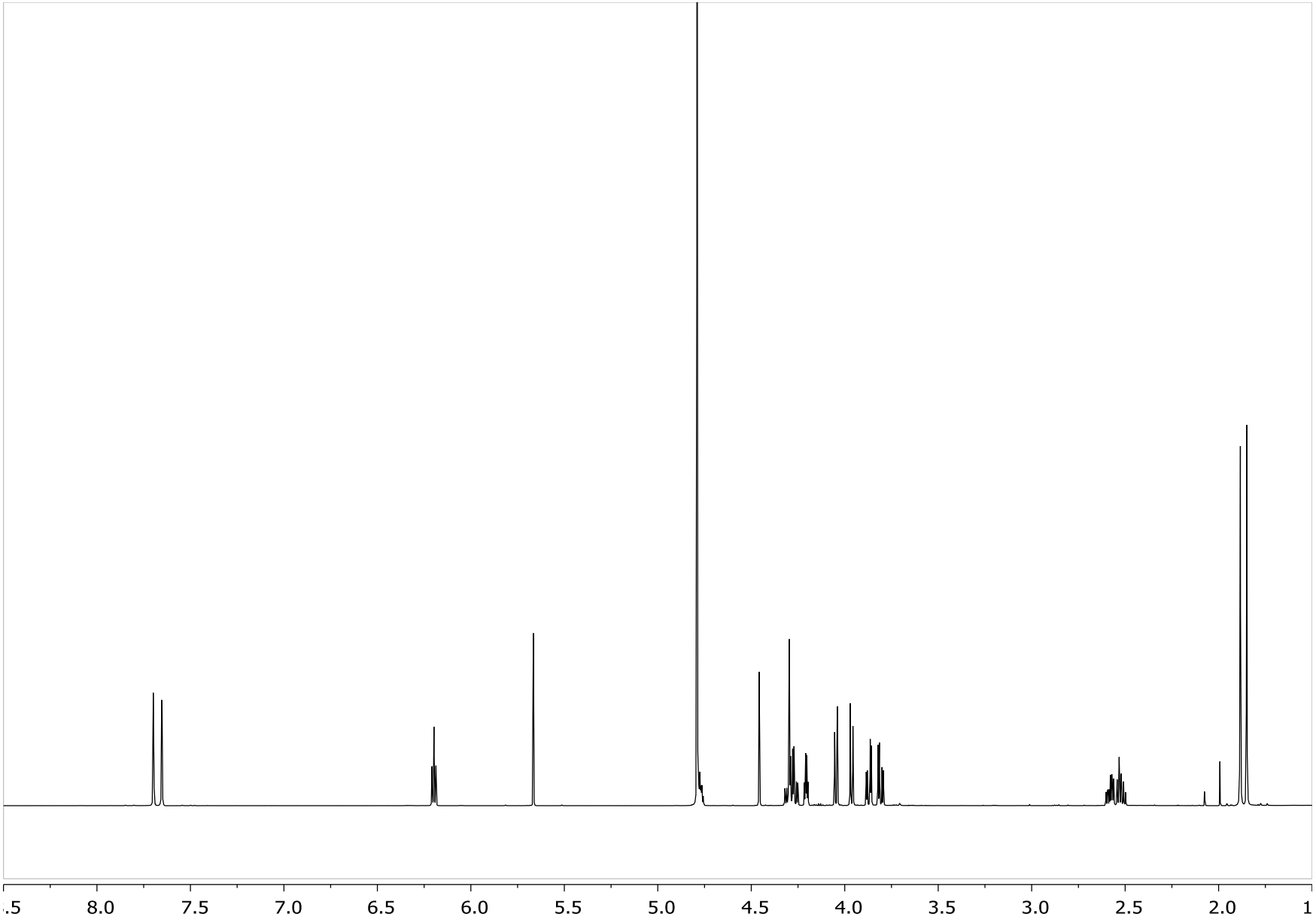


Figure S18: ^{1}H NMR spectrum of **5** (600 MHz, D_2O).

