Supplementary Information accompanying:

A threading receptor for polysaccharides

Tiddo J. Mooibroek[1], Juan M. Casas-Solvas[1,2], Robert L. Harniman[1], Charles M. Renney[1], Tom S. Carter[1], Matthew P. Crump[1] and Anthony P. Davis[1] *

[1] School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK. [2] Department of Chemistry and Physics, University of Almería, Carretera de Sacramento s/n, 04120 Almería, Spain.

*E-mail: Anthony.Davis@bristol.ac.uk

Table of Contents

1. Synthesis and characterisation of receptor 4

All commercially available chemicals were purchased from Sigma-Aldrich, Alfa-Aesar, Acros Organics, ABCR or Carbosynth Ltd., and used without further purification, unless stated otherwise. Pyrene-2-carboxylic acid (**A**)[1] and benzene-1,3-pentafluorophenoxycarbonyl-5-[tris-(2-amido-[tris-(2-tert-butoxycarbonylethyl)methane]-ethyl)amidomethane] (**J**)[2] were prepared according to literature procedures. Solvents were utilized as supplied unless otherwise stated. Anhydrous solvents (e.g. THF or CH_2Cl_2) were dried by passing through a modified Grubbs system[3] manufactured by Anhydrous Engineering. Anhydrous DMF and DMSO were dried by vacuum distillation from P_2O_5 and CaH_2, respectively[4]. Flash column chromatography was carried out using chromatography grade silica 60Å (Sigma-Aldrich, particle size 35-70 micron). TLC analysis was performed using pre-coated silica gel TLC plates (Merck silica gel 60 F254). Spots were visualised by means of UV light (254 or 365 nm) or using solutions of phosphomolybdic acid, alizarin, potassium permanganate, or ninhydrin. HPLC chromatography was performed using a Waters 600 controller with a Waters 2998 photodiode array detector. For analytical runs a XSELECT CSH C_{18} 5 μm (4.6 × 150 mm) column was used, and for preparative runs a XSELECT CSH Prep C_{18} 5 μm OBD (19 × 250 mm) column was utilised.

^{1}H-, ^{13}C-, ^{19}F-, ^{31}P-, and 2D-NMR spectra were acquired at 298 K (unless otherwise specified) at 400, 500, or 600 MHz using the following spectrometers: Jeol ECS 400, Varian 400-MR, Varian VNMR S500a, Varian VNMR S500b, Varian INOVA 600, or Varian VNMR S600 Cryo. Chemical shifts (δ) are reported in parts per million (p.p.m.). Residual solvent resonances were used as internal reference for δ-values in ^{1}H-, and ^{13}C-NMR,[5] while ^{19}F- and ^{31}P-NMR spectra were externally referenced to CF_3COOH (–76.55 p.p.m.) and H_3PO_4 (0 p.p.m.) respectively. IR spectra were measured on a Perkin-Elmer Spectrum One FT-IR. Low resolution mass spectra were recorded on a VG Analytical Quattro or Bruker ESQUIRE 300 for ESI and Nanospray, or an Applied Biosystems 4700 spectrometer for MALDI. High resolution mass spectra were recorded on a Bruker Apex 4e 7.0T FT-MS. The C, H, N elemental composition was analysed at the Microanalytical laboratory of the University of Bristol.

Receptor **4** was prepared as summarised in Supplementary Scheme 1. An important consideration was the maintenance of solubility in pyrenyl intermediates. This explains the use of butyl ester protection in intermediates **B**-**F** (methyl esters tended to insolubility). It also informed the choice of methyl ester protection in **K**-**M**; solubility issues favoured the use of *N*-Boc protection in **F**-**I**, and we therefore needed acid-stable protection for the solubilising group. Subsequently it proved convenient to cyclise **M** with *t*-butyl-protected **J**, then apply a two-stage deprotection procedure in the final step of the synthesis.

The numbering system used to assign and discuss receptor **4** and derivatives is shown in Supplementary Figure 1. The detailed synthetic procedures and characterisation data for compounds in the synthetic sequence are given in the following pages, accompanied by relevant NMR spectra.

---

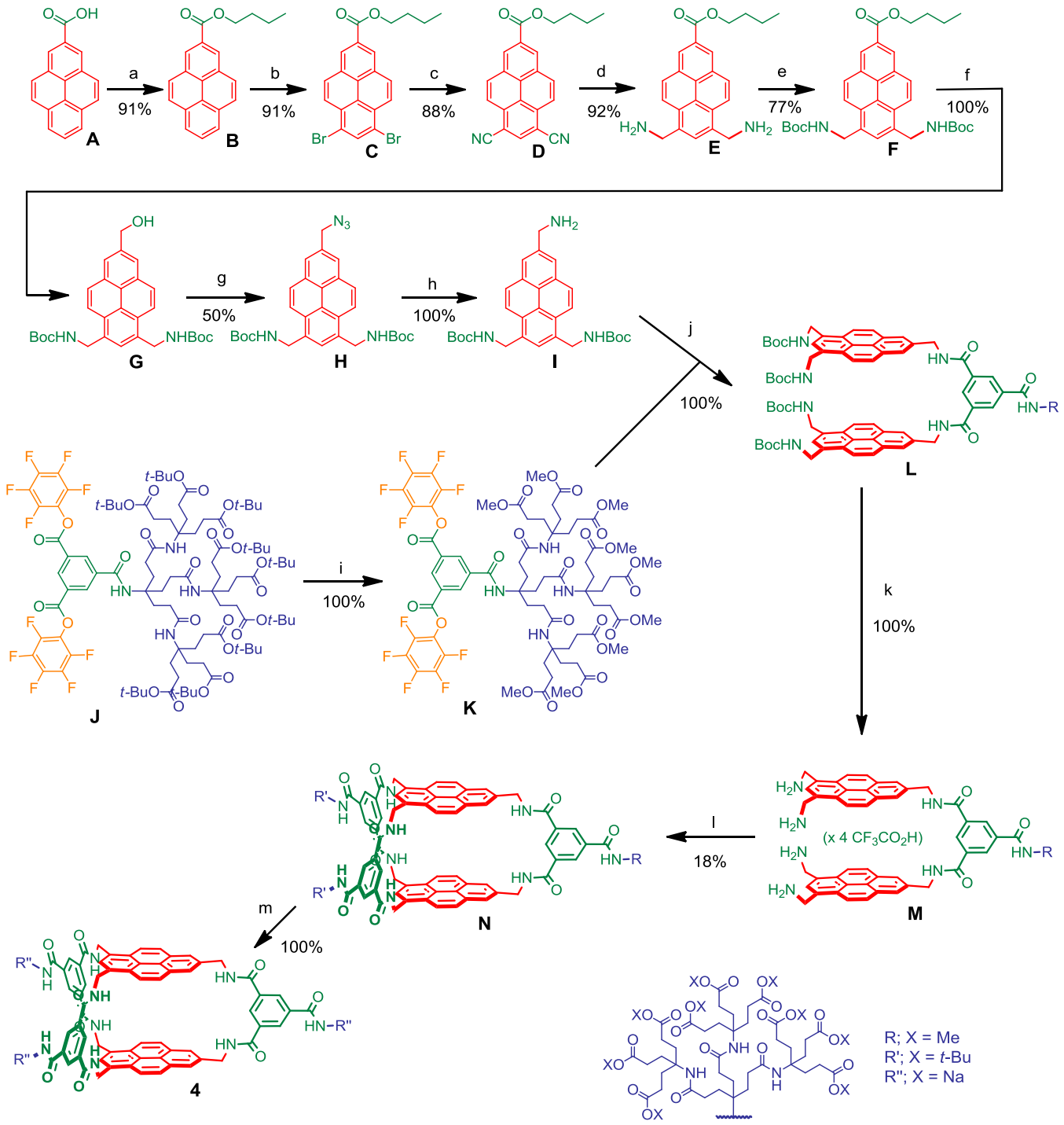
[1] Casas-Solvas, J. M., Mooibroek, T. J., Sandramurthy, S., Howgego, J. D. & Davis, A. P. A practical, large-scale synthesis of pyrene-2-carboxylic acid. *Synlett* **25**, 2591-2594 (2014).

[2] Destecroix, H. *et al.* Affinity enhancements by dendritic side-chains in synthetic carbohydrate receptors. *Angew. Chem., Int. Ed.* **54**, 2057 –2061 (2015).

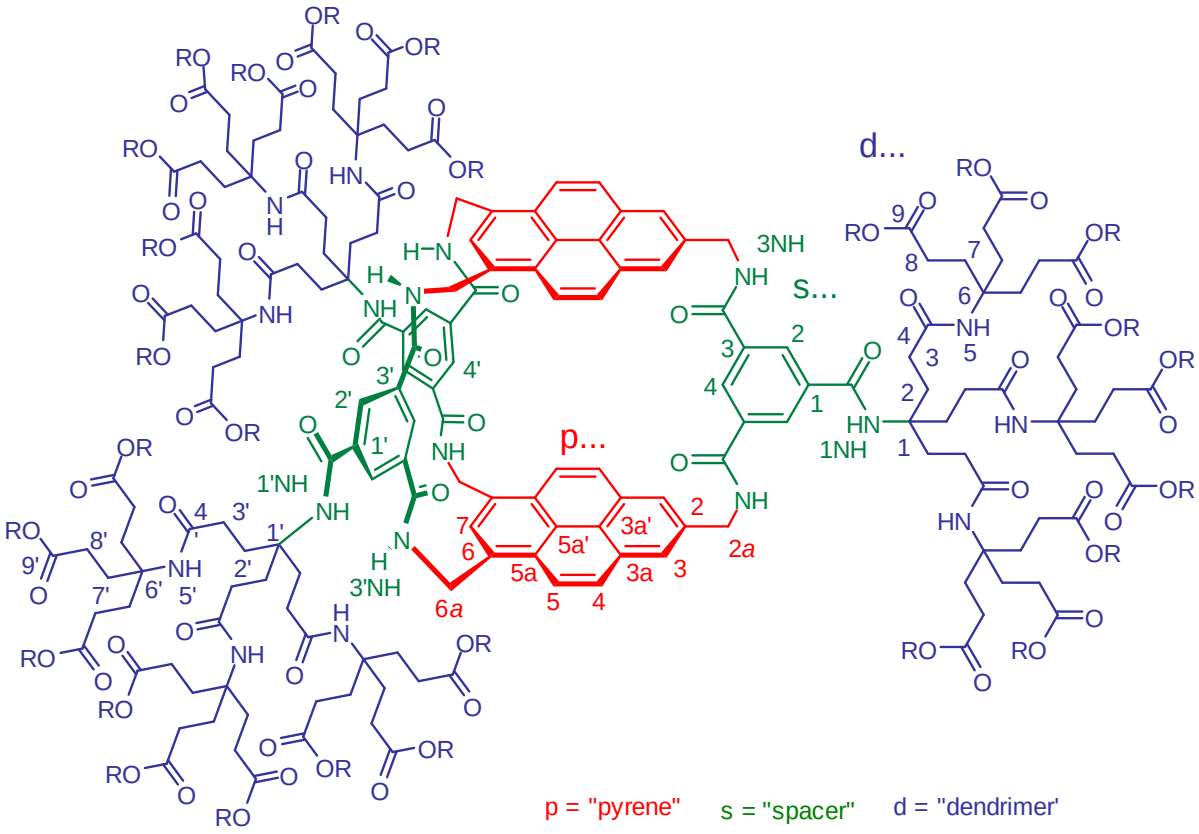
[3] Pangborn, A. B., Giardello, M. A., Grubbs, R. H., Rosen, R. K. & Timmers, F. J. Safe and convenient procedure for solvent purification. *Organometallics* **15**, 1518-1520 (1996).

