Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

Remarkable UV Light Invulnerability of Thymine GNA Dinucleotides

Stéphanie Coantic Castex, Agathe Martinez, Dominique Harakat, Dominique Guillaume and 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. Acetonitrile and dichloromethane were dried by distillation from calcium hydride. Anhydrous pyridine was purchased from Aldrich and used without further purification. Laboratory flasks and small devices were dried before distillation or synthetic reactions. Reactions were carried out under an atmosphere of nitrogen. Phosphoramidites and 3'-acetates were dried overnight at room temperature in a desiccator over P_2O_5 prior to use.

Chromatography was performed on silica gel 60 Geduran® Merck, particle size 35-70 μm, unless otherwise stated. NMR spectra were recorded in D_2O. Observed chemical shift (δ) values are given in ppm and coupling constants (J) in Hz. ^{1}H NMR and ^{13}C NMR spectra were recorded on 300 or 600 MHz spectrometers. The nomenclature T_Gp- and -pT_G represent the 3'-end and the 2'-end GNA nucleoside residues, respectively, of **2** and **5**. ^{1}H NMR chemical shifts were calibrated using the residual HDO signal at δ_H 4.79. ^{13}C NMR spectra were calibrated from dioxane (δ_C 67.8 ppm). In *S* glycol residues, H1' and H1" were assigned to pro-S and pro-R, respectively. In *R* glycol residues, H1' and H1" were assigned to pro-R and pro-S, respectively. Prochiral H3' and H3" protons have not been assigned and the most shielded one has been arbitrarily labeled H3". Coupling constants J_{HH}/J_{HP} and J_{CP} were measured with a precision of ± 0.2 Hz and ± 0.5 Hz, respectively. ^{31}P NMR spectra were recorded on a 250 MHz spectrometer and chemical shifts were reported from an external capillary standard of 85% phosphoric acid (δ_P 0.00 ppm). 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. A flow of 4 mL/min was used with a gradient of 2 to 4% CH_3CN in 0.05 M aqueous ammonium acetate in 35 min then a gradient of 4 to 40% CH_3CN in 0.05 M aqueous ammonium acetate in 5 min. The detection was set at 260 nm.

Irradiation conditions: Aqueous solutions of each GNA dinucleotides **2** and **5** and of TpT (**1**) (4 mg in 8 mL) in a quartz tube were degassed under argon for 30 min then exposed to a 254 nm light source (12 x 8 W, T8C UVC 7H 8W Lamps). Aliquots (200 μL) of the solution were sampled at t = 0, t = 1 h 30 and t = 3 h and analysed by UV and NMR.

Circular Dichroism spectra were recorded from 220 to 330 nm at 50 nm/min and at 10°C on a spectropolarimeter equipped with a Peltier temperature controller. A 0.01 M Na phosphate, 0.1 M NaCl, pH 7.0 buffer was used. CD data are expressed in molar ellipticity *per* residue [θ] (deg.cm^2.decimol^{-1}). The molar extinction coefficients at 267 nm of TpT (**1**) (2 × 9.65 ×

10^3 $M^{-1}cm^{-1}$) and of its nucleoside constituent thymidine (9.65×10^3 $M^{-1}cm^{-1}$)[2] were used for GNA dinucleotides, and for their nucleosides constituent. The spectra were collected with a 1 nm bandwidth, 1 s response, 0.2 nm data acquisition interval and represented an average of three scans. Spectra were corrected for the buffer baseline.

Synthetic procedures and characterization data for all new compounds

(*R*)-**7**. Acetic anhydride (750 μL, 7.90 mmol) was added under nitrogen to a pyridine solution (1.5 mL) of (*R*)-**6**[1] (750 mg, 1.50 mmol). After 1 h, methanol (3 mL) was added and the mixture concentrated. The residue was dissolved in anhydrous dichloromethane (7.5 mL). The mixture was washed with brine (7.5 mL). The organic phase was dried over $MgSO_4$, filtered and concentrated. The residue was dissolved in anhydrous dichloromethane (4.5 mL) cooled at 0°C. Trifluoroacetic acid (TFA, 290 μL, 3.90 mmol) was slowly added. The mixture was stirred for 10 min. Dichloromethane was partially removed and the residue immediately purified by silica gel chromatography using a gradient of methanol in dichloromethane (2 to 5%) to afford (*R*)-**7** (264 mg, 1.09 mmol) in 73% yield. ^{1}H NMR (300 MHz, D_2O): δ 7.47 (d, *J* = 1.2 Hz, 1H), 5.25-5.17 (m, 1H), 4.10 (dd, 1H, *J* = 3.3, 14.7 Hz), 3.83 (dd, 1H, *J* = 9.4, 14.7 Hz), 3.79 (dd, 1H, *J* = 3.5, 12.5 Hz), 3.71 (dd, 1H, *J* = 5.4, 12.5 Hz), 2.05 (s, 3H), 1.85 (d, *J* = 1.2 Hz, 3H). ^{13}C NMR (75 MHz, D_2O): δ 174.2, 167.5, 152.9, 143.7, 111.3, 72.8, 61.1, 49.3, 20.8, 11.7. HRMS ($(M+Na)^+$, CH_3OH): calc. for $C_{10}H_{14}N_2O_5Na$ 265.0800, found 265.0799.

(*S*)-**7**. This compound was obtained in 83% yield from (*S*)-**6**[1] following a similar procedure described for the preparation of (*R*)-**7**. HRMS ($(M+Na)^+$, CH_3OH): calc. for $C_{10}H_{14}N_2O_5Na$ 265.0800, found 265.0807.

General procedure for the synthesis of GNA dinucleotides **2**-**5**.

To an anhydrous acetonitrile solution (2.5 mL) of alcohol **7** (76 mg, 0.31 mmol) and phosphoramidite **8** (200 mg, 0.28 mmol) and under nitrogen was added imidazolium triflate (68 mg, 0.31 mmol). The mixture was stirred for 40 min at room temperature. *t*-Butylhydroperoxide (5M in nonane, 114 μL, 0.56 mmol) was then added. After 45 min of stirring at room temperature, the solution was diluted with dichloromethane (5 mL) and washed with brine (5 mL). Organic phase was dried over anhydrous $MgSO_4$, filtered and concentrated. The residue was dissolved in conc. aqueous NH_4OH (2.5 mL) and the solution stirred at room temperature overnight. The solution was concentrated and the residue dissolved in 80% aqueous acetic acid (3.65 mL). The resulting solution was stirred at room temperature for 4 h and concentrated under reduced pressure. The residue was dissolved in water (5 mL). The aqueous phase was extracted with dichloromethane (5 mL), concentrated

under vacuum and the residue was purified by HPLC to give the corresponding GNA dinucleotide.

(*S*,*S*)-T_GpT_G **2**. This compound was obtained in 2% yield after HPLC purification from (*S*)-**7** and (*S*)-**8**. ^{1}H NMR (600 MHz, D_2O): δ 7.50 (1H, br s, H6 T_Gp-), 7.45 (1H, br s, H6 -pT_G), 4.42 (1H, ddddd, J2'1' = 3.6 Hz, J2'3" = 3.9 Hz, J2'3' = 4.4 Hz, J2'1" = 8.6 Hz, J2'P = 8.8 Hz, H2' T_Gp-), 4.08 (1H, ddd, *J*1'P = 2.3 Hz, J1'2' = 3.6 Hz, J1'1" = 14.5 Hz, H1' T_Gp-), 4.04 (1H, dtd, J2'1' = 3.9 Hz, J2'3' = J2'3" = 4.8 Hz, J2'1" = 8.5 Hz, H2' -pT_G), 3.94 (1H, dd, J1'2' = 3.9 Hz, J1'1" = 14.4 Hz, H1' -pT_G), 3.88 (1H, dd, J1"2' = 8.6 Hz, J1"1' = 14.5 Hz, H1" T_Gp-), 3.80 (2H, dd, J2'3' = J2'3" = 4.8 Hz; J3'P = J3"P = 6.1 Hz, H3' and H3" -pT_G), 3.81 (1H, dd, J3'2' = 4.4 Hz, J3'3" = 12.4 Hz, H3' T_Gp-), 3.73 (1H, dd, J3"2' = 3.9 Hz, J3"3' = 12.4 Hz, H3" T_Gp-), 3.70 (1H, dd, J1"2' = 8.5, J1"1' = 14.4 Hz, H1" -pT_G), 1.88 (6H, br s, Me T_Gp- and Me -pT_G). ^{13}C NMR (D_2O, 150 MHz): δ 167.0 (C4 T_Gp- and C4 -pT_G), 152.4 (C2 T_Gp- and C2 -pT_G), 144.1 (C6 T_Gp-), 143.7 (C6 -pT_G), 110.4 (C5 -pT_G), 110.3 (C5 T_Gp-), 73.8 (d, *J*C2'P = 5.7 Hz, C2' T_Gp-), 68.0 (d, *J*C2'P = 7.8 Hz, C2' -pT_G), 66.7 (d, *J*C3'P = 5.9 Hz, C3' -pT_G), 61.7 (d, *J*C3'P = 2.2 Hz, C3' T_Gp-), 50.6 (C1' -pT_G), 49.8 (d, *J*C1'P = 5.7 Hz, C1' T_Gp-), 11.2 (CH_3 T_Gp- and -pT_G). ^{31}P NMR (101 MHz, D_2O): δ -0.91. HRMS ((M+Na)$^+$, MeOH): calc. for $C_{16}H_{23}N_4O_{10}NaP$ 485.1050, found 485.1045.

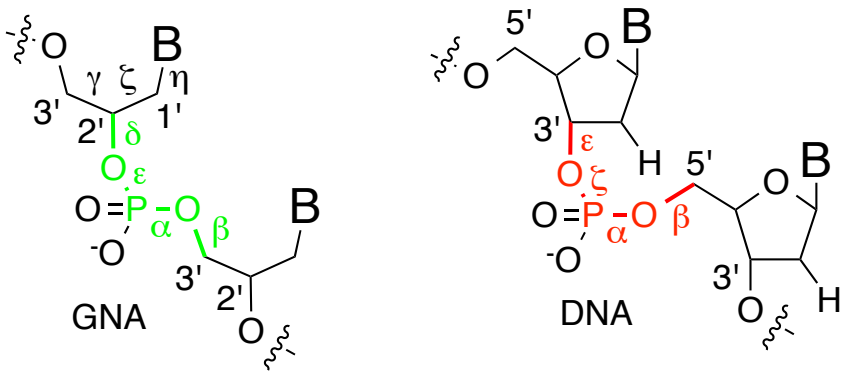
(*R*,*R*)-T_GpT_G **3**. This compound was obtained in 13% yield after HPLC purification from (*R*)-**7** and (*R*)-**8**. HRMS ((M+Na)$^+$, MeOH): calc. for $C_{16}H_{23}N_4O_{10}NaP$ 485.1050, found 485.1057.