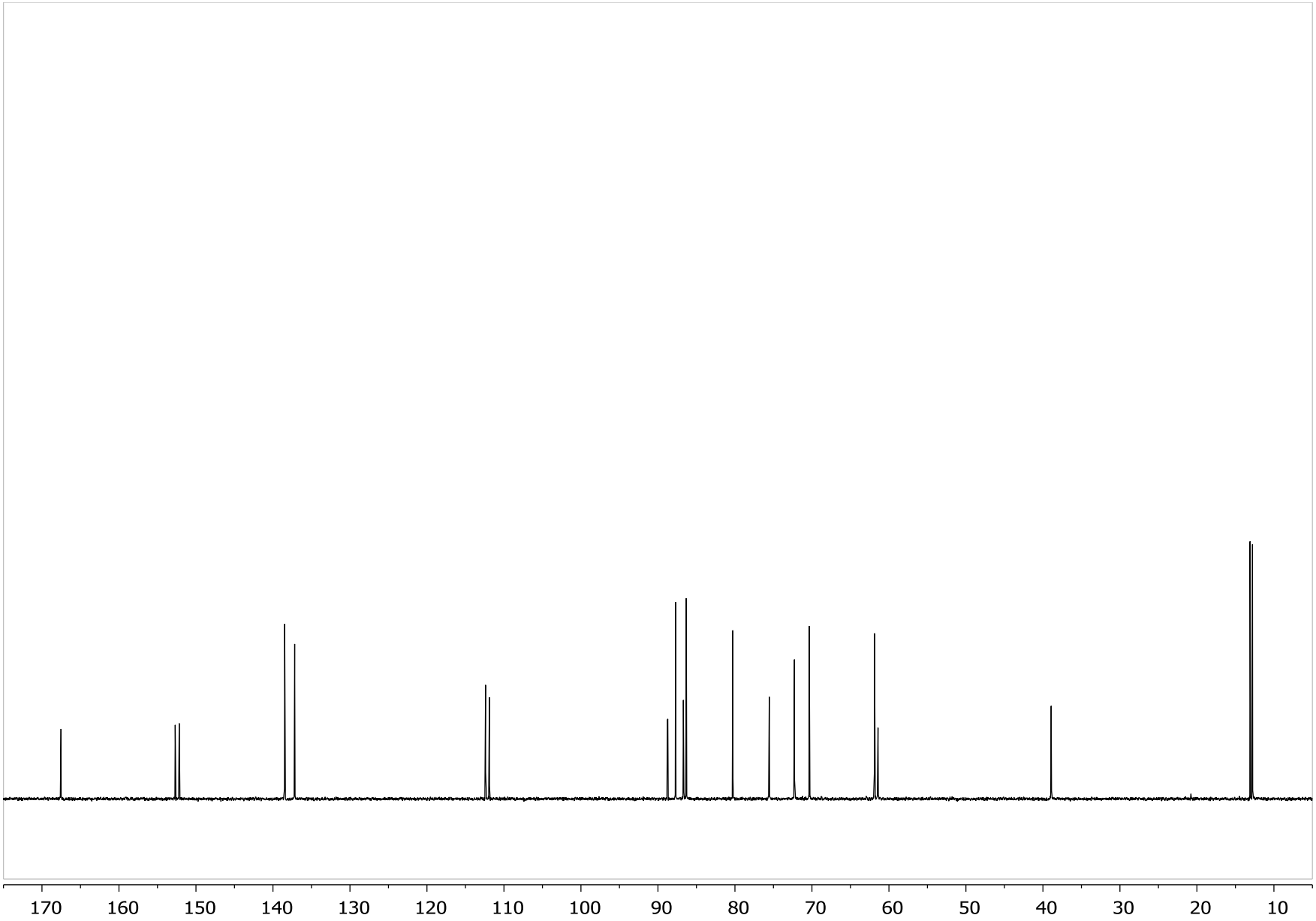


Figure S19: ^{13}C NMR spectrum of **5** (150 MHz, D_2O).

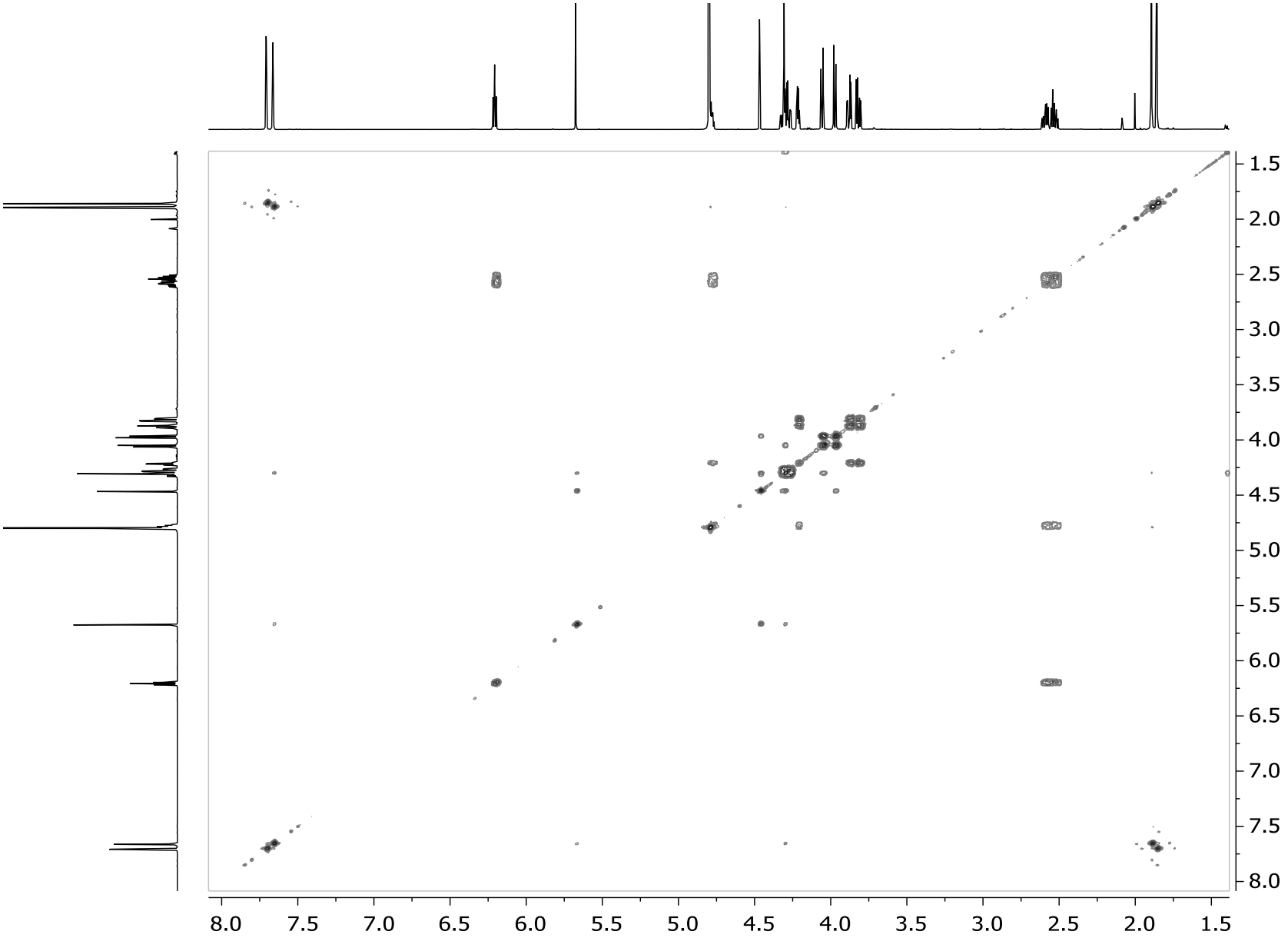


Figure S20: COSY spectrum of **5** (600 MHz, D_2O).