[4] Armarego, W. F. & Chai, C. L. L. *Purification of Laboratory Chemicals* (Elsevier, Oxford, 2009).

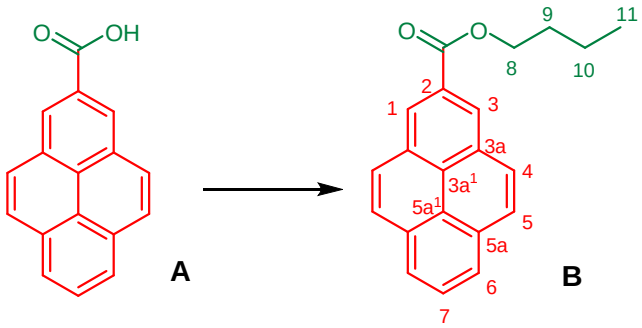
[5] Fulmer, G. R. *et al.* NMR chemical shifts of trace impurities: common laboratory solvents, organics, and gases in deuterated solvents relevant to the organometallic chemist. *Organometallics* **29**, 2176-2179 (2010).



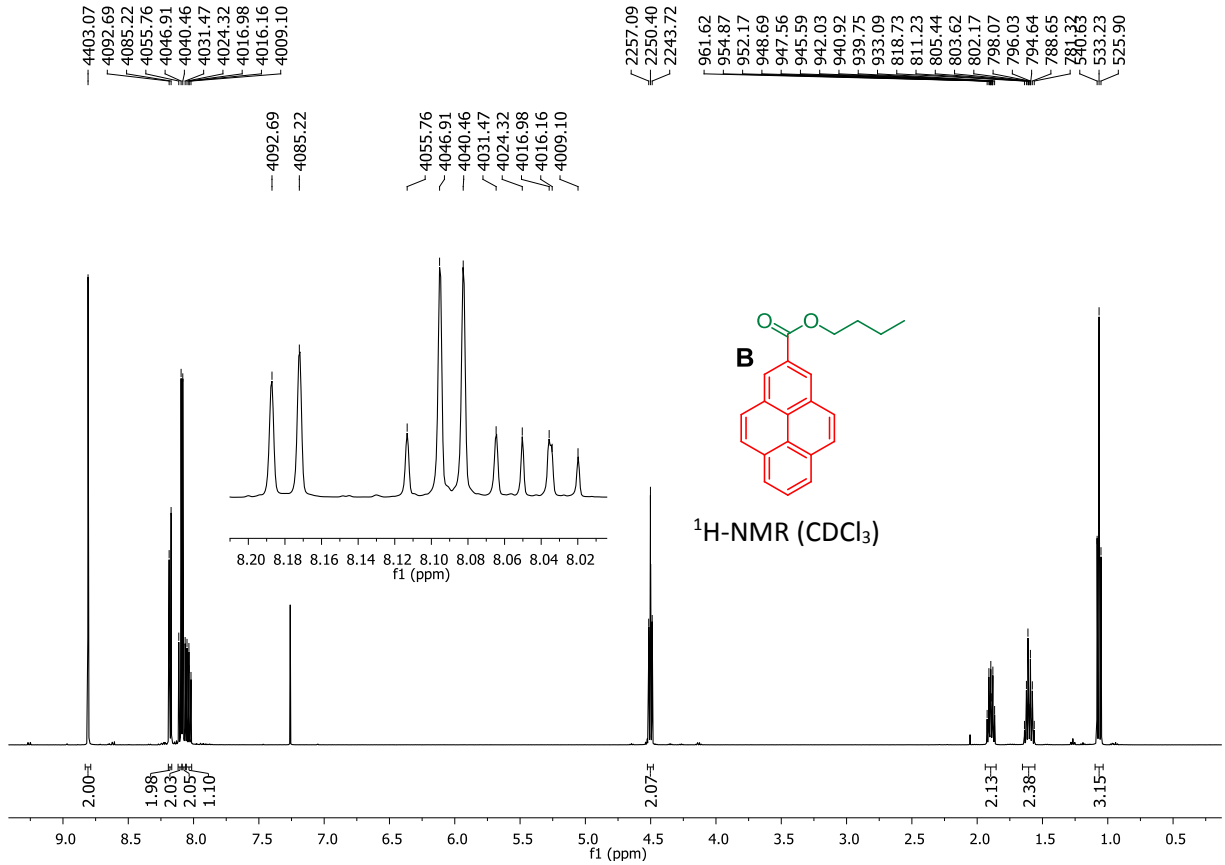
Supplementary Scheme 1 | Synthesis of receptor 4. a) *n*-BuOH, 2% (v/v) H_2SO_4, reflux, 22 h, 91%; b) CH_2Cl_2, 2.1 eq. Br_2 (dropwise, 45 min.) 25 °C, 24 hrs, 91%; c) *N*,*N*'-dimethylformamide (DMF), 2.1 eq. $Zn(CN)_2$, 16 % $Pd(Ph_3)_4$, 60 °C, 12 hrs, 88%; d) *n*-BuOH/NH_3OH(aq) 15:1, Raney Nickel, 50 bar H_2, 25 °C, 24 hrs, 92%; e) 2.9 eq. di-*tert*-butyl dicarbonate , 2.9 eq. *N*,*N*-diisopropylethylamine (DIPEA), 25 °C, 16 hrs, 77%; f) Tetrahydrafuran (THF), 6 eq. CH_3OH, 6 eq. $NaBH_4$, 50 °C, 16 hrs, 100%; g) DMF, 2 eq. CBr_4, 2 eq. PPh_3, 10 eq. NaN_3, 25 °C, 16 hrs, 50%; h) THF, 6 eq. PMe_3, 0 °C, 1h, then > 100 eq. H_2O, 40 °C, 1 hrs, 100%; i) CH_2Cl_2, dropwise trifluoroacetic acid (TFA) to 50% v/v), 0 °C, then 25 °C, 16 hrs; then CH_3OH/trimethyl orthoformate 1:1, 5% (v/v) conc. HCl, 25 °C, 24 hrs, 100% over two steps; j) THF, 1 eq. **K**, 2.5 eq. **J**, 20 eq. DIPEA, 25 °C, 12 hrs, 100%; k) CH_2Cl_2, dropwise TFA to 50% v/v, 0 °C, then 25 °C, 16 hrs, 100%; l) THF, 1 eq. **M** (0.1 mM), 150 eq. DIPEA, dropwise 2 eq. **J** over 48 hrs, then 25 °C, 72 hrs, 18%; m) CH_2Cl_2, dropwise TFA to 50% v/v), 0 °C, then 25 °C, 16 hrs; then –volatiles and + 0.5 M NaOH, 25 °C, 15 min., to pH < 4 with Amberlyst 15 resin, then to pH = 7 with NaOH, 100% over two steps.

