Supporting Information

Synthesis and NMR Elucidation of Novel Octa-Amino Acid Resorcin[4]arenes derivatives.

Iman Elidrisi,[a] Pralav V. Bhatt,[b] Thavendran Govender,[b] Hendrik G. Kruger,[b] and

Glenn E. M. Maguire,[a]*

[a] *School of Chemistry, University of KwaZulu--Natal, Westville Campus, Private Bag X54001, Durban 4000 South Africa,*

[b] *School of Pharmacology, University of KwaZulu--Natal, Westville Campus, Private Bag X54001, Durban 4000 South Africa*

**Email*: maguireg@ukzn.ac.za

Note that the NMR elucidation of the remainder of the compounds (i.e. those compounds that were not elucidated in the main paper) follows at the end of this document. The carbon 13 data are summarised in Table 1.

The following spectra are included:

S1: ^{1}H NMR spectrum of 2,3 and 4a-i

S2: ^{13}C NMR spectrum of 2,3 and 4a-i

S3: COSY NMR spectrum of 4a-i

S4: HSQC NMR spectrum of 4a-i

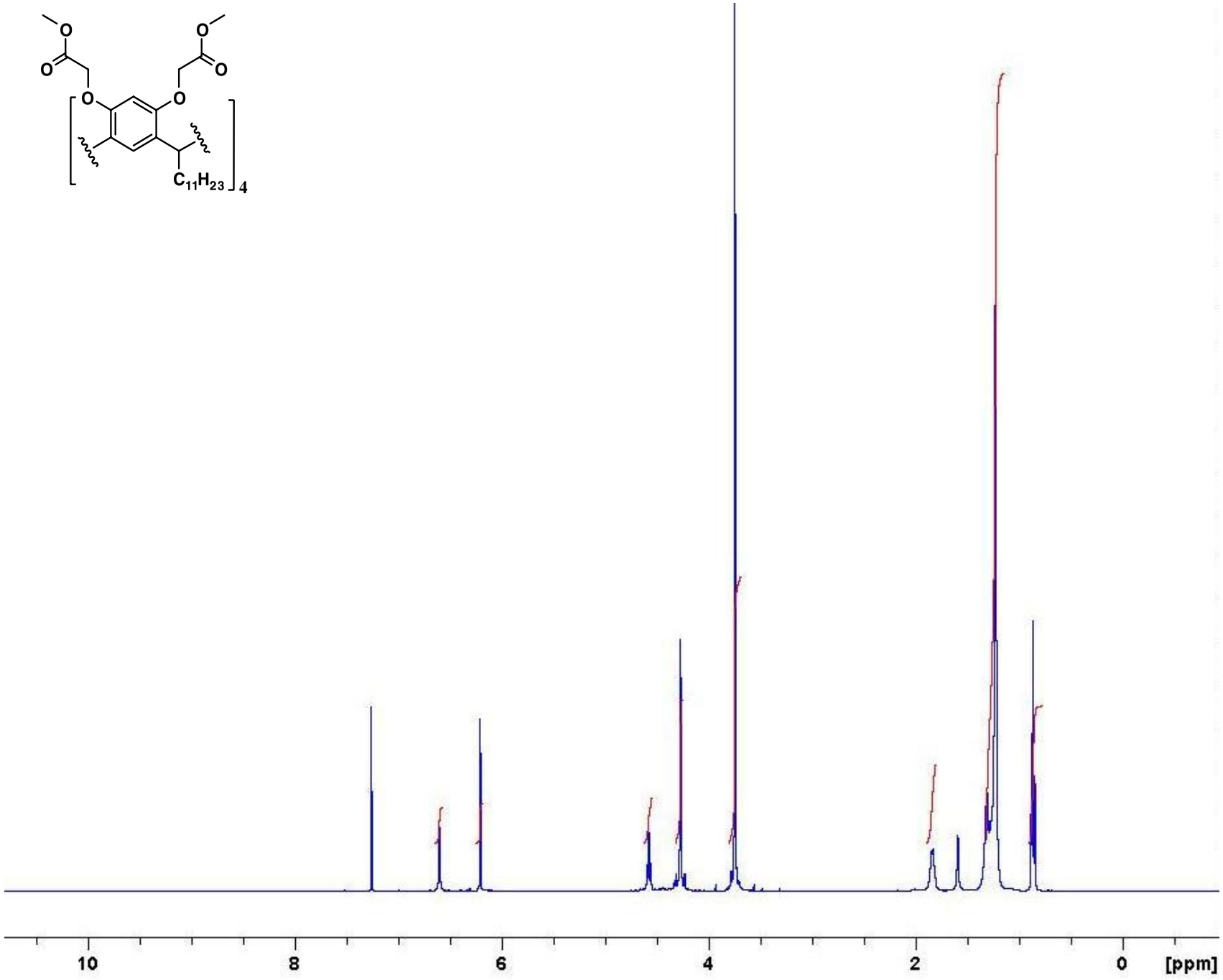
S5: Infrared spectrum of 2,3 and 4a-i

Discussion about the NMR elucidation of the compounds not presented in the manuscript.

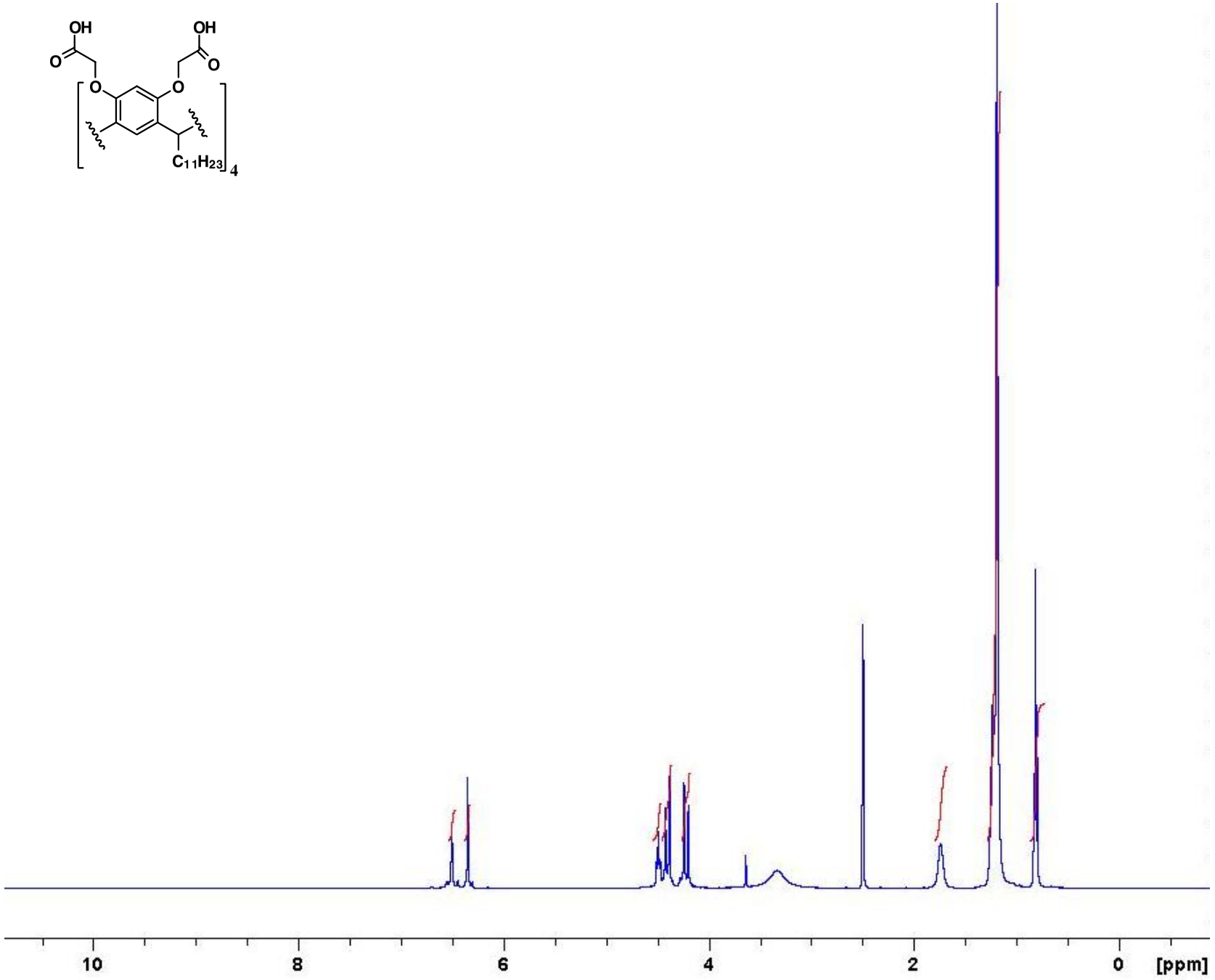
Table-1: ^{13}C NMR shifts for compounds 4a-i

Carbon	**4a**	**4b**	**4c**	**4d**	**4e**	**4f**	**4g**	**4h**	**4i**
C1/C5	154.16	154.40	154.16	154.32,	154.85	154.46	154.69,154.4	154.47,	154.76,
C2/C4	127.08	127.19,	127.07	127.38,	-	127.67,	127.69,	127.26	127.96,
C3	100.84	101.43	100.43	100.99	99.99	101.37	101.57	101.26	101.92
C6	126.24	126.27	126.35,	126.41	126.46	126.32	126.4	126.38	126.49
C7	35.14	35.39	34.78	34.81	34.89	34.8	35.72	35.11	35.51
C8	68.6	68.8,	68.33	68.81,	68.91,	68.85,	69.43,	68.86,	69.35,
C9(C=O)	168.82	168.09	168.32,	168.36,	167.0	168.44,	168.22,	168.54	168.52,
AAC*									
(C=O)	169.83	172.84,	171.77,	173.83,	172.61	172.91,	169.19	172.89,172.8	171.84,171.82,
Aromatic	-	-	137.27,137.07,	-	-	139.53,136.68,	-	-	137.87,136.32,
t-BuO	-	-	-	-	-	-	81.44,81.35	-	-
-OCH_2	60.89	-	-	-	-	-	-	-	-
-OCH_2Ph	-	-	-	-	-	-	-	-	66.63,66.6,
-OCH_3	-	52.27	52.23,	52.18	51.95	52.12,	-	52.33,52.32,	-
Ot-Bu	-	-	-	-	-	-	53.32	-	-
α-C	40.99,	47.83,	53.59,	50.42	59.18	53.41,	53.46,	51.45,	52.43,
β-C	-	17.46,	37.39,	40.81,	41.08	28.02	62.37,	26.74,	31.63,
Y-C	-	-	-	24.83,	24.94	-	-	30.12,	29.99,
δ-C				22.85	46.39	-	-	-	29.46,
ε-C	-	-	-	-	-	-	-	-	49.05
-CH_3	14.41		-	14.15	-	-	22.33	-	-
"Feet"	35.14,31.7, 29.9,29.55, 29.51,29.44, 29.1,27.84, 22.44,14.19	35.13,31.69, 29.88,29.5, 29.41,29.08, 28.09,22.44, 14.18	36.07,34.59, 31.71,30.59, 29.96,29.59, 29.52,29.43, 29.11,27.86, 22.44,21.42, 14.17	35.82,31.68, 29.46,29.55, 29.52,29.5, 29.39,29.08, 28.05,22.98, 22.92,22.42, 21.80,21.72	35.87,31.65, 29.8,29.5, 29.49,29.44, 29.37,29.01, 28.85,28.0, 22.43,14.01	35.39,35.26, 31.85,30.96, 29.99,29.65, 29.53,29.44, 29.11,27.87, 27.84,22.44, 21.42,14.19	35.25,31.61, 30.7,29.58, 29.49,29.44, 29.32,29.0 27.89,22.31, 14.01	31.39,29.95, 29.58,29.51, 29.42,29.09, 28.02,26.74, 26.7,22.44, 14.18	34.75,31.02, 29.5,29.02. 28.05,22.99, 22.33,14.0

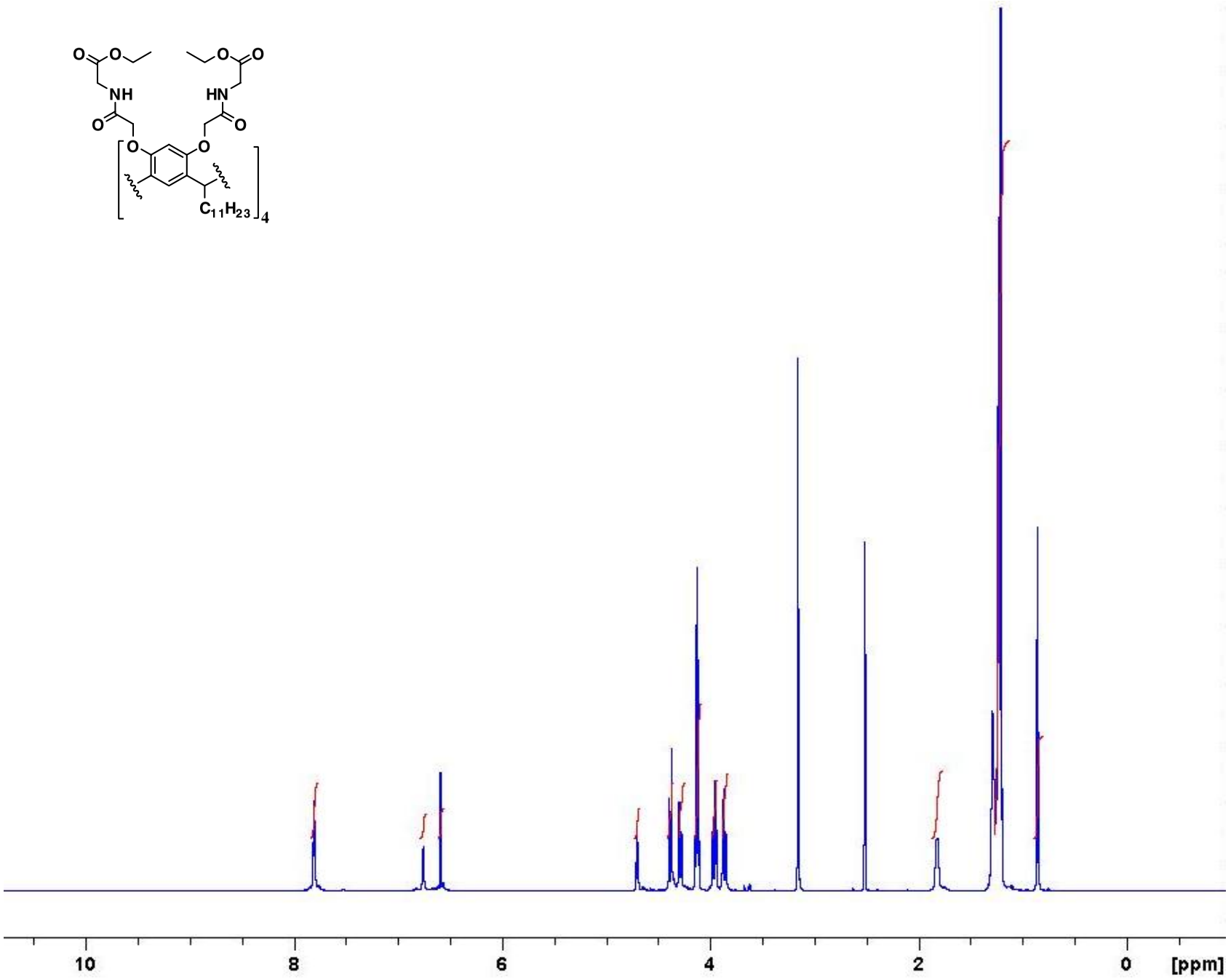
AAC* denotes amino acid carbon peaks



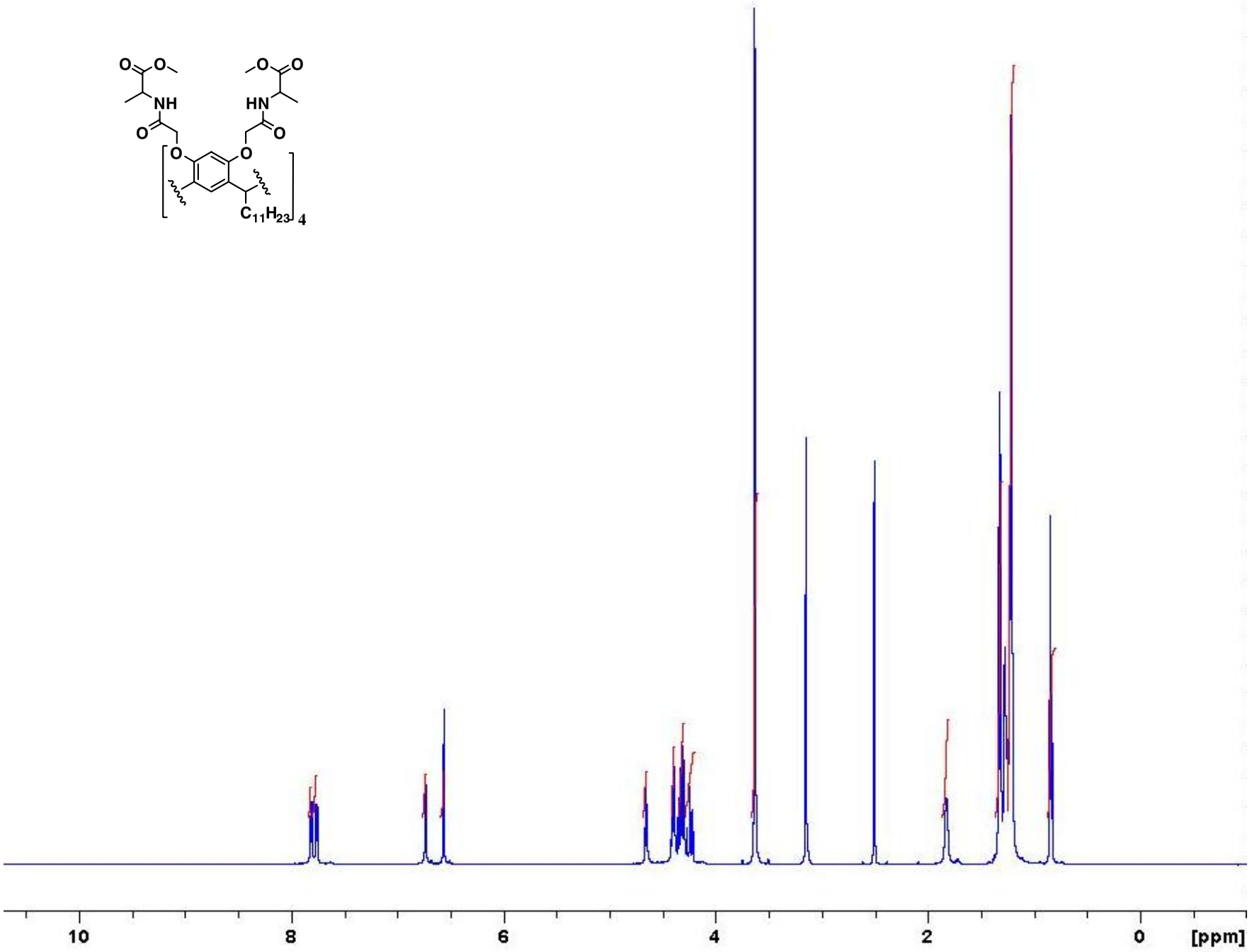
S1: ^{1}H NMR spectrum of 2 in $CDCl_3$ (400 MHz)



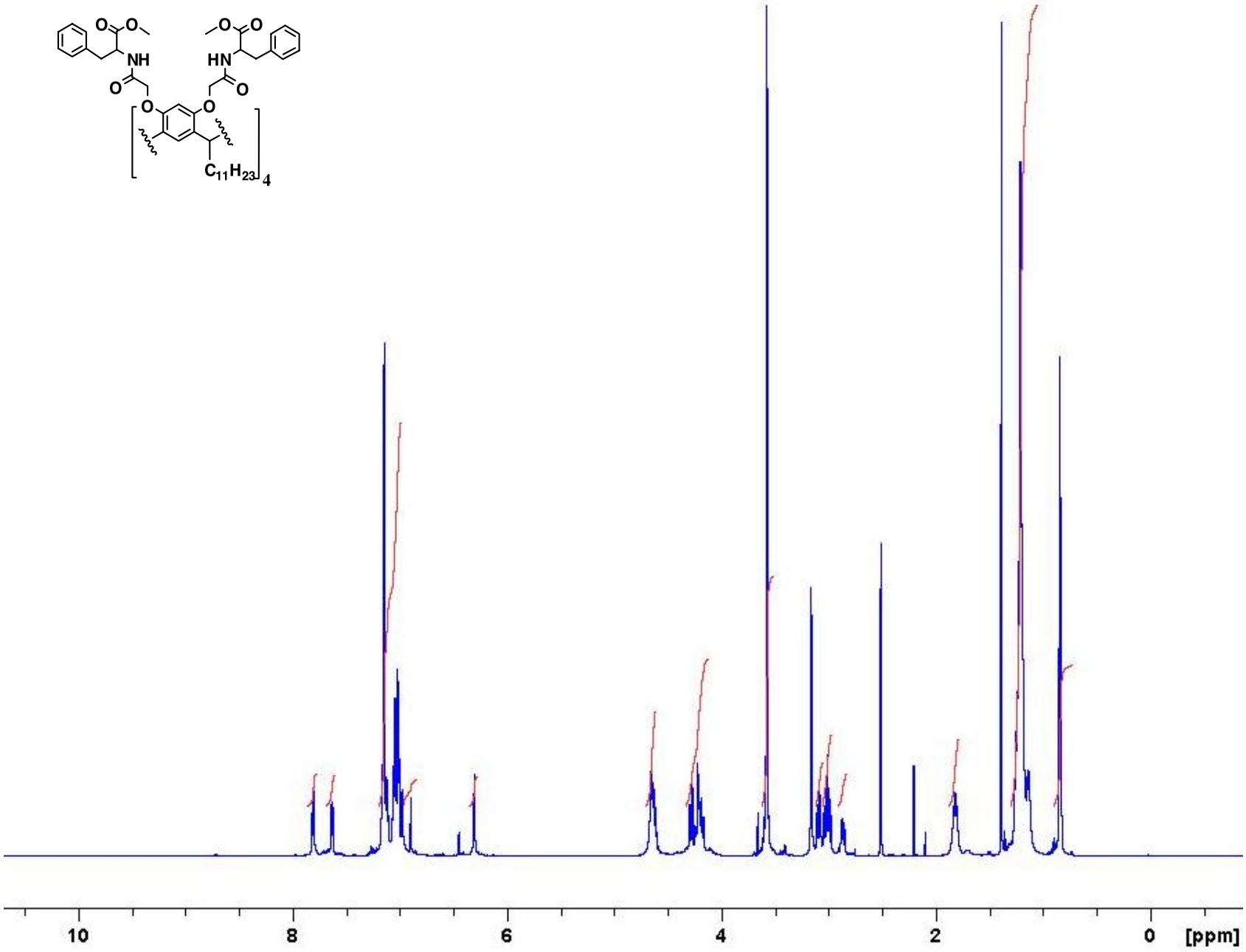
S1: ^{1}H NMR spectrum of 3 in DMSO-d_6 (400 MHz)



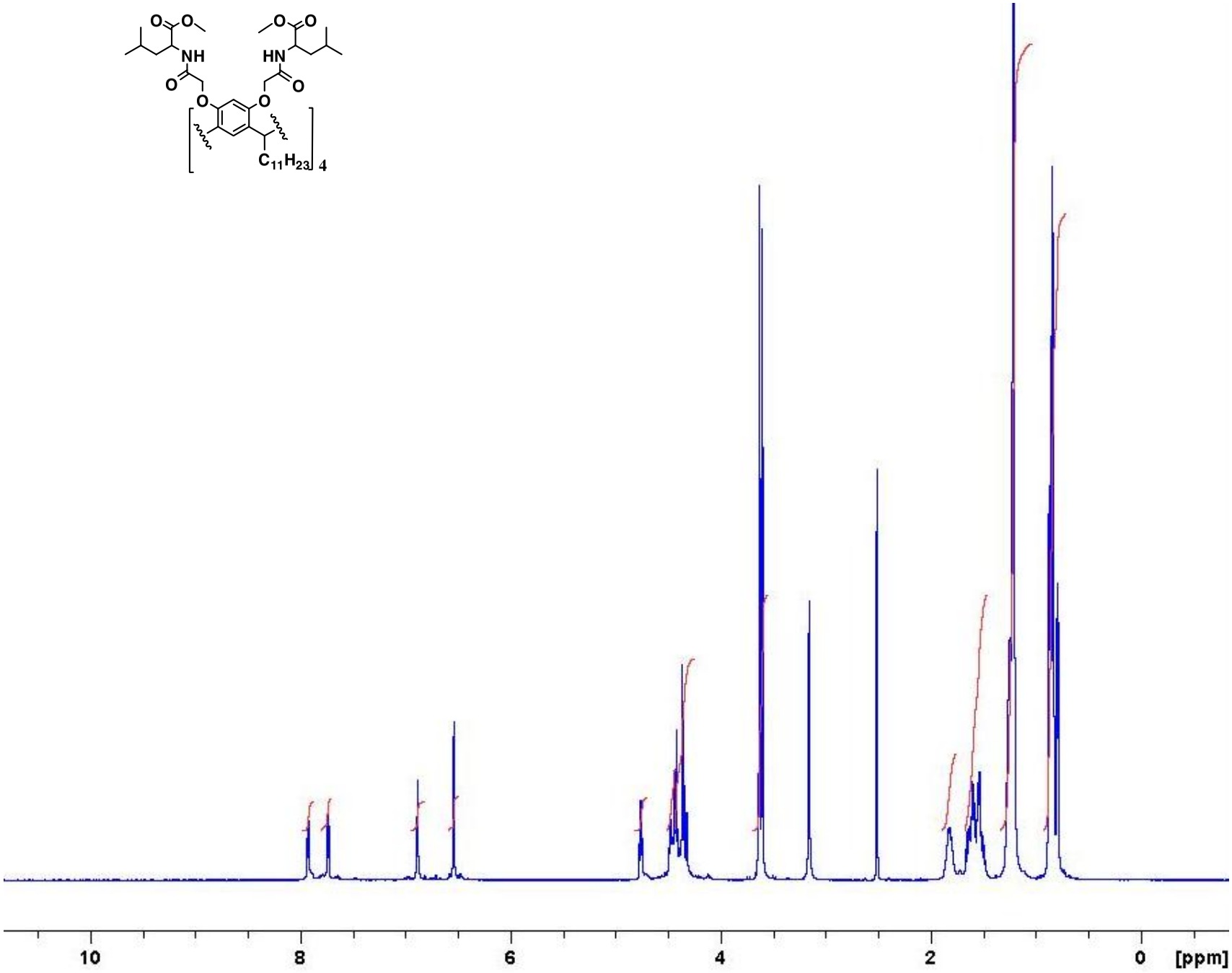
S1: ^{1}H NMR spectrum of 4a in DMSO-d_6 (600 MHz)