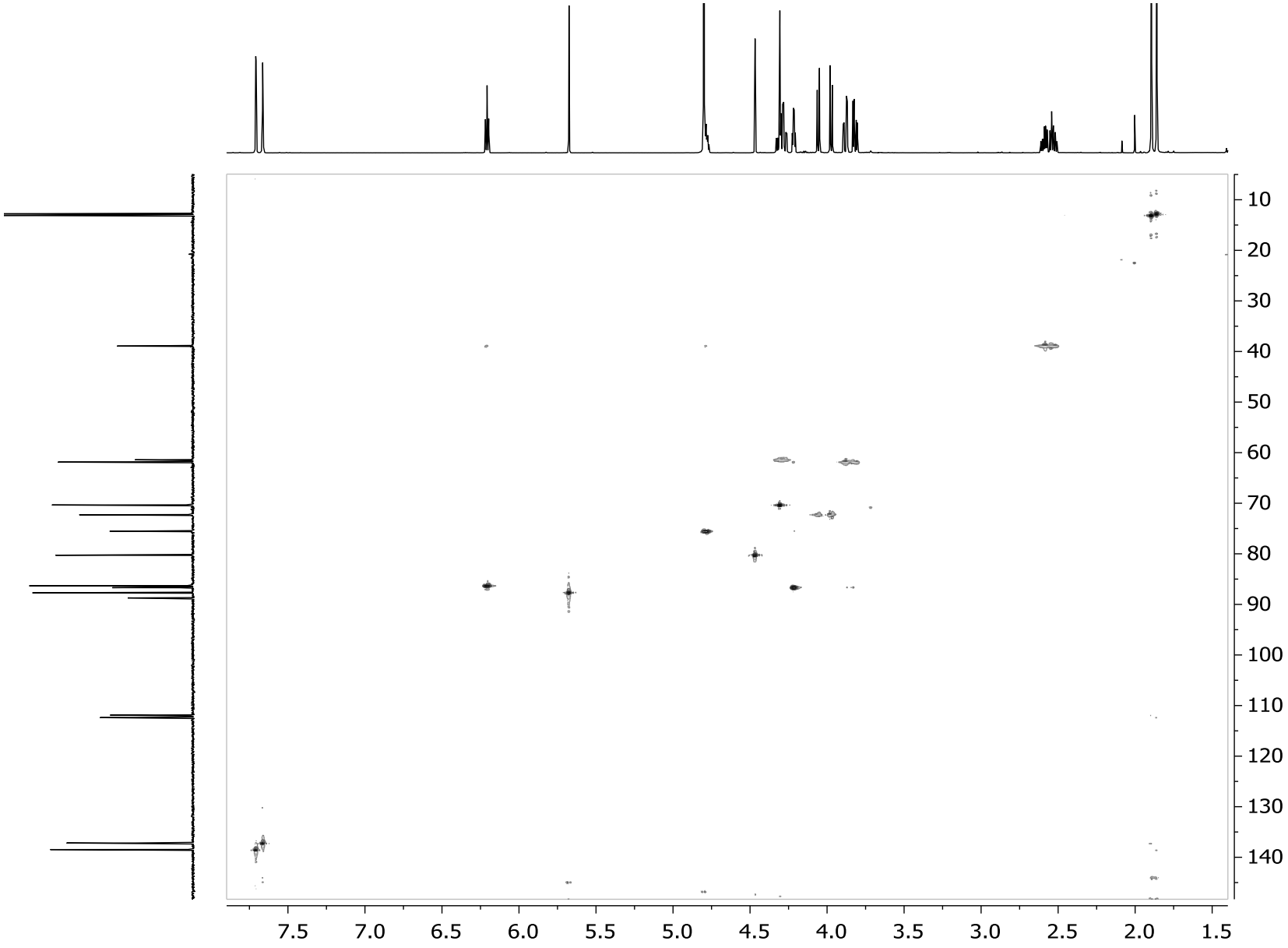


Figure S21: HSQC spectrum of **5** (600 MHz, D_2O).