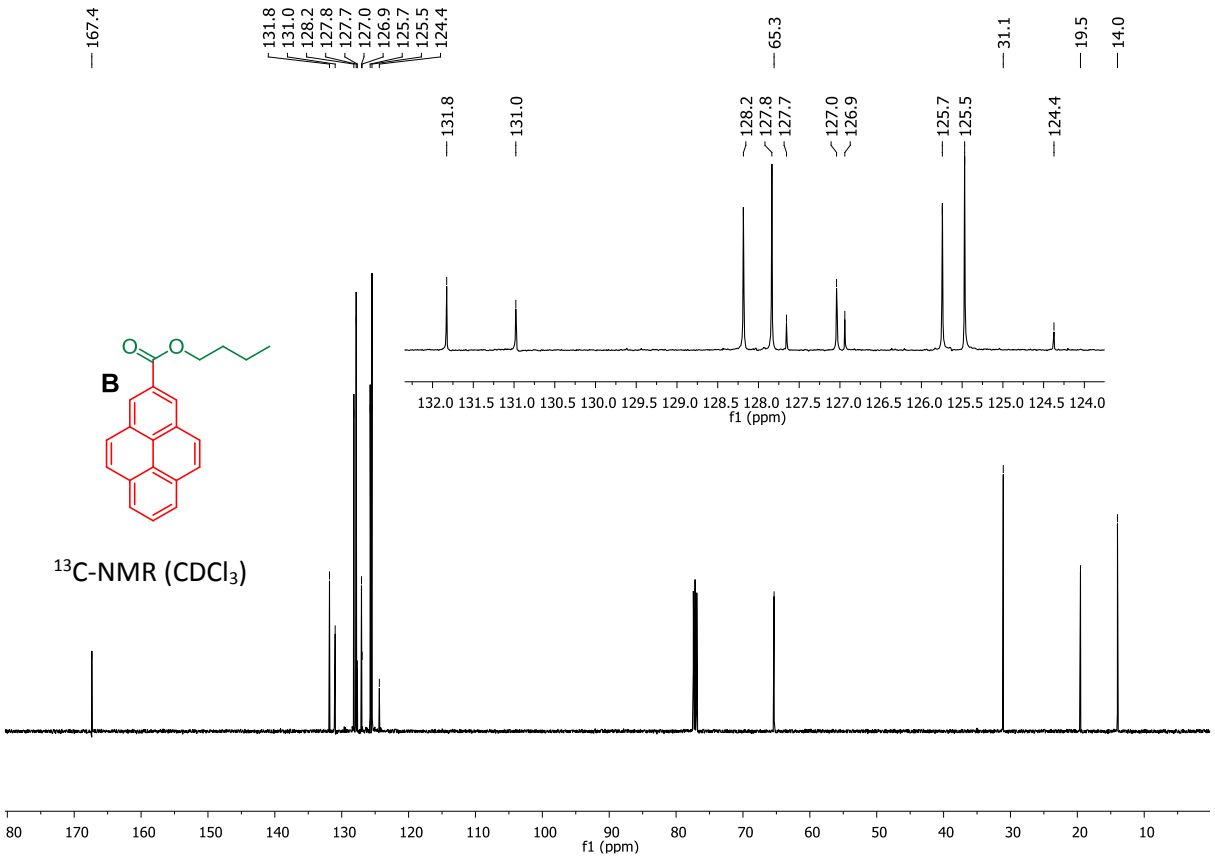


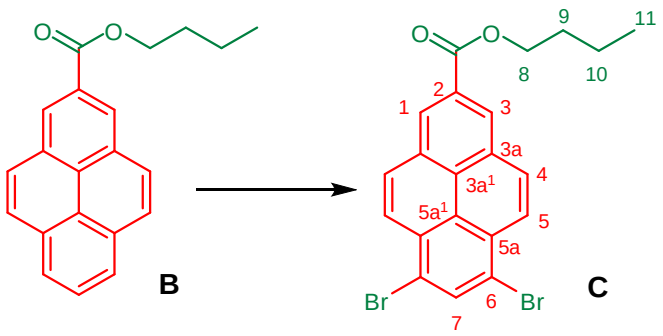
Supplementary Figure 1 | Numbering system used for the assignment and discussion of NMR spectra of receptor **4** and its precursors. Pyrenyl proton designations (red) are prefixed with "p" (i.e. p4, p2α etc.), spacer protons (green) are prefixed with "s" , and protons from the dentritic side-chains are prefixed with "d". R can be H, Na, CH_3 (d10/d10') or $C(CH_3)_3$ (d10/d10' and d11/d11').



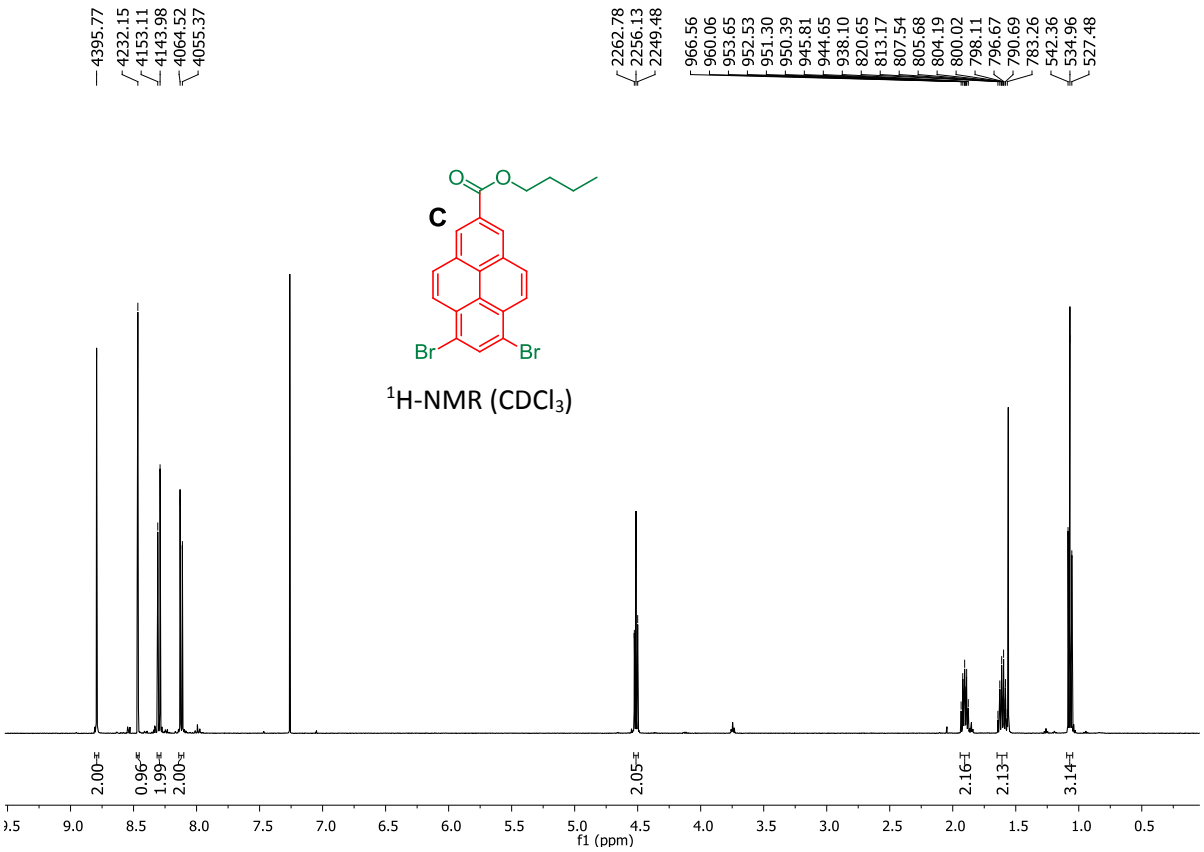
***n*-Butyl pyrene-2-carboxylate (B).** Pyrene-2-carboxylic acid (**A**)[1] (2.13 g, 8.64 mmol) was suspended in *n*-BuOH (1 L) and conc. H_2SO_4 (2 % v/v, 20 mL) was added. The mixture was heated under reflux (~118 °C) for 22 h and the solvent was reduced under vacuum to *ca.* 10 mL. The solution was diluted with CH_2Cl_2 (200 mL), washed with saturated $NaHCO_3$ solution (200 mL) and water (200 mL) and dried ($MgSO_4$). The solvent was evaporated and the residue was purified by column chromatography (EtOAc/Hexane 1:20) to yield *n*-butyl pyrene-2-carboxylate **B** (2.38 g, 7.86 mmol, 91 %) as a yellow solid. R_f = 0.55 (EtOAc/Hexane 1:20); IR: n_{max} = 3042, 2962, 2933, 2893, 1857, 1708, 1299, 1216, 902, 845, 826, 759 and 710 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$): *d* = 8.82 (s, 2H, H-1), 8.18 (d, 2H, ^3J = 7.6 Hz, H-6), 8.11 (d, 2H, ^3J = 9.1 Hz, H-4), 8.08 (d, 2H, ^3J = 9.2 Hz, H-5), 8.04 (t, 1H, ^3J = 7.6 Hz, H-7), 4.51 (t, 2H, ^3J = 6.7 Hz, H-8), 1.93-1.88 (m, 2H, H-9), 1.65-1.57 (m, 2H, H-10) and 1.08 (t, 3H, ^3J = 7.3 Hz, H-11) p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3$): *d* = 167.2 (CO), 131.7 (C-5a), 130.8 (C-3a), 128.0 (C-5), 127.7 (C-4), 127.5 (C-2), 126.9 (C-7), 126.8 (C-3a[1]), 125.6 (C-1,3), 125.3 (C-6), 124.2 (C-5a[1]), 65.2 (C-8), 30.9 (C-9), 19.4 (C-10) and 13.8 (C-11) p.p.m.; HRMS (ESI^+): *m/z* calculated for $C_{21}H_{18}O_2Na$ $[M + Na]^+$: 325.1199, found 325.1205; $C_{21}H_{19}O_2$ $[M + H]^+$: 303.1380, found 303.1385; $C_{17}H_{11}O_2$ $[M - C_4H_8]^+$: 247.0754, found 247.0758; $C_{16}H_{11}$ $[M - C_5H_8O_2]^+$: 203.0855, found 203.0860.