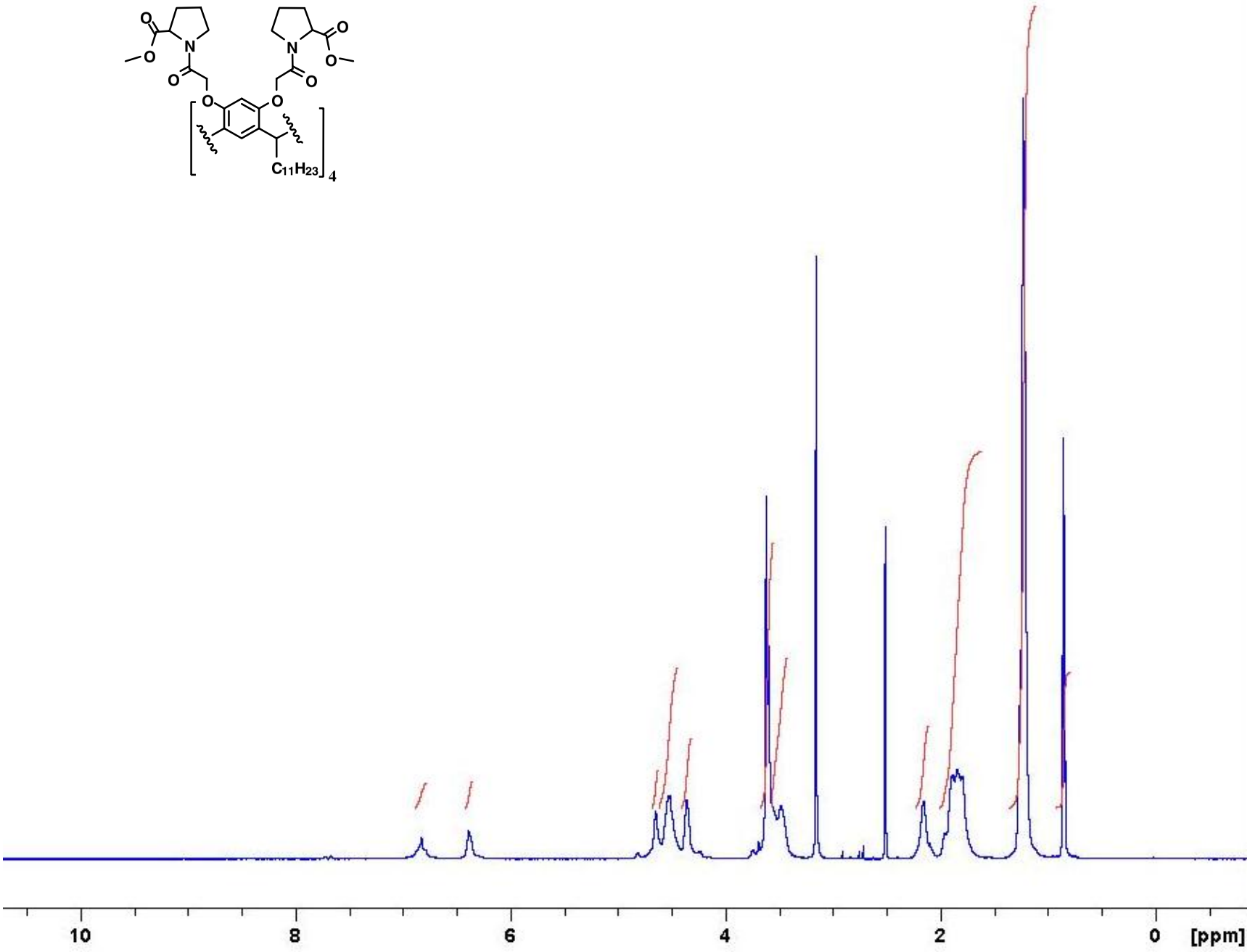
S1: ^{1}H NMR spectrum of 4b in DMSO-d_6 (600 MHz)



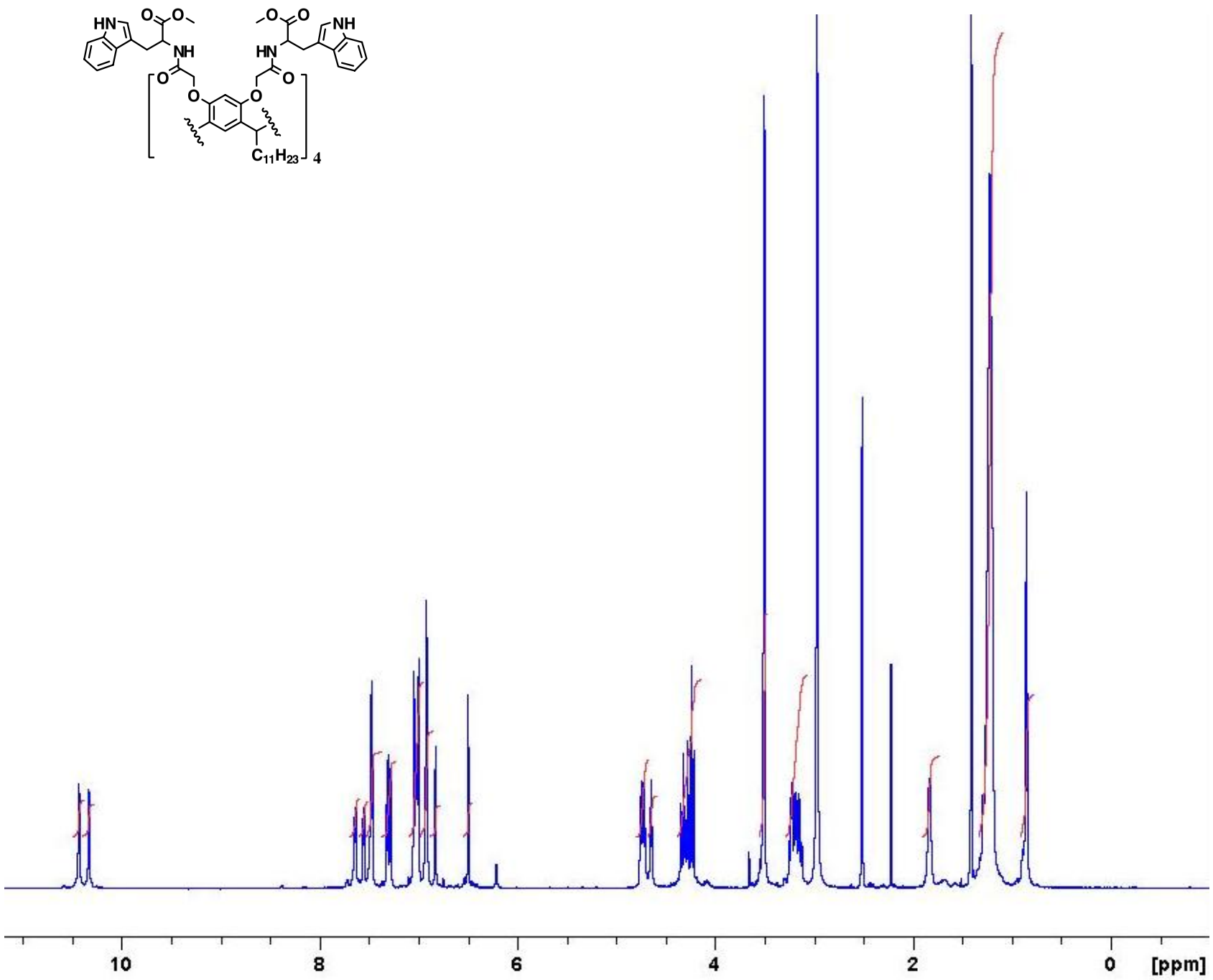
S1: ^{1}H NMR spectrum of 4c in DMSO-d_6 (600 MHz)



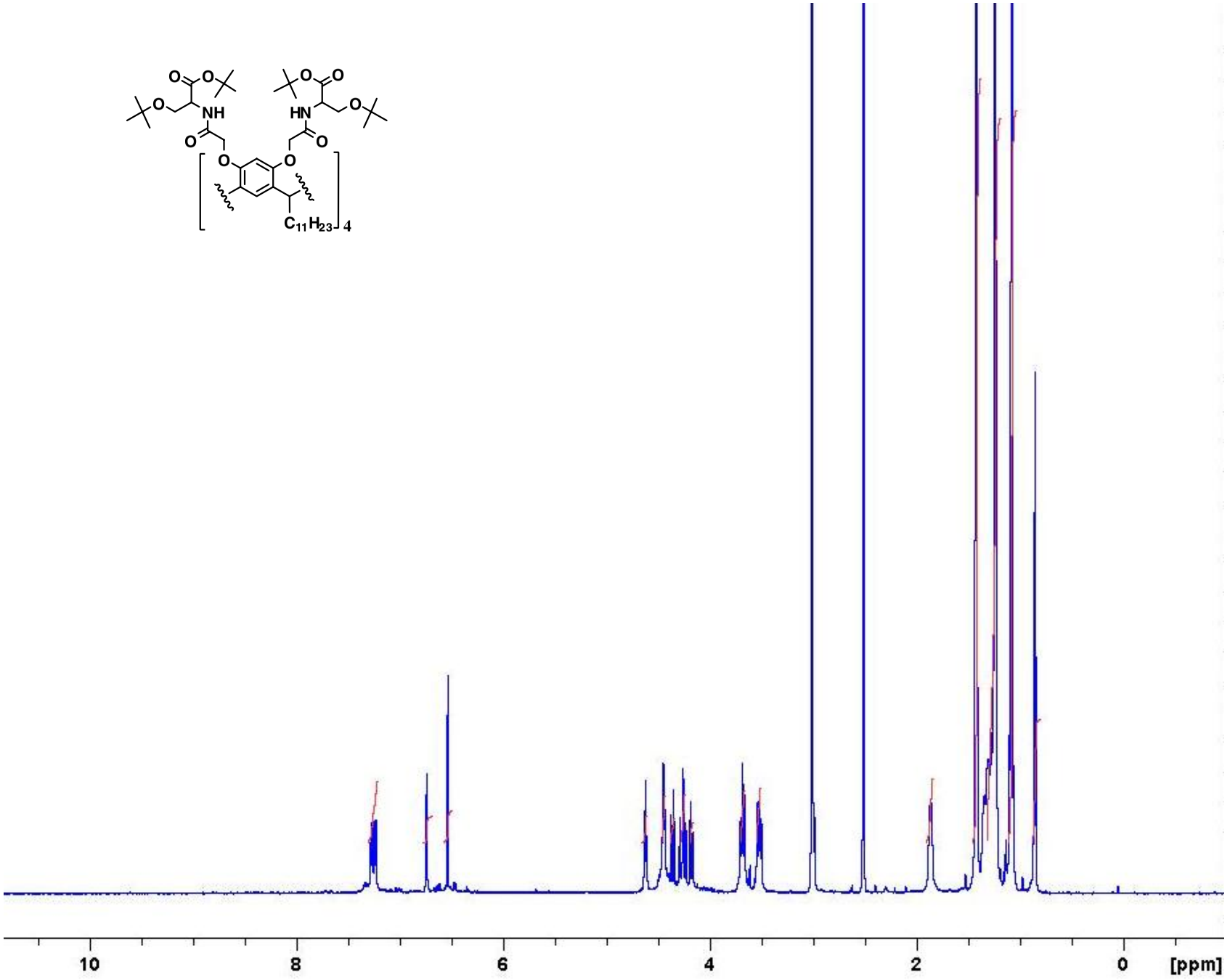
S1: ^{1}H NMR spectrum of 4d in DMSO-d_6 (600 MHz)



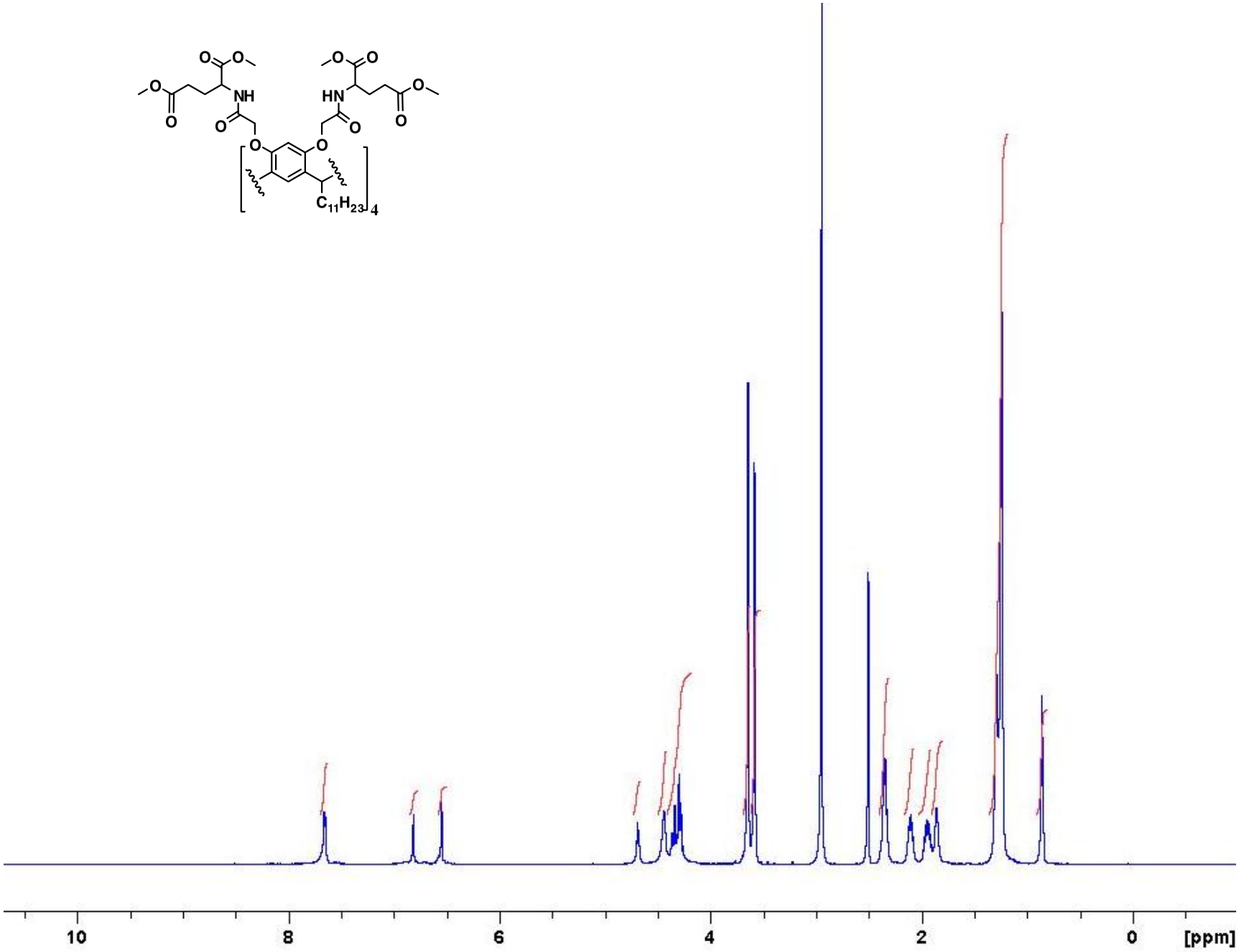
S1: ^{1}H NMR spectrum of 4e in DMSO-d_6 (600 MHz)



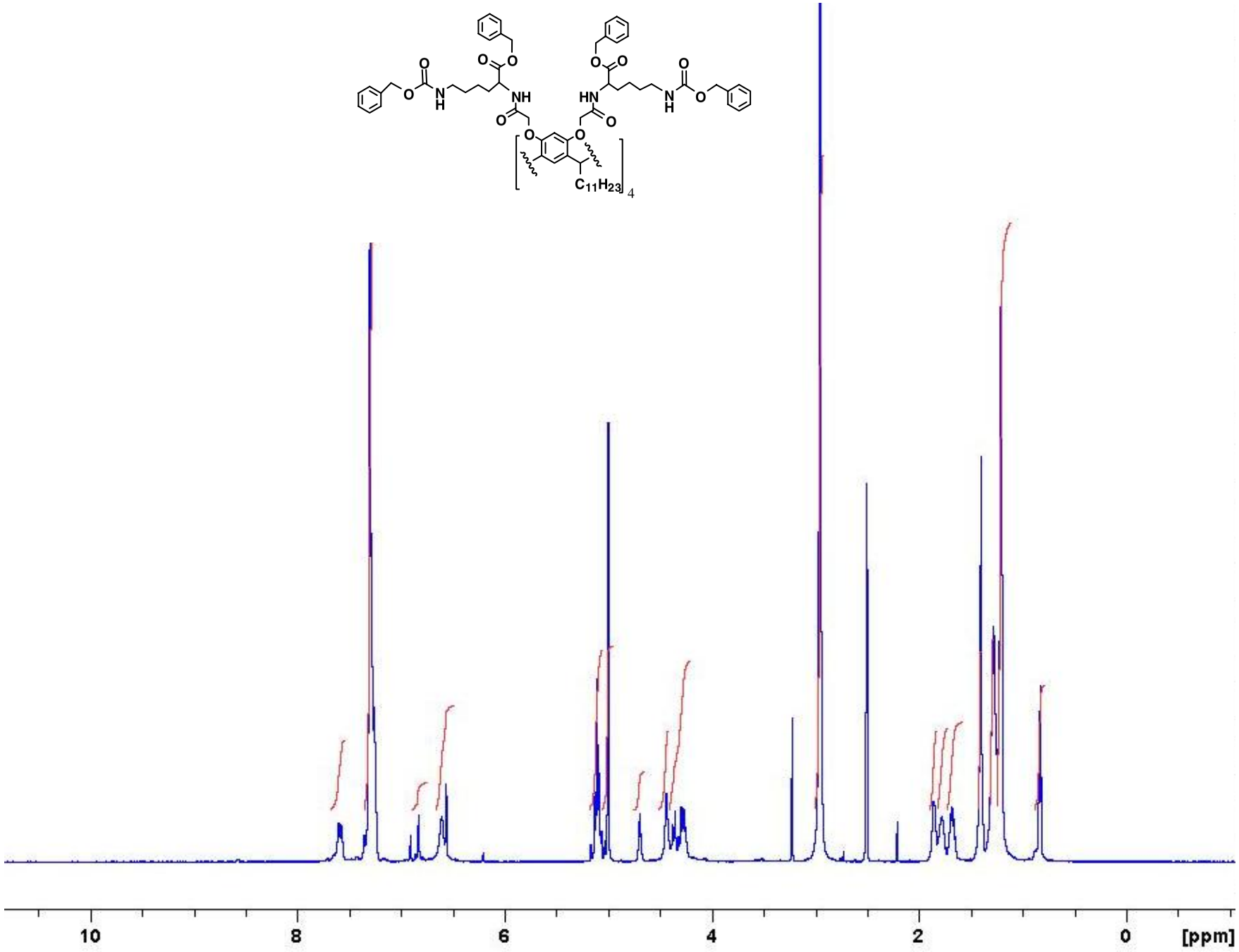
S1: ^{1}H NMR spectrum of 4f in DMSO-d_6 (600 MHz)



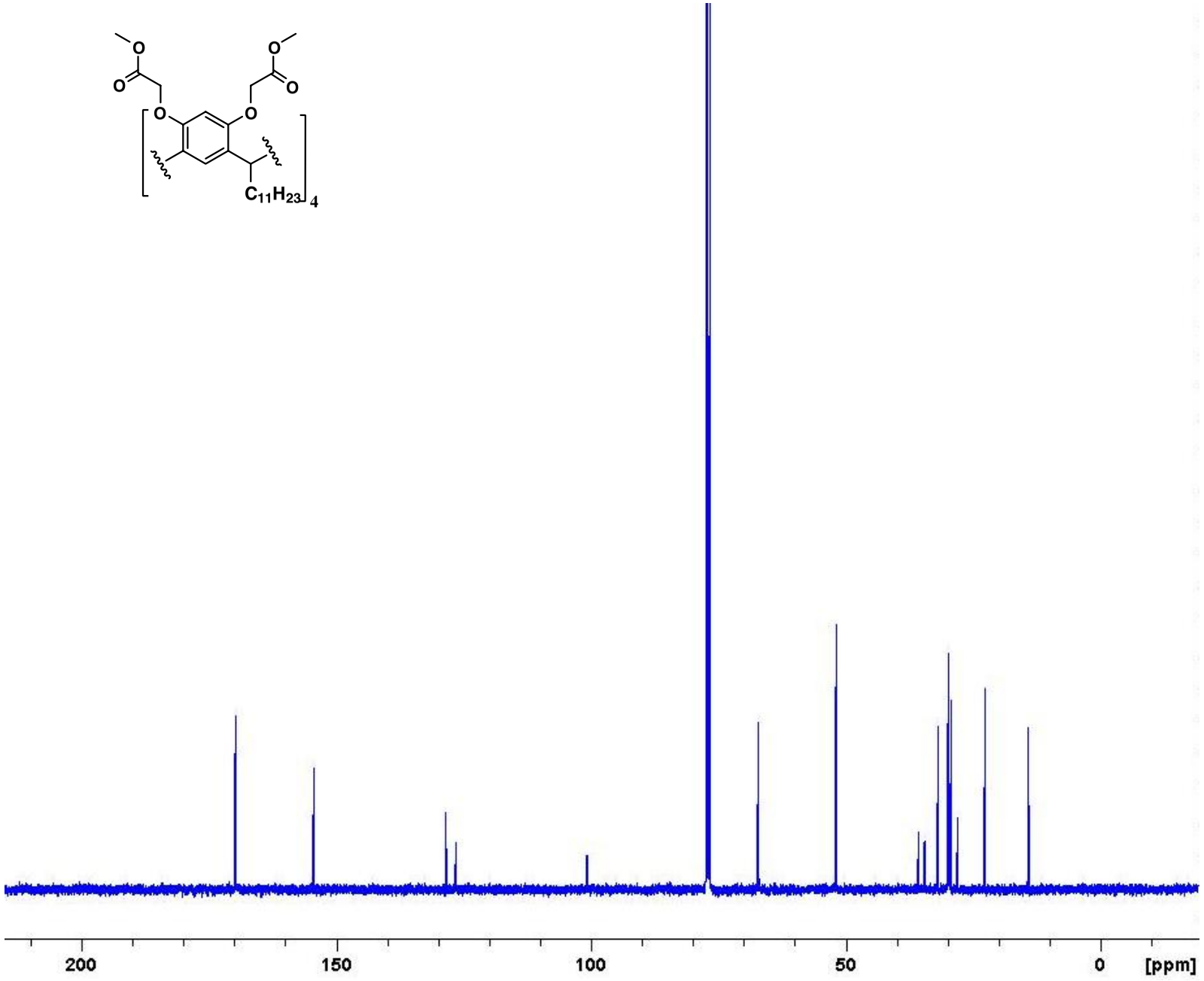
S1: ^{1}H NMR spectrum of 4g in DMSO-d_6 (600 MHz)



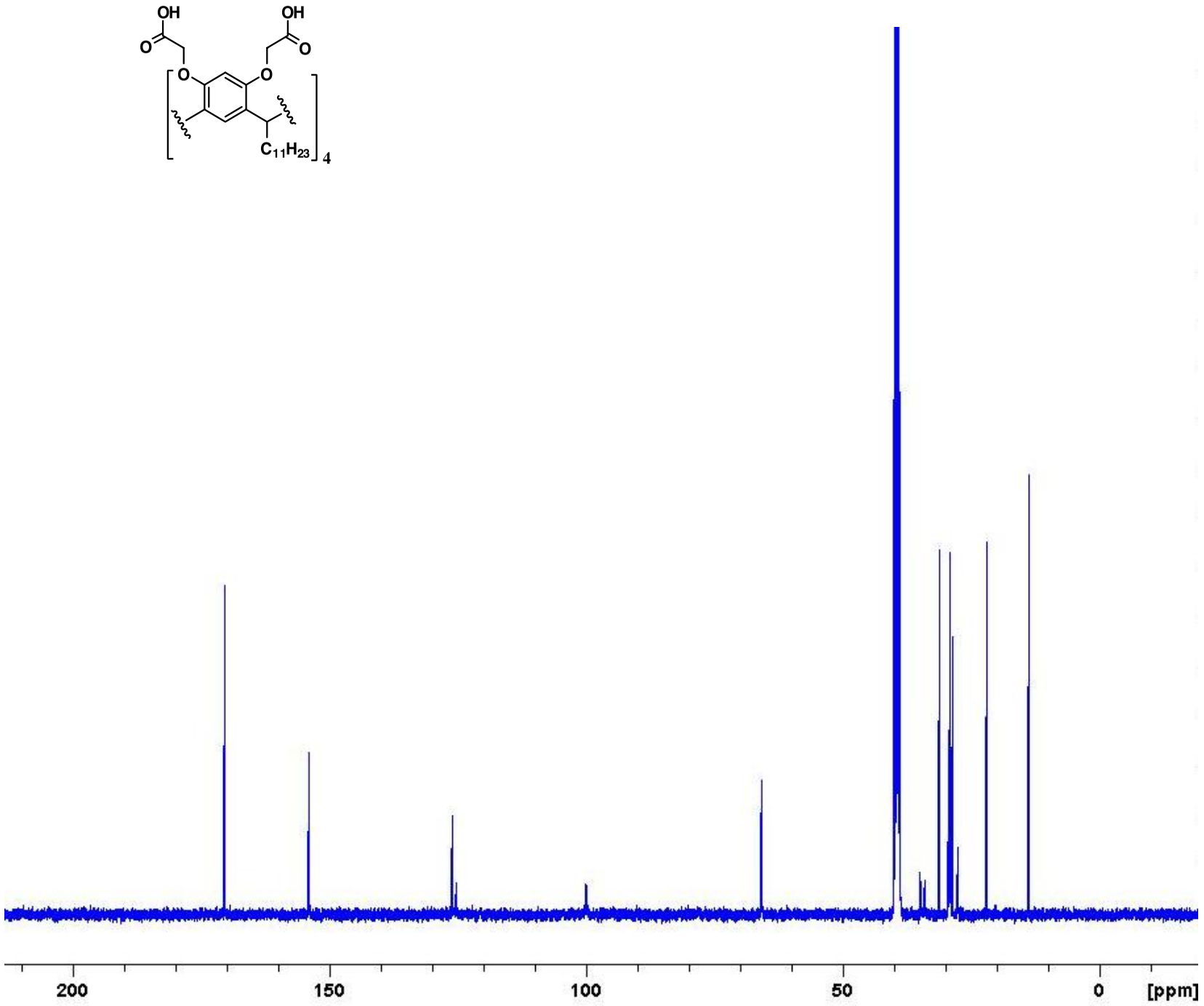
S1: ^{1}H NMR spectrum of 4h in DMSO-d_6 (600 MHz)