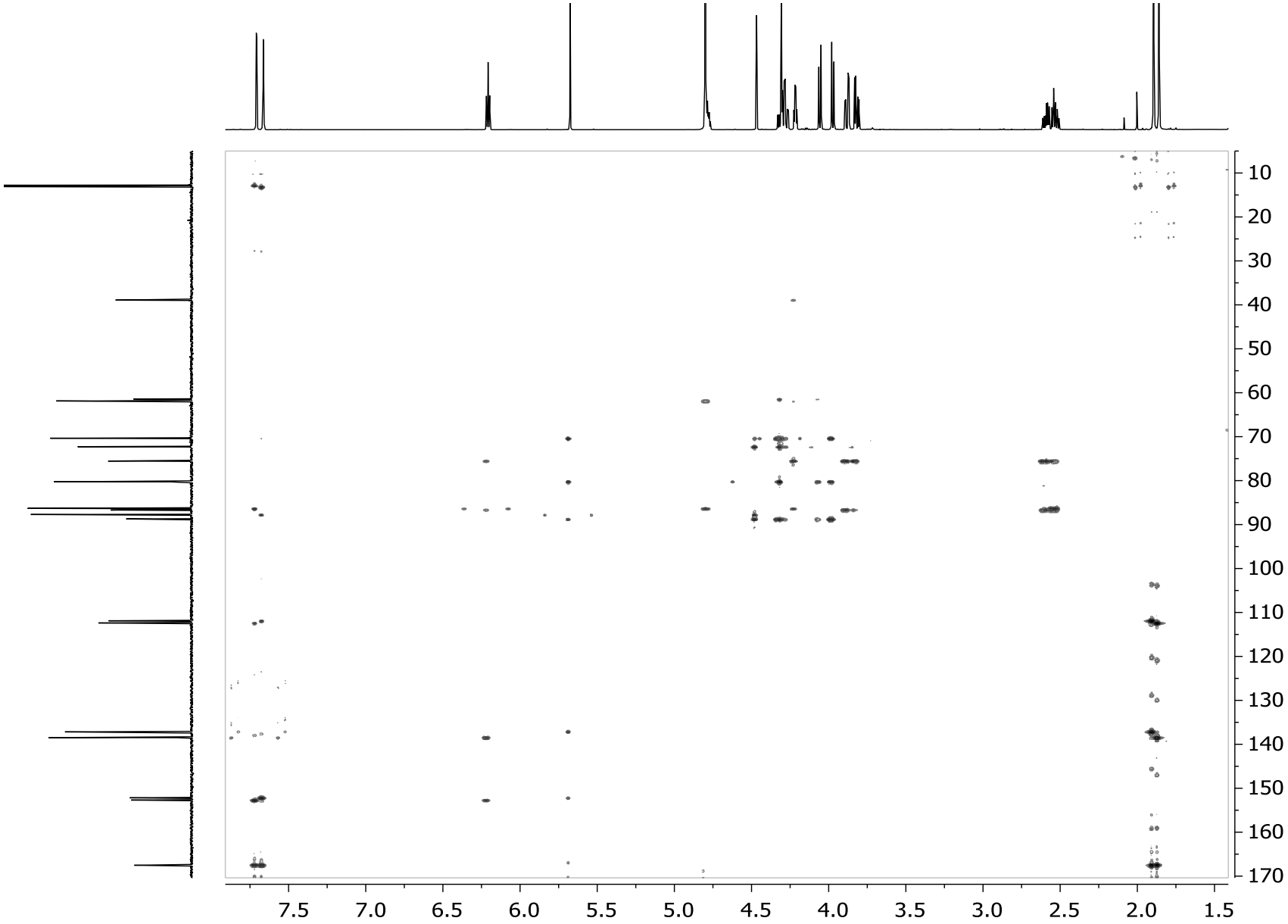


Figure S22: HMBC spectrum of **5** (600 MHz, D_2O).

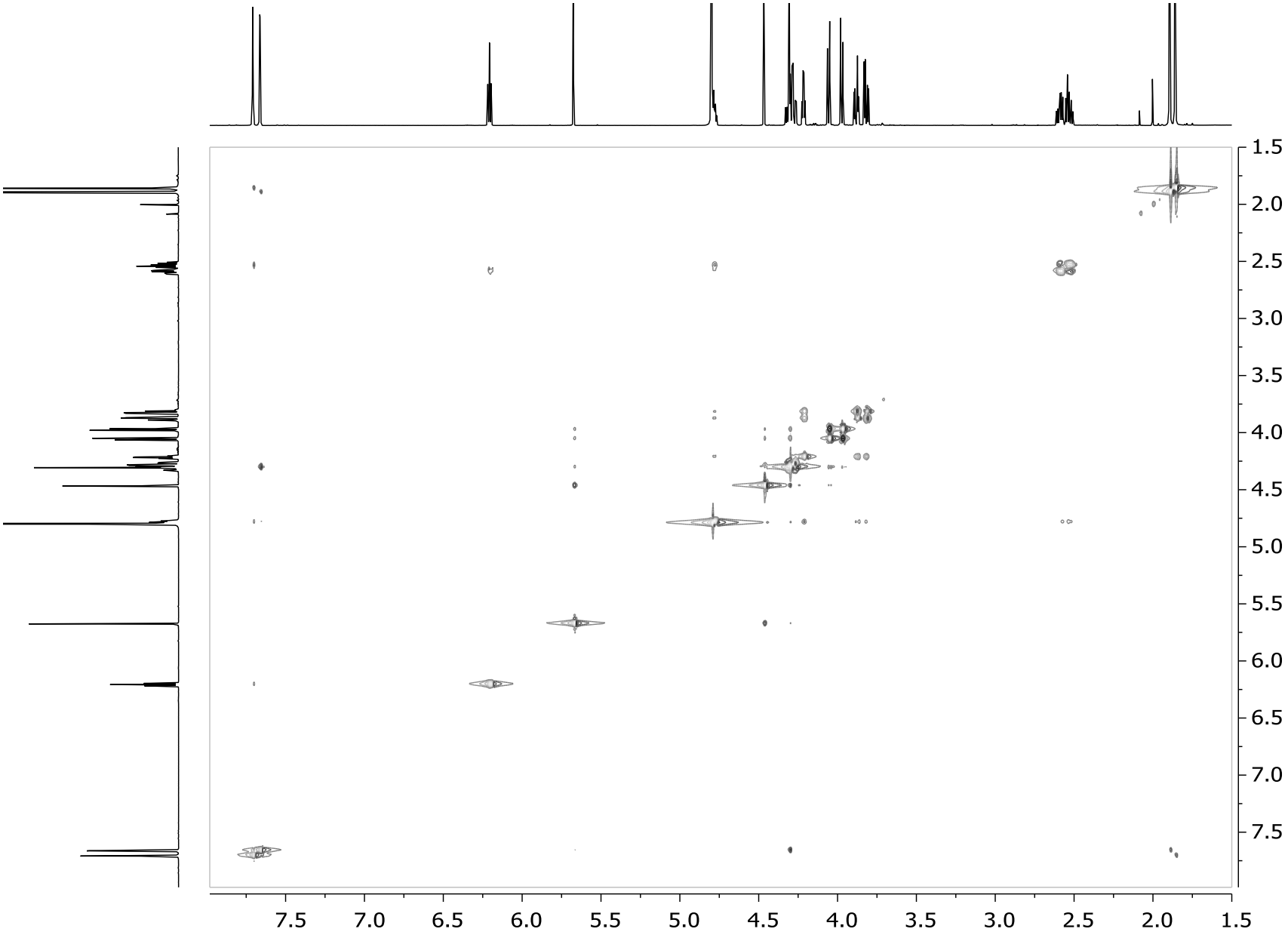


Figure S23: ROESY spectrum of **5** (600 MHz, D_2O, mixing time 800 ms).