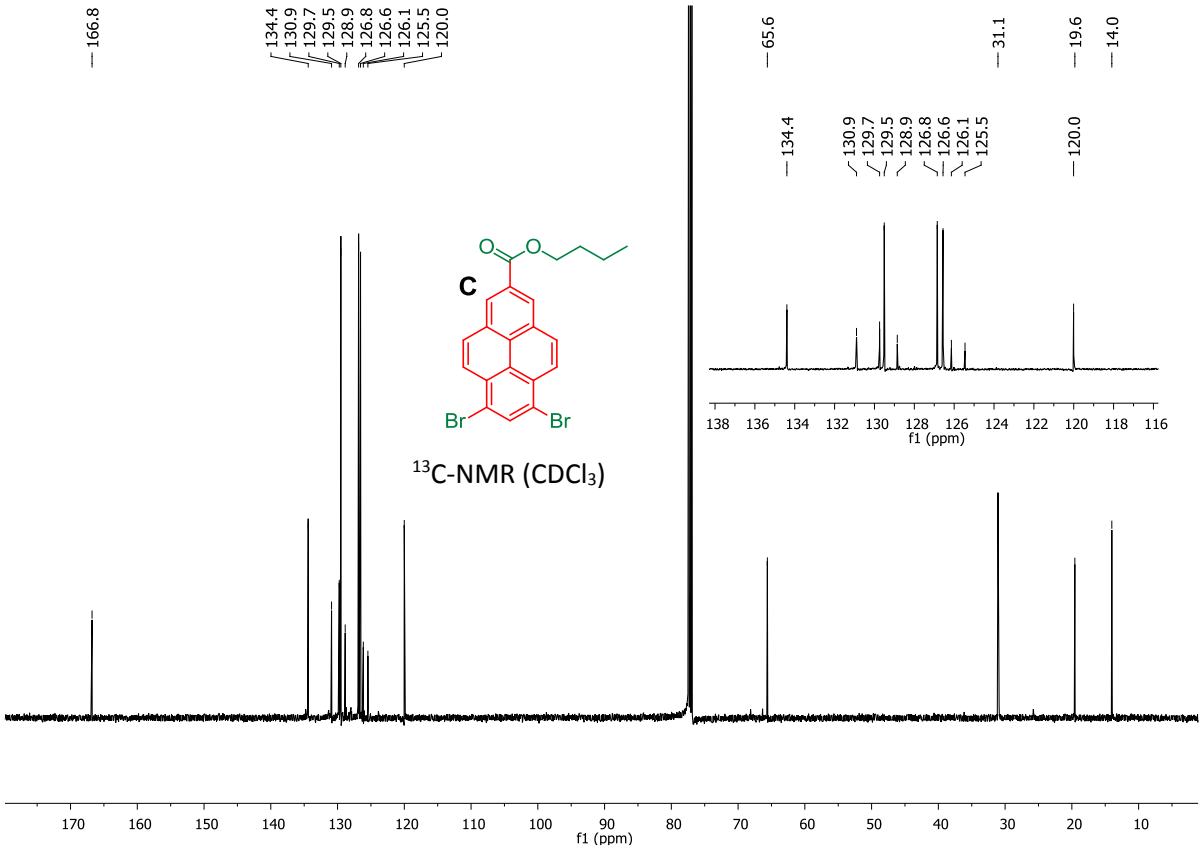


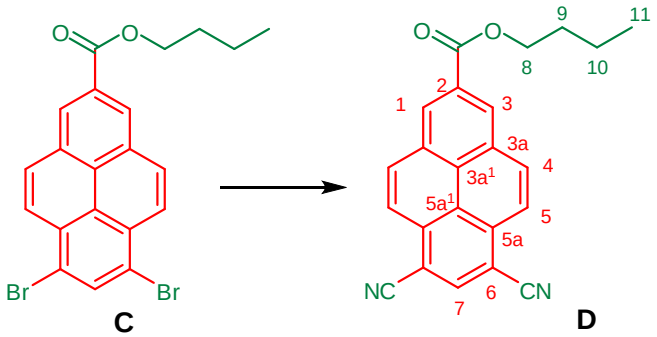

167.4
131.8
131.0
128.2
127.8
127.7
127.0
126.9
125.7
125.5
124.4
65.3
31.1
19.5
14.0
131.8
131.0
128.2
127.8
127.7
127.0
126.9
125.7
125.5
124.4
B
132.0 131.5 131.0 130.5 130.0 129.5 129.0 128.5 128.0 127.5 127.0 126.5 126.0 125.5 125.0 124.5 124.0
f1 (ppm)
13C-NMR (CDCl3)
80 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)



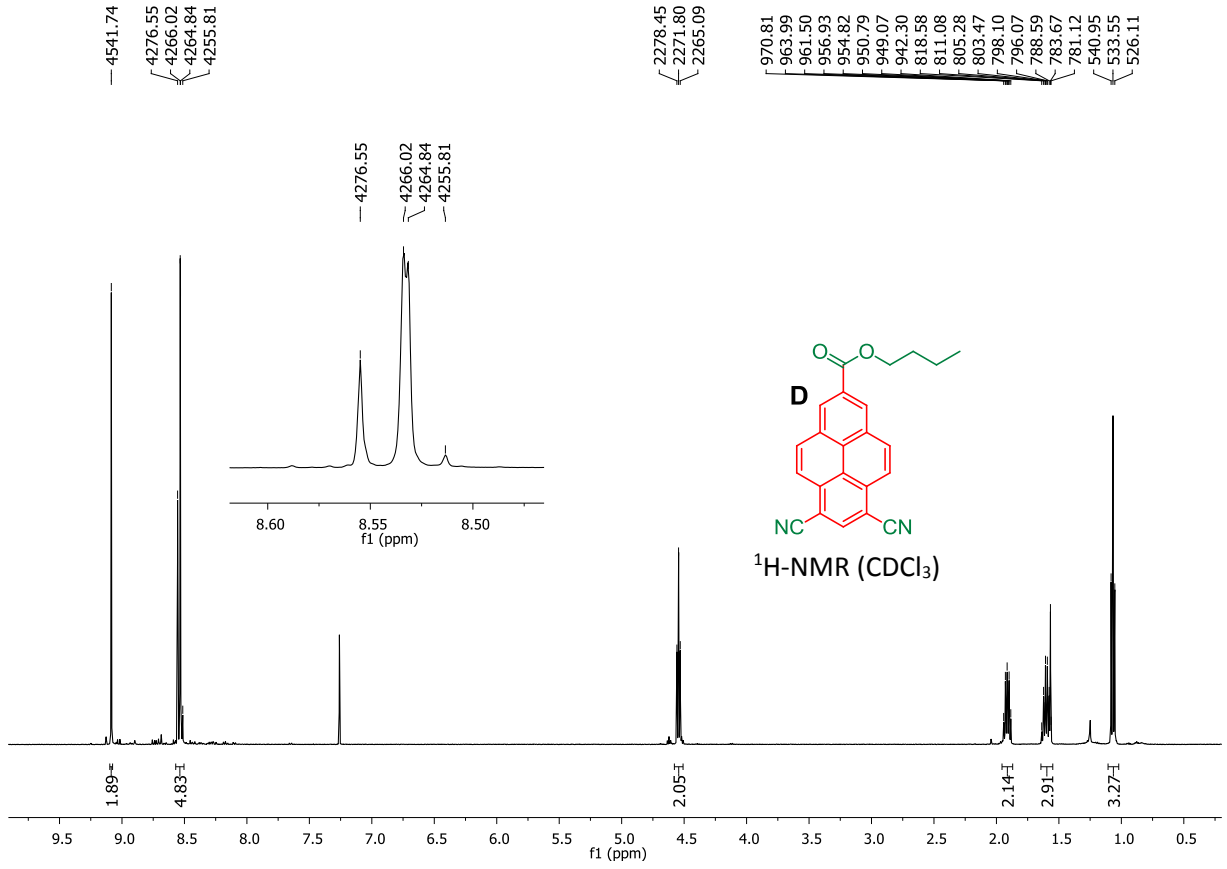
***n*-Butyl 6,8-dibromopyrene-2-carboxylate (C).** *n*-Butyl pyrene-2-carboxylate (**B**) (3.83 g, 12.7 mmol) was dissolved in dry CH_2Cl_2 (100 mL) and the solution was cooled to 0 °C under N_2 atmosphere. A solution of Br_2 (4.24 g, 1.36 mL, 26.57 mmol) in dry CH_2Cl_2 (100 mL) was then added dropwise over 45 min. The mixture was stirred overnight at room temperature and the solvent was evaporated. The residue was crystallised from EtOAc (190 mL) to yield *n*-butyl 6,8-dibromopyrene-2-carboxylate **C** (5.32 g, 11.56 mmol, 91 %) as bright yellow needles. R_f = 0.57 (EtOAc/Hexane 1:15); IR: n_{max} = 2958, 2933, 2871, 1714, 1313, 1289, 1224, 814 and 764 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$): *d* = 8.76 (s, 2H, H-1,3), 8.43 (s, 1H, H-7), 8.25 (d, 2H, ^3J = 9.2 Hz, H-5), 8.07 (d, 2H, ^3J = 9.2 Hz, H-4), 4.52 (t, 2H, ^3J = 6.6 Hz, H-8), 1.95-1.89 (m, 2H, H-9), 1.66-1.58 (m, 2H, H-10) and 1.08 (t, 3H, ^3J = 7.5 Hz, H-11) p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3$): *d* = 166.6 (CO), 134.2 (C-7), 130.6 (C-3a), 129.5 (C-5a), 129.3 (C-4), 128.6 (C-2), 126.6 (C-1,3), 126.3 (C-5), 125.8 (C-5a^1), 125.2 (C-3a^1), 119.8 (C-6), 65.4 (C-8), 30.9 (C-9), 19.4 (C-10) and 13.9 (C-11) p.p.m.; HRMS (EI$^+$): *m/z* calculated for $C_{21}H_{17}Br_2O_2$ $[M + H]^+$: 458.9590, found 458.9397.

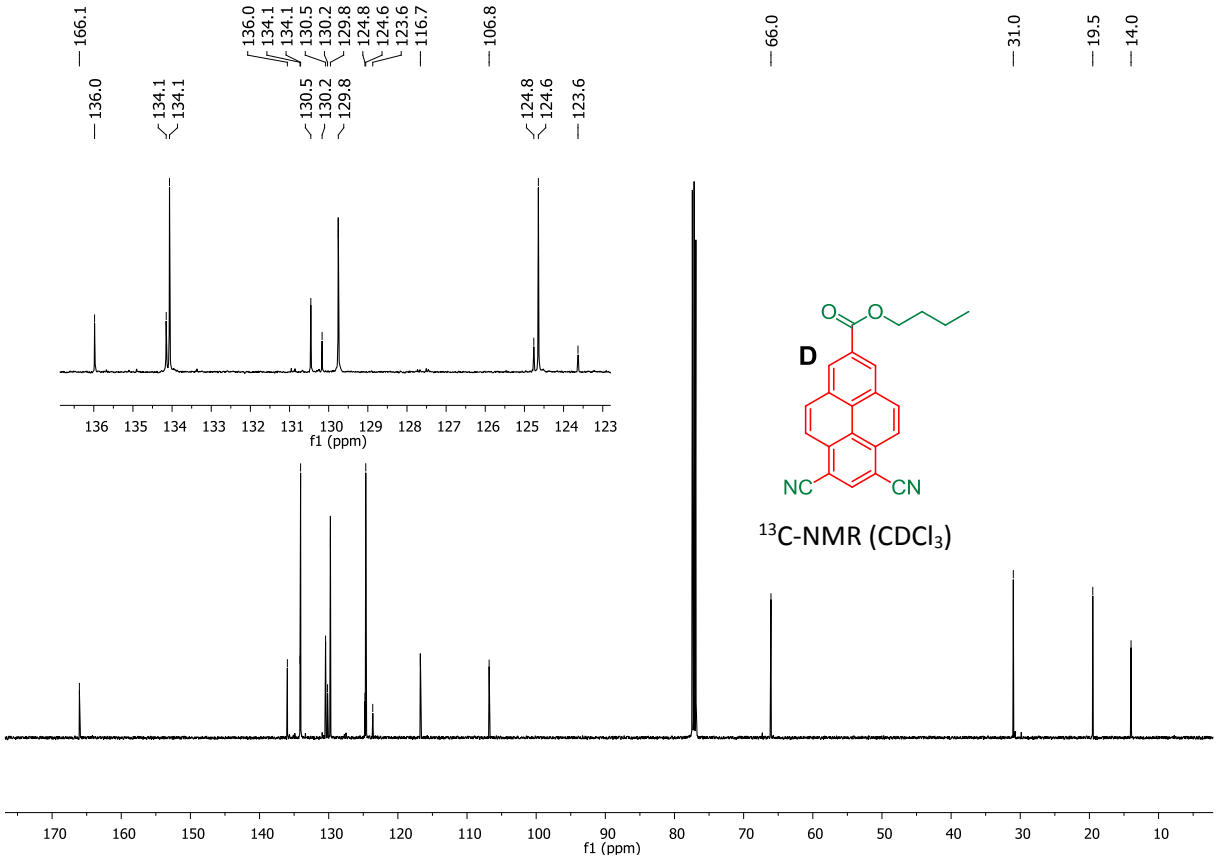



166.8
134.4
130.9
129.7
129.5
128.9
126.8
126.6
126.1
125.5
120.0
65.6
31.1
19.6
14.0
C
Br
Br
^{13}C-NMR ($CDCl_3$)
f1 (ppm)