(*S*,*R*)-T_GpT_G **4**. This compound was obtained in 7% yield after HPLC purification from (*R*)-**7** and (*S*)-**8**. HRMS ((M+Na)$^+$, MeOH): calc. for $C_{16}H_{23}N_4O_{10}NaP$ 485.1050, found 485.1053.

(*R*,*S*)-T_GpT_G **5**. This compound was obtained in 7% yield after HPLC purification from (*S*)-**7** and (*R*)-**8**. ^{1}H NMR (600 MHz, D_2O): δ 7.45 (1H, br s, H6 T_Gp), 7.40 (1H, br s, H6 pT_G), 4.37 (1H, ddddd, J2'1' = 3.5 Hz, J2'3" = 3.7 Hz, J2'3' = 4.3 Hz, J2'1" = 8.7 Hz, J2'P = 9.0 Hz, H2' T_Gp-), 4.04 (1H, ddd, *J*1'P = 3.1 Hz, J1'2' = 3.5 Hz, J1'1" = 14.5 Hz, H1' T_Gp-), 3.98 (1H, dddd, J2'1' = 3.9 Hz, J2'3' = 4.3 Hz, J2'3" = 5.1 Hz, J2'1" = 8.5 Hz, H2' -pT_G), 3.88 (1H, dd, J1'2' = 3.9 Hz, J1'1" = 14.3 Hz, H1' -pT_G), 3.82 (1H, dd, J1"2' = 8.7 Hz, J1"1' = 14.5 Hz, H1" T_Gp-), 3.80 (1H, ddd, J3'2' = 4.3 Hz, J3'P = 6.3 Hz, J3'3" = 11.0 Hz, H3' -pT_G), 3.77 (1H, dd, J3'2' = 4.3 Hz, J3'3" = 12.4 Hz, H3' T_Gp-), 3.71 (1H, ddd, J3"2' = 5.1 Hz, J3'P = 6.4 Hz, J3"3'

= 11.0 Hz, H3" -pT_G) 3.68 (1H, dd, J3"2' = 3.7 Hz, J3"3' = 12.4 Hz, H3" T_Gp-) 3.63 (1H, dd, J1"2' = 8.5, J1"1' = 14.3 Hz, H1" -pT_G), 1.83 (6H, s, Me T_Gp- and Me -pT_G). ^{13}C NMR (D_2O, 75 MHz): δ 167.0 (C4 T_Gp- and C4 -pT_G), 152.3 (C2 T_Gp- and C2 -pT_G), 144.1 (C6 T_Gp-), 143.7 (C6 -pT_G), 110.3 (C5 T_Gp-), 110.2 (C5 -pT_G), 73.8 (d, *J*C2'P = 5.8 Hz, C2' T_Gp-), 67.9 (d, *J*C2'P = 8.0 Hz, C2' -pT_G), 66.6 (d, *J*C3'P = 5.6 Hz, C3' -pT_G), 61.7 (d, *J*C3'P = 2.4 Hz, C3' T_Gp-), 50.6 (C1' -pT_G), 49.8 (d, *J*C1'P = 5.9 Hz, C1' T_Gp-), 11.2 (CH_3 T_Gp- and CH_3 -pT_G). ^{31}P NMR (101 MHz, D_2O) δ -0.13. HRMS ($(M+Na)^+$, MeOH): calc. for $C_{16}H_{23}N_4O_{10}NaP$ 485.1050, found 485.1058.

NMR conformational analysis



Glycol nucleoside residue conformation (η, ζ, γ):

Torsion angle η (C2-N1-C1'-C2') was evaluated from the observed favoured intraresidue noe between H6 and H1".

Torsion angle ζ (N1-C1'-C2'-O2') was assessed from the magnitude of $^{3}J_{1'2'}$ and $^{3}J_{1"2'}$ couplings.

Torsion angle γ (O2'-C2'-C3'-O3') was determined from the magnitude of $^{3}J_{2'3'}$ and $^{3}J_{2'3"}$ couplings.

Phosphodiester backbone conformation (*δ*, *ε*, *α*, *β*):

The phosphodiester torsion angles *δ* (C3'-C2'-O2'-P) and β (P-O3'-C3'-C2') of GNA dimers were estimated using the Karplus equations developed in nucleic acids for torsion angles *ε* and *β*, respectively.[3] The phosphodiester torsion angles ε (C2'-O2'-P-O3') and α (O2'-P-O3'-C3') were estimated as are torsion angles ζ and α in nucleic acids.[4]

(*S*,*S*) GNA **2**

Torsion angle η (C2-N1-C1'-C2') of each residue was estimated to be preferentially around -90°. Torsion angle ζ (N1-C1'-C2'-O2') of each residue was found to prefer a g- orientation from the small (3.6 -3.9 Hz) and large (8.5 Hz) values of $^{3}JH_{1'2'}$ and $^{3}J_{1"2'}$, respectively. A g- orientation of torsion angle γ (O2'-C2'-C3'-O3') of the 3'-end residue was attested by the similar magnitude of $^{3}J_{2'3'}$ and $^{3}J_{2'3"}$ (4.4 and 3.8 Hz, respectively). Accidental isochronicity of H3' and H3" protons of the 2'-end residue precluded conformational information about γ. The torsional angle δ (C3'-C2'-O2'-P) was evaluated to be around -104° using $^{3}J_{H2'P}$ (8.8 Hz), $^{3}J_{C3'P}$ (2.2 Hz) and $^{3}J_{C1'P}$ (5.7 Hz). Torsion angle β (P-O3'-C3'-C2') was estimated to be in the -

143°/+143° range using $^{3}J_{C2'P}$ (7.8 Hz). The magnitude of $^{2}J_{PC2'}$ (5.7 Hz) and $^{2}J_{PC3'}$ (5.9 Hz) suggested a gauche orientation for ε (C2'-O2'-P-O3') and α (O2'-P-O3'-C3') torsion angles. A W-type long-range coupling between H1' of the 3'-end residue and phosphorus ($^{4}J_{1'P}$ = 2.8 Hz) is in line with the preferred g- orientation of ζ and δ torsion angles of this residue.

(*R*,*S*) GNA **4**

Torsion angle η (C2-N1-C1'-C2') was estimated to be preferentially around +90° and -90° in 3'-end *R* and 2'-end *S* nucleoside residues, respectively. Torsion angle ζ (N1-C1'-C2'-O2') was found to be in the g+ domain for the 3'-end *R* nucleoside residue and in the g- domain for 2'-end *S* nucleoside residue from the small magnitude of $^{3}J_{1'2'}$ (3.5 and 3.9 Hz) and large magnitude of $^{3}J_{1''2'}$ (8.7 and 8.5 Hz) couplings of each residue. The similar magnitude of $^{3}J_{2'3'}$ and $^{3}J_{2'3''}$ (4.3 and 3.7 Hz, respectively) of the 3'-end *R* glycol residue signed a g+ orientation of γ (O2'-C2'-C3'-O3'). The values of $^{3}J_{2'3'}$ and $^{3}J_{2'3''}$ (4.3 and 5.1 Hz, respectively) of the 2'-end *S* glycol residue attested of a g- orientation of γ. Torsion angle δ (C3'-C2'-O2'-P) was evaluated to be around + 105° using $^{3}J_{H2'P}$ (9.0 Hz), $^{3}J_{C3'P}$ (2.4 Hz) and $^{3}J_{C1'P}$ (5.6 Hz). Torsion angle β (P-O3'-C3'-C2') was estimated to be in the -143°/143° domain using $^{3}J_{C2'P}$ (8.0 Hz). This quasi anti orientation of β is confirmed by the identical value of $^{3}J_{H3'P}$ and $^{3}J_{H3''P}$ (6.4 Hz) of the 2'-end nucleoside residue. The value of $^{2}J_{PC2'}$ (5.8 Hz) and $^{2}J_{PC3'}$ (5.6 Hz) coupling constants suggested a gauche orientation for ε (C2'-O2'-P-O3') and α (O2'-P-O3'-C3') torsion angles, respectively. A W-type long-range coupling between H1' of the 3'-end residue and phosphorus ($^{4}J_{1'P}$ = 3.1 Hz) attested the preferred g+ orientation of ζ and δ torsion angles of this residue.

References

1 L. Zhang, A. Peritz and E. Meggers, *J. Am. Chem. Soc.*, 2005, **127**, 4174.

2 H. A. Sober, *CRC Handbook of Biochemistry*, ed. H. A. Sober, Chemical Rubber Co, Cleveland, 1968, p. G-50

3 (a) P. P. Lankhorst, C. A. G. Haasnoot, C. Erkelens and C. Altona, *J. Biomol. Struct. Dyn.*, 1984, **1**, 1387; (b) P. P. Lankhorst, C. A. G. Haasnoot, C. Erkelens and C. Altona, *Nucleic Acids Res.*, 1984, **12**, 5419.