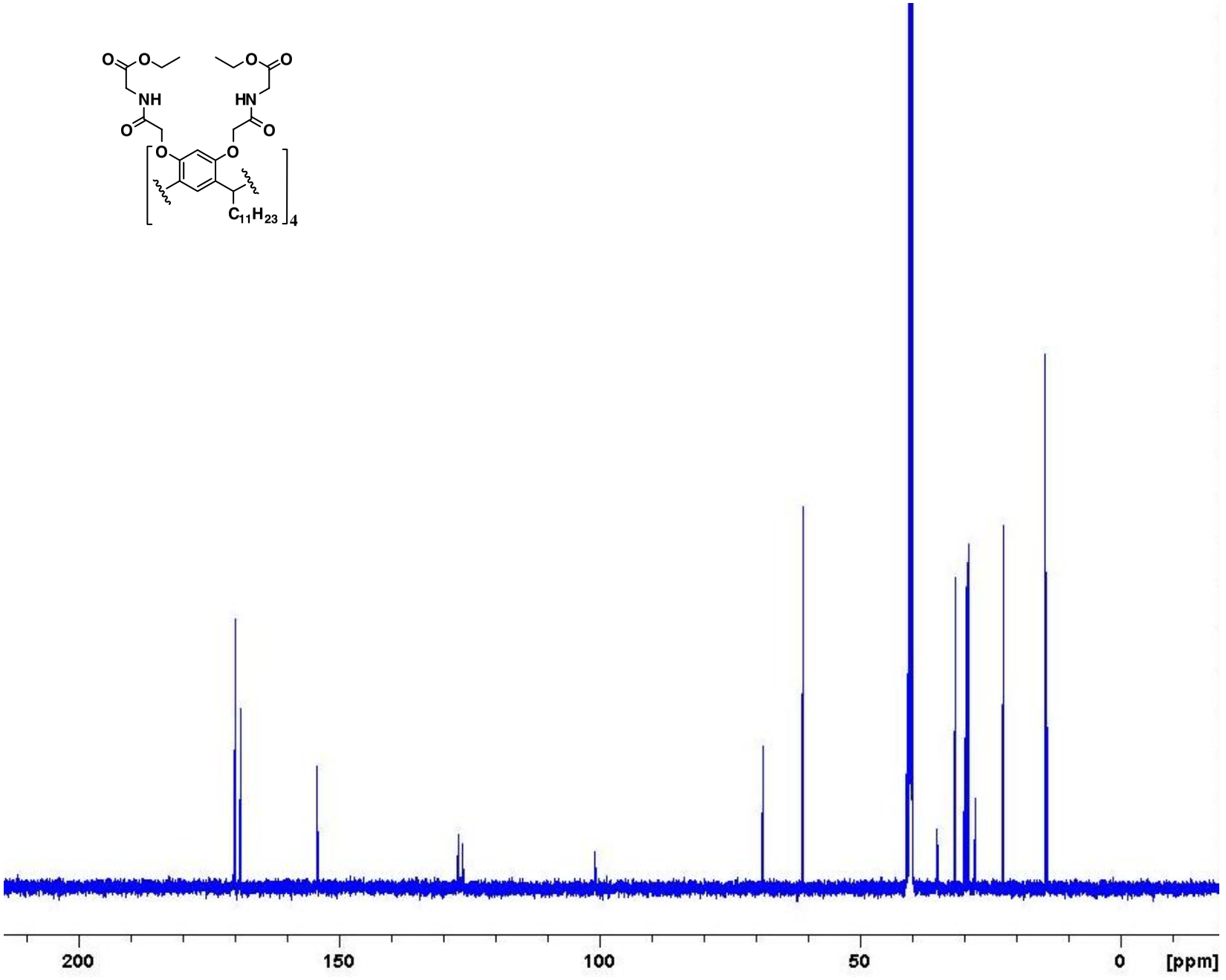
S1: ^{1}H NMR spectrum of 4i in DMSO-d_6 (600 MHz)



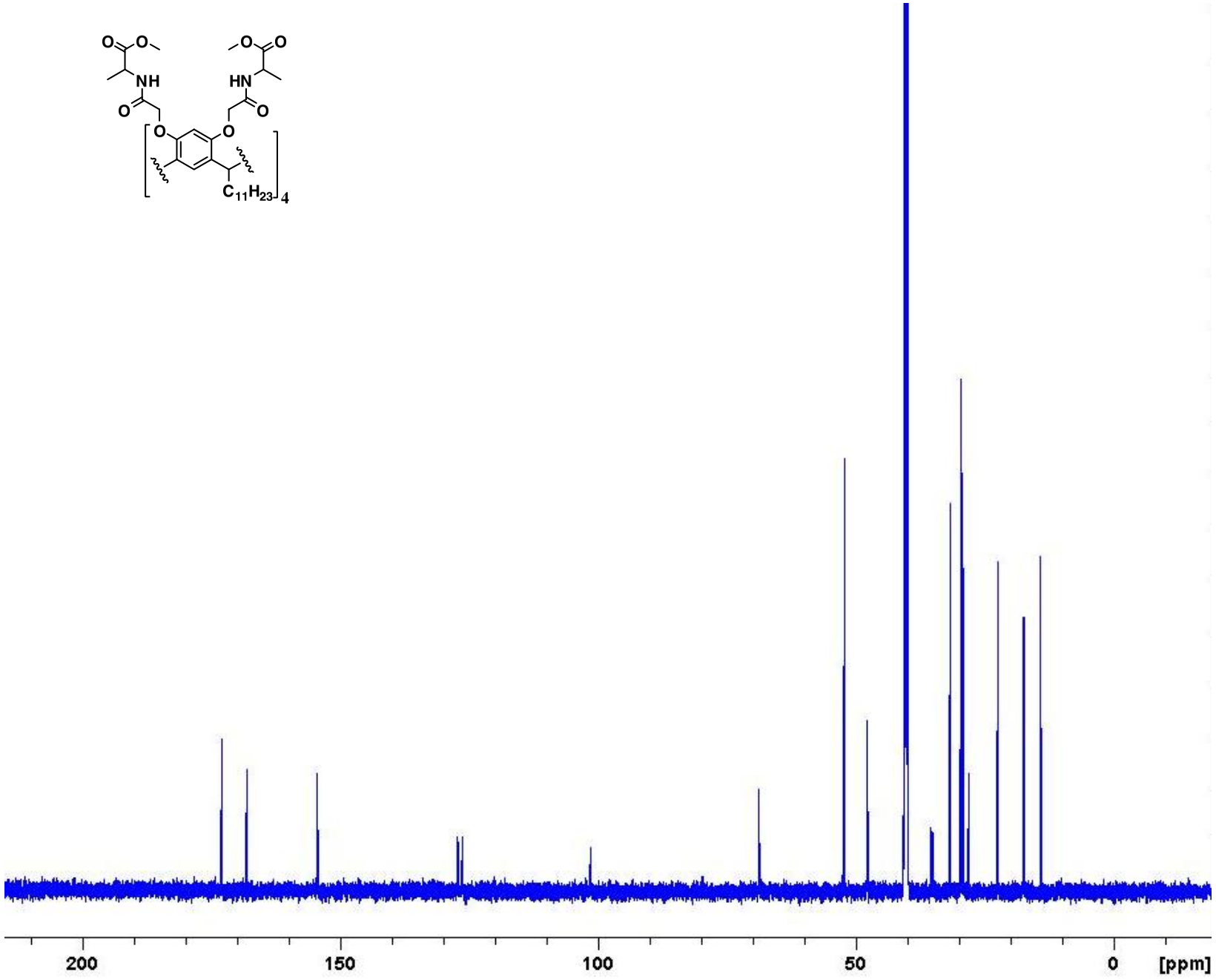
S2: ^{13}C NMR spectrum of 2 in $CDCl_3$ (100 MHz)



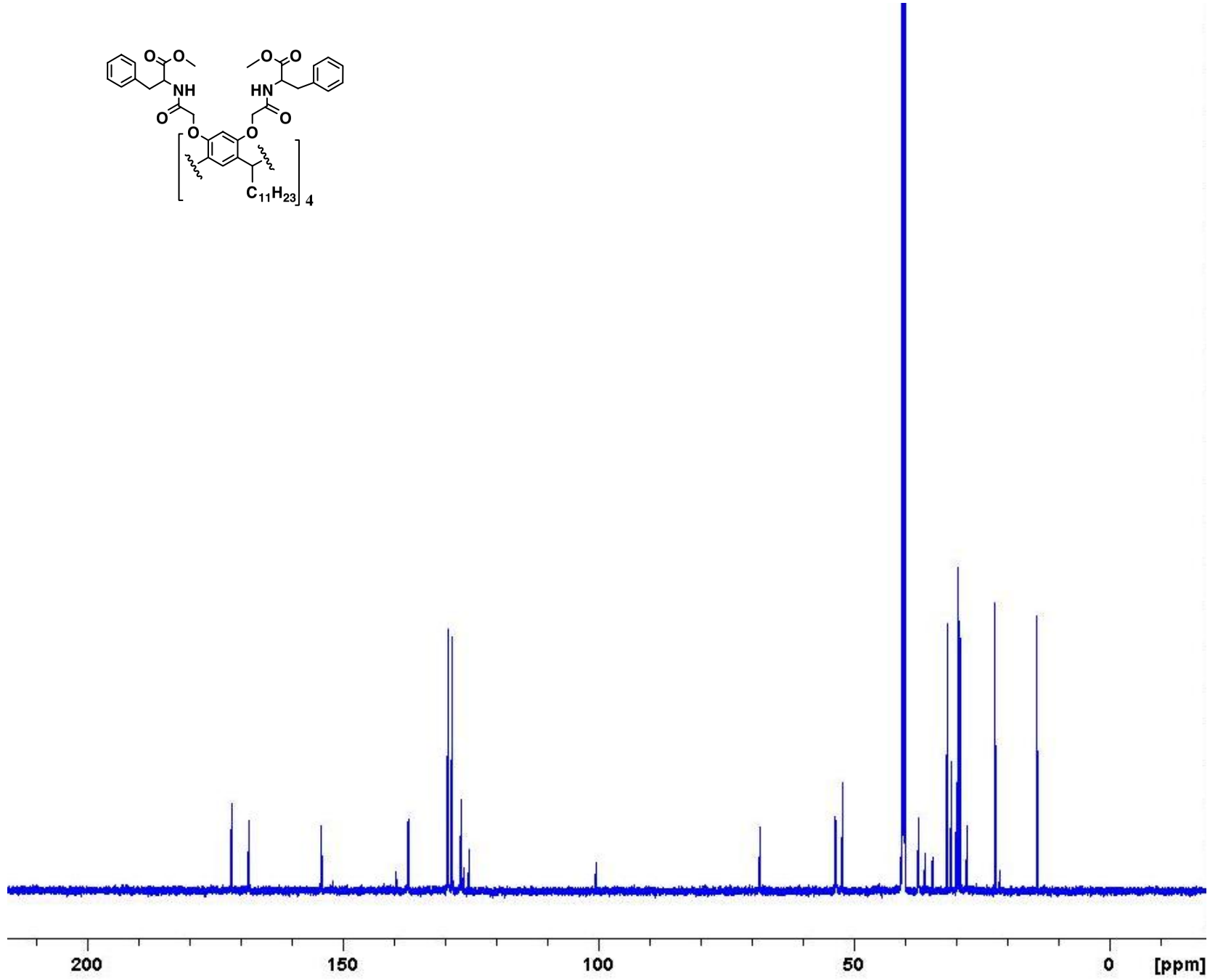
S2: ^{13}C NMR spectrum of 3 in DMSO-d$_6$ (100 MHz)



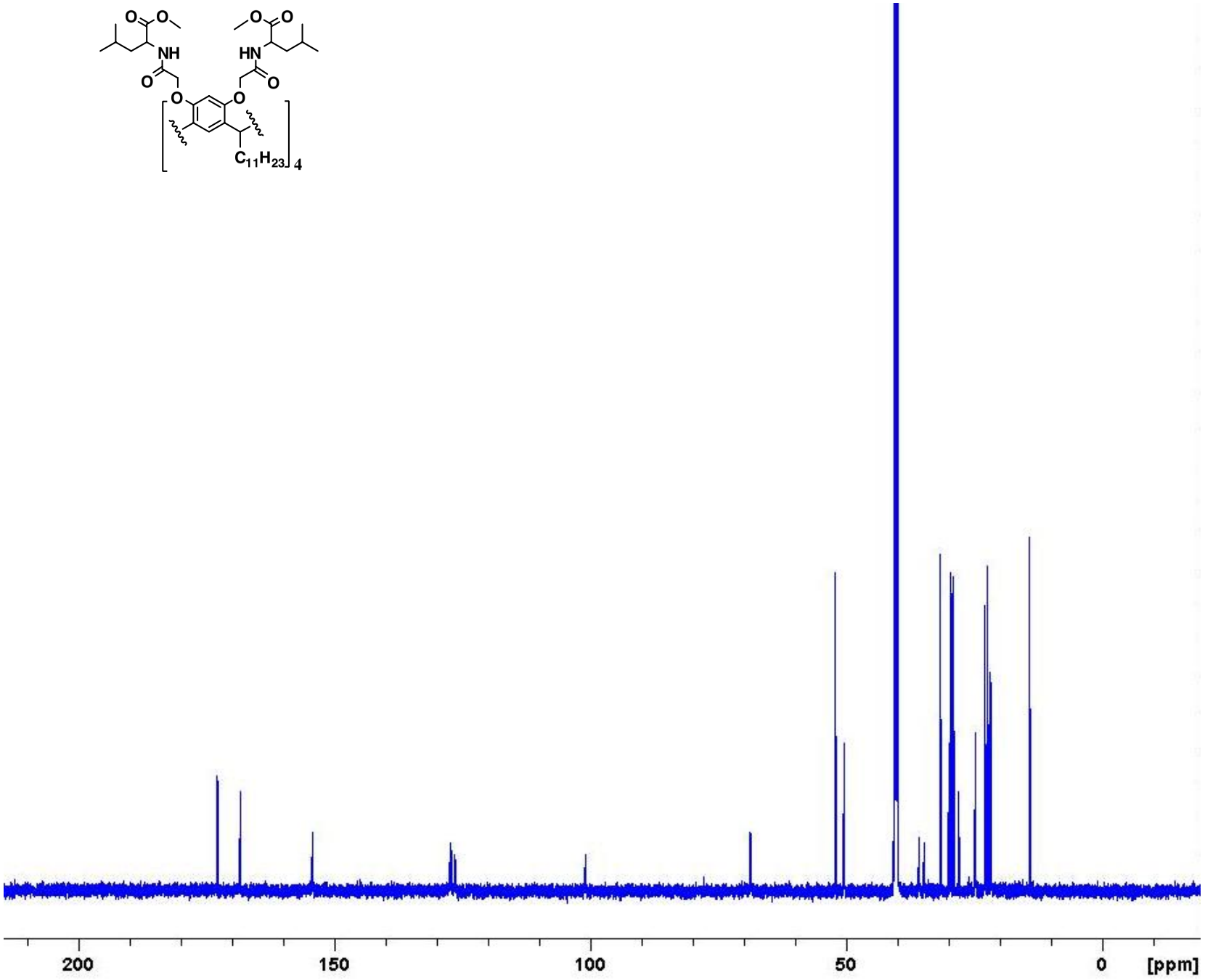
S2: ^{13}C NMR spectrum of 4a in DMSO-d_6 (150 MHz)



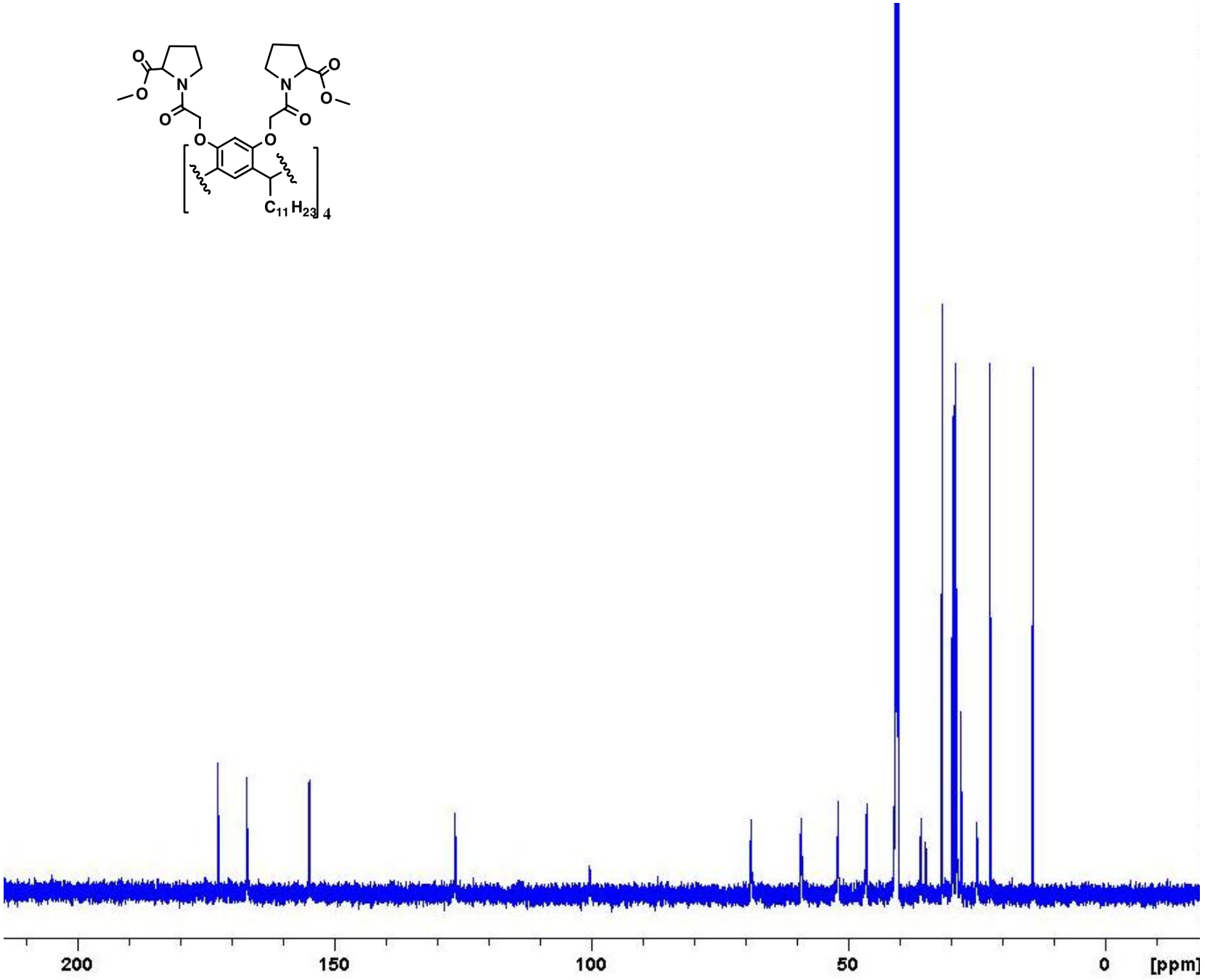
S2: ^{13}C NMR spectrum of 4b in DMSO-d$_6$ (150 MHz)



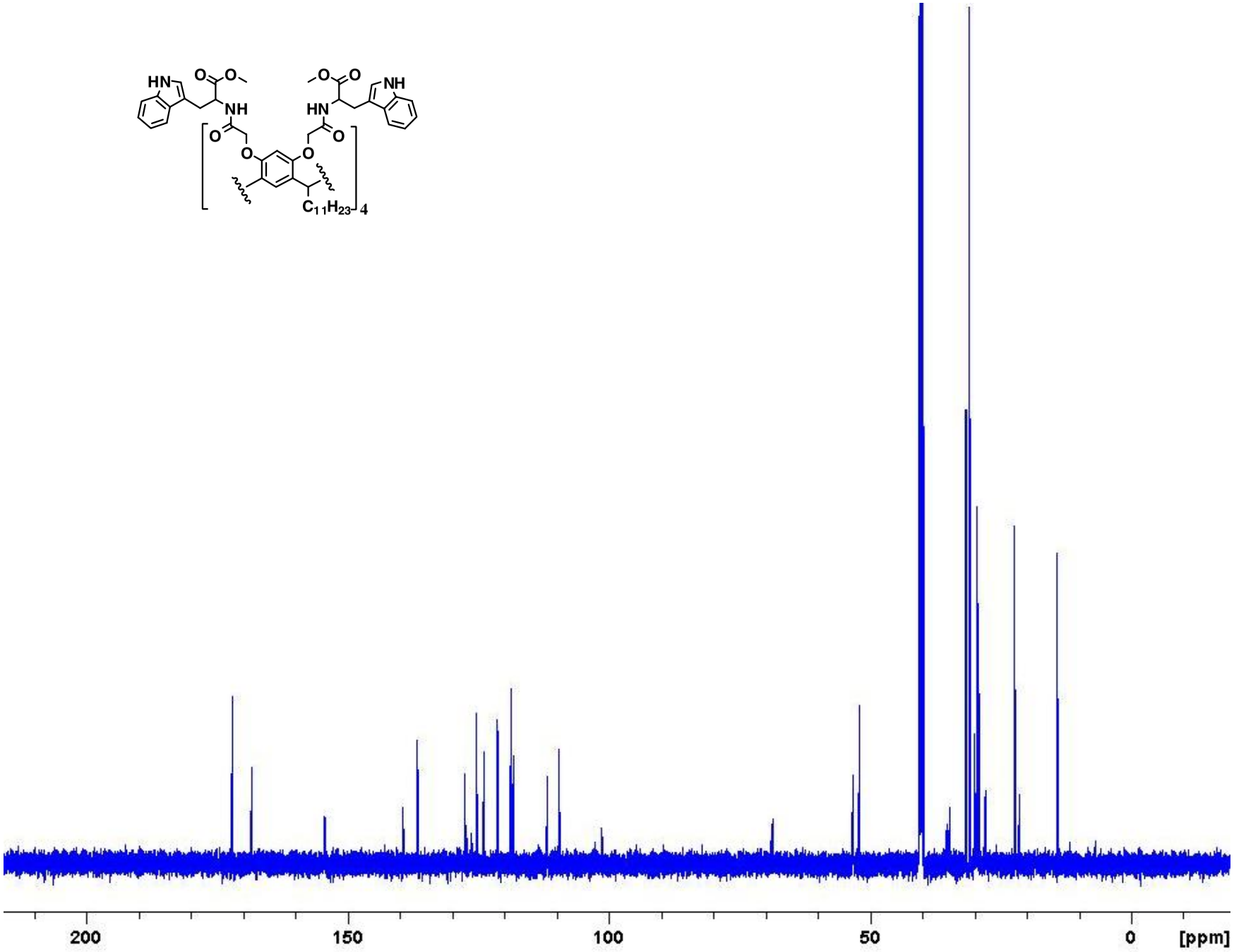
S2: ^{13}C NMR spectrum of 4c in DMSO-d_6 (150 MHz)



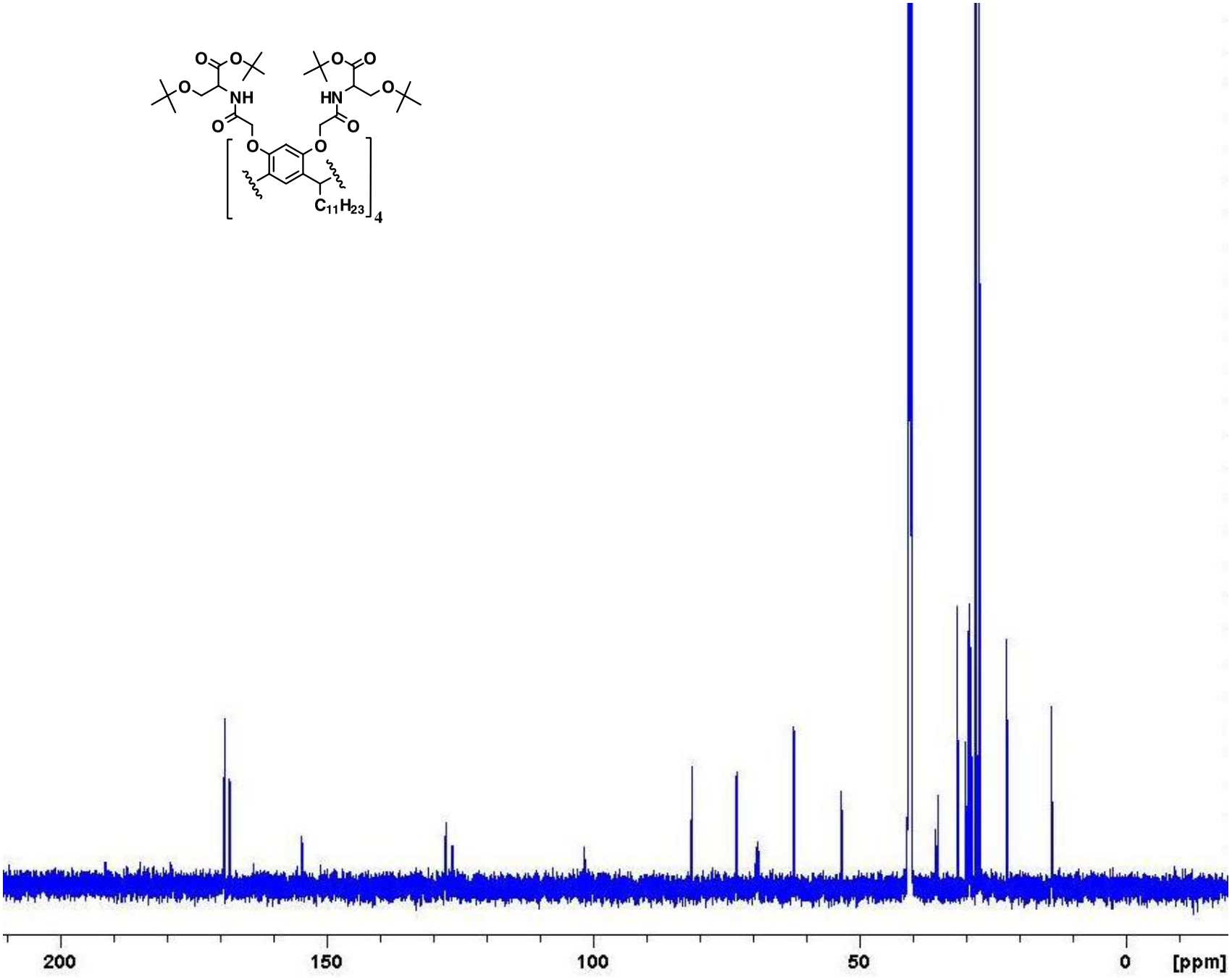
S2: ^{13}C NMR spectrum of 4d in DMSO-d_6 (150 MHz)



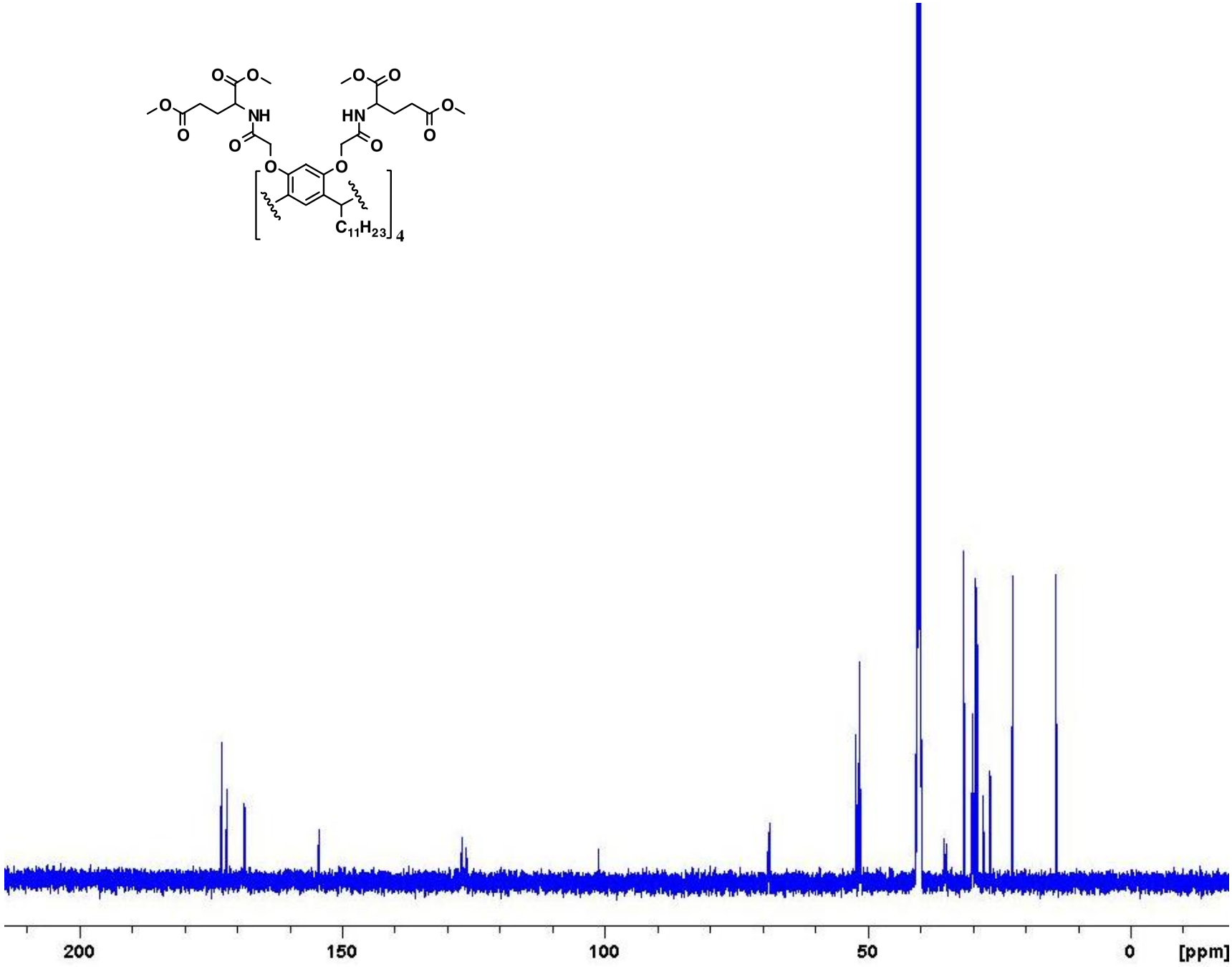
S2: ^{13}C NMR spectrum of 4e in DMSO-d$_6$ (150 MHz)