CD Experiments

CD spectra were recorded on a spectropolarimeter equipped with a Peltier temperature controller. Spectra were recorded between 10 to 80°C (10°C increment) in a 0.01 M Na phosphate, 0.1 M NaCl, pH 7.0 buffer as previously reported.[3] CD data are expressed in molar ellipticity *per* residue [θ] ($deg.cm^2.decimol^{-1}$). The molar extinction coefficient at 267 nm of **TpT** and of its nucleoside constituent thymidine (**T**) ($2 \times 9.65 \times 10^3$ and 9.65×10^3 $M^{-1}cm^{-1}$, respectively) were used for the dimers, and for their nucleosides constituent. The wavelength was varied from 220 to 330 nm at 50 nm/min. The spectra were collected with a 1 nm bandwith, 1 s response, 0.2 nm data acquisition interval and represented an averaged of three scans. The spectra were corrected by substraction of the background spectrum with buffer. The instrumental parameters for the CD melting experiments were: 10°C data pitch, 5°C/min ramp, 180 s delay after equilibration, ± 0.1°C equilibration tolerance.

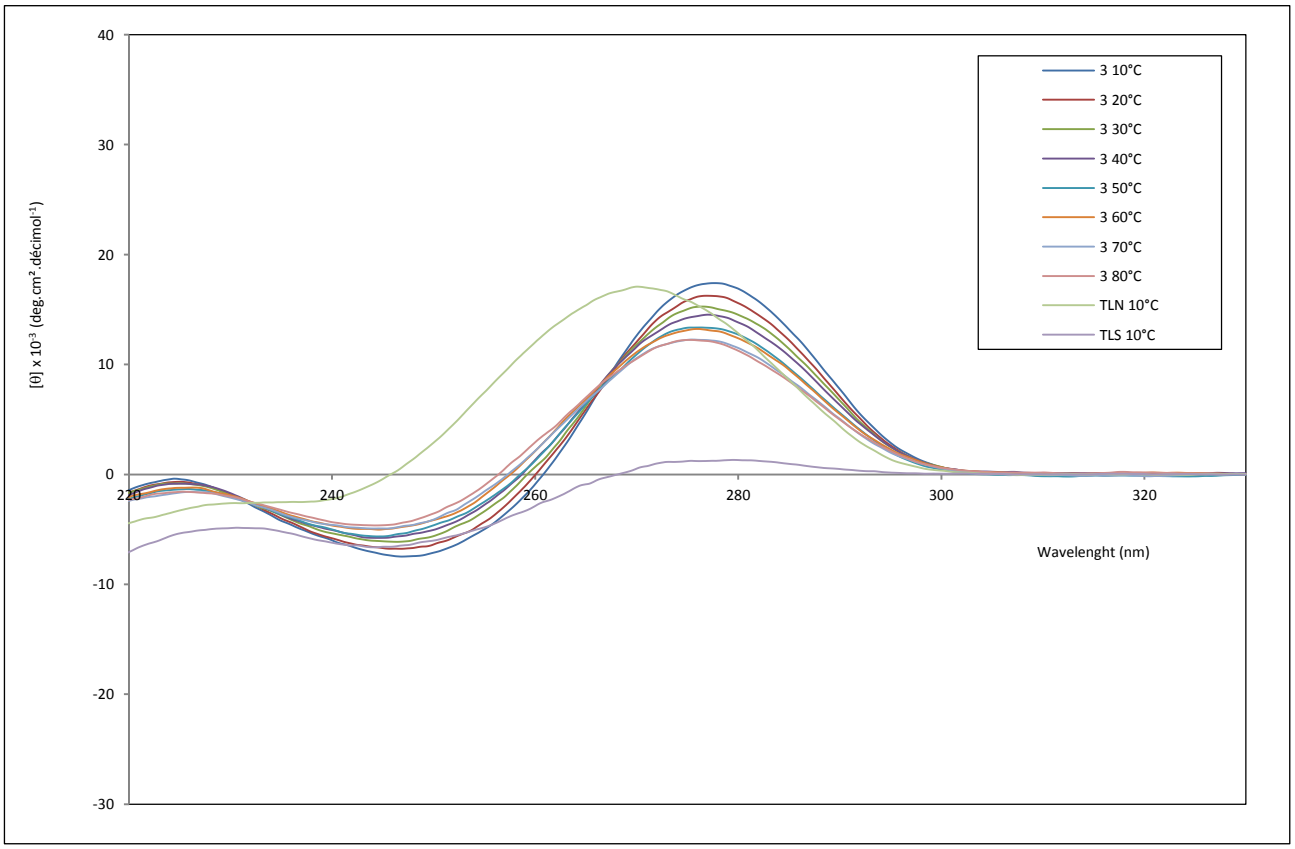


Figure S24. CD spectra of **3** ($\mathbf{T_{LN}pT_{LS}}$) between 10 to 80°C and CD spectra at 10°C of 2′-*O*-4′-*C*-methylene-5-methyluridine ($\mathbf{T_{LN}}$) and 3′-*O*-4′-*C*-methylene-5-methyluridine ($\mathbf{T_{LS}}$).