4 L. Benda, Z. Sochorová Volkácová, M. Straka and V. J. Sychrovský, *Phys. Chem.*, 2012, **116**, 3823.

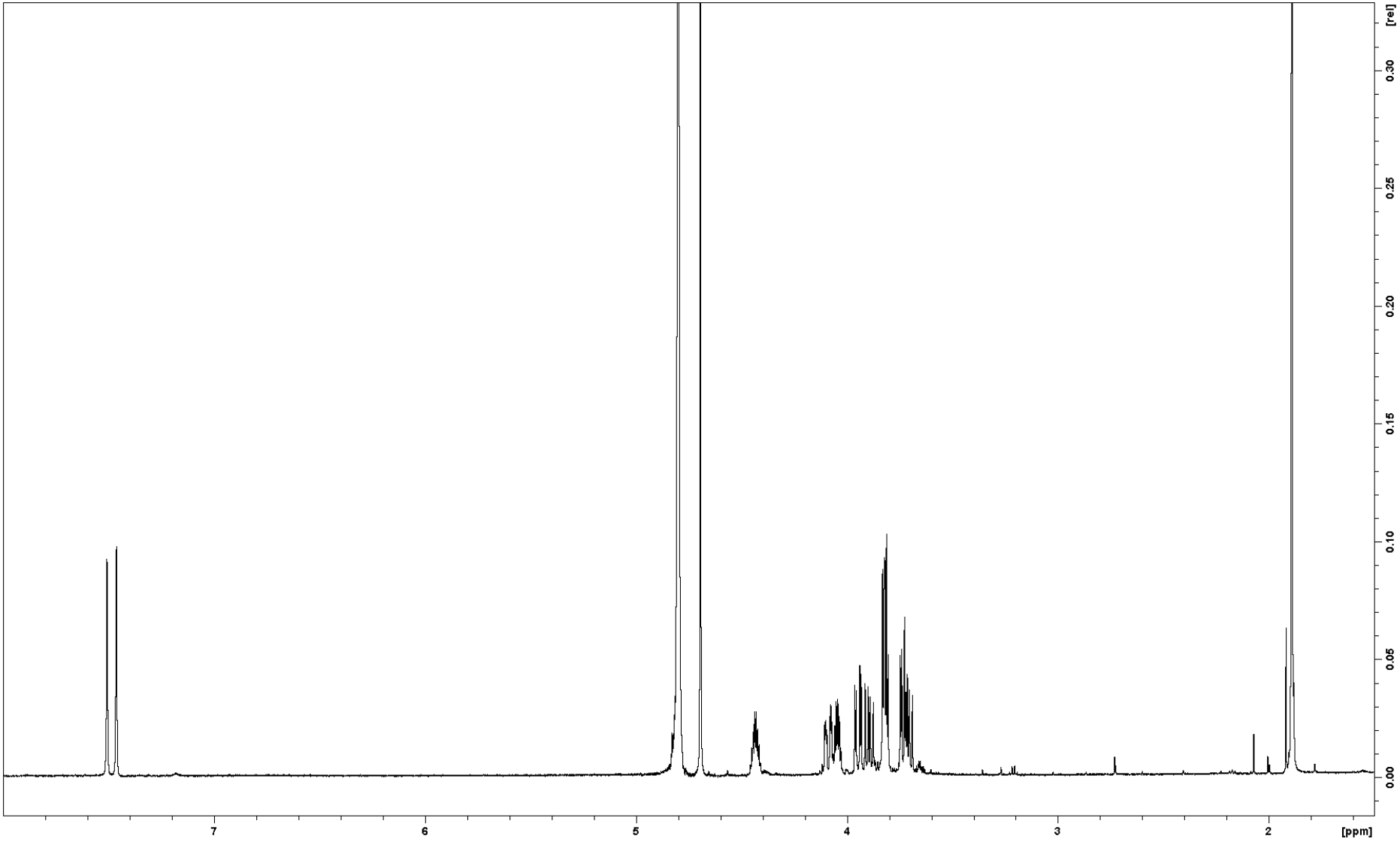


Figure S10: ^{1}H NMR spectrum of (*S*,*S*) GNA **2** (600 MHz, D_2O).

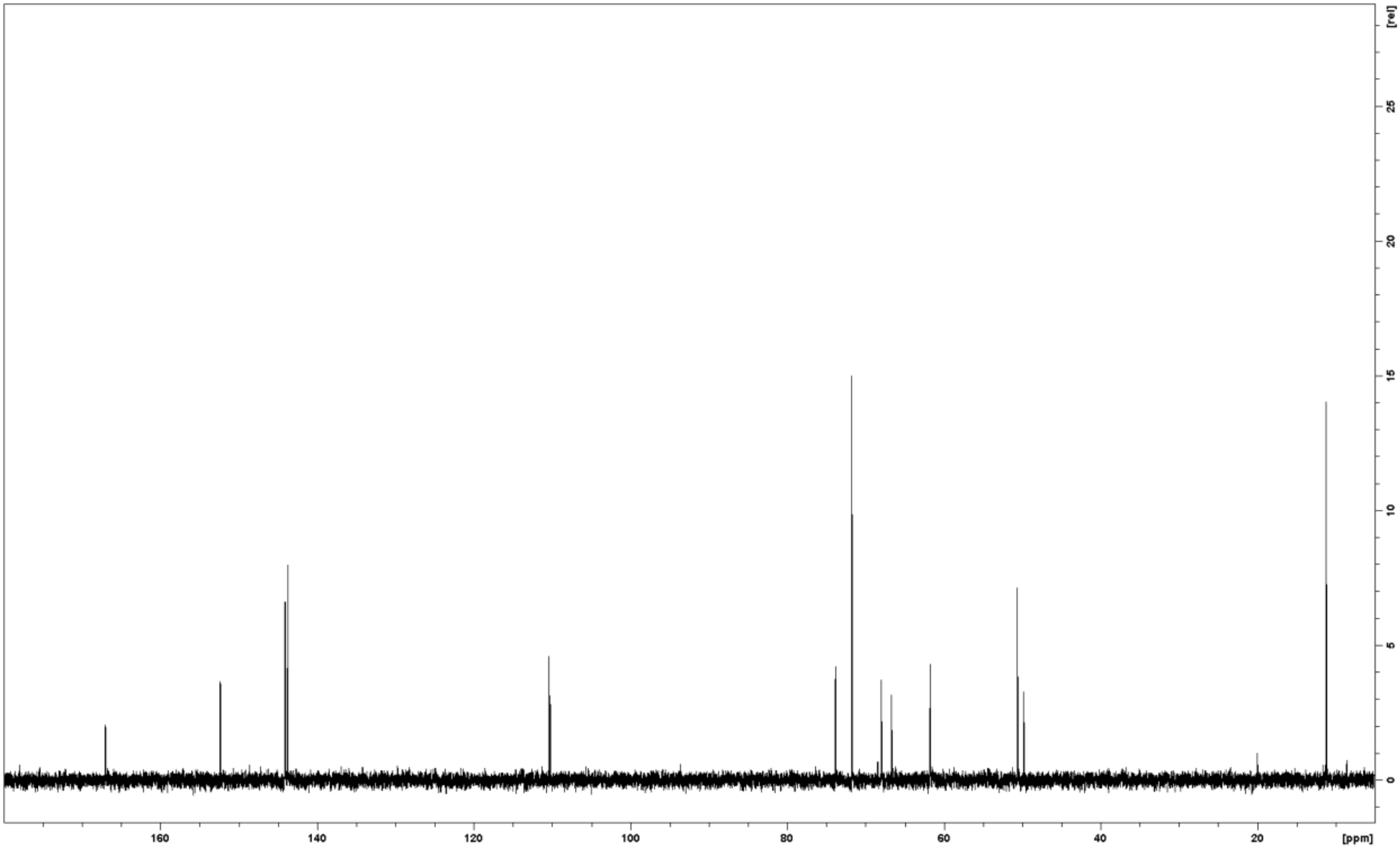


Figure S11: ^{13}C NMR spectrum of (*S*,*S*) GNA **2** (150 MHz, D_2O).

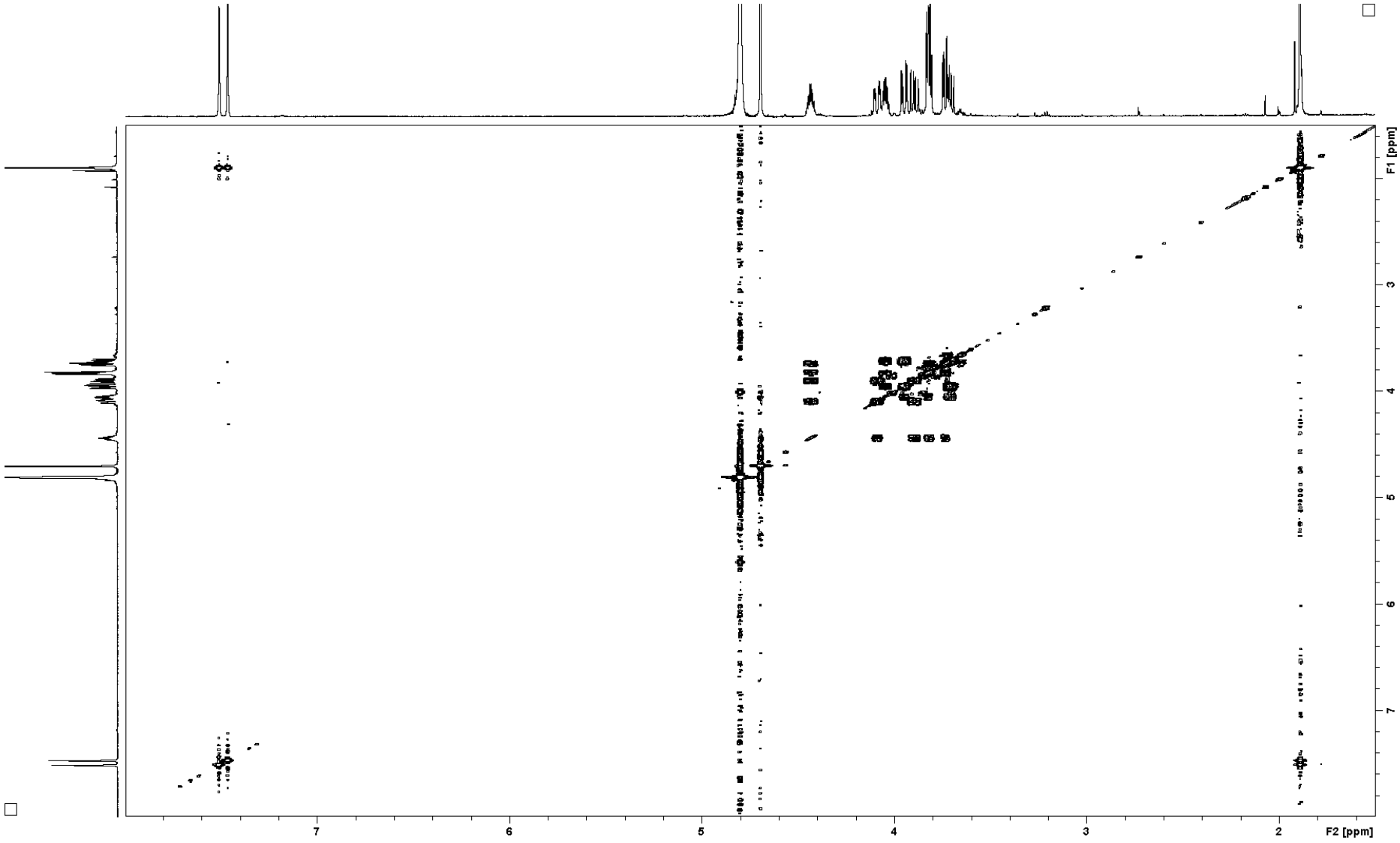


Figure S12: COSY spectrum of (*S*,*S*) GNA **2** (600 MHz, D_2O).