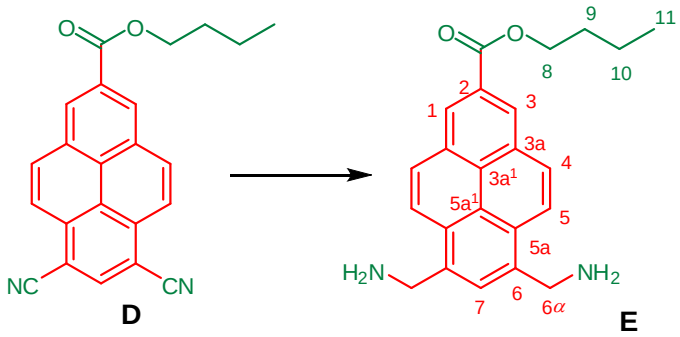


***n*-Butyl 6,8-dicyanopyrene-2-carboxylate (D).** Into a 100 mL two-necked round bottom flask containing *n*-butyl 6,8-dibromopyrene-2-carboxylate (**C**) (2.06 g, 4.48 mmol), $Zn(CN)_2$ (1.10 g, 9.37 mmol), $Pd(PPh_3)_4$ (0.83 g, 0.72 mmol), and a stirring bar, was distilled approximately 40-50 mL of DMF from P_2O_5 at around 60 °C under reduced pressure. The resulting solution was stirred and heated at 90 °C for 12 h. The solvent was then removed under vacuum and the resulting solid was purified by column chromatography (CH_2Cl_2) to yield *n*-butyl 6,8-dicyanopyrene-2-carboxylate (**D**) (1.38 g, 8.25 mmol, 88 %) as a yellow solid. R_f = 0.44 (CH_2Cl_2); IR: n_{max} = 3056, 2957, 2932, 2873, 2224, 1716, 1298, 1211, 822, 766 and 710 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$): *d* = 9.09 (s, 2H, H-1,3), 8.56 (s, 1H, H-7), 8.55 (d, 2H, ^3J = 9.2 Hz, H-5), 8.53 (d, 2H, ^3J = 9.2 Hz, H-4), 4.55 (t, 2H, ^3J = 6.7 Hz, H-8), 1.95-1.89 (m, 2H, H-9), 1.65-1.59 (m, 2H, H-9 overlapped with HDO signal) and 1.08 (t, 3H, ^3J = 7.5 Hz, H-10) p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3$): *d* = 165.9 (CO), 135.8 (C-5a), 140.0 (C-7), 133.9 (C-4), 130.3 (C-3a), 130.0 (C-2), 129.6 (C-1,3), 124.6 (C-3a^1), 124.5 (C-5), 123.5 (C-5a^1), 116.6 (CN), 106.6 (C-6), 65.9 (C-8), 30.8 (C-9), 19.4 (C-10) and 13.8 (C-11) p.p.m.; HRMS (EI^+): *m/z* calculated for $C_{23}H_{17}N_2O_2$ $[M + H]^+$: 353.1285, found 353.1128.