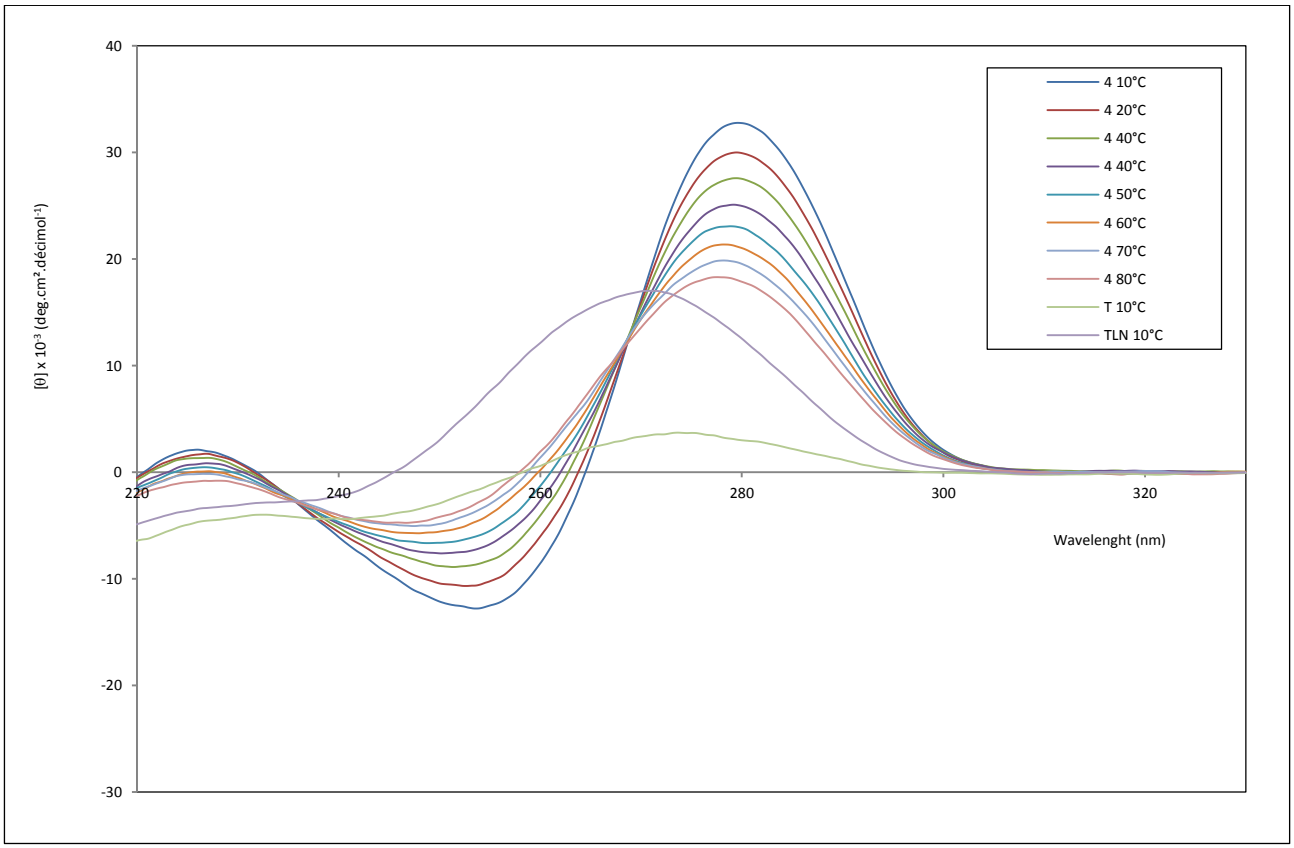


Figure S25. CD spectra of **4** ($\mathbf{T_{LN}pT}$) between 10 to 80°C and CD spectra at 10°C of $\mathbf{T_{LN}}$ and **T**.