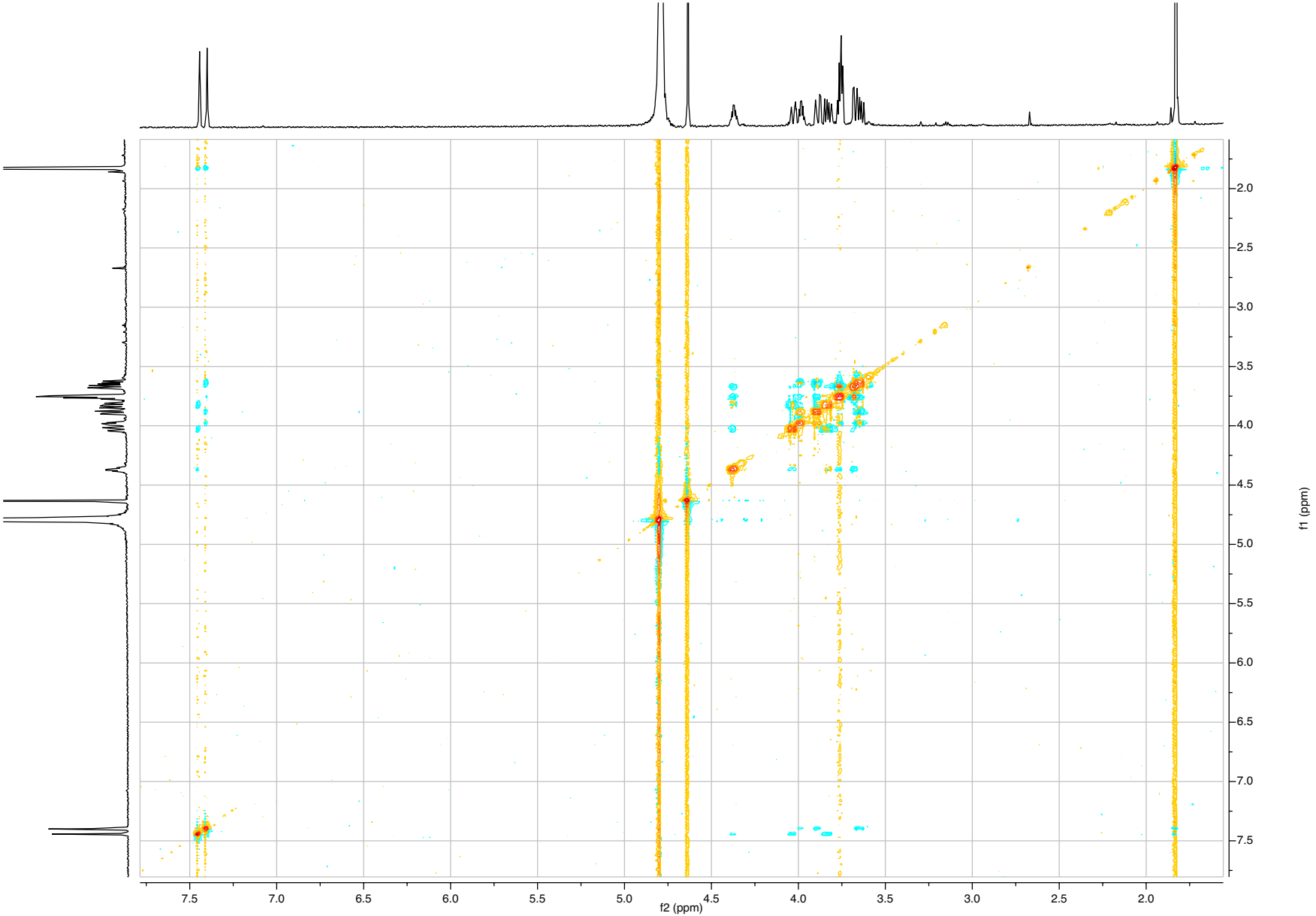


Figure S13: ROESY spectrum of (*S*,*S*) GNA **2** (600 MHz, D_2O, mixing time 500 ms).

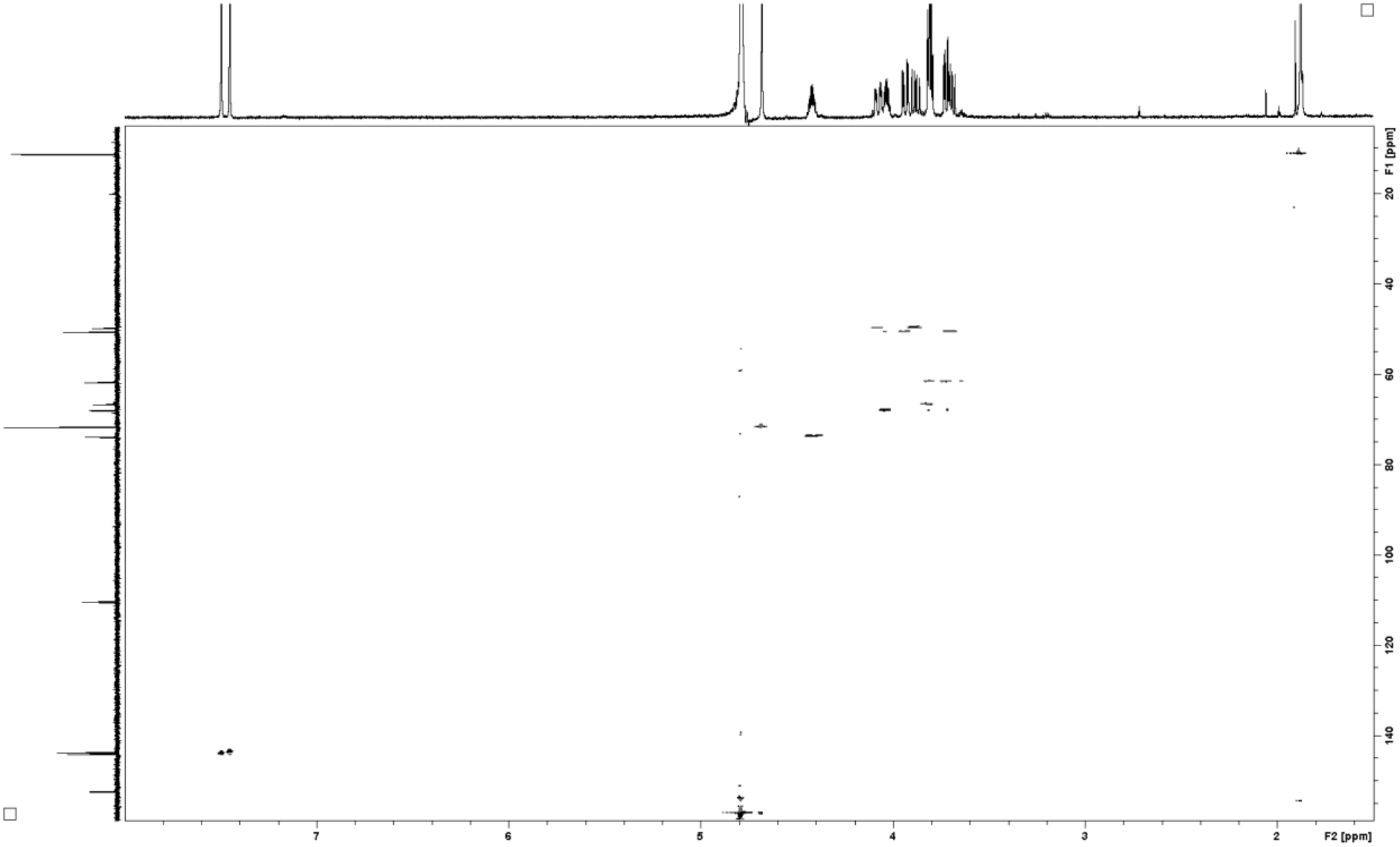


Figure S14: HSQC spectrum of (*S*,*S*) GNA **2** (600 MHz, D_2O).