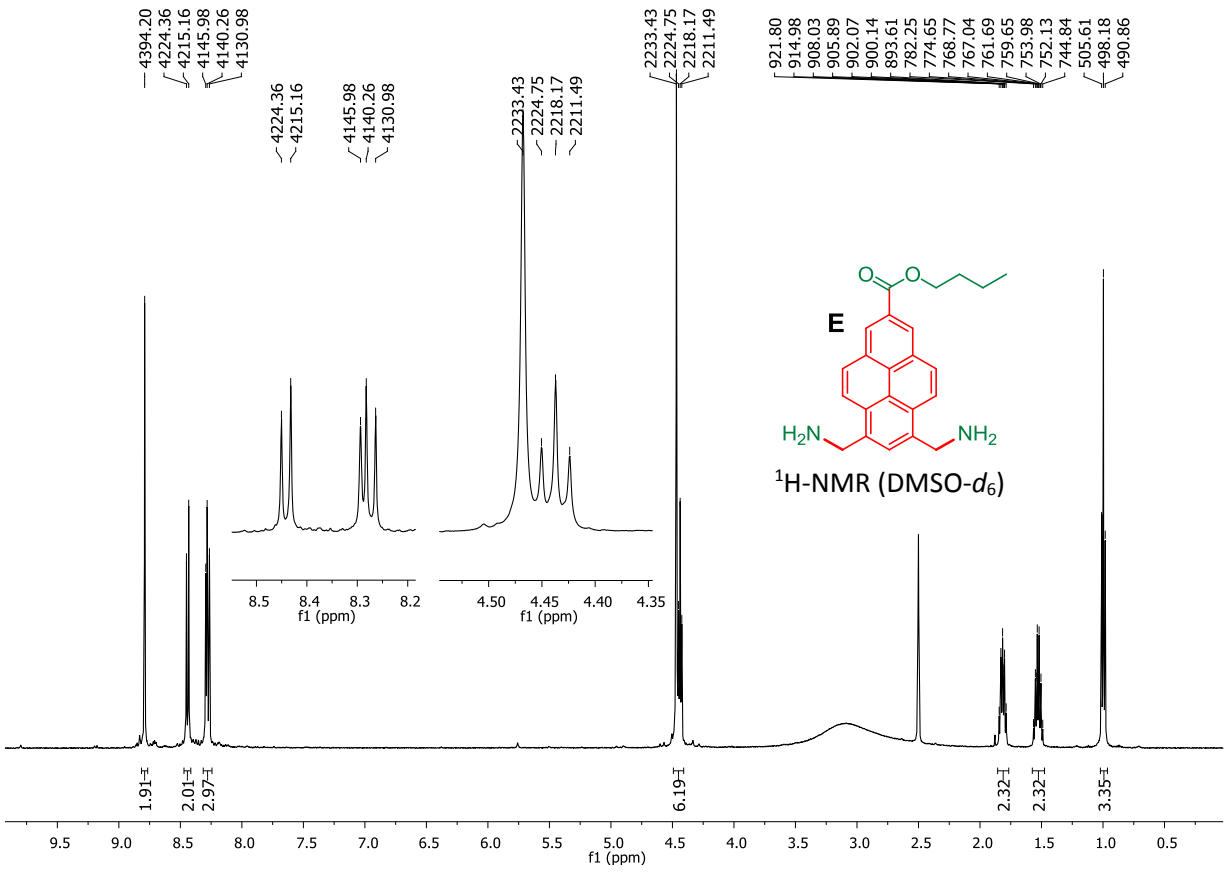


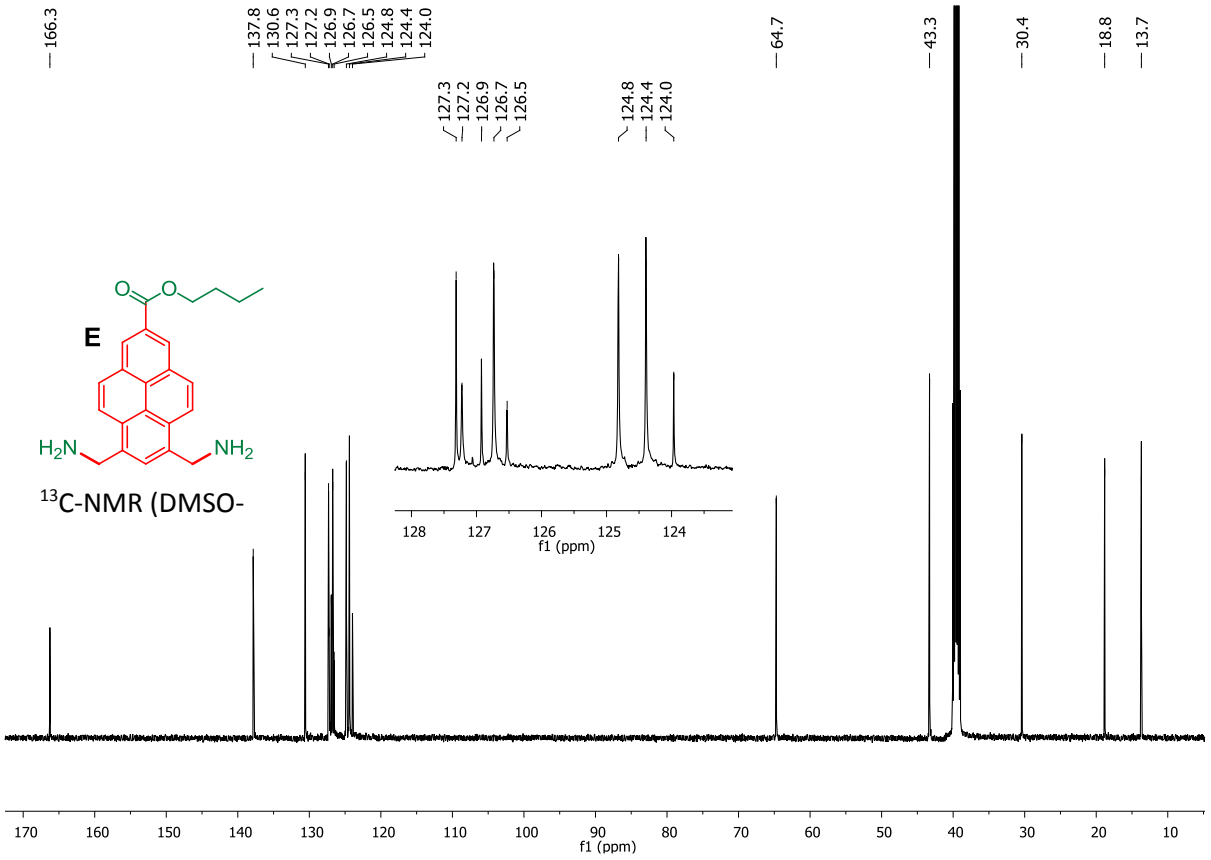


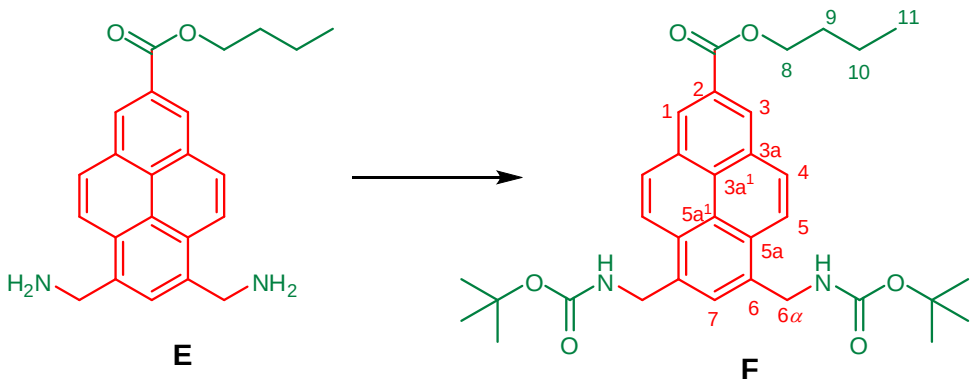
^{13}C-NMR ($CDCl_3$)



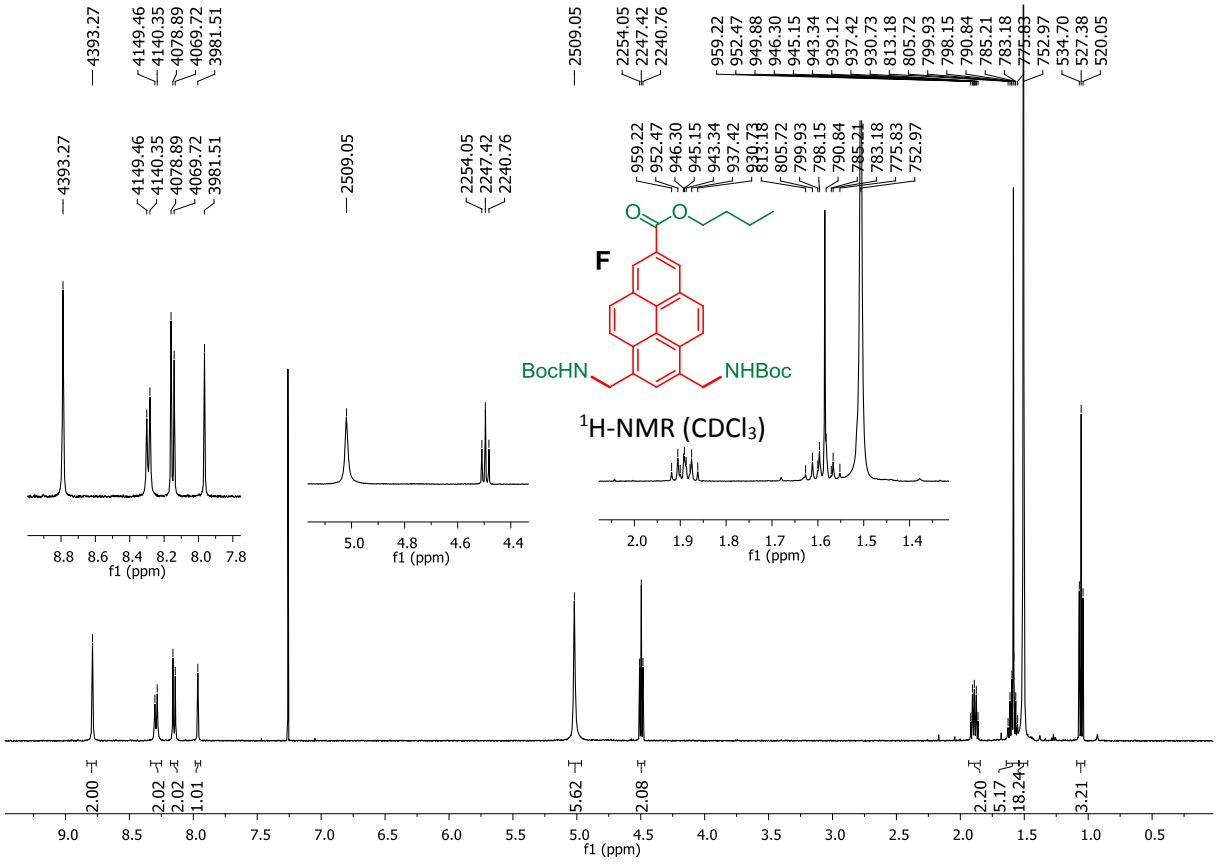
***n*-Butyl 6,8-diaminomethylpyrene-2-carboxylate (E).** *n*-Butyl 6,8-dicyanopyrene-2-carboxylate (**D**) (0.73 g, 2.07 mmol) was suspended in a mixture of *n*-butanol (75 mL), 6.7 % v/v NH_3 aqueous solution (5 mL), and ± 5 ml Raney Nickel slurry (50 % w/w in water). The mixture was stirred at room temperature under H_2 atmosphere (50 bar) for 23 h, after which the liquid was pipetted into a round bottom flask. The remaining solid Raney nickel was washed with methanol (50 mL), methanol/DCM (50 mL), and DCM (50 mL). The combined solutions were evaporated to dryness, yielding *n*-butyl 6,8-diaminomethylpyrene-2-carboxylate **E** as a pale off-white solid (0.69 g, 1.91 mmol, 92 %). R_f = 0.20 (CH_2Cl_2/MeOH-NH_3 10:1); IR: n_{max} = 3356, 2958, 2933, 2873, 1714, 1296, 1213, 897 and 806 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6): d = 8.79 (s, 2H, H-1,3), 8.44 (d, 2H, ^3J = 9.3 Hz, H-5), 8.29 (s, 1H, H-7), 8.27 (d, 2H, ^3J = 9.3 Hz, H-4), 4.47 (s, 4H, H-6α), 4.44 (t, 2H, ^3J = 6.7 Hz, H-8), 1.84-1.79 (m, 2H, H-9), 1.56-1.49 (m, 2H, H-10) and 0.99 (t, 3H, ^3J = 7.4 Hz, H-11) p.p.m.; ^{13}C NMR (125 MHz, DMSO-d_6): d = 166.3 (CO), 137.8 (C-6), 130.6 (C-3a), 127.3 (C-5a), 127.2 (C-7), 126.9 (C-2), 126.7 (C-4), 126.5 (C-3a^1), 124.8 (C-1,3), 124.4 (C-5), 124.0 (C-5a^1), 64.7 (C-8), 43.3 (C-6α), 30.4 (C-9), 18.8 (C-10) and 13.7 (C-11) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{23}H_{22}NO_2$ $[M + H - NH_3]^+$: 344.1645, found 344.1642.