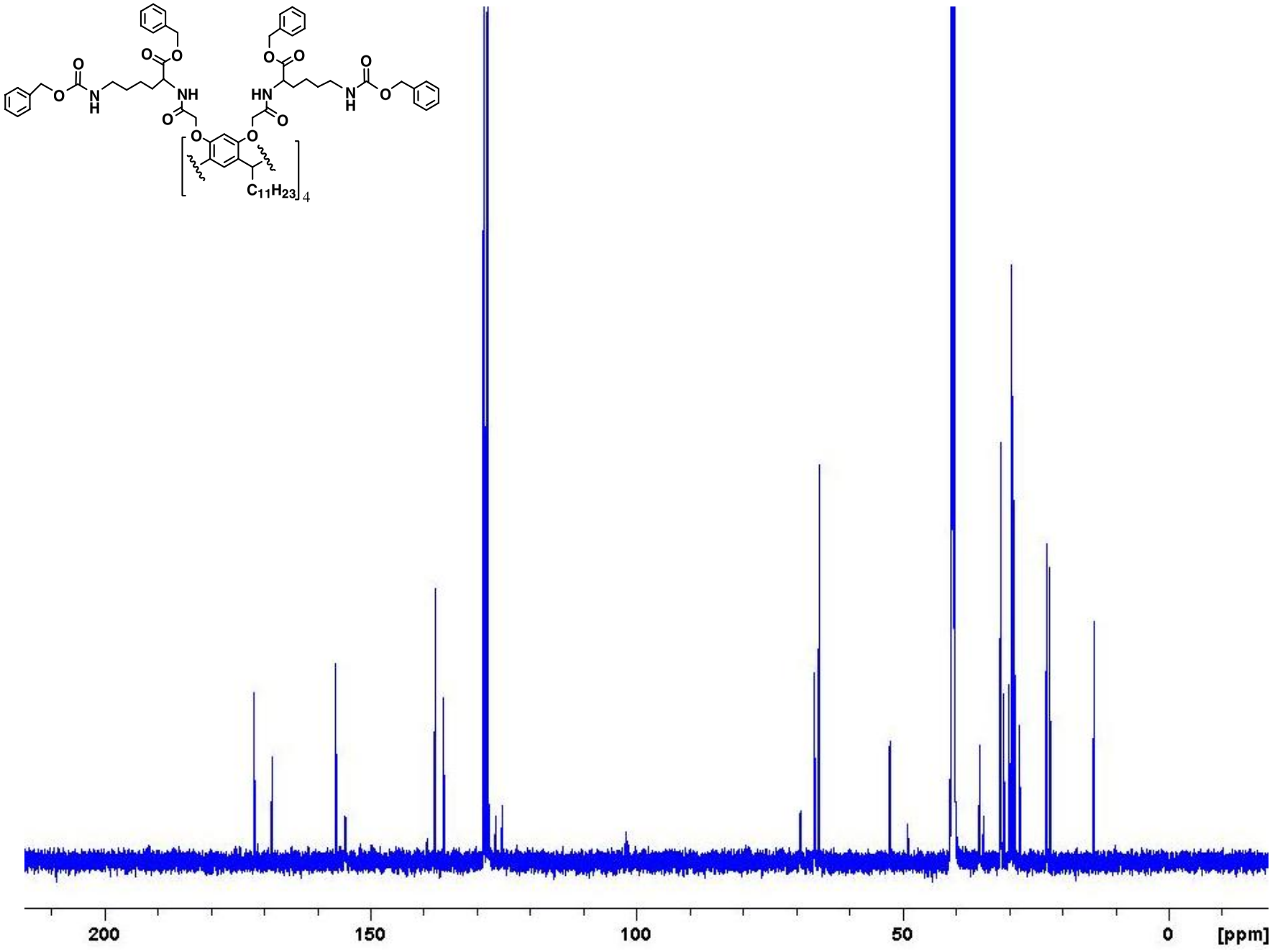
S2: ^{13}C NMR spectrum of 4f in DMSO-d$_6$ (150 MHz)



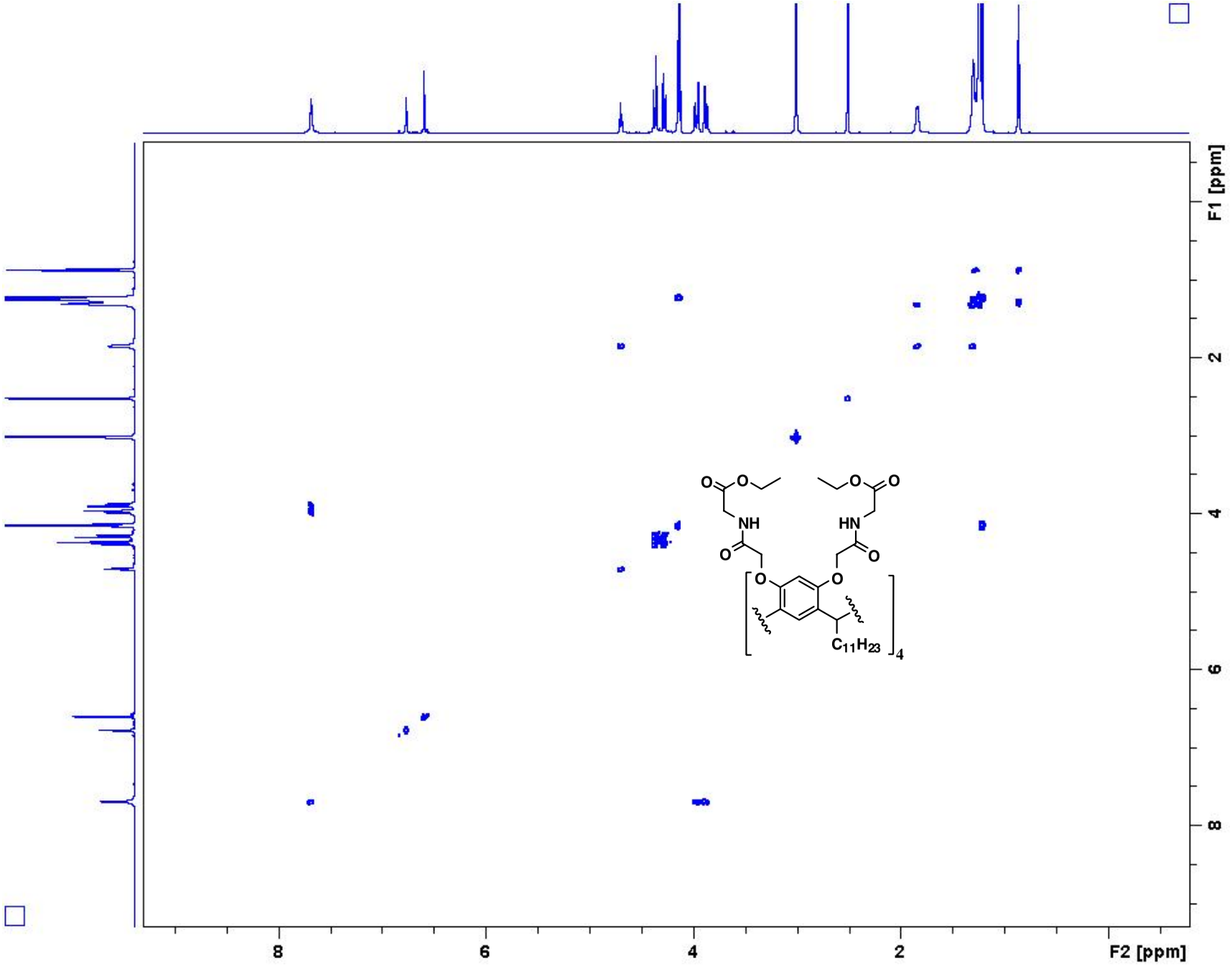
S2: ^{13}C NMR spectrum of 4g in DMSO-d$_6$ (150 MHz)



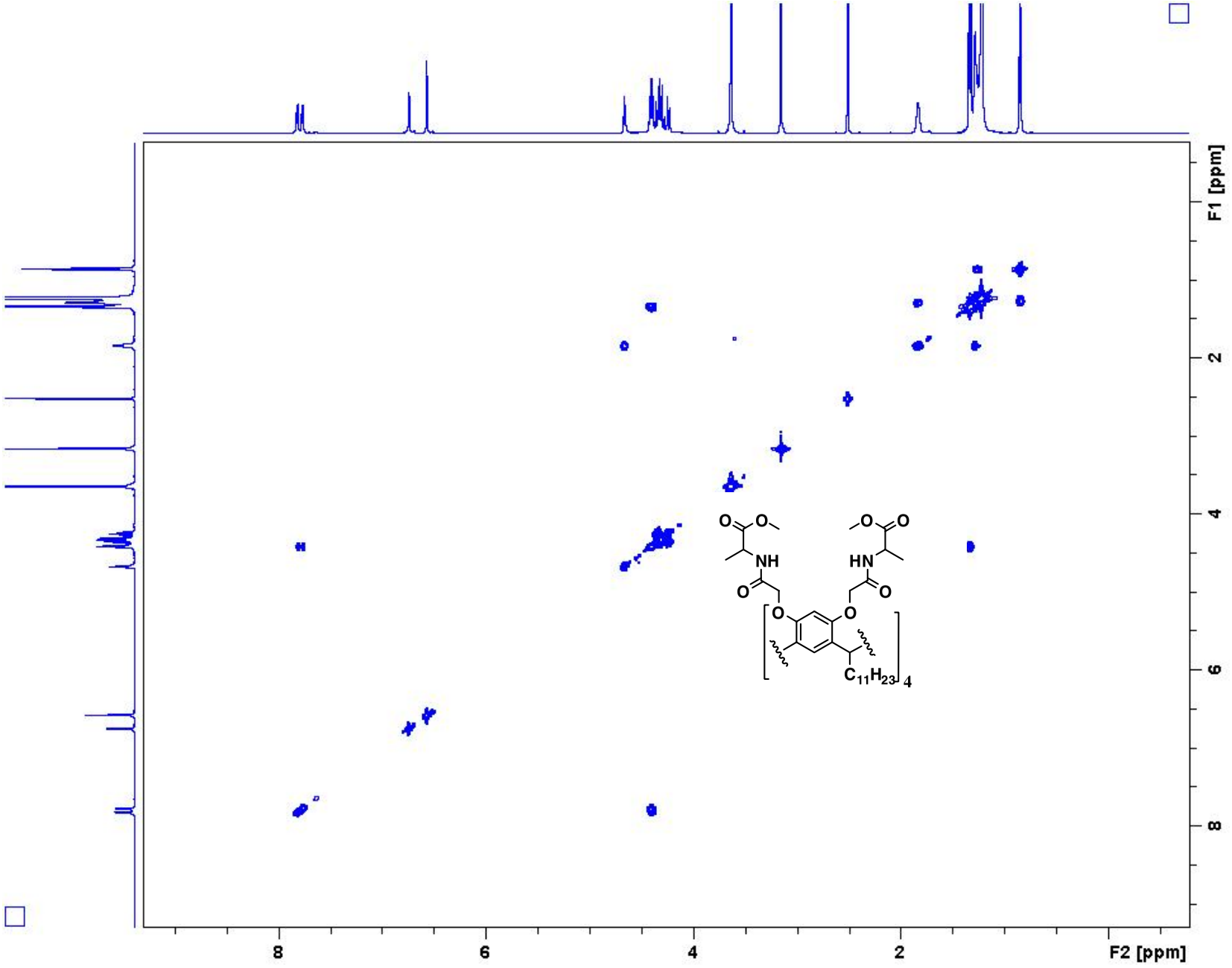
S2: ^{13}C NMR spectrum of 4h in DMSO-d$_6$ (150 MHz)



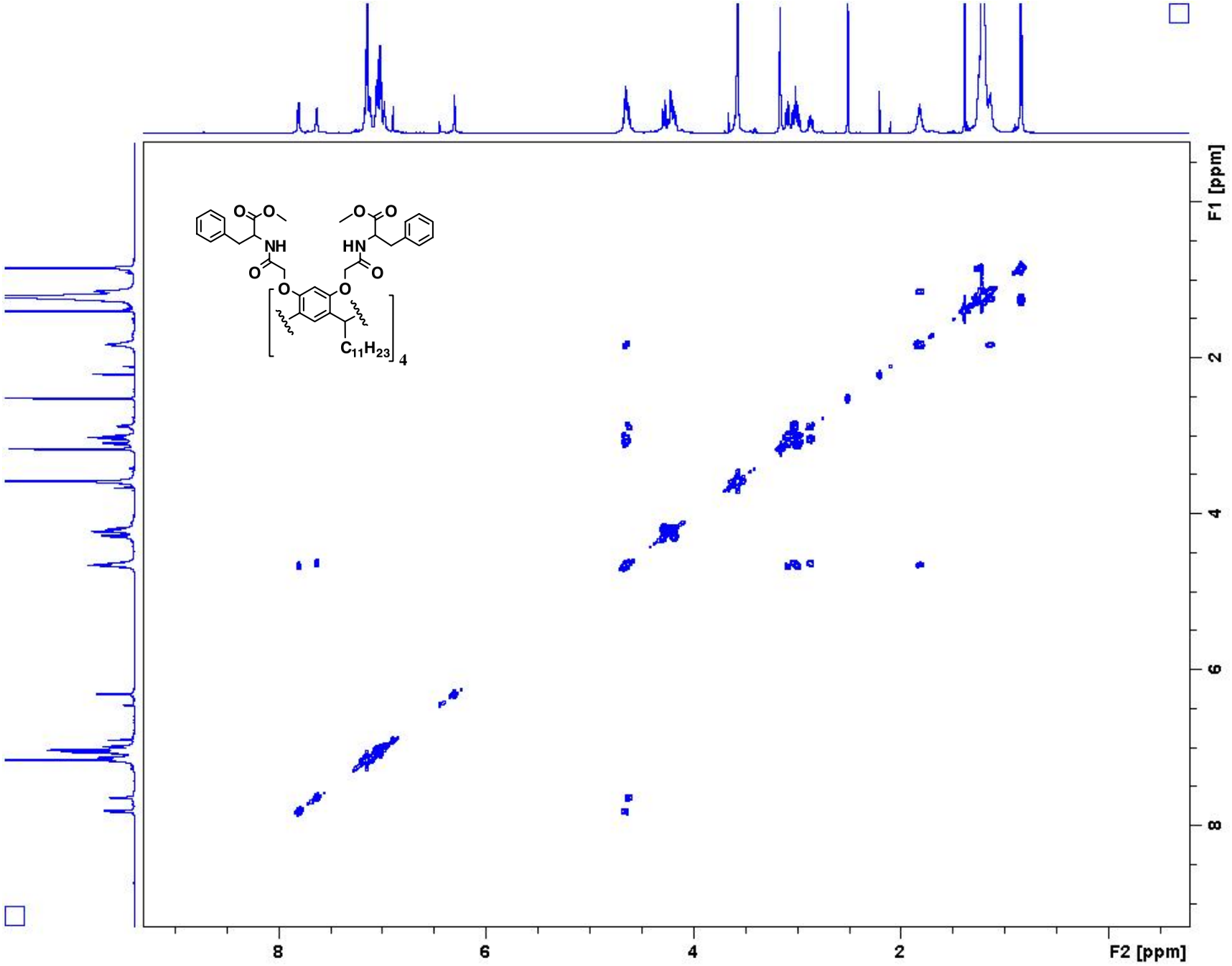
S2: ^{13}C NMR spectrum of 4i in DMSO-d$_6$ (150 MHz)



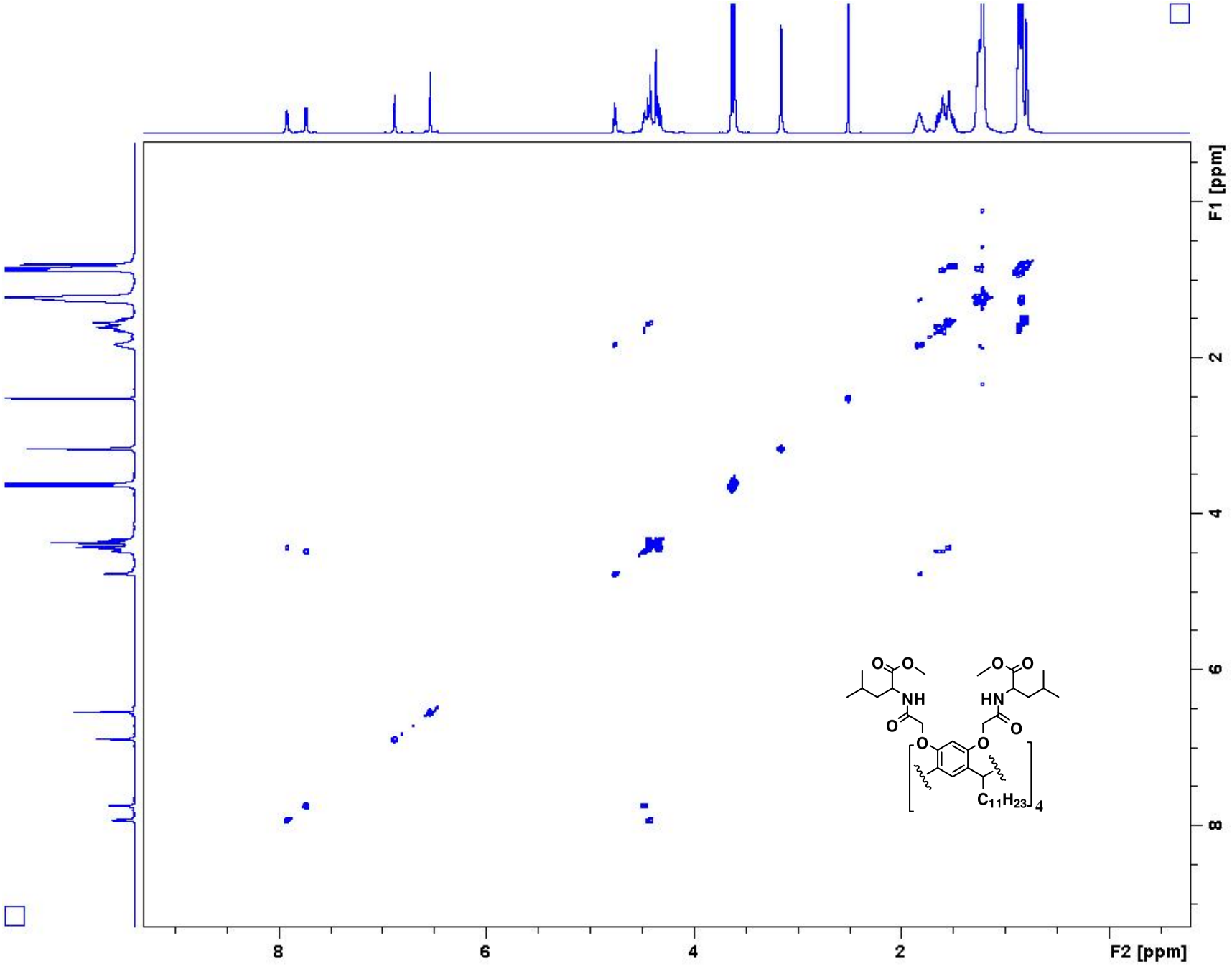
S3: COSY NMR spectrum of 4a in DMSO-d_6 (600 MHz)



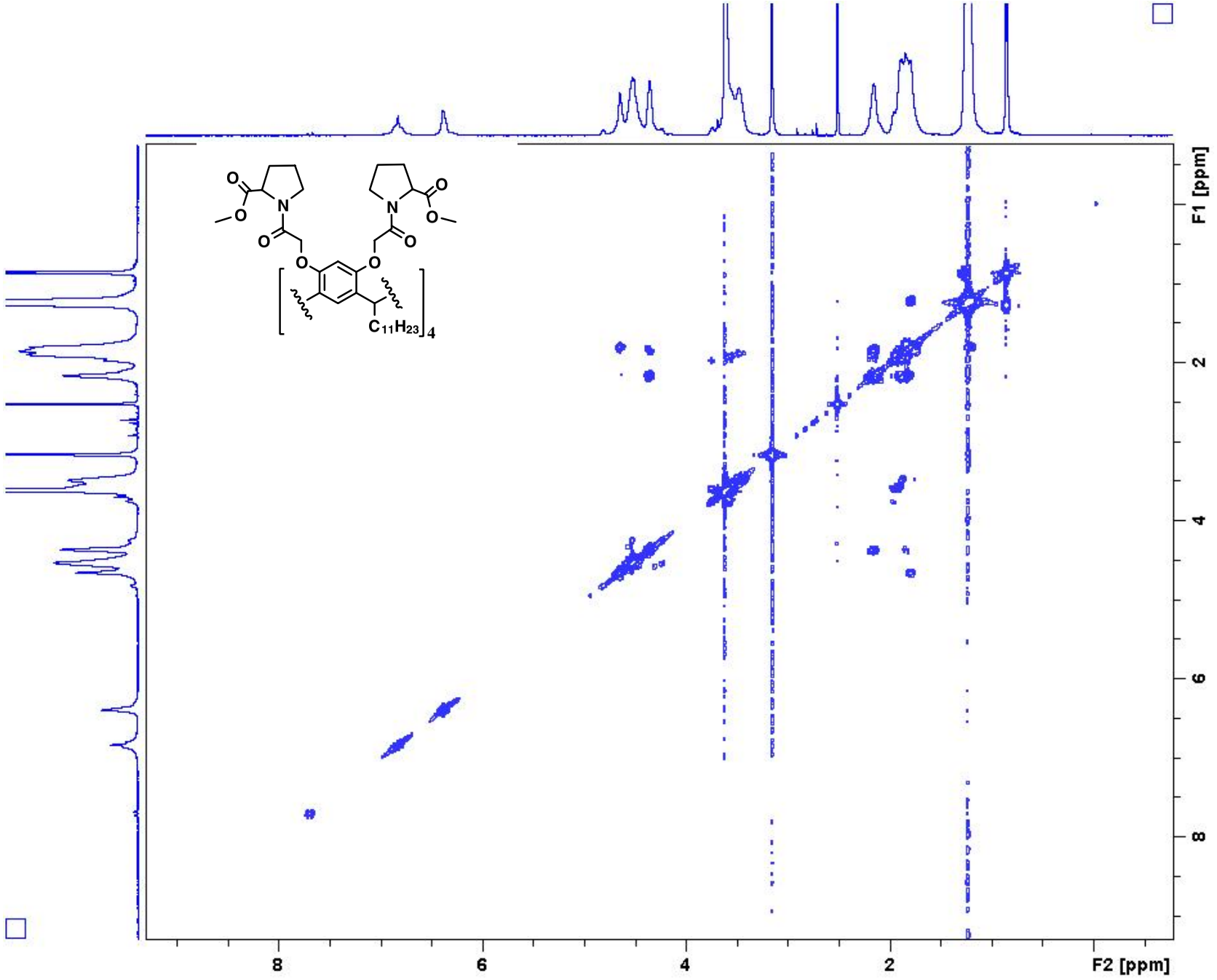
S3: COSY NMR spectrum of 4b in DMSO-d_6 (600 MHz)



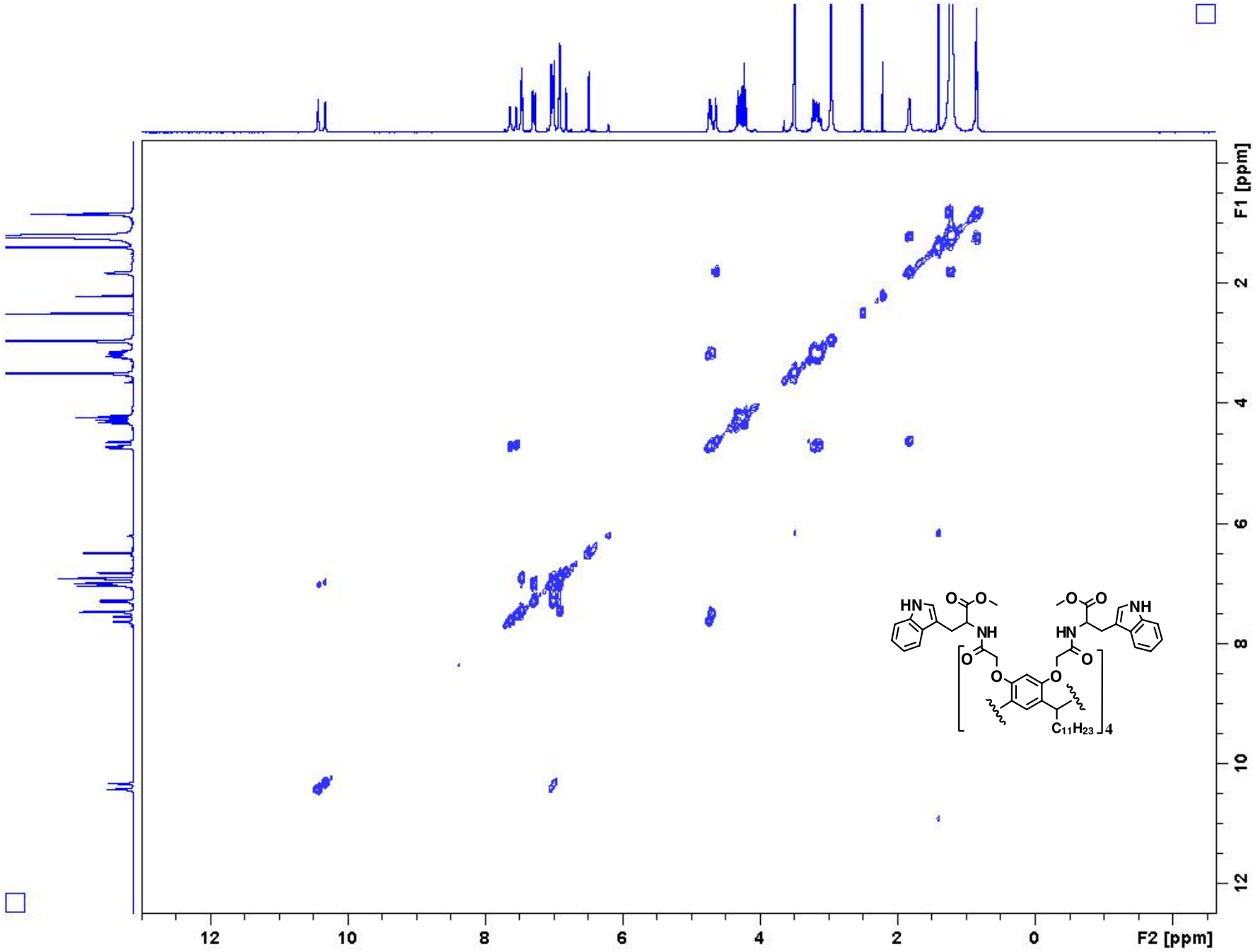
S3: COSY NMR spectrum of 4c in DMSO-d_6 (600 MHz)