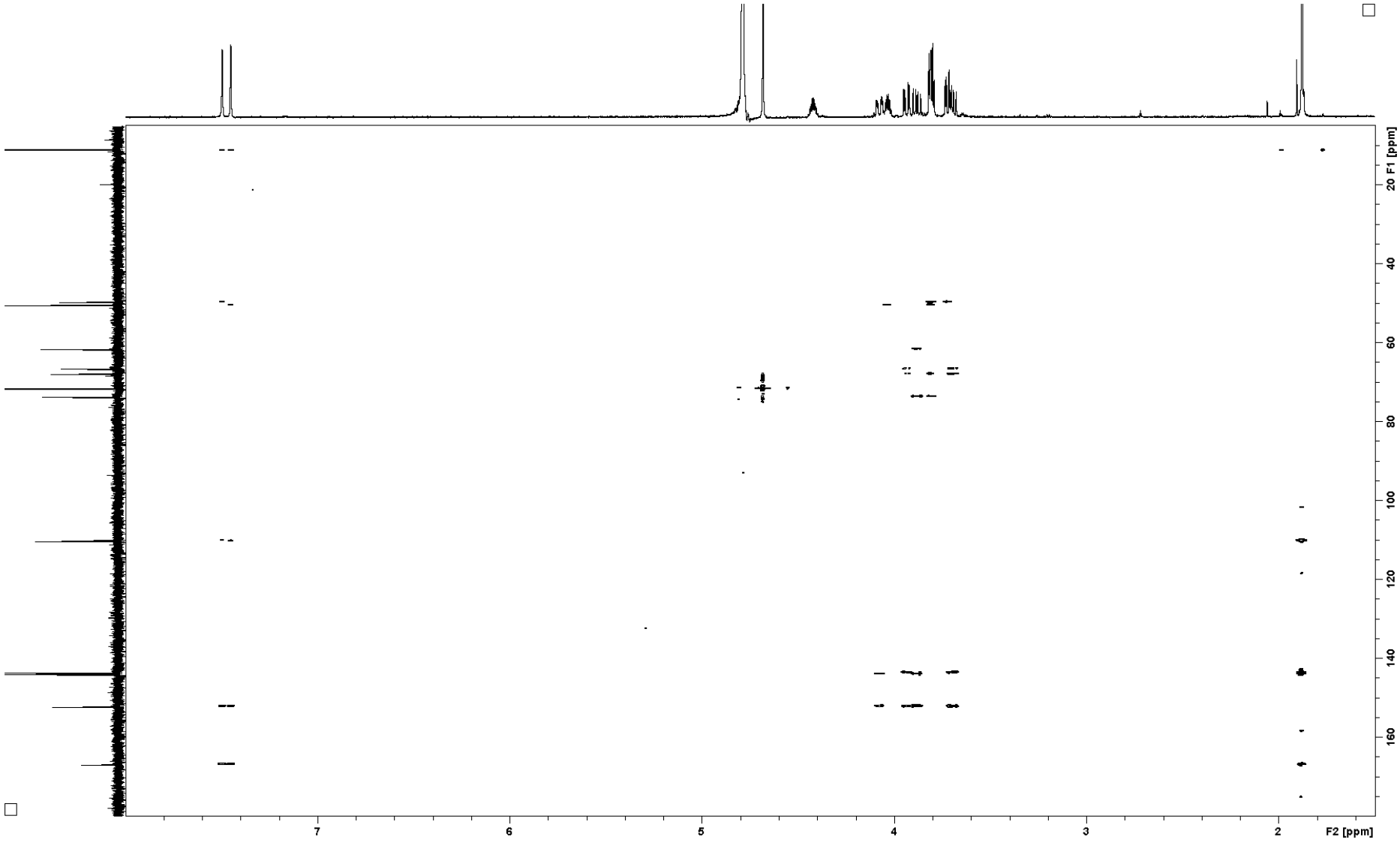


Figure S15: HMBC spectrum of (*S*,*S*) GNA **2** (600 MHz, D_2O).

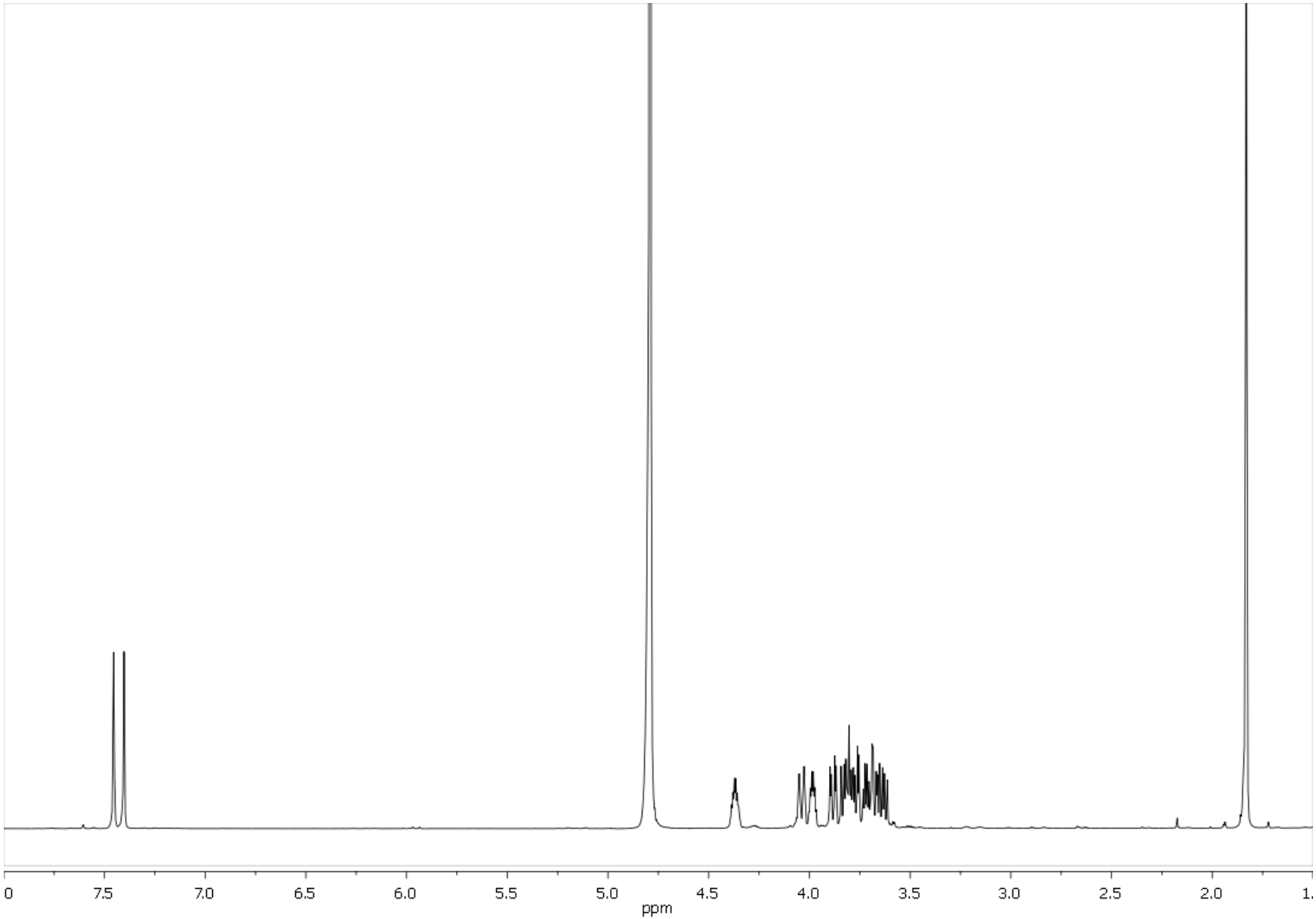


Figure S16: ^{1}H NMR spectrum of (*R*,*S*) GNA **5** (600 MHz, D_2O).

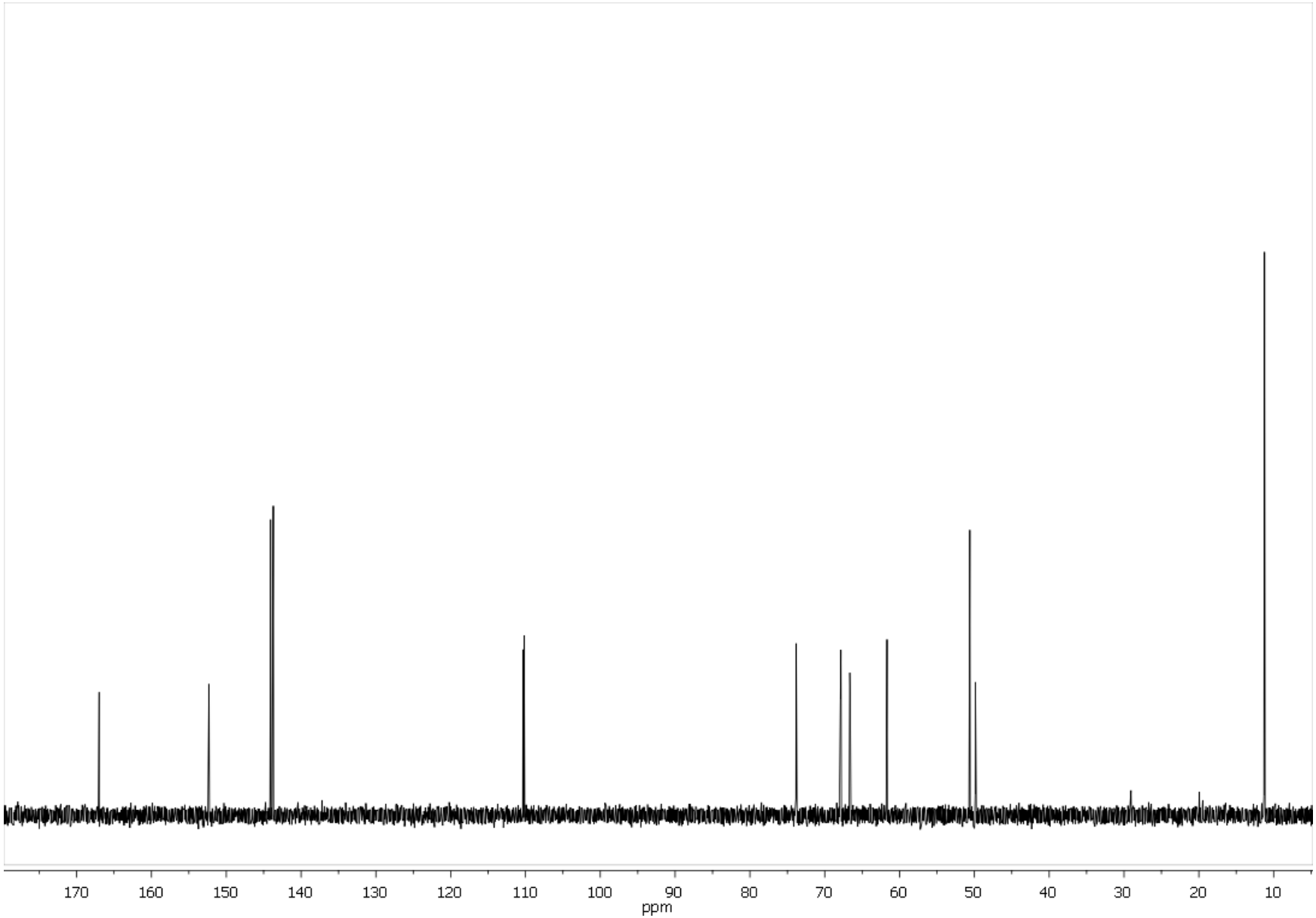


Figure S17: ^{13}C NMR spectrum of (*R*,*S*) GNA **5** (150 MHz, D_2O).