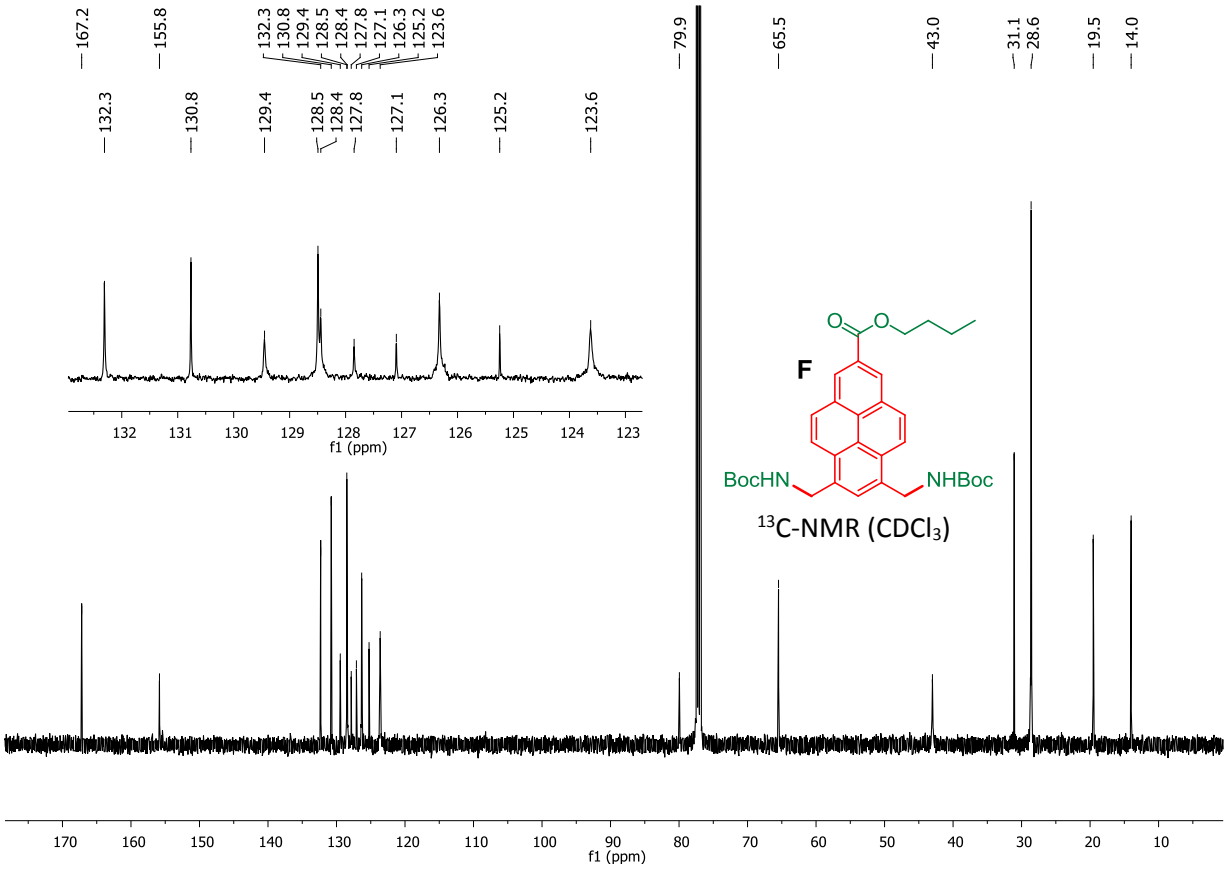


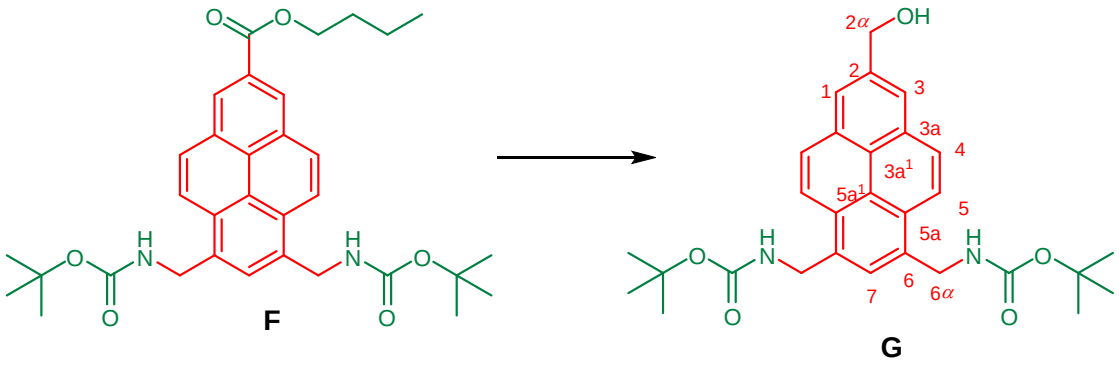

166.3
137.8
130.6
127.3
127.2
126.9
126.7
126.5
124.8
124.4
124.0
64.7
43.3
30.4
18.8
13.7
127.3
127.2
126.9
126.7
126.5
124.8
124.4
124.0
E
H2N
NH2
13C-NMR (DMSO-
128
127
126
125
124
f1 (ppm)
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)



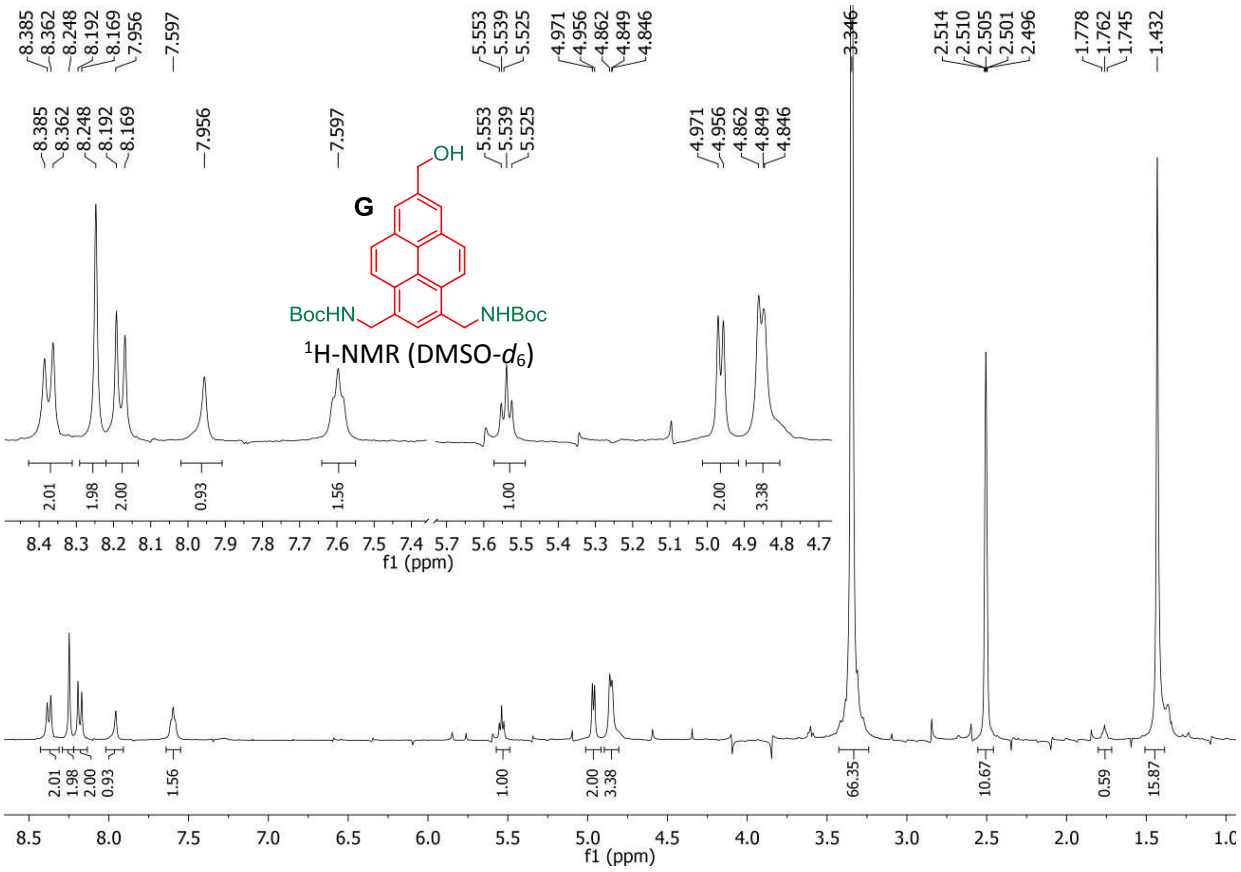
***n*-Butyl 6,8-bis[(*N*-*tert*-butyloxycarbonyl)aminomethyl]pyrene-2-carboxylate (F).** *n*-Butyl 6,8-diaminomethylpyrene-2-carboxylate (**E**) (0.23 g, 0.64 mmol) was suspended in dry THF (15 mL) under N_2 atmosphere. DIPEA (326 *m*L, 1.87 mmol) and di-*tert*-butyl dicarbonate (432 *m*L, 1.87 mmol) were added. The mixture was stirred at room temperature for 16 h and the solvent was removed under vacuum. The residue was purified by column chromatography (CH_2Cl_2/EtOAc 10:1) to yield *n*-butyl 6,8-bis[(*N*-*tert*-butyloxycarbonyl)aminomethyl]pyrene-2-carboxylate (**F**) (275 mg, 0.49 mmol, 77 %) as a white solid. R_f = 0.43 (CH_2Cl_2/EtOAc 10:1); IR: n_{max} = 3353, 2966, 2932, 2877, 1709, 1686, 1682, 1526, 1250, 1161, 1049 and 647 cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$): *d* = 8.79 (s, 2H, H-1,3), 8.29 (d, 2H, ^3J = 9.2 Hz, H-5), 8.15 (d, 2H, ^3J = 9.2 Hz, H-4), 7.96 (s, 1H, H-7), 5.02 (bs, 6H, H-6*α*, N*H*), 4.49 (t, 2H, ^3J = 6.7 Hz, H-8), 1.92-1.86 (m, 2H, H-9), 1.63-1.55 (m, 2H, H-10 overlapped with H_2O), 1.51 (s, 18H, *t*-Bu-CH_3) and 1.05 (t, 3H, ^3J = 7.3 Hz, H-11) p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3$): *d* = 167.2 (COO), 155.8 (CONH), 132.3 (C-6), 130.8 (C-3a), 129.5 (C-5a), 128.5 (C-4), 128.4 (C-7), 127.9 (C-2), 127.1 (C-3a^1), 126.3 (C-1,3), 125.3 (C-5a^1), 123.6 (C-5), 80.0 (*C*$(CH_3)_3$), 65.5 (C-8), 43.0 (C-6*α*), 31.1 (C-9), 28.6 (*t*-Bu-*C*H_3), 19.5 (C-10) and 14.0 (C-11) p.p.m.; HRMS (ESI^+): *m/z* calculated for $C_{33}H_{40}N_2O_6Na$ $[M + Na]^+$: 583.2784, found 583.2764.