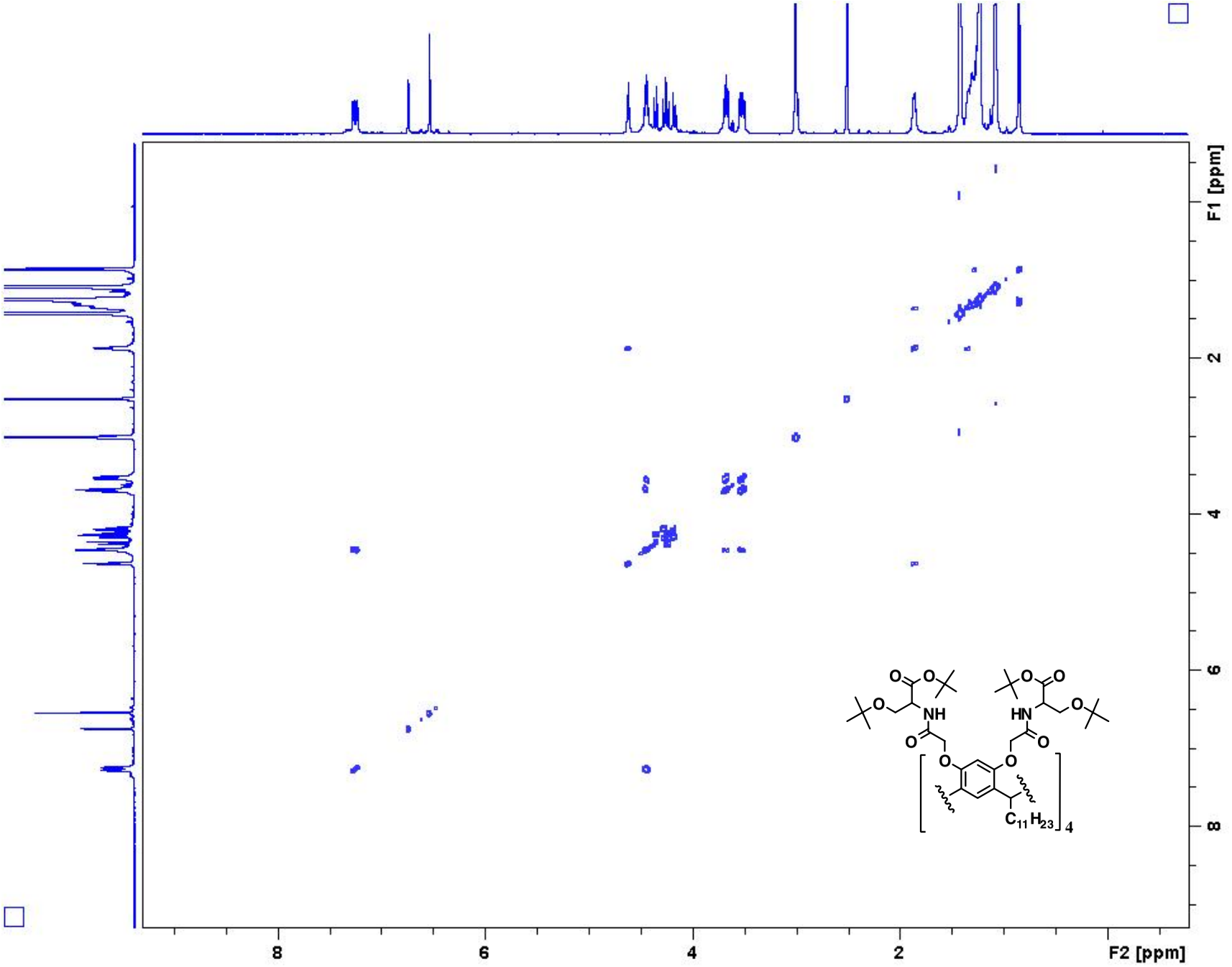
S3: COSY NMR spectrum of 4d in DMSO-d_6 (600 MHz)



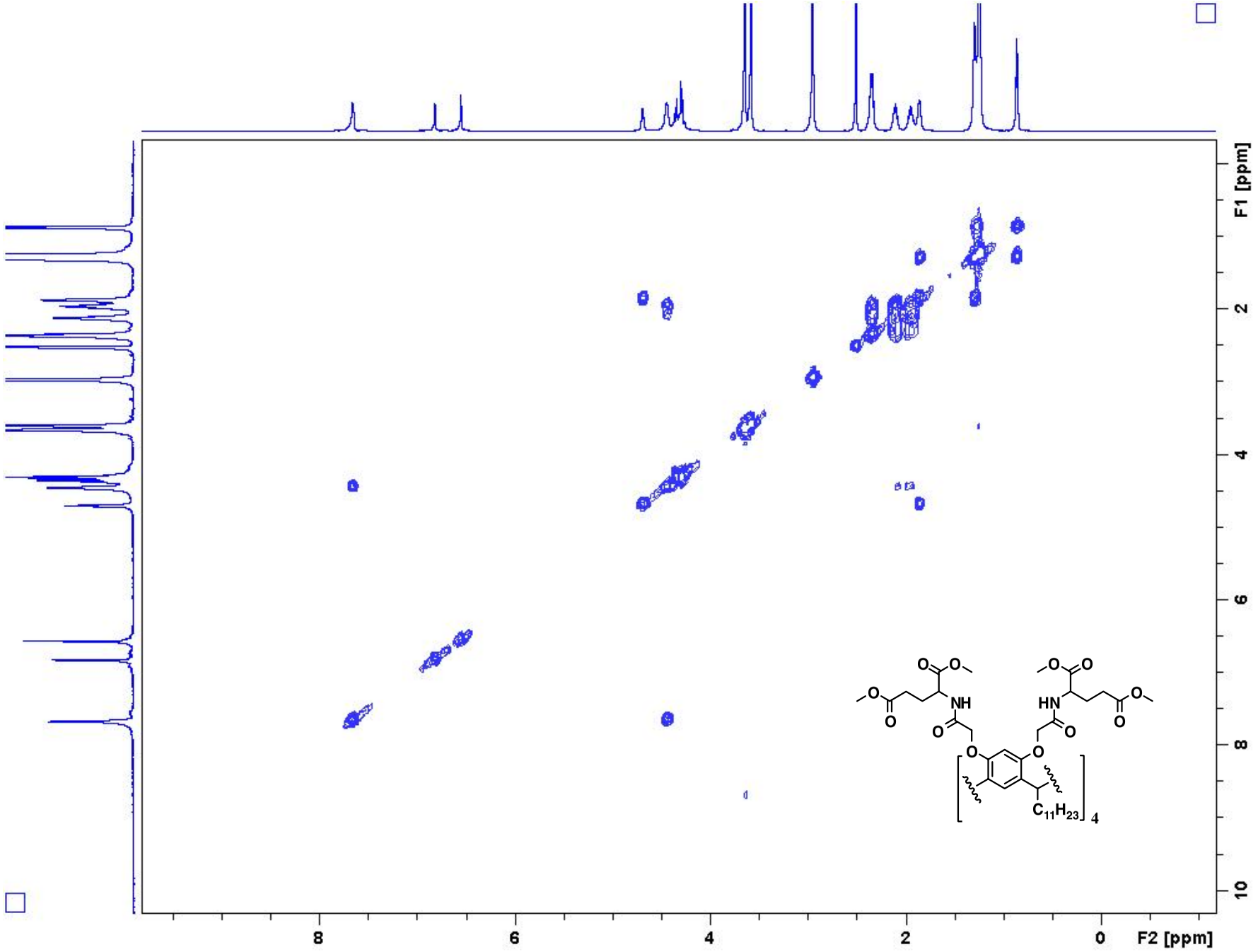
S3: COSY NMR spectrum of 4e in DMSO-d_6 (600 MHz)



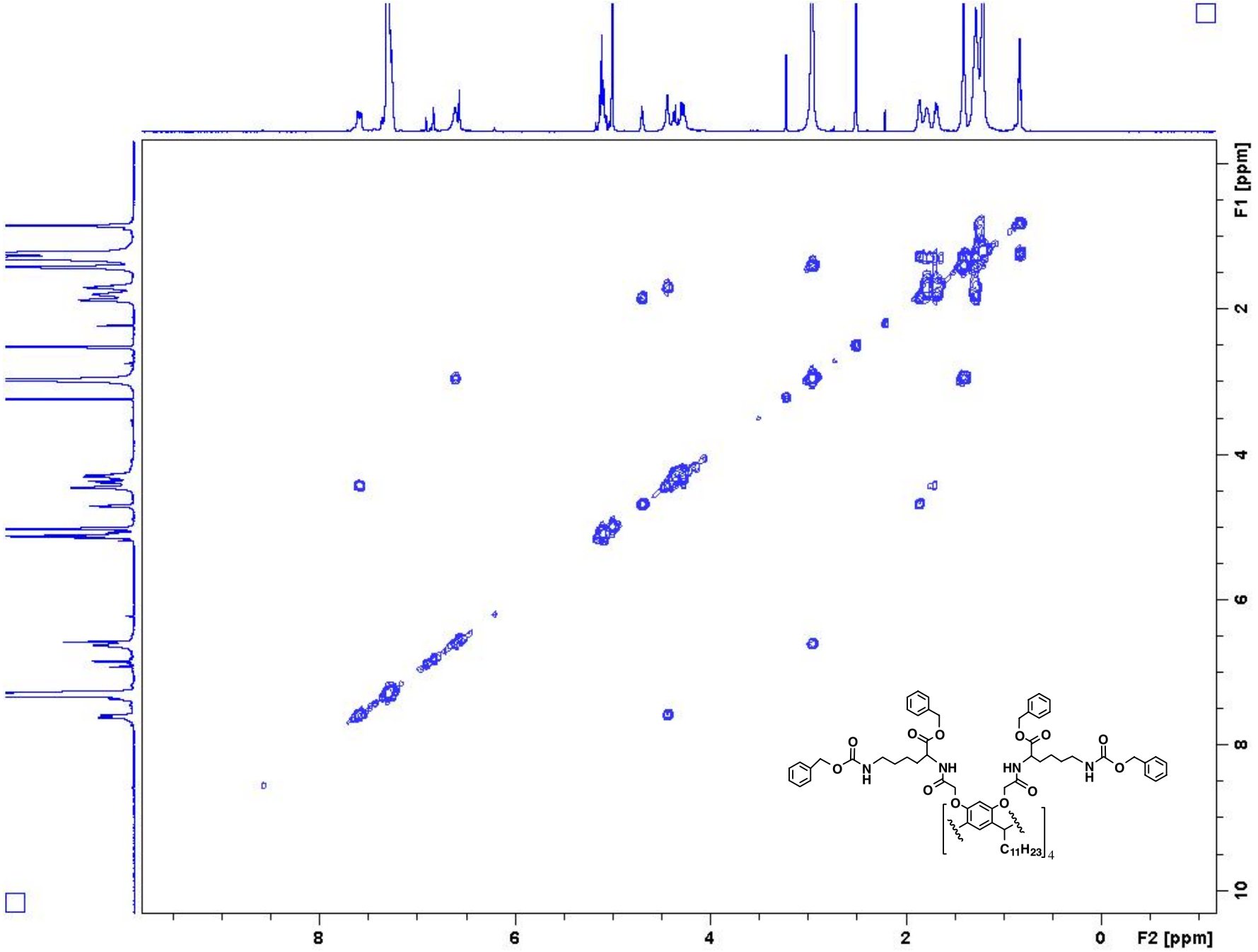
S3: COSY NMR spectrum of 4f in DMSO-d_6 (600 MHz)



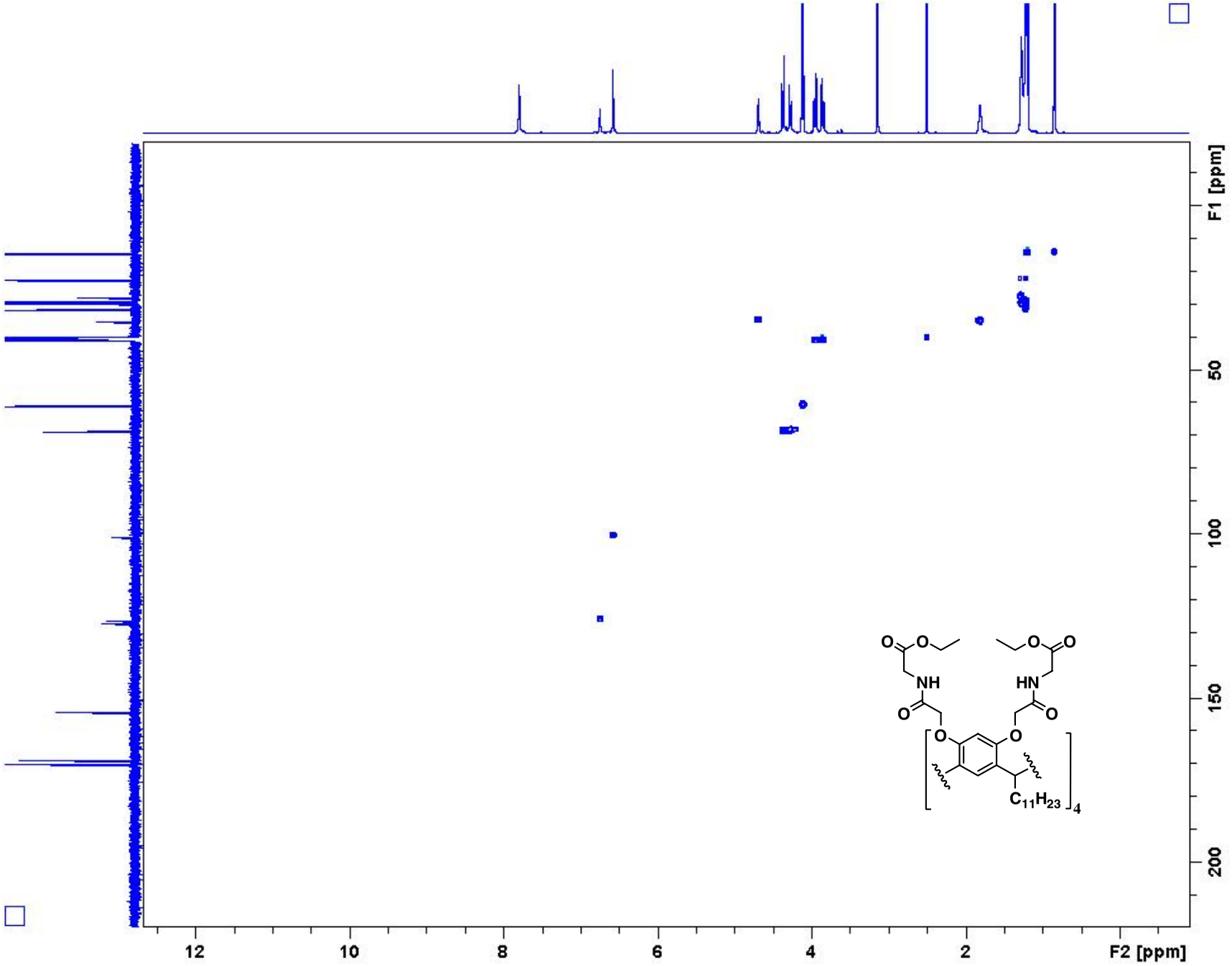
S3: COSY NMR spectrum of 4g in DMSO-d_6 (600 MHz)



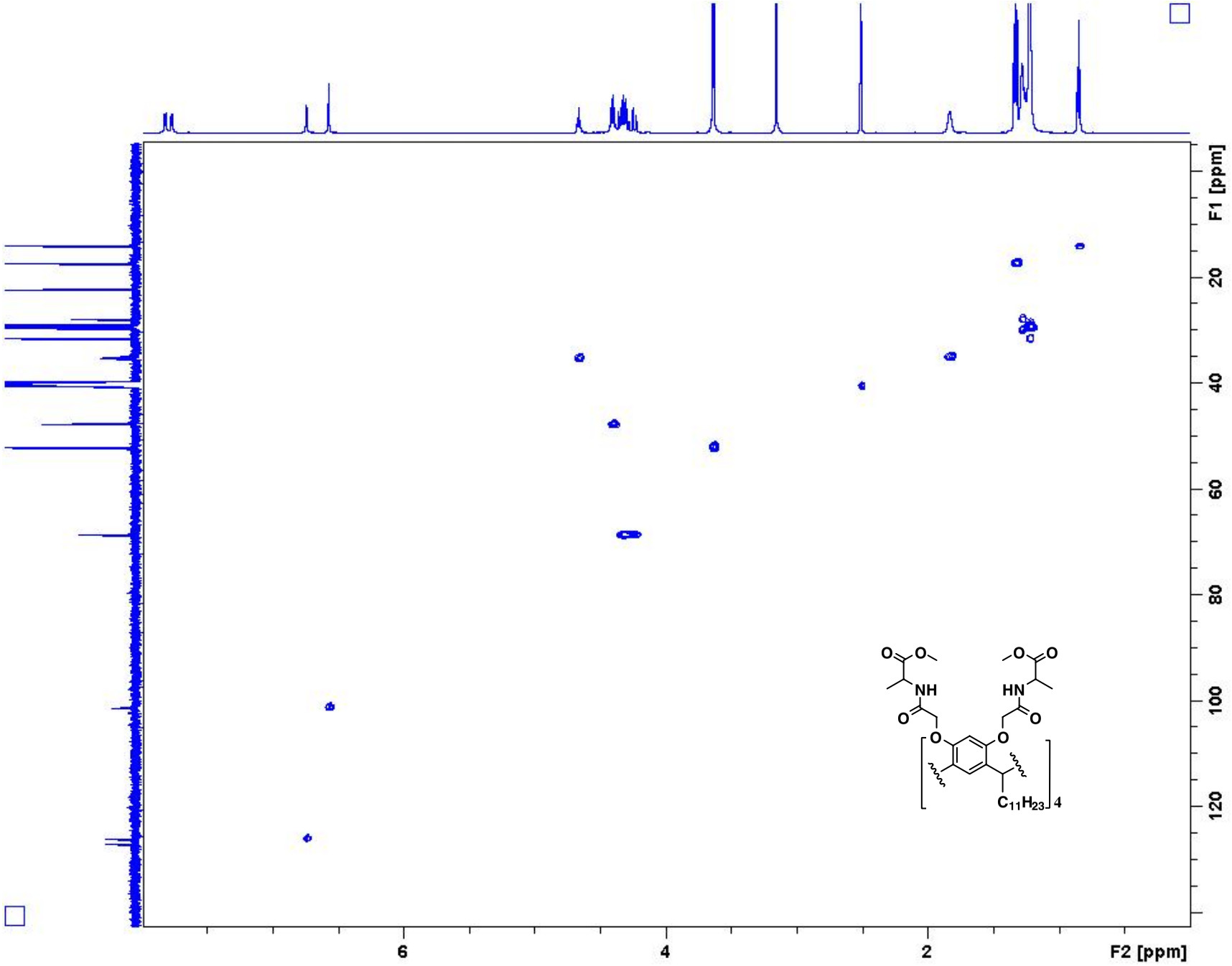
S3: COSY NMR spectrum of 4h in DMSO-d_6 (600 MHz)



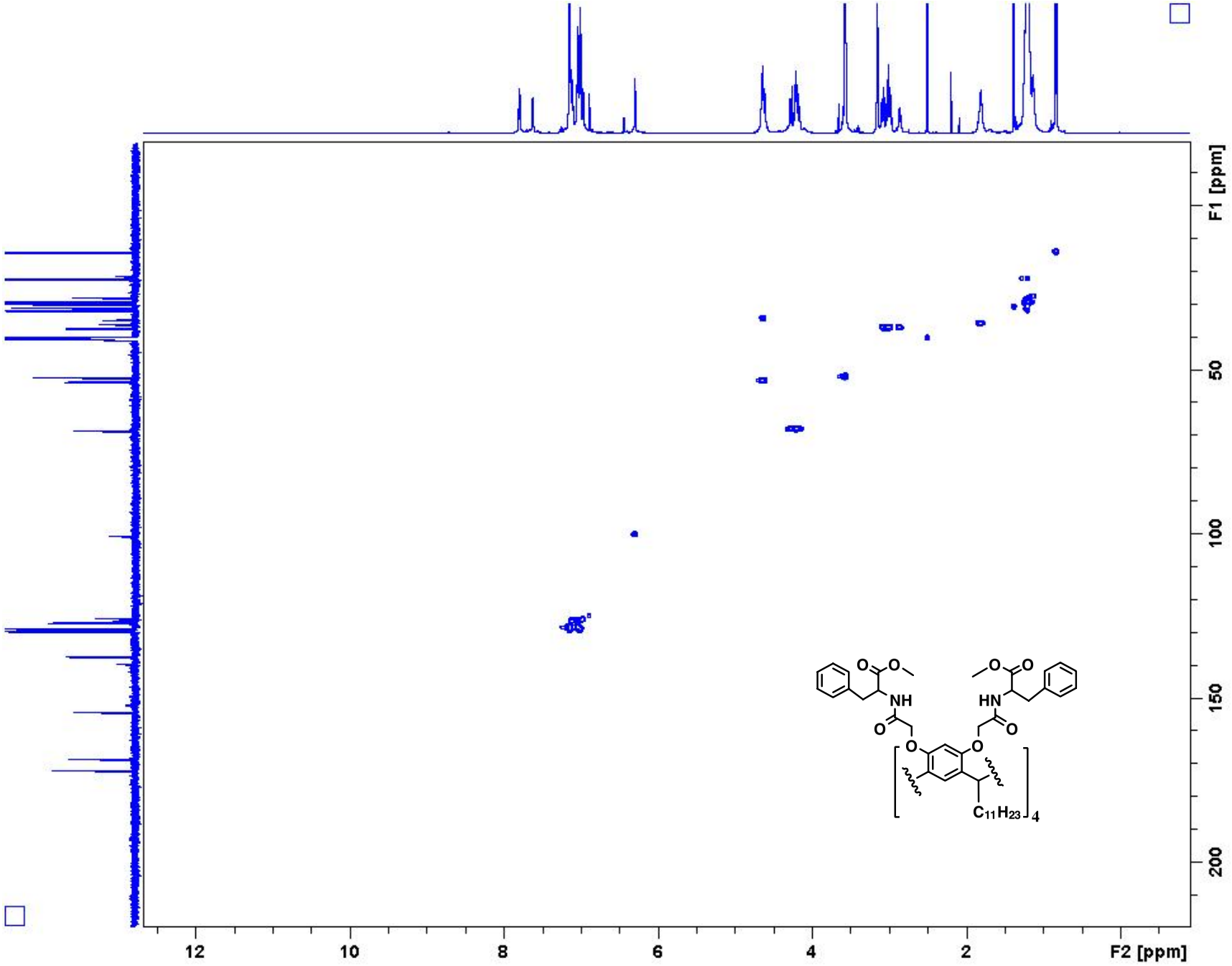
S3: COSY NMR spectrum of 4i in DMSO-d_6 (600 MHz)



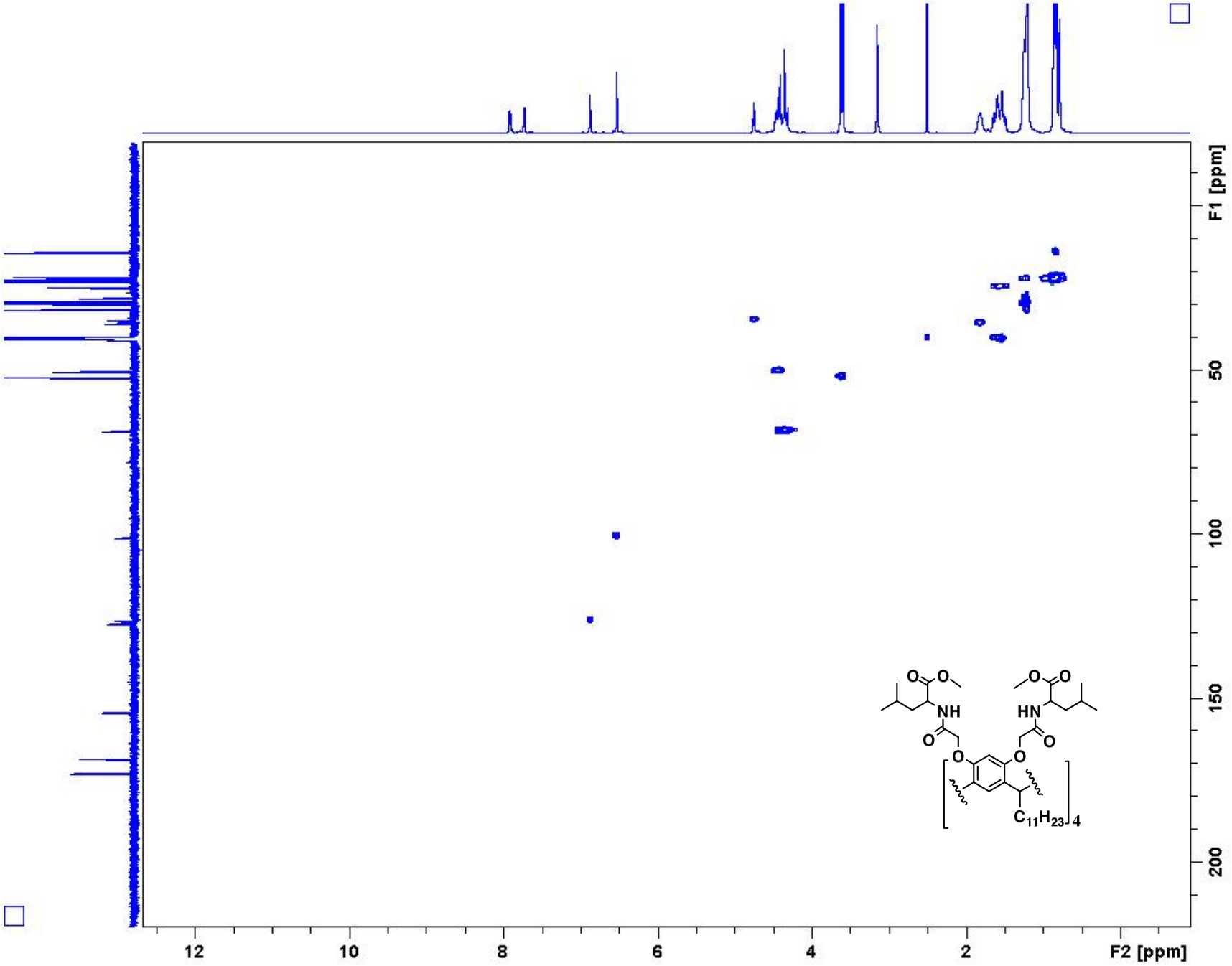
S4: HSQC NMR spectrum of 4a in DMSO-d_6 (600 MHz)