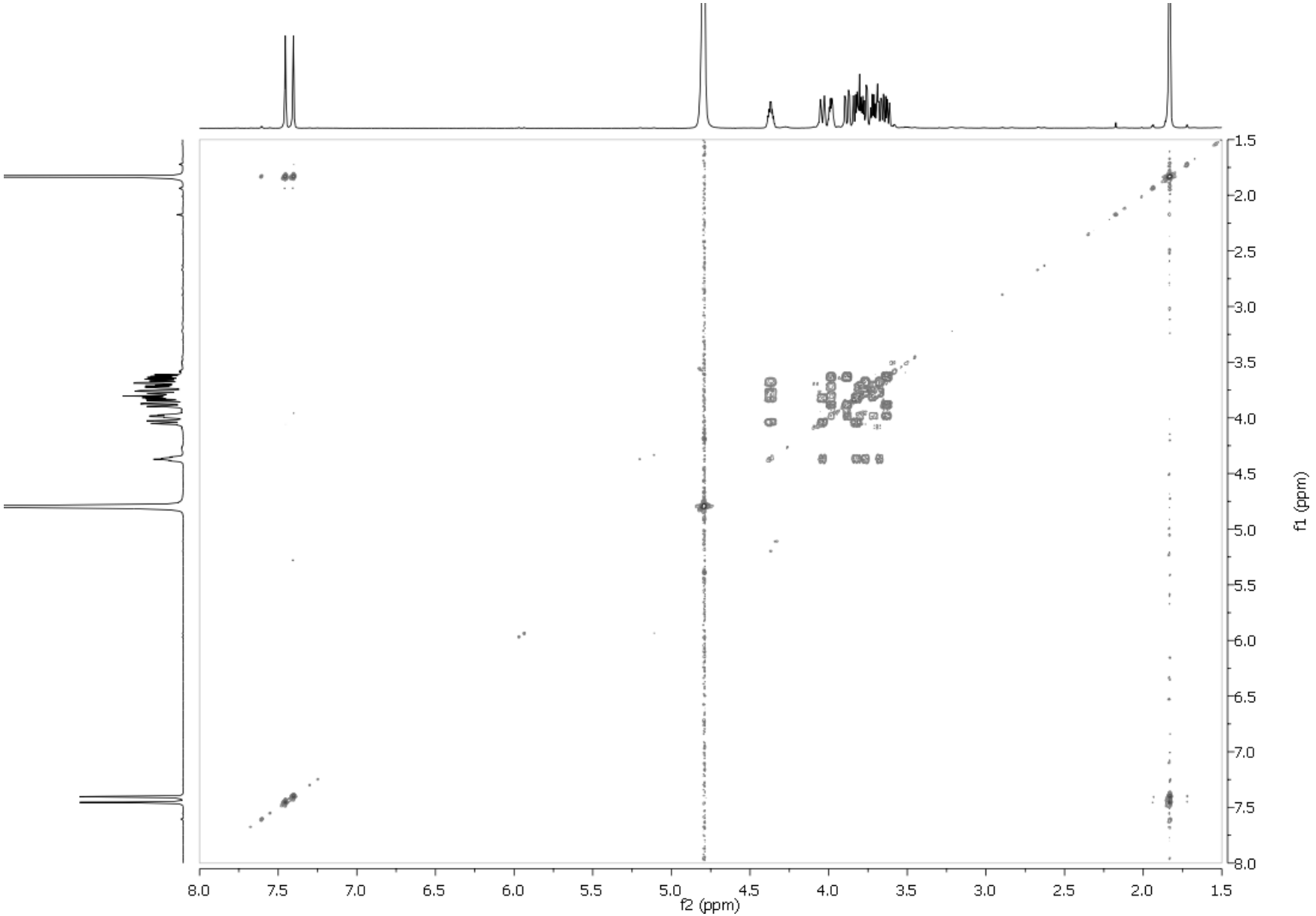


Figure S18: COSY spectrum of (*R*,*S*) GNA **5** (600 MHz, D_2O).

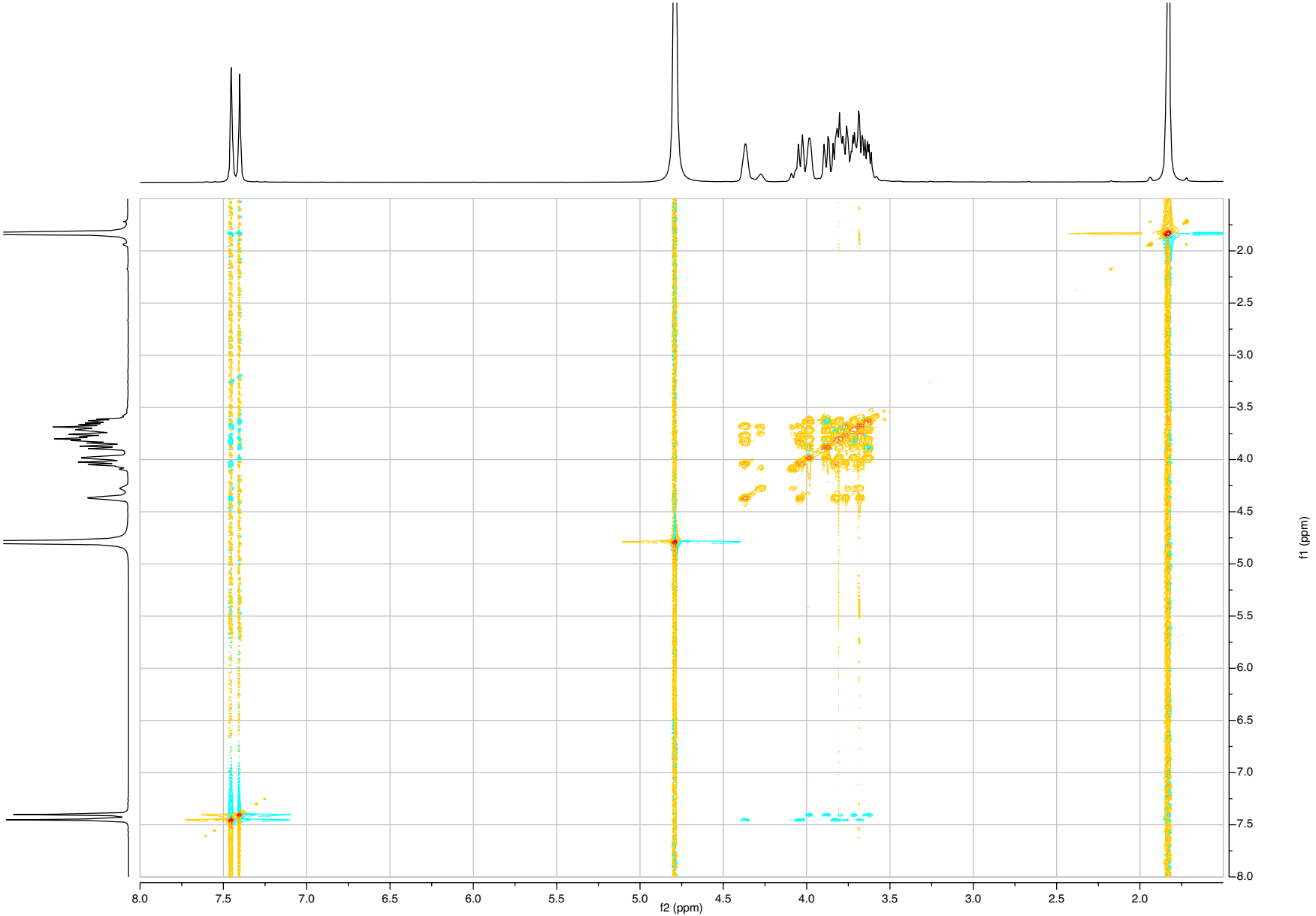


Figure S19: ROESY spectrum of (*S*,*R*) GNA **4** (600 MHz, D_2O, mixing time 500 ms).

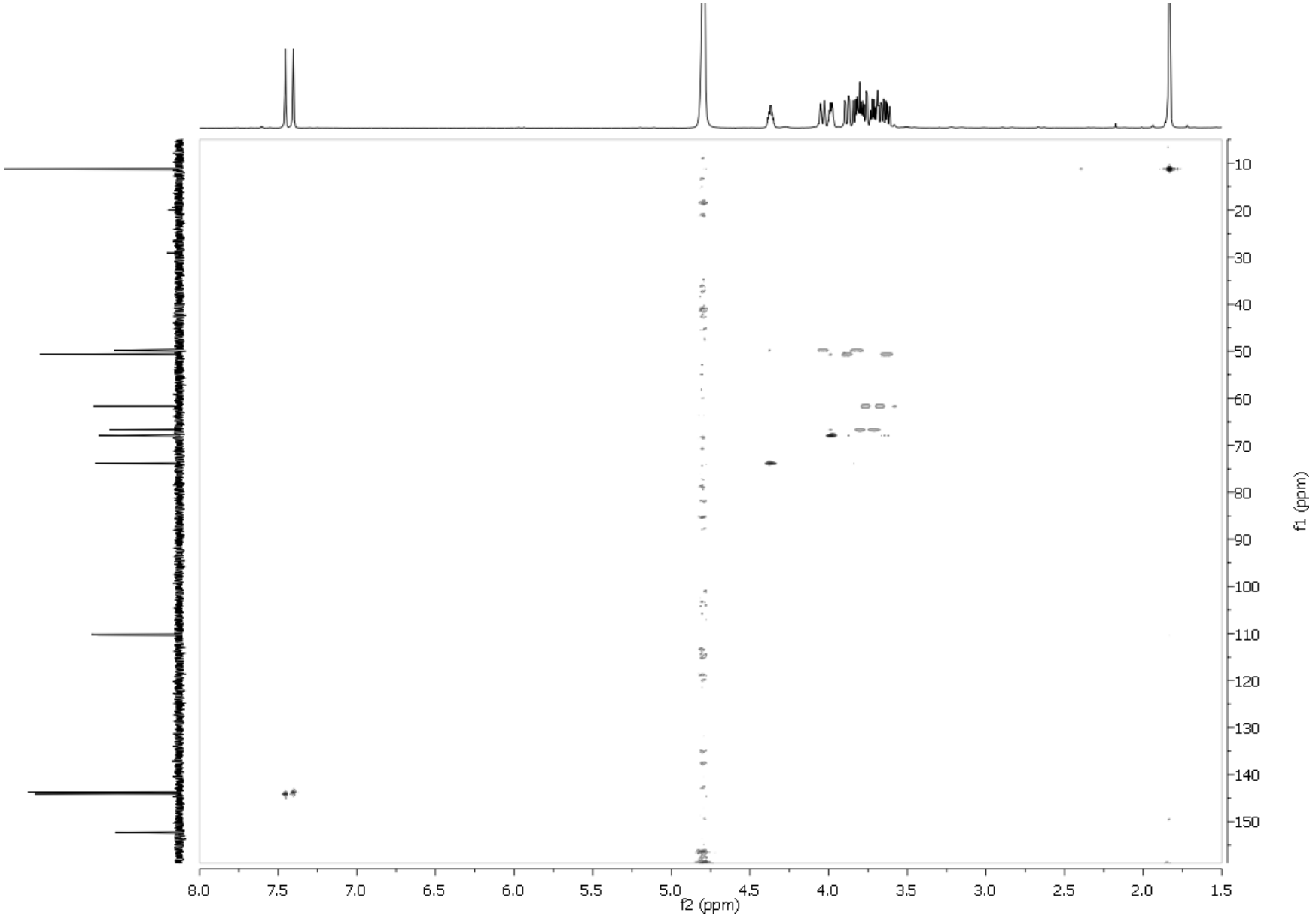


Figure S20: HSQC spectrum of (*R*,*S*) GNA **5** (600 MHz, D_2O).