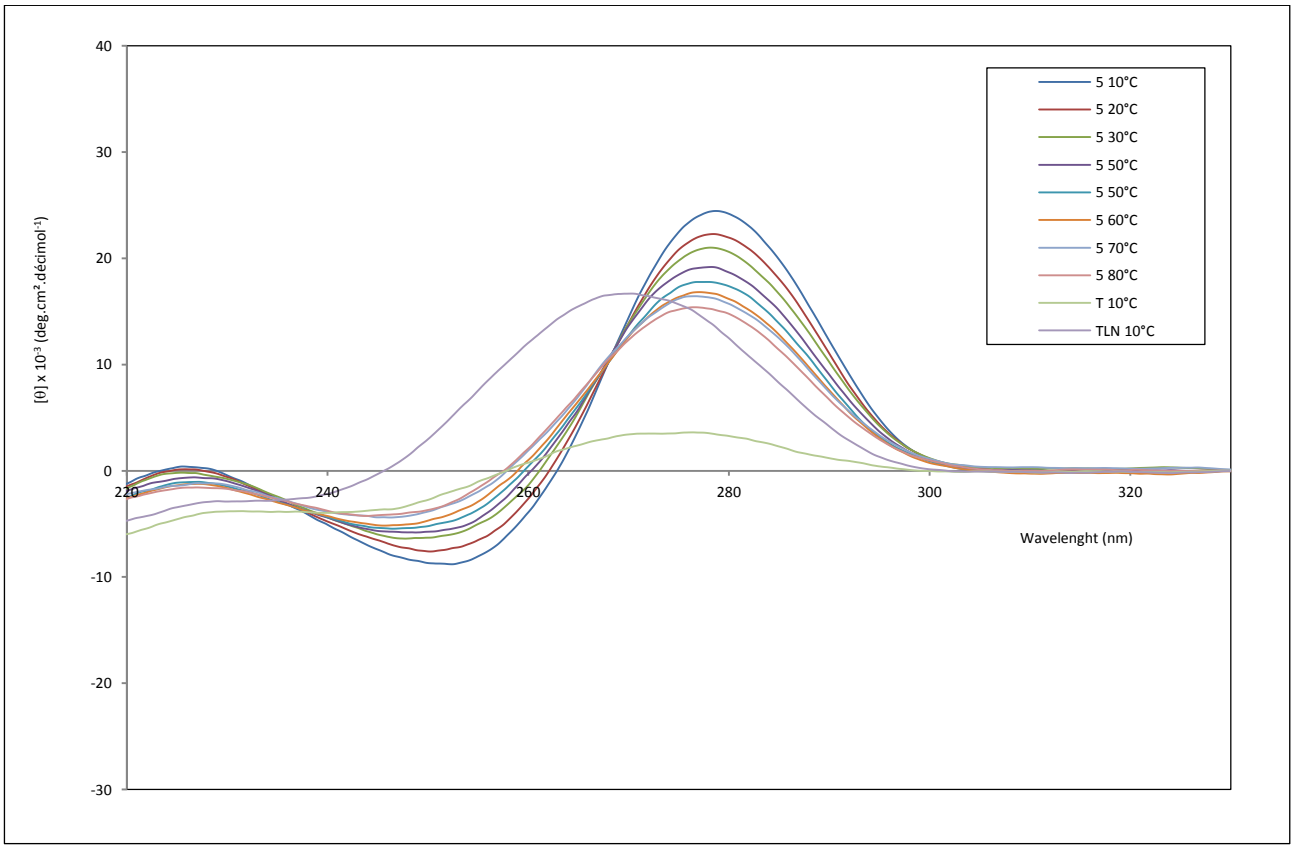


Figure S26. CD spectra of **5** (**TpT$_{LN}$**) between 10 to 80°C and CD spectra at 10°C of **T** and **T$_{LN}$**.

CD Data Analysis by SVD and dinucleotide stacking level

The CD differential spectra versus temperature of **1**-**5** were analyzed by singular value decomposition (SVD) to determine the number of spectral species involved in the equilibrium melting. CD melting data were collected in a matrix **D** (*M* x *N*) consisting of *M* rows of CD differential spectra recorded over in 10°C increments from 10 to 80°C and *N* columns being the number of wavelengths measured in every spectrum (Figure S28). Generated **D** matrices were analyzed by SVD as previously described[4-6] to produce the **S**, **U**, **V** matrices using R software (version 3.1.1). **S** contains the singular values (weights of the component comprising the data set), **U** contains the basis spectra that make up the data set, and **V** contains the amplitudes of each component as a function of temperature. The first and second wavelength basis vectors (V1 and V2, respectively) generated from matrix D for each compounds are represented in Figure S29. The number of significant spectral components contained in **D** was determined by evaluating the relative magnitude of the singular values, their contribution to the total variance of the signal, and the values of the autocorrelation coefficients of the amplitude (**V**) vectors.

Stacking level of compounds **1**-**5** at 298.15K was determined by fitting of the first temperature basis vector from SVD,[5] and the temperature dependence of the Taylor coefficient B[5,6] to the Van't Hoff equation using R software (version 3.1.1).

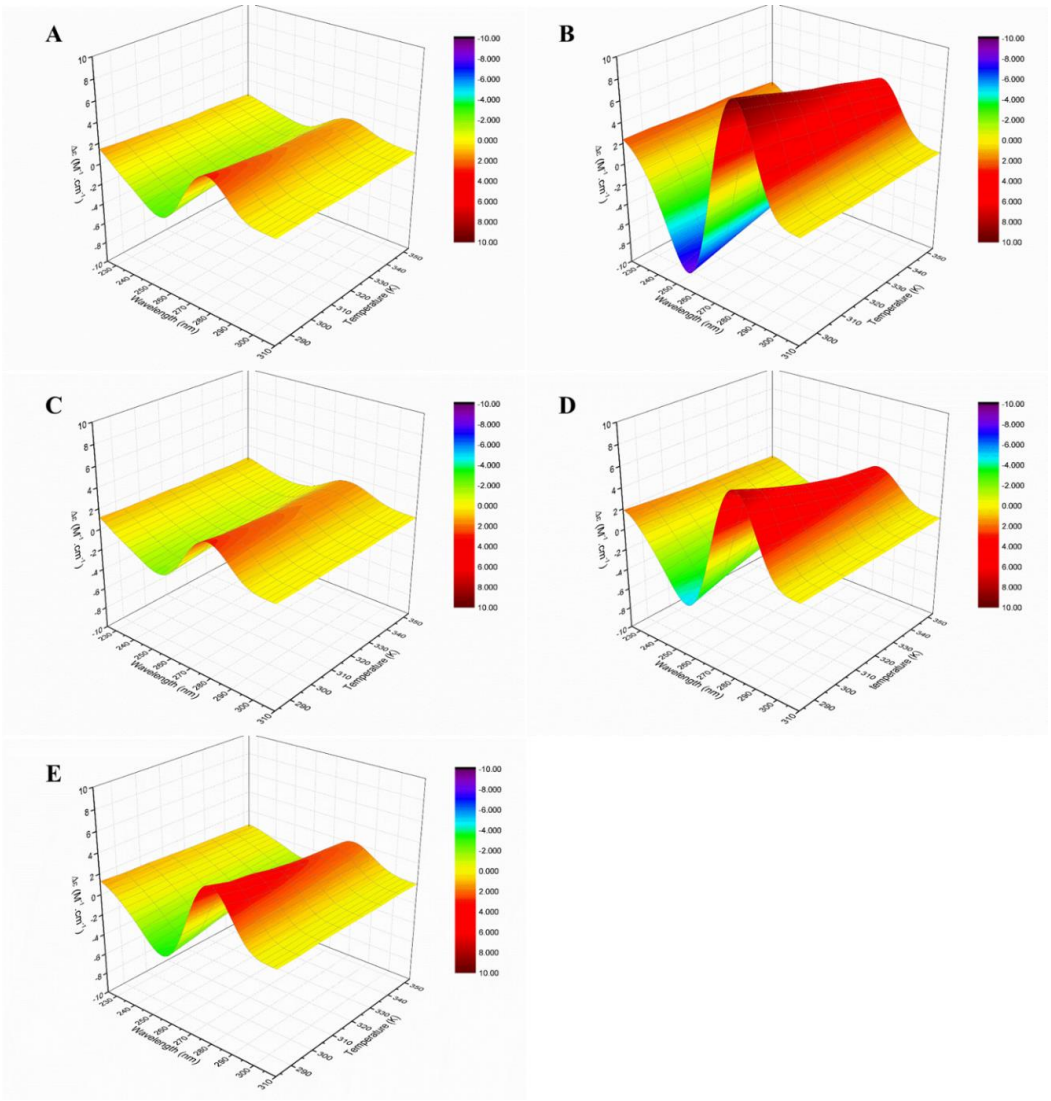


Figure S28. Tridimensional representation of matrix **D** generated from CD differential melting spectra of compound **1** (**TpT**, A), **2** (**T_{LN}pT_{LN}**, B), **3** (**T_{LN}pT_{LS}**, C), **4** (**T_{LN}pT**, D) and **5** (**TpT_{LN}**, E).