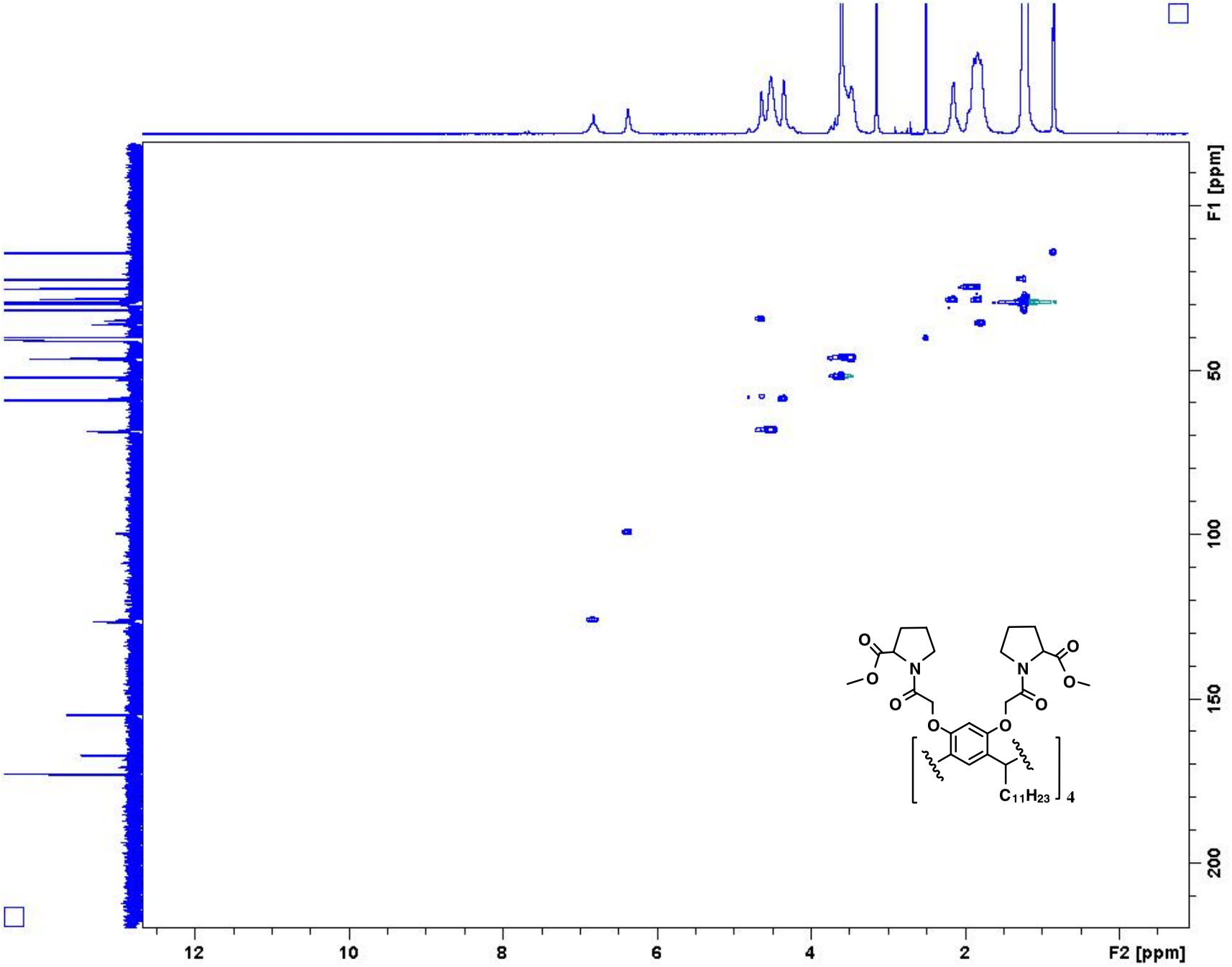
S4: HSQC NMR spectrum of 4b in DMSO-d_6 (600 MHz)



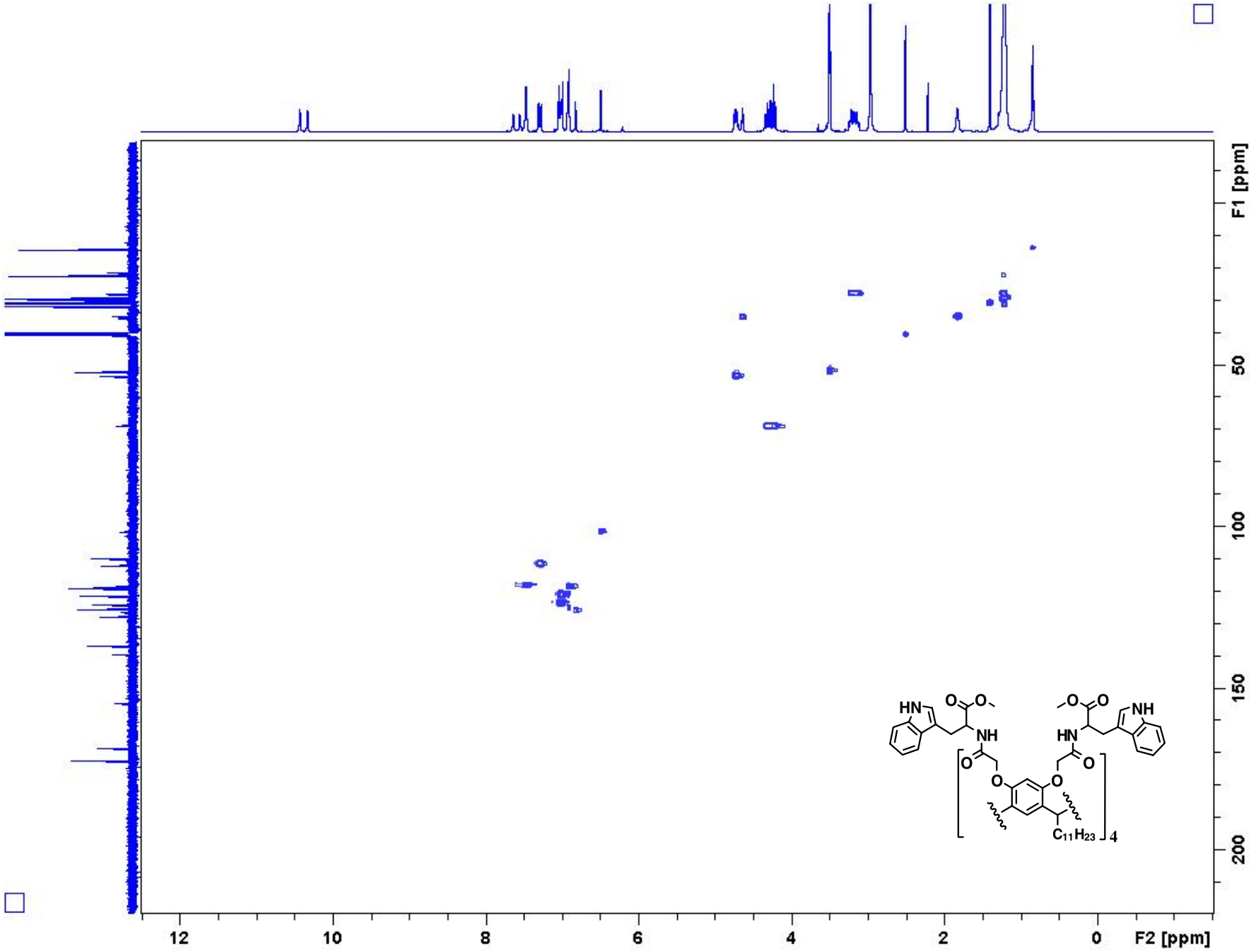
S4: HSQC NMR spectrum of 4c in DMSO-d_6 (600 MHz)



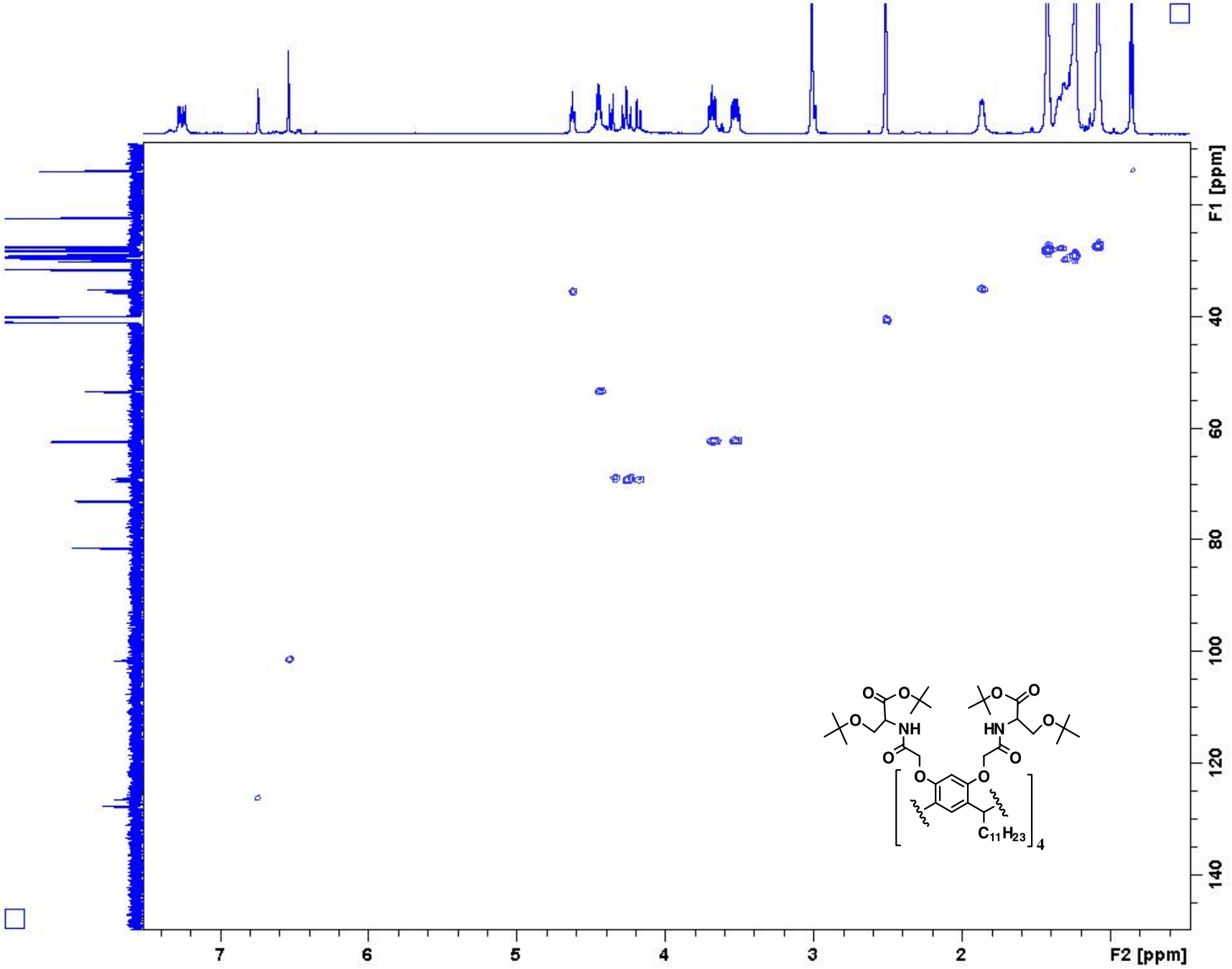
S4: HSQC NMR spectrum of 4d in DMSO-d_6 (600 MHz)



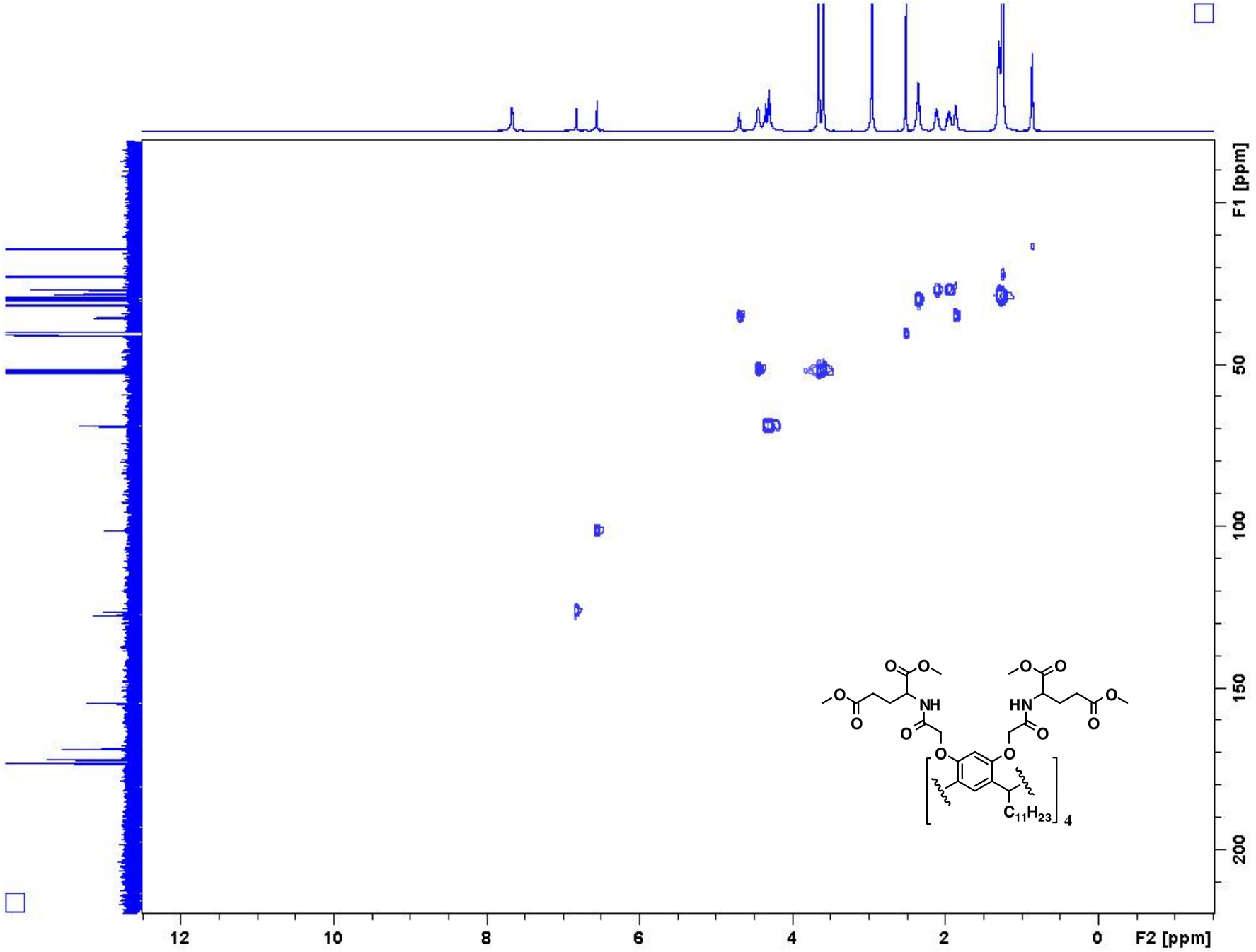
S4: HSQC NMR spectrum of 4e in DMSO-d_6 (600 MHz)



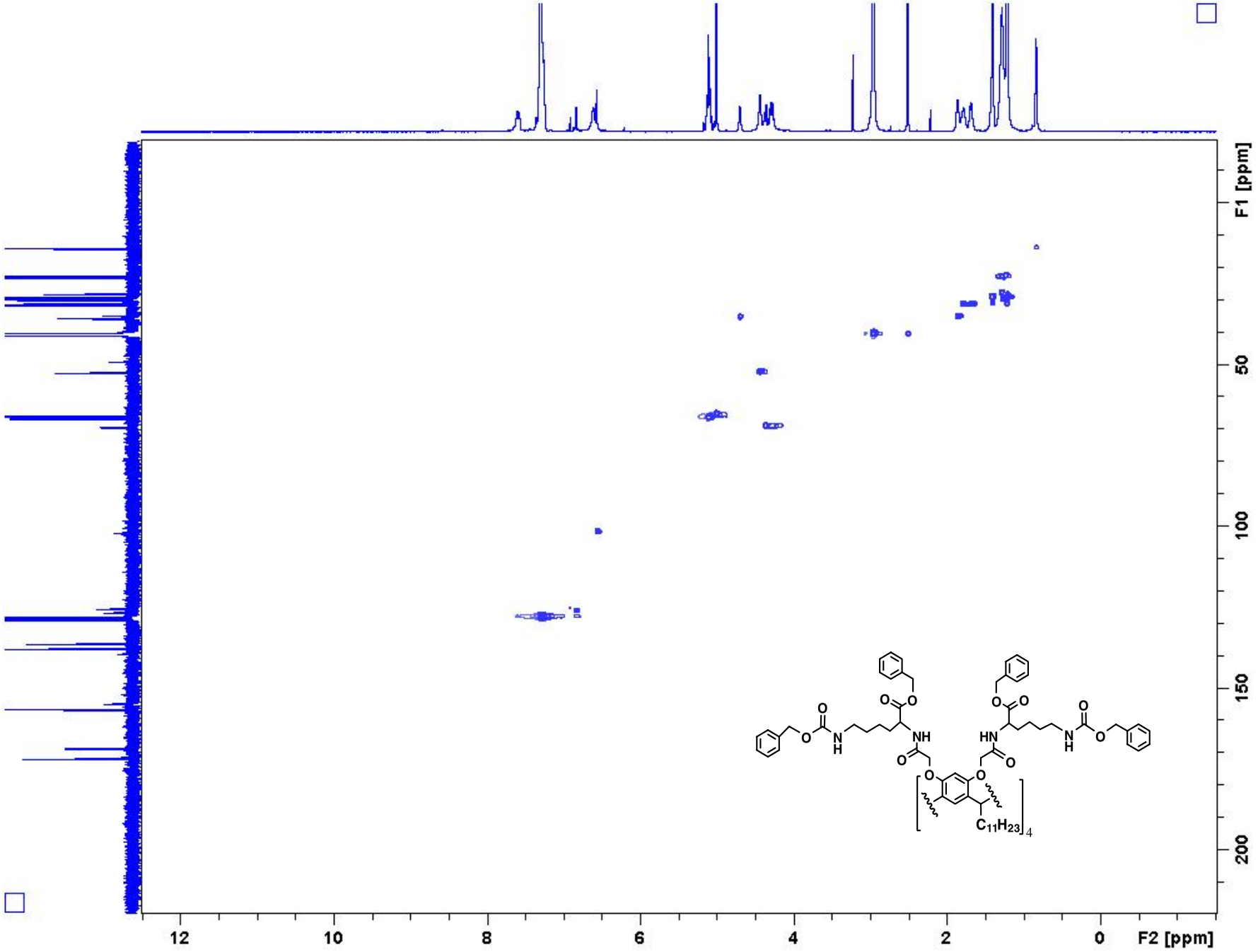
S4: HSQC NMR spectrum of 4f in DMSO-d_6 (600 MHz)



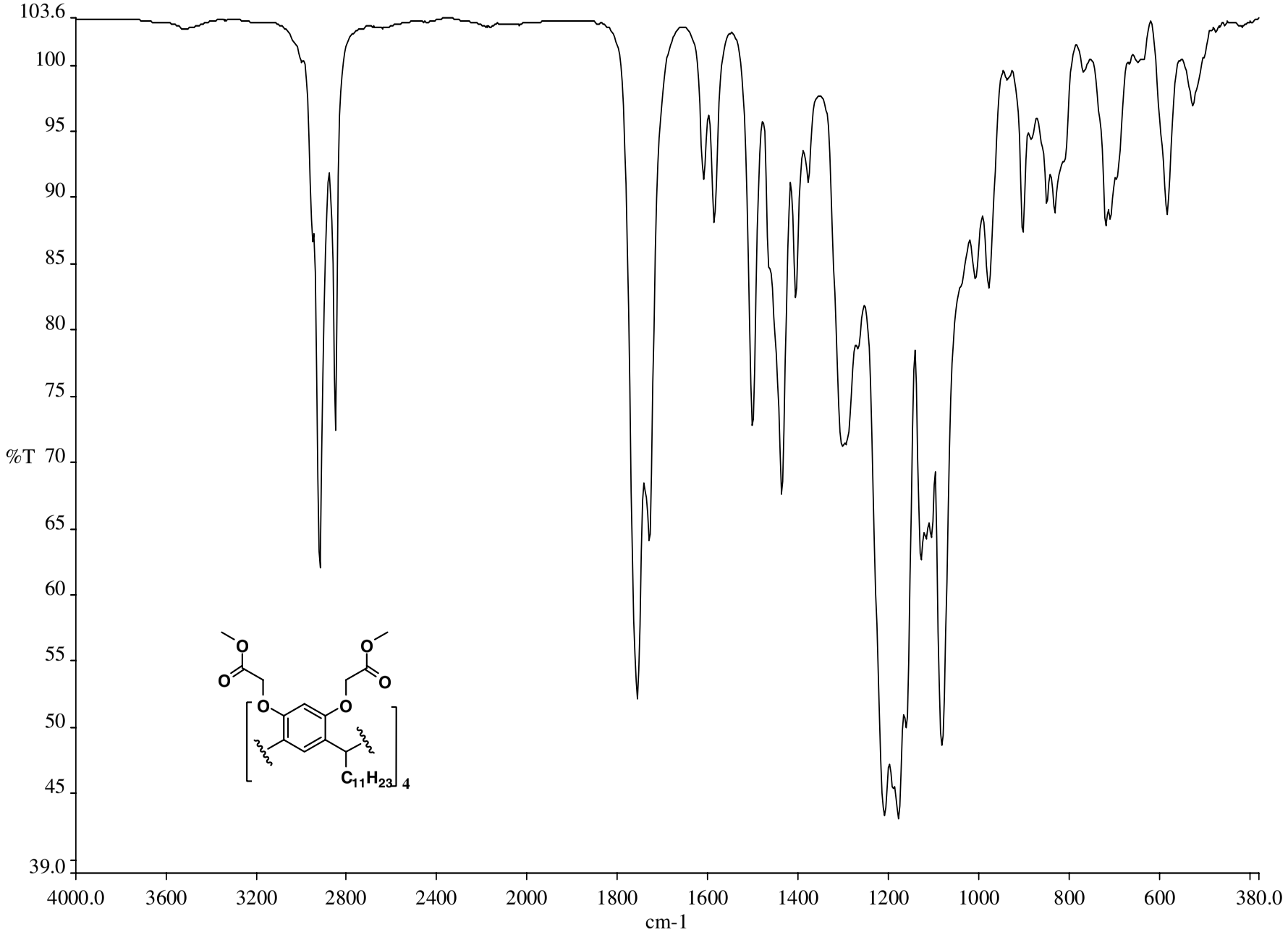
S4: HSQC NMR spectrum of 4g in DMSO-d_6 (600 MHz)



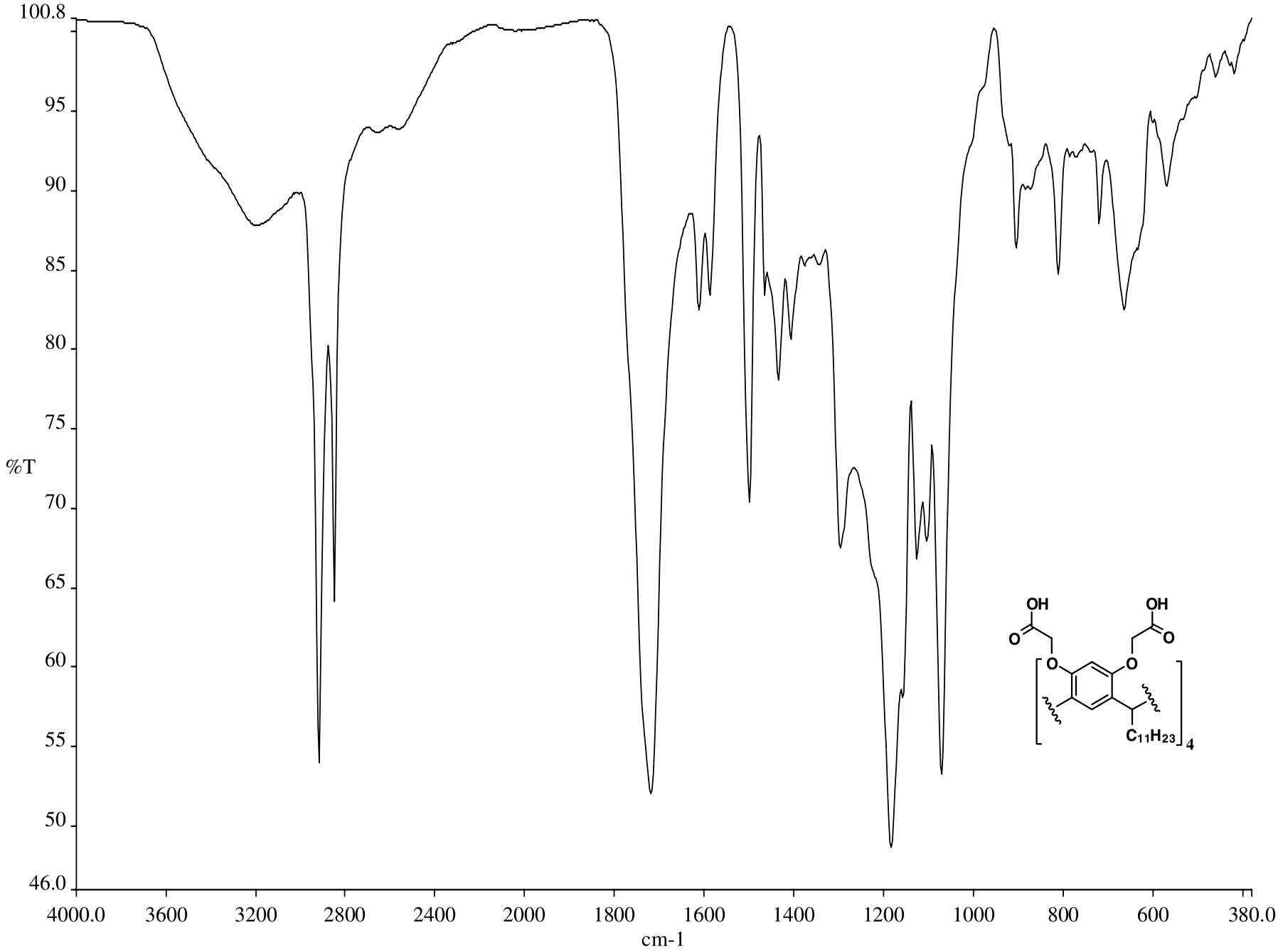
S4: HSQC NMR spectrum of 4h in DMSO-d_6 (600 MHz)