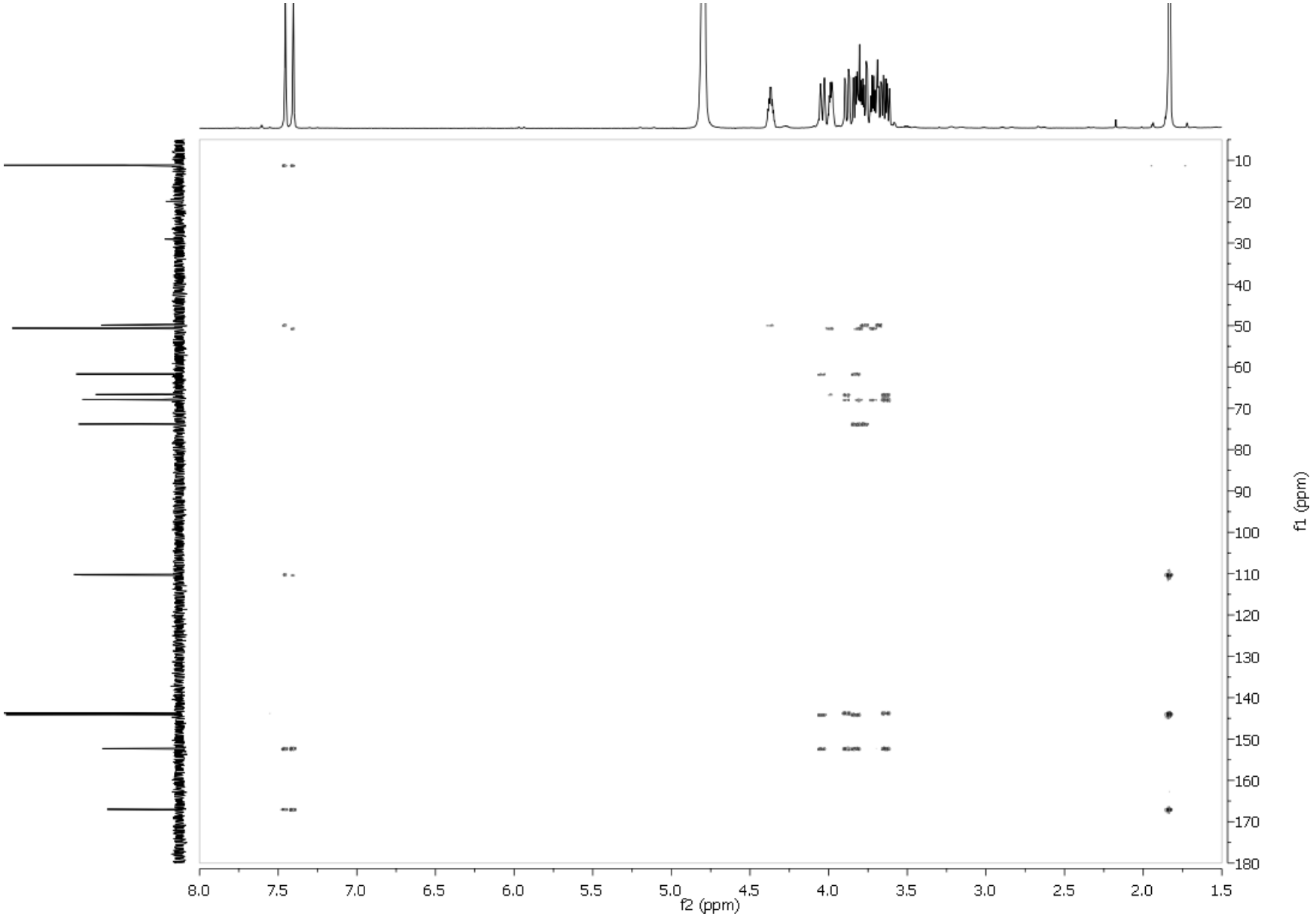


Figure S21: HMBC spectrum of (*R*,*S*) GNA **5** (600 MHz, D_2O).

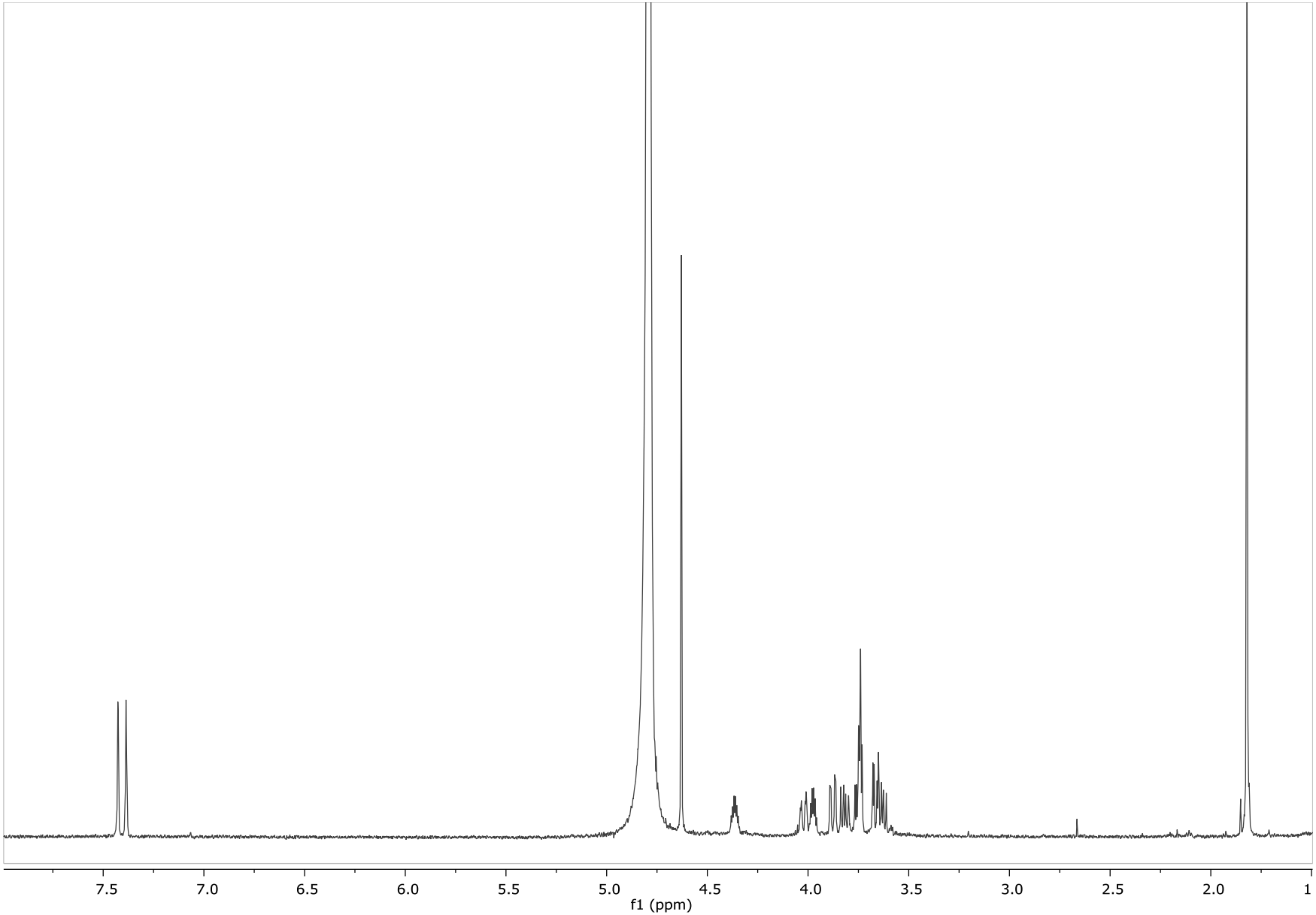


Figure S22: ^{1}H NMR spectrum of the crude irradiation mixture (t = 3 h) of (*S*,*S*) GNA **2** (600 MHz, D_2O).