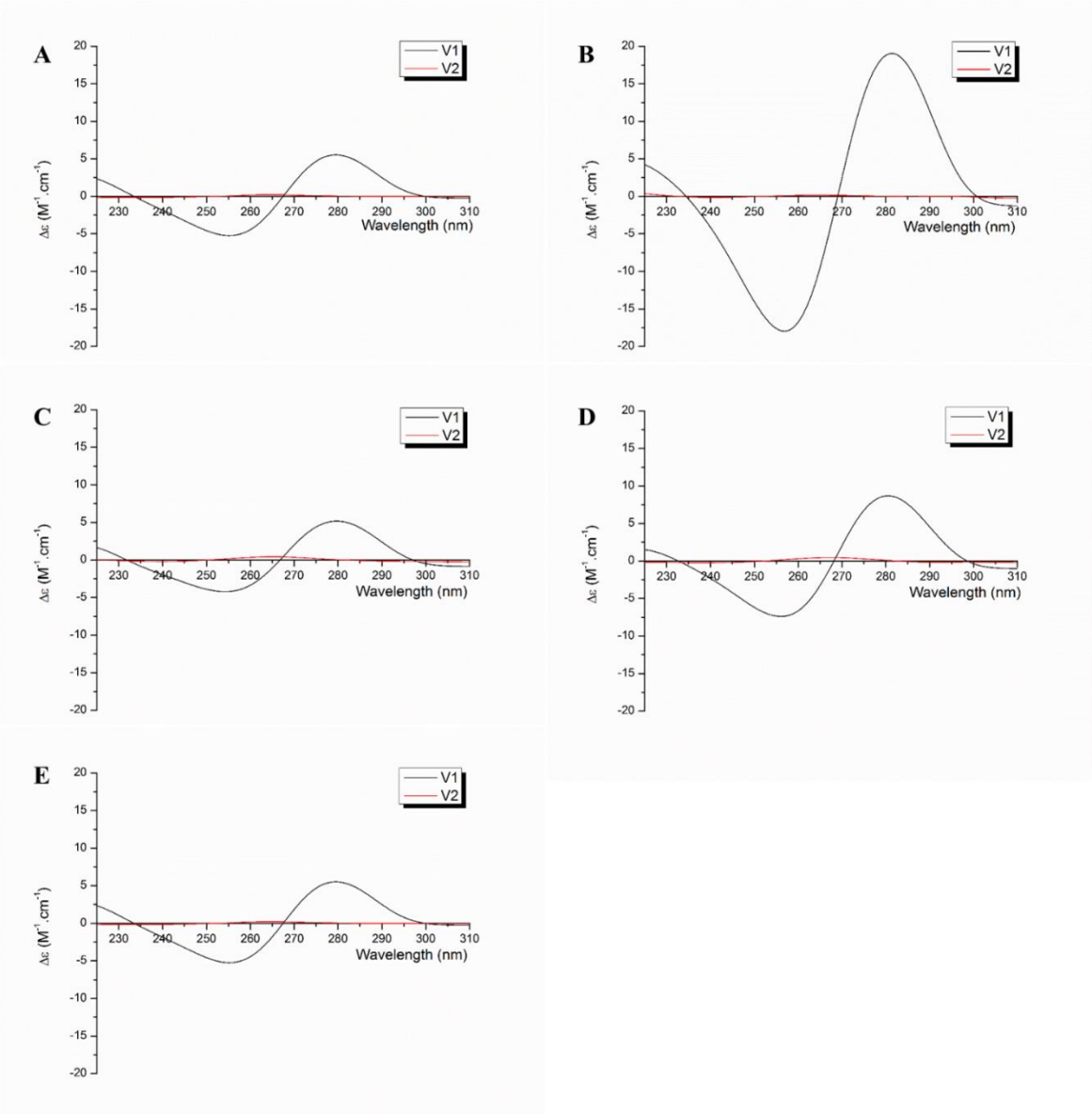


Figure S29. Representation of the first (V1) and second (V2) wavelength basis vectors determined from SVD of CD differential melting spectra of compound **1** (**TpT**, A), **2** (**T_{LN}pT$_{LN}$**, B), **3** (**T_{LN}pT$_{LS}$**, C), **4** (**T_{LN}pT**, D) and **5** (**TpT$_{LN}$**, E).

References

(1) A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607.

(2) A. Marchand, D. Marchand, R. Busson, P. Marliere and P. Herdewijn, *Chemistry and Biodiversity*, 2007, **4**, 1450.

(3) (*a*) C. Desnous, B. R. Babu, C. Moriou, J. U. Ortiz Mayo, A. Favre, J. Wengel and P. Clivio, *J. Am. Chem. Soc.*, 2008, **130**, 30. (*b*) C. Moriou, C. Denhez, O. Plashkevych, S. Coantic-Castex, J. Chattopadhyaya, D. Guillaume and P. Clivio, *J. Org. Chem.*, 2015, **80**, 615.

(4) J. T. Powell, E. G. Richards and W. B. Gratzer, *Biopolymers*, 1972, **11**, 235.

(5) G. C. Causley, P. W. Staskus and W. C. Johnson Jr., *Biopolymers*, 1983, **22**, 945.

(6) H. Kang, P.-J. Chou, W. C. Johnson Jr., D. Weller, S.-B. Huang and J. E. Summerton, *Biopolymers*, 1992, **32**, 1351.