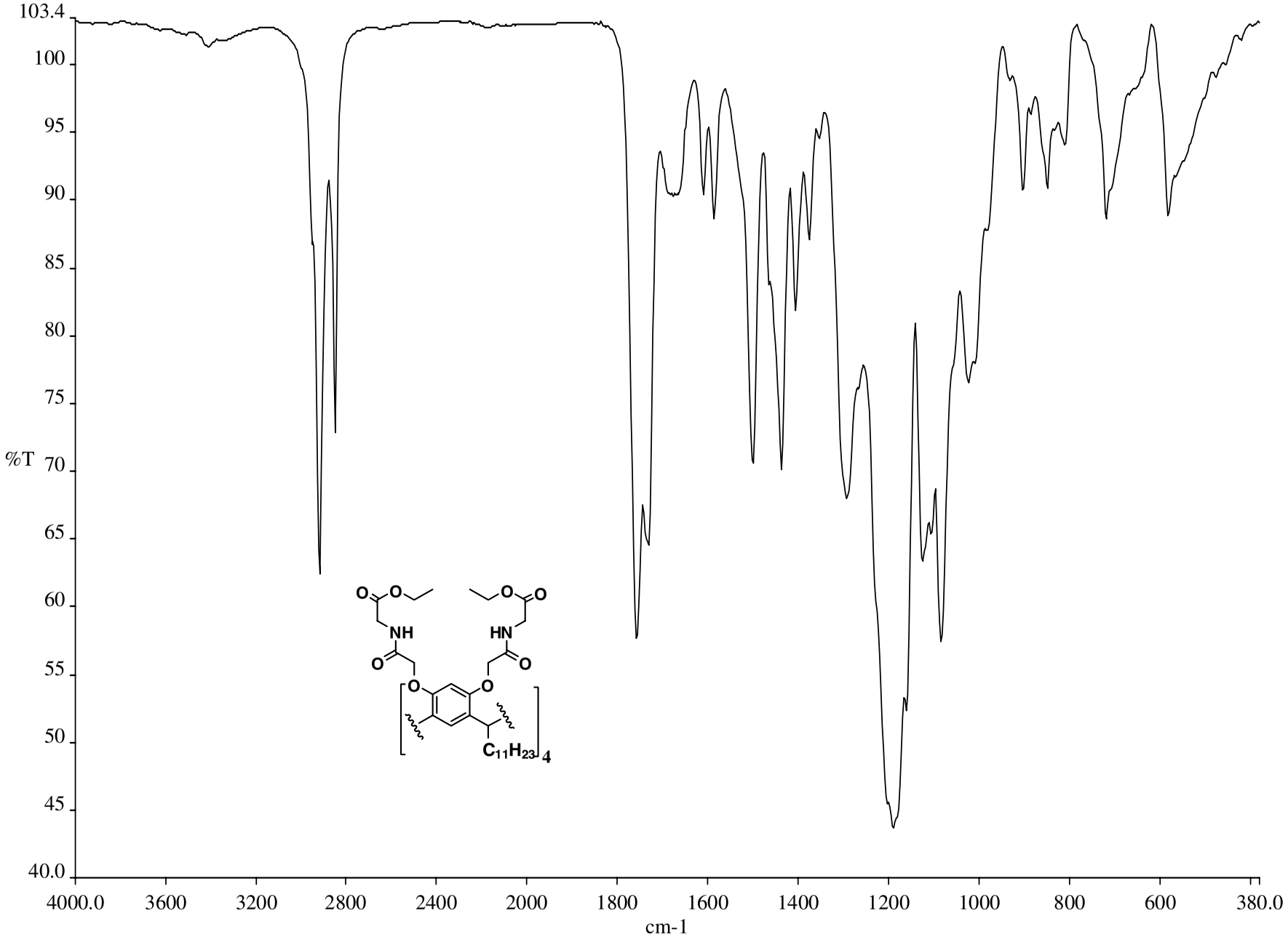
S4: HSQC NMR spectrum of 4i in DMSO-d_6 (600 MHz)