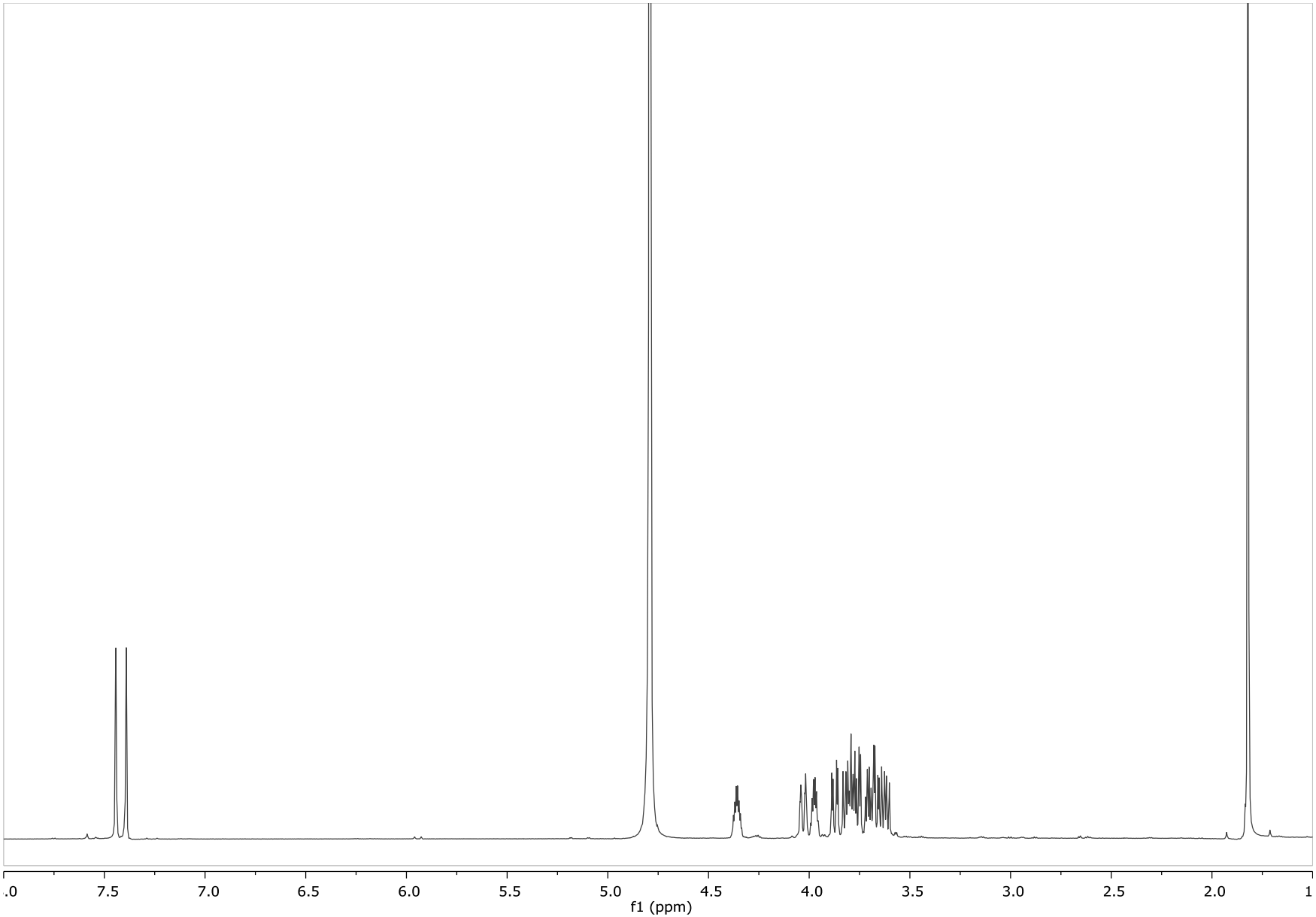


Figure S23: ^{1}H NMR spectrum of the crude irradiation mixture (t = 3 h) of (*R*,*S*) GNA **5** (600 MHz, D_2O).

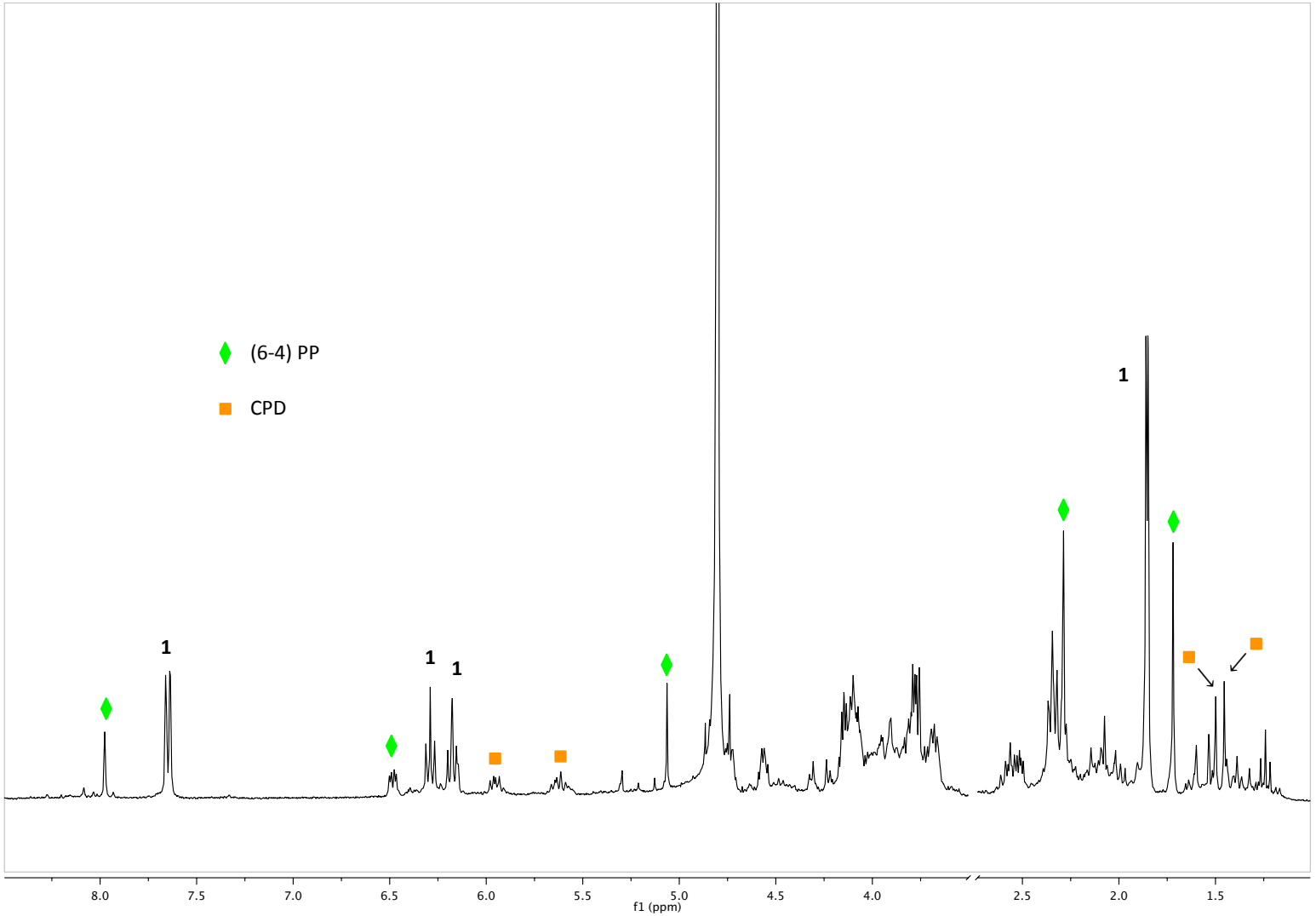
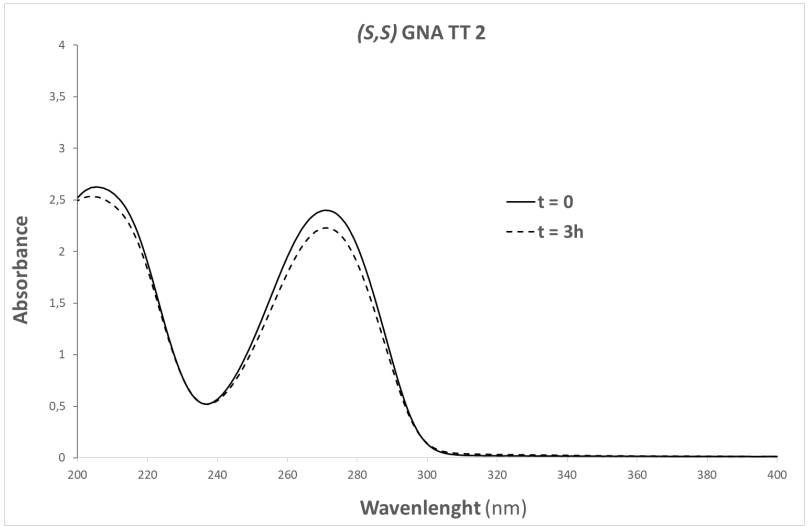
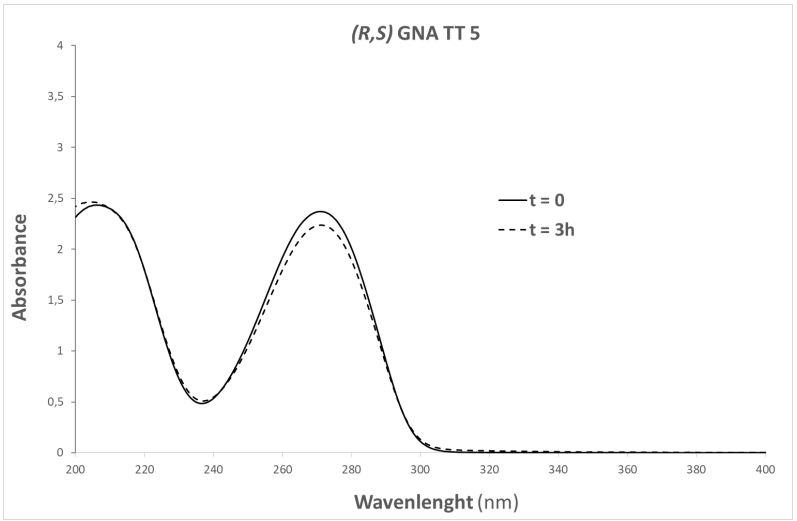


Figure S24: ^{1}H NMR spectrum of the crude irradiation mixture (t = 3 h) of TpT **1** (300 MHz, D_2O).





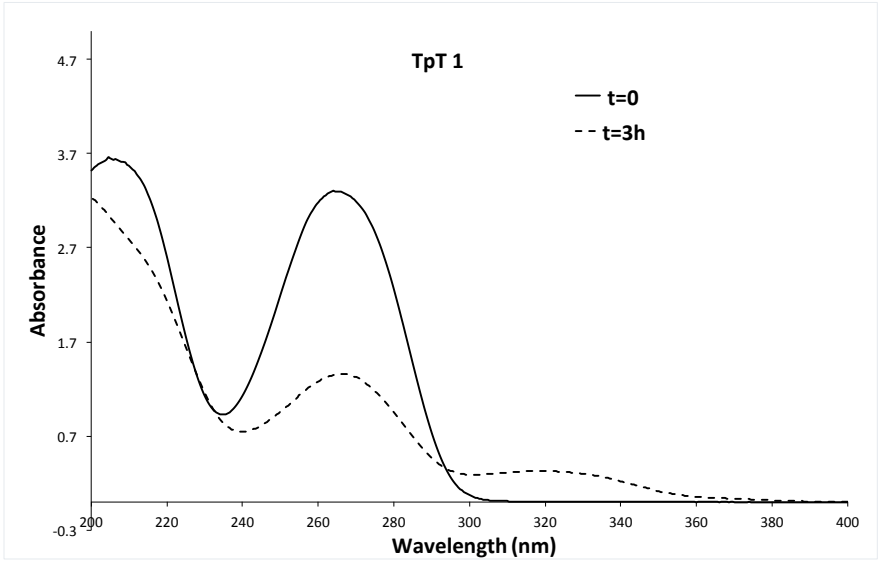


Figure S25: UV spectra of (*S*,*S*) GNA **2**, (*R*,*S*) GNA **5** and TpT **1** at t = 0 and t = 3 h of 254 nm irradiation.