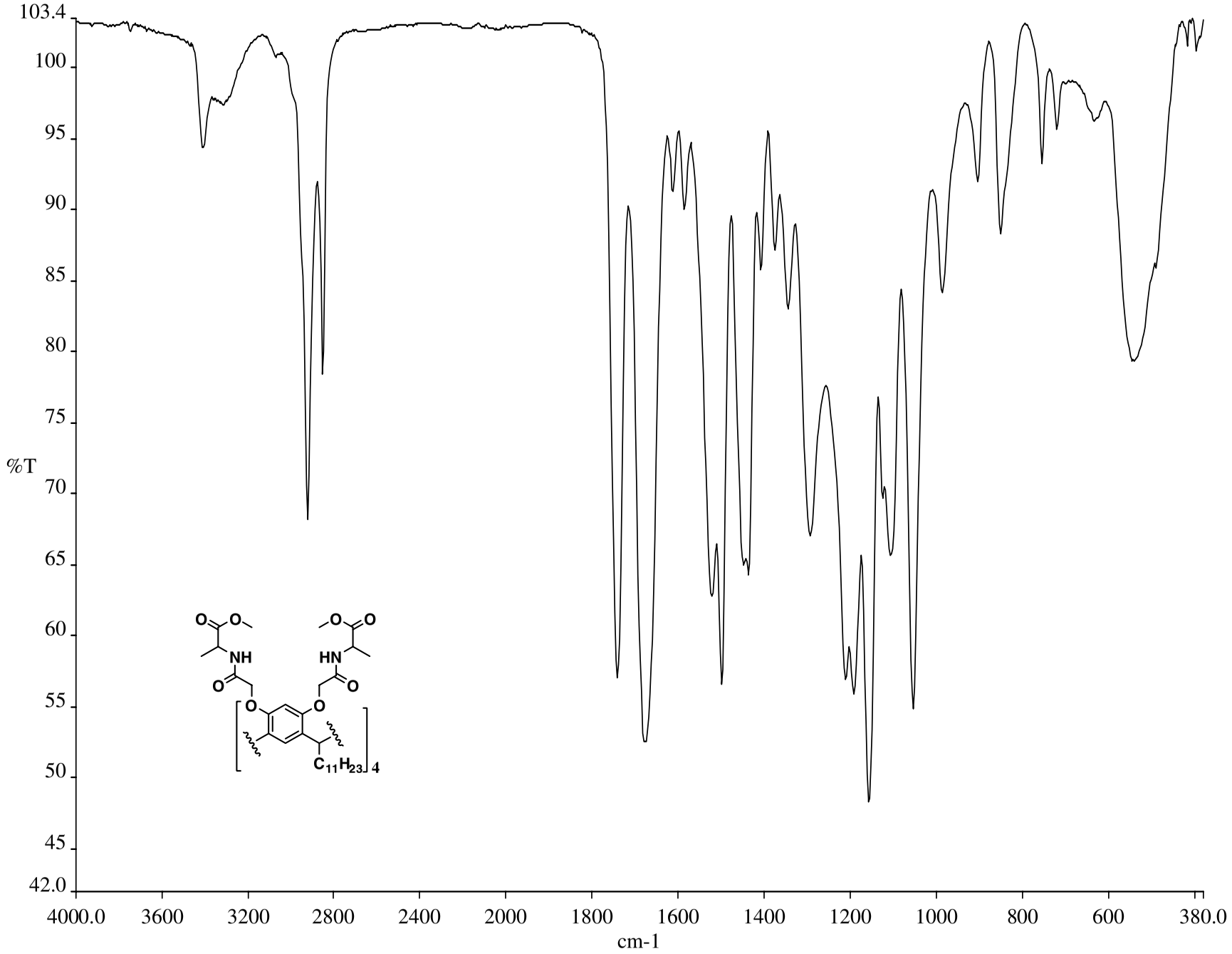
S5: Infrared spectrum of 2



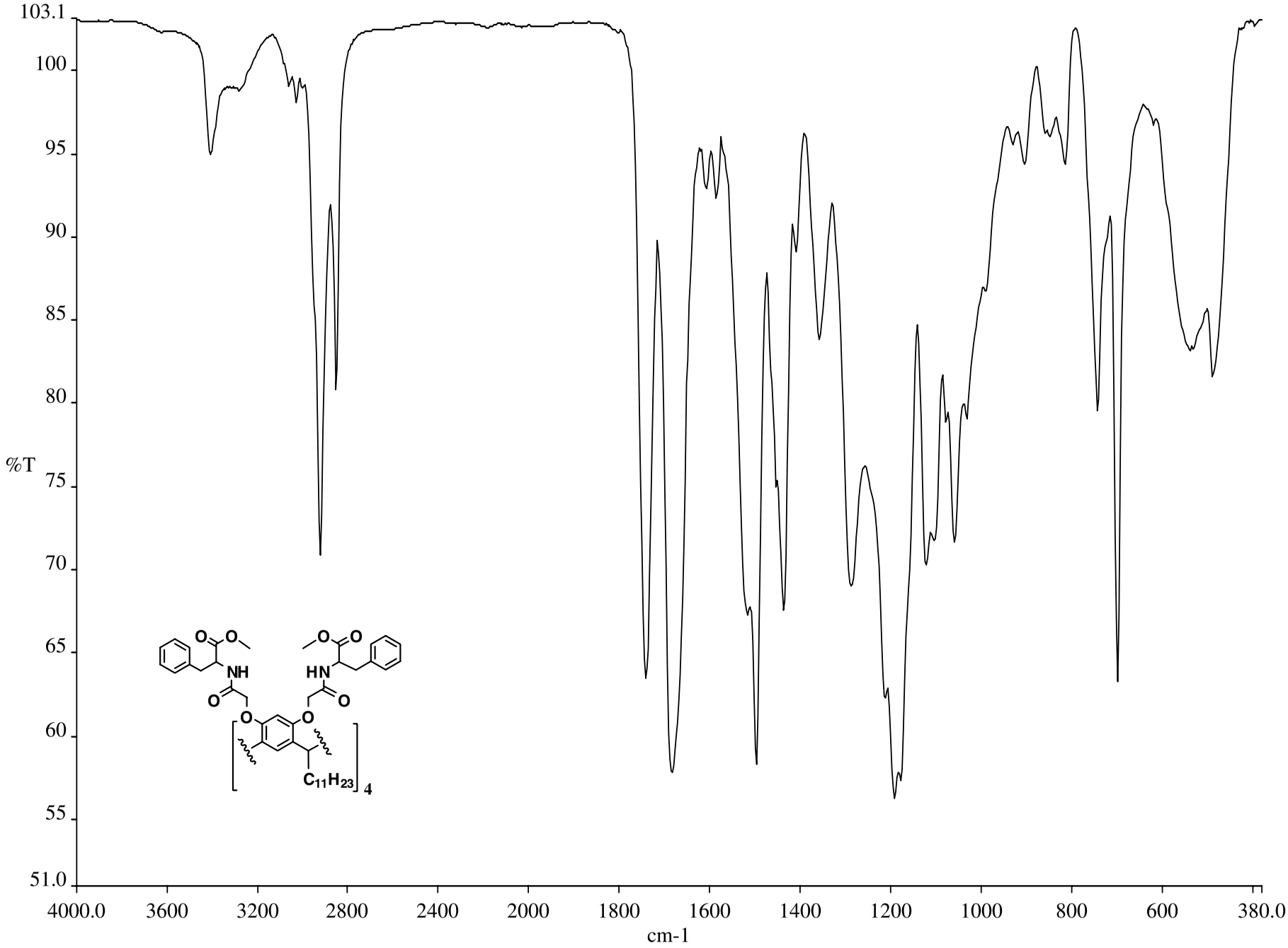
S5: Infrared spectrum of 3



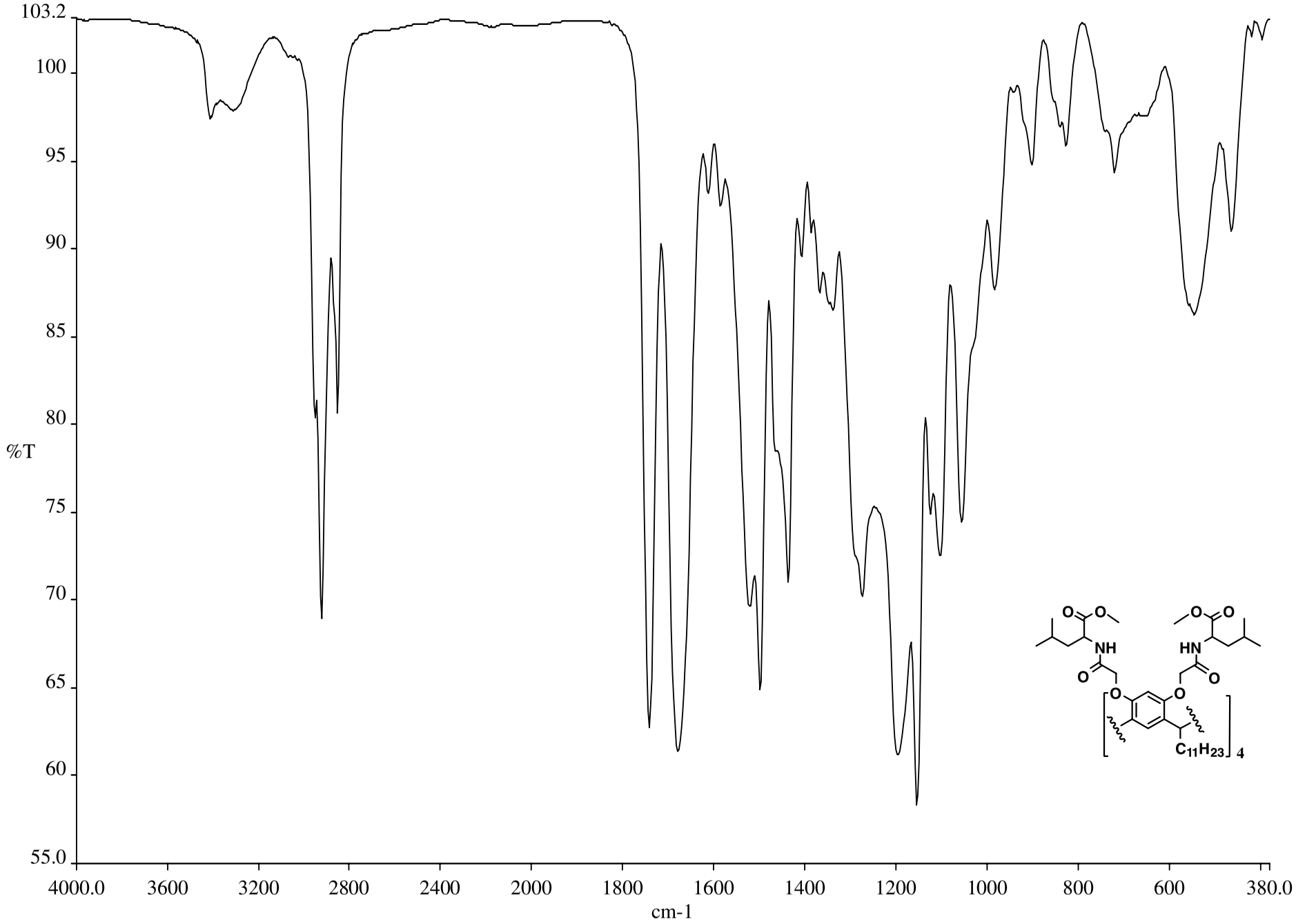
S5: Infrared spectrum of 4a



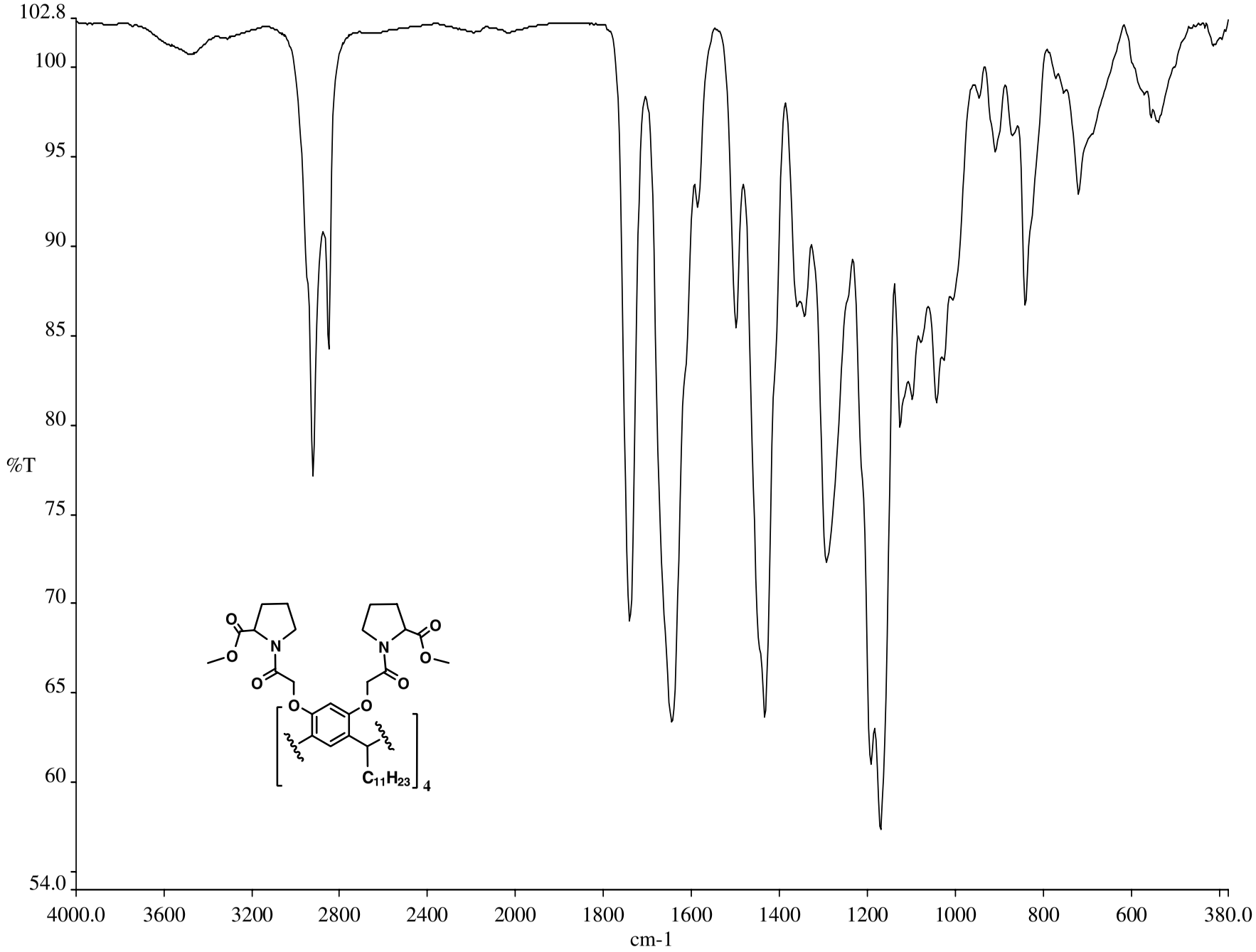
S5: Infrared spectrum of 4b



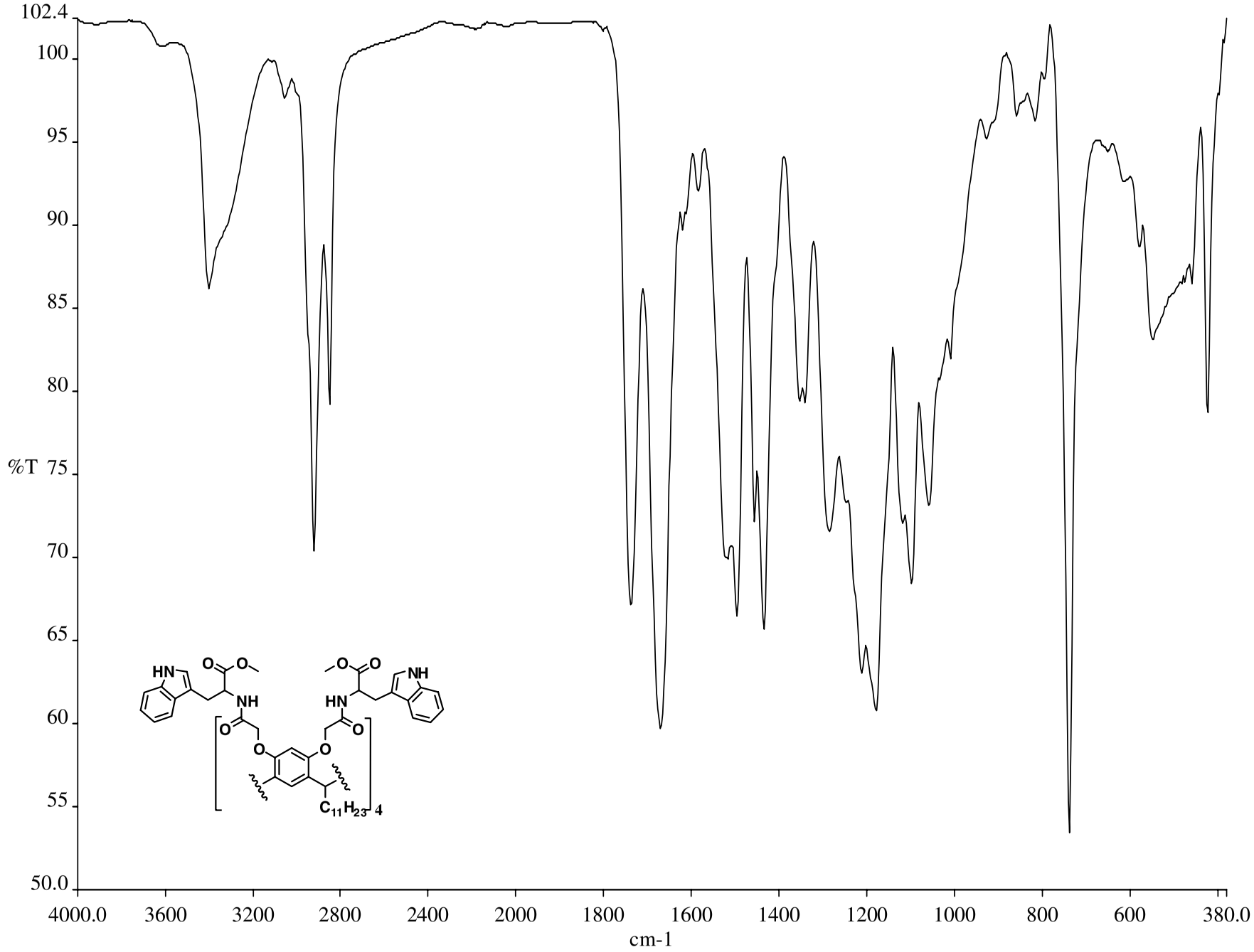
S5: Infrared spectrum of 4c



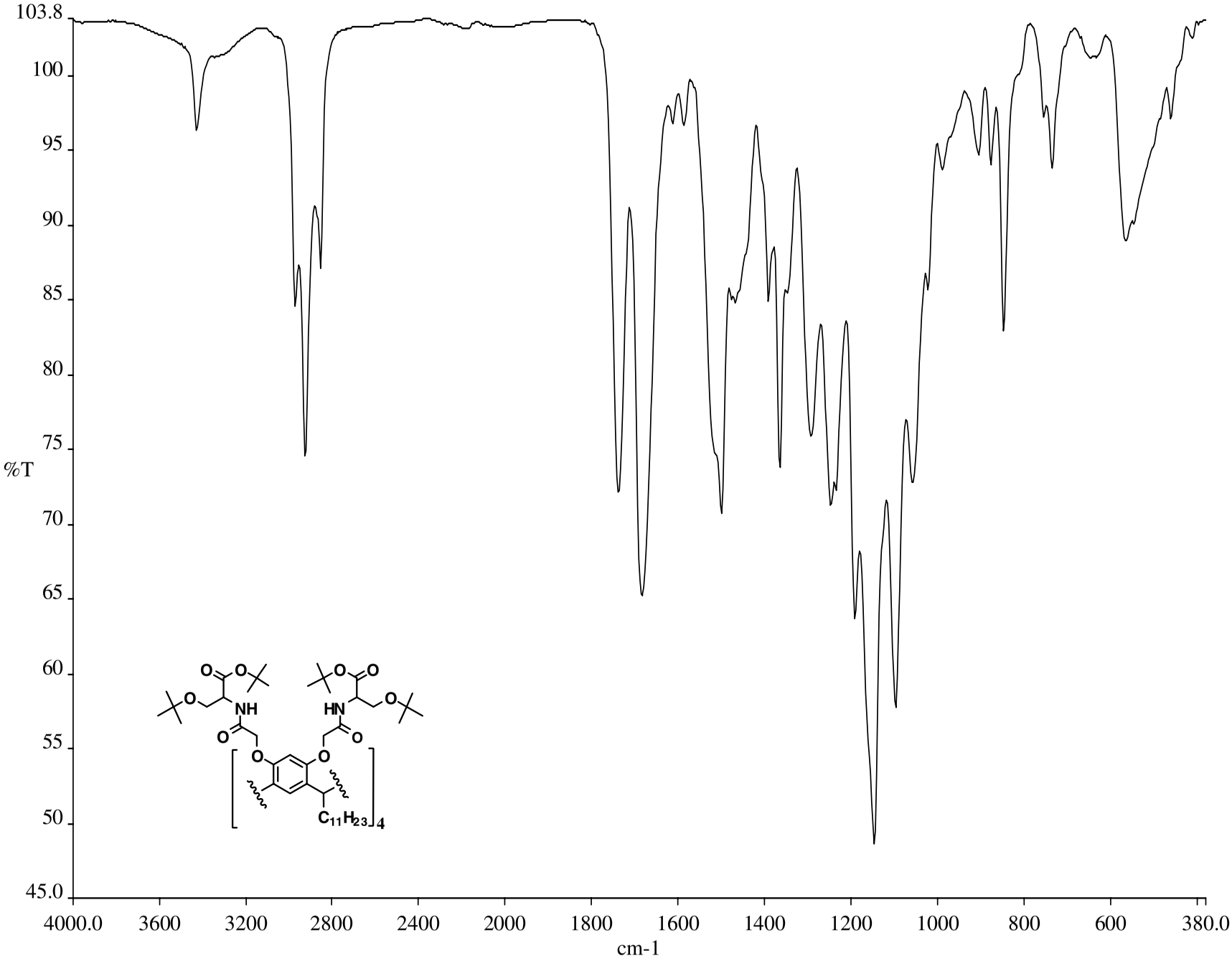
S5: Infrared spectrum of 4d



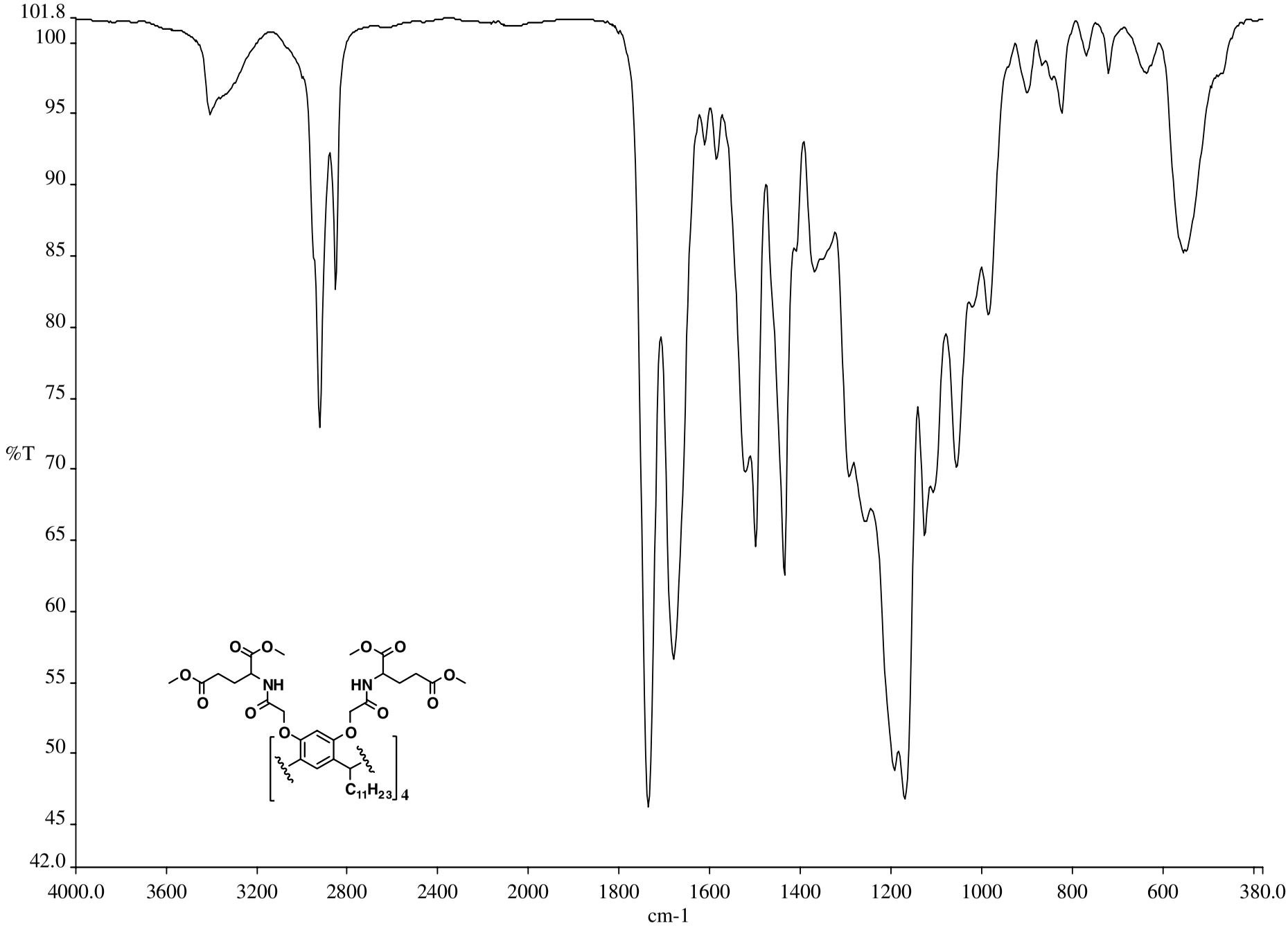
S5: Infrared spectrum of 4e



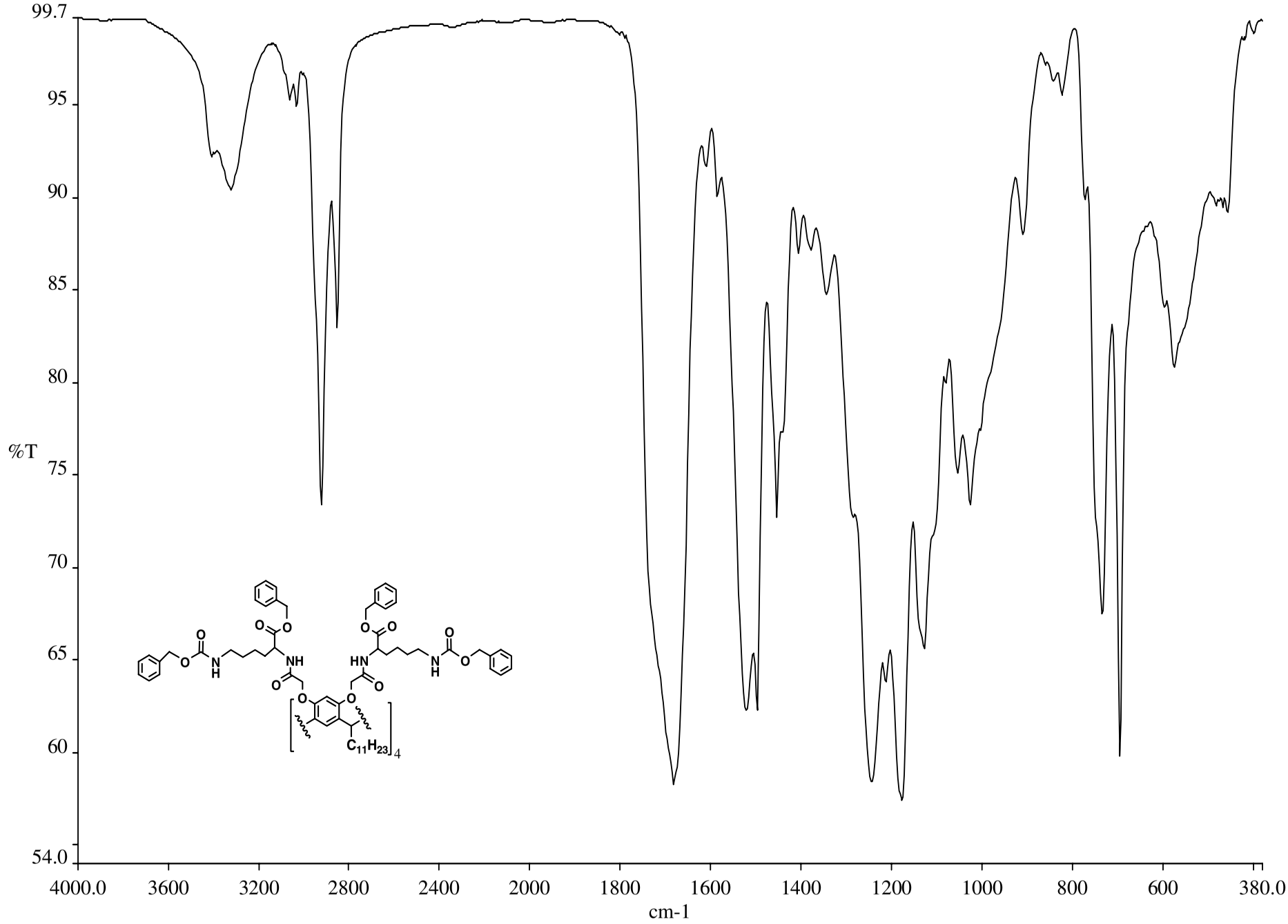
S5: Infrared spectrum of 4f



S5: Infrared spectrum of 4g



S5: Infrared spectrum of 4h



S5: Infrared spectrum of 4i

NMR Elucidation of compounds

Characterization of compound **1** was completed using proton (^{1}H) and carbon (^{13}C) NMR. The ^{1}H NMR chemical shifts for this compound were assigned with reference to Figure 1.

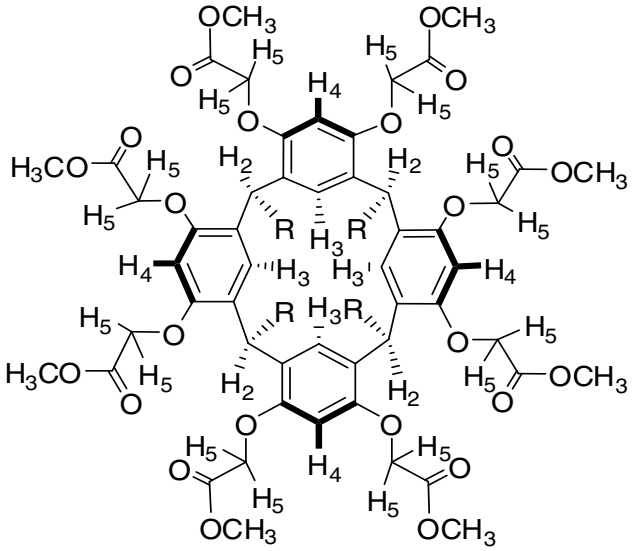


Figure 1: Expanded structure of **2**, showing distinctive protons.

Alkylation of the eight hydroxyl groups in **1** with methyl-2-bromoacetate afforded **2**, which result in the appearance of a singlet at 4.30 ppm due to the methylene protons of OCH_2CO group (H_5, Figure 2), integrating to sixteen. This signal appears at a lower frequency due to the deshielding nature of the neighbouring oxygen atom and carbonyl group. The signals associated with the methoxy groups appear as a singlet at 3.77 ppm, integrating to twenty four. The signals related to the two aromatic resorcin[4]arene protons appear as a singlet at 6.58 ppm for H_3 protons (*meta* to the hydroxyl groups) and at 6.20 ppm due to the H_4 protons (*ortho* to the hydroxyl group), each of these signals integrating to four. The proton signal of H_3 appears at a lower frequency compared to the H_4 protons. The signal related to the methine protons (H_2) which bridge the aromatic moieties, appears as a triplet at 4.57 ppm, integrating to four. The signals associated with the undecyl "feet" (R) give rise to a quartet at 1.90 ppm (integrating to eight), a multiplet at 1.20-1.30 ppm (integrating to seventy two), and a triplet at 0.87 ppm (integrating to twelve) due to the terminal methyl groups of the "feet". The two singlet peaks at 6.58 ppm and 6.20 ppm for H_3 and H_4 protons, respectively, indicate the symmetric positions of these protons in a *crown* (C_{4v}) conformation.(1, 2)

Subsequent hydrolysis of compound **2** using a potassium hydroxide solution (2 M KOH) in ethanol under reflux for 3 hours (Scheme 1, main paper), gave **3** in 93% yield after re-crystallisation from methanol/water in a 1:1 ratio.(3) Compound **3** was characterised from its proton and carbon NMR spectra. The ^{1}H NMR chemical shifts for this compound were assigned with reference to Figure 2.

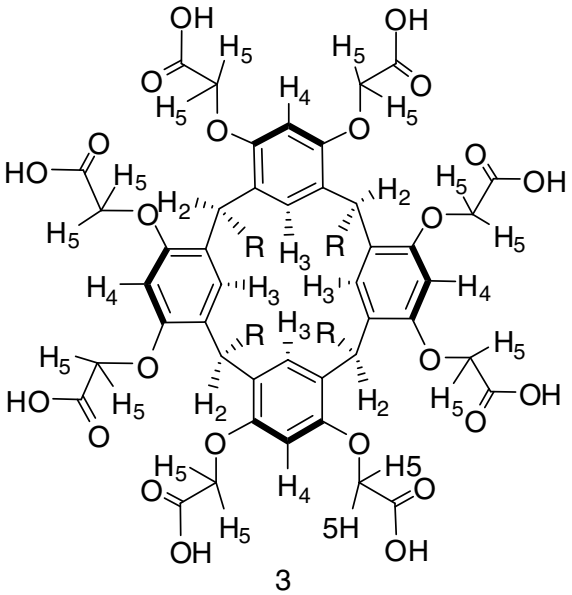


Figure 2: Expanded structure of **3**, showing distinctive protons.

Hydrolysis of **2** confirmed by the disappearance of the methoxy group signal (singlet at 3.77 ppm) in **3**. The signal associated with the methylene protons of OCH_2CO (H_5, Figure 3) appears as a pair of doublets at 4.23 ppm and 4.41 ppm, each of these signals integrates to eight. Compared to the singlet signal for the methylene protons for **2** in non-polar solvent ($CDCl_3$), this splitting clearly shows that in polar organic solvent (DMSO), there is weak intramolecular hydrogen bonding. The signals related to the two aromatic resorcin[4]arene protons (H_3 and H_4) appear as a slightly broad singlet (compared to the aromatic protons signals for **2** in $CDCl_3$). One appears at 6.49 ppm for H_3 protons and the other at 6.20 ppm due to the H_4 protons, each of these integrates to four. The signal associated with the methine protons (H_2) appears as a triplet at 4.48 ppm, integrating to four. The signals related to the "feet" have resolved into three signals: a quartet at 1.75 ppm, multiplets at 1.20-1.29 ppm, and a triplet at 0.82ppm. These signals maintain their associated integration.

The appearance of the methylene protons (H_5) as two doublets and the slight broadening of the two aromatic protons (H_3 and H_4), show that compound **3** is flexible and mainly exists in a *boat* conformation with C_{4v} symmetry on NMR time scale.

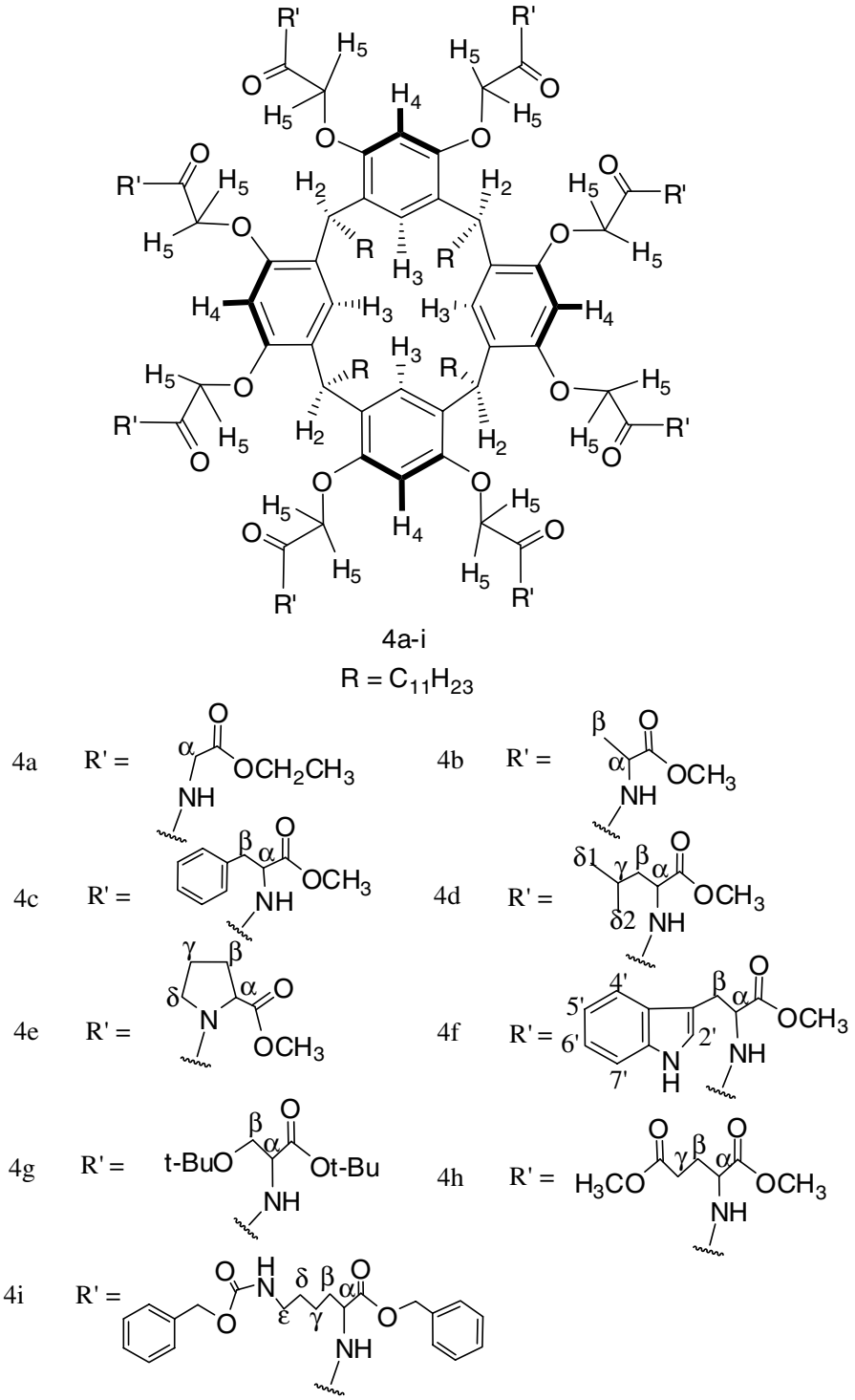


Figure 3: Expanded structure of **4a-i**, showing distinctive protons.

The ^{1}H NMR spectrum for **4b** displays signals characteristic for both units. The signal associated with the methyl protons of the ester group appears as a doublet at 3.64 ppm, integrating to twenty four. The signal related to the alanine α-protons appears as a quartet at 4.42 ppm, integrating to eight. The signal due to the methyl group attached to the β-carbon appears as a triplet at 1.34 ppm, integrating to twenty four. The amide NH protons signal for this derivative appears as two doublets at 7.82 ppm and 7.80 ppm, each integrating to four.

The signal for the methylene protons of the OCH_2CO groups (H_5) appears as a pair of quartets at 4.35 ppm and 4.24 ppm, each of these integrates to eight. This splitting could be attributed to the presence of two chiral amino acid units on each

aromatic ring making these protons (H_5) diastereotopic.(4, 5) The signals for the aromatic ring protons (H_3 and H_4) appear as two singlets in the ^{1}H NMR spectrum, one at 6.73 ppm (for H_3), and the other at 6.56 ppm (for H_4). The signal for the H_2 protons at 4.67 ppm is a triplet, integrating to four. The signal related to the undecyl "feet" (R, Figure 3) is largely unchanged from **4a**, and maintains the associated multiplicity and integration.

Coupling of L-phenyl alanine methyl ester to the octa-acyl chloride resorcin[4]arene **3** afforded compound **4c** in 71 % yield. The ^{1}H NMR spectrum exhibits signals for the phenyl alanine residue and the resorcin[4]arene scaffold. The signal associated with the methyl protons of the ester group at 3.59 ppm is a doublet, integrating to twenty four. The signal associated with the α-protons at 4.64 ppm appears as a quartet, integrating to eight. The signal associated with the β-protons splits into four pairs of doublets at 3.11 ppm, 3.04 ppm, 3.03 ppm, and 2.87 ppm due to coupling with the α-protons. Each of these integrates to four. The signals related to the phenyl rings at the side chains appear as a multiplet at 7.04-7.15 ppm, integrating to 40. The amide NH protons signal for this compound appears as two doublets at 7.77 ppm and 7.60 ppm, each of these integrates to four.

The signal for the diastereotopic methylene protons of the OCH_2CO groups (H_5) appears as a pair of quartets at 4.28 ppm and 4.20 ppm, each integrating to eight. The signals related to the aromatic protons appear as two singlets, one at 6.89 ppm related to H_3 protons in Figure 7 and the other at 6.33 ppm due to the H_4 protons. Each of these integrates to four. The signal related to the H_2 protons appears as a triplet at 4.64 ppm and integrates to four. The signals associated with the undecyl "feet" are unchanged.

The ^{1}H NMR spectrum for **4d** derivative displays signals characteristic for both units. The signal related to the methyl protons of the ester group appears as a doublet at 3.62 ppm, integrating to 24. The signal related to the α-protons (Figure 3) appears as a quartet at 4.47 ppm, integrating to eight. The signals associated with the β- and γ-protons appear as multiplets at 1.45-1.67 ppm, integrating to twenty four protons. The signals for the methyl groups attached to the δ-carbon atoms appear at a higher frequency as two triplets at 0.87 ppm and 0.77 ppm, each of these integrates to 24. The signal associated with the amide NH protons appears as two doublets at 8.02 ppm and 7.79 ppm, each of these integrates to four.

The signal for the diastereotopic methylene protons of the OCH_2CO groups (H_5) appears as a pair of quartets at 4.36 ppm and 4.46 ppm, integrating to eight each. The signals related to the H_3 and H_4 protons appear as two singlets at 6.89 ppm for H_3 and at 6.53 ppm for H_4. Each integrates to four. The signal associated with the H_2 protons appears as a triplet at 4.77 ppm, integrating to four. The signals related to the undecyl "feet" are unchanged.

Reaction of the octa-acyl chloride resorcin[4]arene **3** with L-proline methyl ester afforded compound **4e** in 69 % yield. The ^{1}H NMR spectrum exhibits signals for both moieties. The signal related to the methyl protons of the ester group appears as a doublet at 3.69 ppm, integrating to twenty four. The signal related to the α-protons at 4.38 ppm is a quartet, integrating to eight. The signals associated with the β- and γ-protons appear as multiplets at 1.90-2.15 ppm, integrating to thirty two. The signal related to the δ-protons appears as a multiplet at 3.56 ppm, integrating to sixteen.

The signal for the methylene protons of the OCH_2CO groups (H_5, Figure 3) appears as a broad doublet at 4.48 ppm, integrating to 16. This broadening could be attributed to the slow rate of conformational interchange of the proline-pyrrolidine ring.(6) The signals related to the aromatic resorcin[4]arene protons (H_3 and H_4) appear as a broad singlet,

one at 6.82 ppm for H_3 protons and the other at 6.38 ppm for H_4 protons. Each of these integrates to four. The signal for H_2 at 4.68 ppm is a triplet and integrates to four. The signals related to the undecyl "feet" (R), remain essentially unchanged.

Compound **4f** was synthesised in 64 % yield by reacting L-tryptophan methyl ester with the octa-acid resorcin[4]arene **3**. The ^{1}H NMR spectrum for this derivative displays signals characteristic for both units. The signal related to the methyl protons of the ester group at 3.68 ppm appears as a doublet, integrating to twenty four. The signal associated with the α-protons at 4.71 ppm appears as a quartet, integrating to eight. The signal assigned to the β-protons at 3.11-3.25 ppm appears as a multiplet due to coupling to the α-protons and integrates to sixteen. The signal assigned to the NH protons of the indole ring (Figure 3) appears as two singlets at 10.42 ppm and 10.32 ppm, each of these integrates to four. The signal assigned to the tryptophan-7′-protons at 7.48 ppm appears as a triplet. The signal assigned to the tryptophan-4′-protons at 7.29 ppm appears as pair of doublets. The signal related to the tryptophan-5′-protons at 7.02 ppm is a triplet. The signal assigned to the tryptophan-6′-protons at 7.00 ppm appears as a triplet. The signal assigned to the tryptophan-2′-protons at 6.90 ppm is a singlet. Each of these signals integrates to eight protons. The signal related to the amide NH protons appears as two doublets at 7.63 ppm and 7.55 ppm, each integrates to four.

The signal for the diastereotopic methylene protons of the OCH_2CO groups (H_5, Figure 3) appears as a pair of quartets at 4.26 ppm and 4 32 ppm, each of these signals integrates to eight. The signals associated with the aromatic resorcin[4]arene protons appear as two singlets at 6.80 ppm for H_3 and at 6.47 ppm for H_4 protons, each of these signals integrates to four. The signal related to the methine protons (H_2) appears as a triplet at 4.63 ppm, integrating to four. The signals related to the undecyl "feet" are unchanged.

The ^{1}H NMR spectrum for **4g** exhibits signals for both residues. The signal related to the *t*-butyl protons, which protects the carboxylic group, is a doublet at 1.40 ppm, integrating to seventy two. The signal for the α-protons of this amino acid is a quartet at 4.46 ppm, integrating to eight. The signal for the β-protons appears as two pairs of doublets at 3.70 ppm and 3.54 ppm due to coupling to the α-protons and each integrates to eight. The signal related to the *t*-butyl protons at the side chain, which protects the hydroxyl groups, at 1.09 ppm is a doublet, integrating to seventy two. The amide NH protons signal appears as two doublets at 7.27 ppm and 7.24 ppm, each integrates to four (Figure 3).

The signal related to the diastereotopic methylene protons of the OCH_2CO groups (H_5, Figure 3) appears as four doublets at 4.36 ppm, 4.28 ppm, 4.25 ppm, and 4.18 ppm, each of these integrate to four. The presence of two serine units with bulky *t*-butyl groups at side chains per each aromatic ring could affect this splitting. The signals associated with the aromatic protons H_3 and H_4 appear as two singlets: one at 6.73 ppm (for H_3), and the other at 6.53 ppm (for H_4), each integrates to four. The signal related to H_2 appears as a triplet at 4.63 ppm, integrating to four. The signals related to the undecyl "feet" are unchanged.

The ^{1}H NMR spectrum for **4h** derivative displays signals characteristic for both moieties. The signal associated with the methyl protons of the ester group, is a singlet at 3.70 ppm, integrating to 24. The signal assigned to the α-protons for this amino acid at 4.44 ppm is a quartet, integrating to eight. The signal related to the β-protons appears as two multiplets at 1.96 ppm and 2.13 ppm due to coupling to the α-protons, each integrates to eight. The signal related to the γ-protons appears as a multiplet at 2.36 ppm, integrating to sixteen. The signal related to the methyl protons of the ester group at

the side chain, appears as a singlet at 3.58 ppm, integrating to twenty four. The amide NH protons signal appears as two doublets at 7.66 ppm, integrating to eight.

The signal for the diastereotopic methylene protons of the OCH_2CO groups (H_5, Figure 3) appears as a pair of quartets at 4.30 ppm and 4.32 ppm. Each of these integrates to eight. The signals associated with the H_3 and H_4 (aromatic ring protons) appear as two singlets, at 6.81 ppm for H_3 and at 6.55 ppm for H_4, each integrates to four. The signal for H_2 appears as a triplet at 4.69 ppm, integrating to four. The signals related to the undecyl "feet" are essentially unchanged.

The ^{1}H NMR spectrum for **4i** displays signals characteristic for both moieties. The signals related to the aromatic protons of carboxybenzyl group (Cbz) and the benzyl ester group (Bn) appear as a multiplet at 7.26-7.36 ppm, integrating to 80. The signal for the methylene protons of Cbz group appears at 5.13 ppm as a multiplet due to coupling with the α-protons, and the one associated with the benzyl ester group (Bn) appears at 4.77 ppm, as a singlet, each integrates to 16. The signal related to the α-protons appears as a quartet at 4.44 ppm, integrating to eight. The signal for the β-protons appears as two multiplets at 1.82 ppm and 1.78 ppm due to coupling with the α-protons, each integrates to eight. The signals related to the γ- and δ-protons appear as a multiplet at 1.39 ppm, integrating to 32. The signal for the ε-protons appears as a quartet at 3.10 ppm, integrating to 16. The signal for the εNH-protons appears as a broad triplet at 6.61 ppm, integrating to eight. The amide NH protons signal appears as a pair of doublets at 7.61 ppm, and integrates to eight.

The signal for the diastereotopic methylene protons of the OCH_2CO groups (H_5, Figure 3) appears as a pair of quartets at 4.27 ppm and 4.35 ppm. Each integrates to eight. The signals related to the aromatic ring protons (H_3 and H_4) appear as two singlets, each integrates to four: one at 6.84 ppm (for H_3) and the other at 6.59 ppm (for H_4). The signal for H_2 appears as a triplet at 4.68 ppm, integrating to four protons.

References

(1) Hogberg, A.G.S., Stereoselective Synthesis And Dnmr Study Of 2 1,8,15-22-Tetraphenyl 14 Metacyclophan-3,5,10,12,17,19,24,26-Octols. *Journal of the American Chemical Society* **1980,** *102*, (19), 6046-6050.

(2) Abis, L.; Dalcanale, E.; Duvosel, A.; Spera, S., Structurally New Macrocycles From The Resorcinol Aldehyde Condensation - Configurational And Conformational-Analyses By Means Of Dynamic NMR, NOE, And T1 Experiments. *Journal of Organic Chemistry* **1988,** *53*, (23), 5475-5479.

(3) Bazzanella, A.; Morbel, H.; Bachmann, K.; Milbradt, R.; Bohmer, V.; Vogt, W., Highly efficient separation of amines by electrokinetic chromatography using resorcarene-octacarboxylic acids as pseudostationary phases. *Journal of Chromatography A* **1997,** *792*, (1-2), 143-149.

(4) Sansone, F.; Barboso, S.; Casnati, A.; Fabbi, M.; Pochini, A.; Ugozzoli, F.; Ungaro, R., Synthesis and structure of chiral cone calix 4 arenes functionalized at the upper rim with L-alanine units. *European Journal of Organic Chemistry* **1998**, (5), 897-905.

(5) Frkanec, L.; Visnjevac, A.; Kojic-Prodic, B.; Zinic, M., Calix 4 arene amino acid derivatives. Intra- and intermolecular hydrogen-bonded organisation in solution and the solid state. *Chemistry-a European Journal* **2000,** *6*, (3), 442-453.

(6) Deslauri.R; Smith, I.C.P.; Walter, R., Conformational Mobility Of Pyrrolidine Ring Of Proline In Peptides And Peptide Hormones As Manifest In C-13 Spin-Lattice Relaxation-Times. *Journal of Biological Chemistry* **1974,** *249*, (21), 7006-7010.