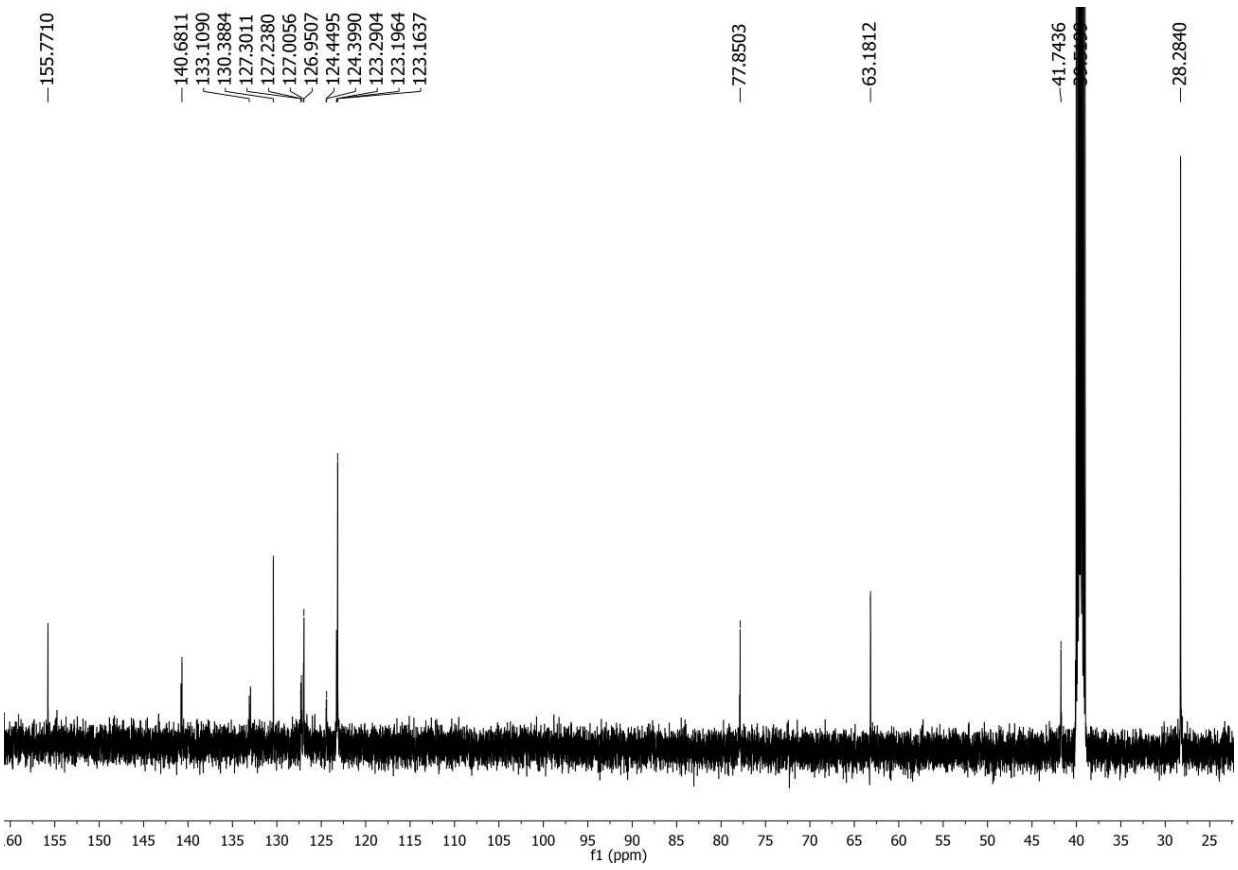


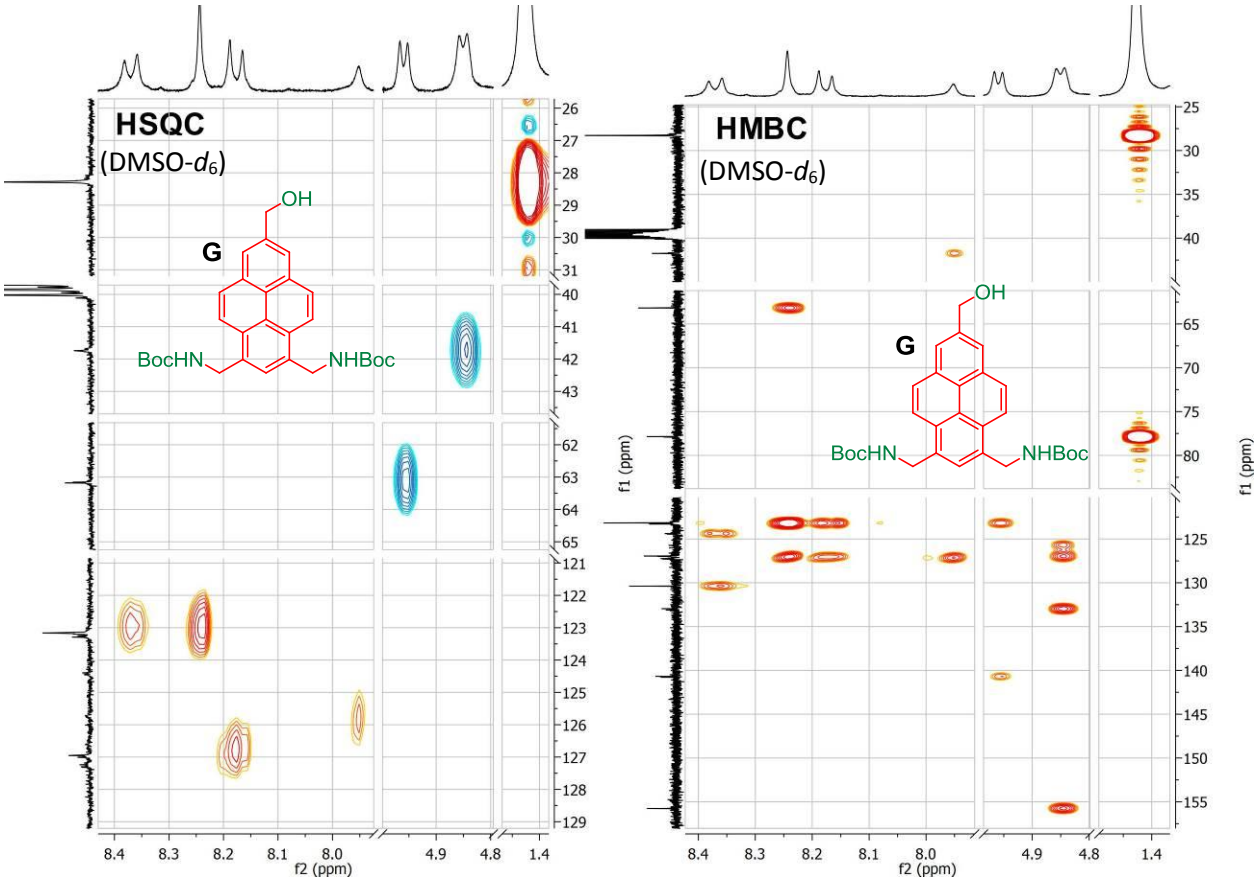

167.2
155.8
132.3
130.8
129.4
128.5
128.4
127.8
127.1
126.3
125.2
123.6
79.9
65.5
43.0
31.1
28.6
19.5
14.0
f1 (ppm)
F
BocHN
NHBoc
13C-NMR (CDCl3)

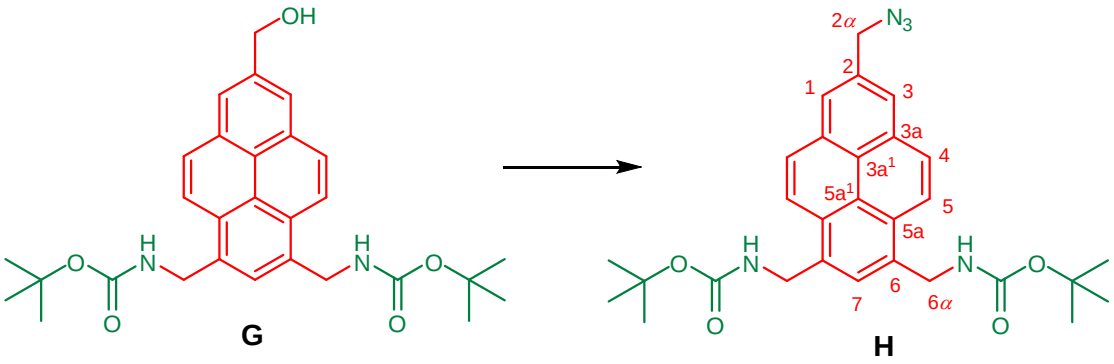


2-Hydroxymethyl-6,8-bis[(*N*-*tert*-butyloxycarbonyl)aminomethyl]pyrene (G). Ester **F** (0.54 g, 0.47 mmol) was suspended in dry THF (16 mL) under N_2 atmosphere. MeOH (115 mL, 2.85 mmol) and a solution of $LiBH_4$ in THF (4 M, 713 mL, 2.85 mmol) were added. The mixture was stirred at 50 °C for 16 h and cooled to 0 °C before water (2.5 mL) was added carefully. The solvent was then removed under vacuum and the residue was suspended in water (40 mL), isolated by filtration and washed thoroughly with water (40 mL). The resulting white solid was dissolved from the filter using CH_2Cl_2 (2×75 mL), which was then was removed under vacuum. The remaining solid was thoroughly dried by addition-evaporation of toluene (2×75 mL), affording alcohol **G** (0.53 g, 0.47 mmol, 100%) as a pale yellow solid. R_f = 0.38 (EtOAc); IR: n_{max} = 3272, 2978, 1687, 1520, 1244, 1161, 1025 and 854 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6): *d* = 8.37 (d, 2H, ^3J = 9.2 Hz, H-5), 8.25 (s, 2H, H-1,3), 8.18 (d, 2H, ^3J = 9.2 Hz, H-4), 7.96 (s, 1H, H-7), 7.60 (t, 2H, ^3J = 4.8 Hz, CON*H*), 5.54 (t, 1H, ^3J = 5.6 Hz, O*H*), 4.96 (d, 2H, ^3J = 6.0 Hz, H-2α), 4.70 (d, 2H, ^3J = 6.4 Hz, H-6α), and 1.43 (s, 18H, *t*-Bu-CH_3) p.p.m.; ^{13}C NMR (125 MHz, DMSO-d_6): d = 155.8 (CO), 140.7 (C-2) 133.1 (C-6), 130.4 (C-3a), 127.3 (C-5a), 127.2 (C-3a), 127.0 (C-4), 127.0 (C-7), 124.5 (C-5a^1), 124.4 (C-3a^1), 123.3 (C-2), 123.2 (C-1), 123.2 (C-5), 77.89 (*C*$(CH_3)_3$), 63.2 (C-2α), 41.7 (C-6α) and 28.2 (*t*-Bu-*C*H_3) p.p.m.; {^1H-^{13}C}-HSQC (500 MHz, DMSO-d_6): *d* = 127.2 (C-4), 126.3 (C-7), 123.5 (C-1,3), 123.4 (C-5), 63.6 (C-2α), 41.7 (C-6α), 28.8 (*t*-Bu-*C*H_3) p.p.m.; {^1H-^{13}C}-HMBC (500 MHz, DMSO-d_6): 155.7 (*C*O), 140.7 (C-2), 133.0 (C-6), 130.4 (C3a), 127.1 (C-5a), 126.9 (C-4), 125.7 (C-5a^1), 124.4 (C-3a^1), 123.2 (C-2), 77.9 (*C*$(CH_3)_3$), 63.2 (C-2α), 41.7 (C-6α), 28.3 (*t*-Bu-*C*H_3) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{29}H_{35}N_3O_4Na$ $[M + Na]^+$: 513.2365, found 513.2385.

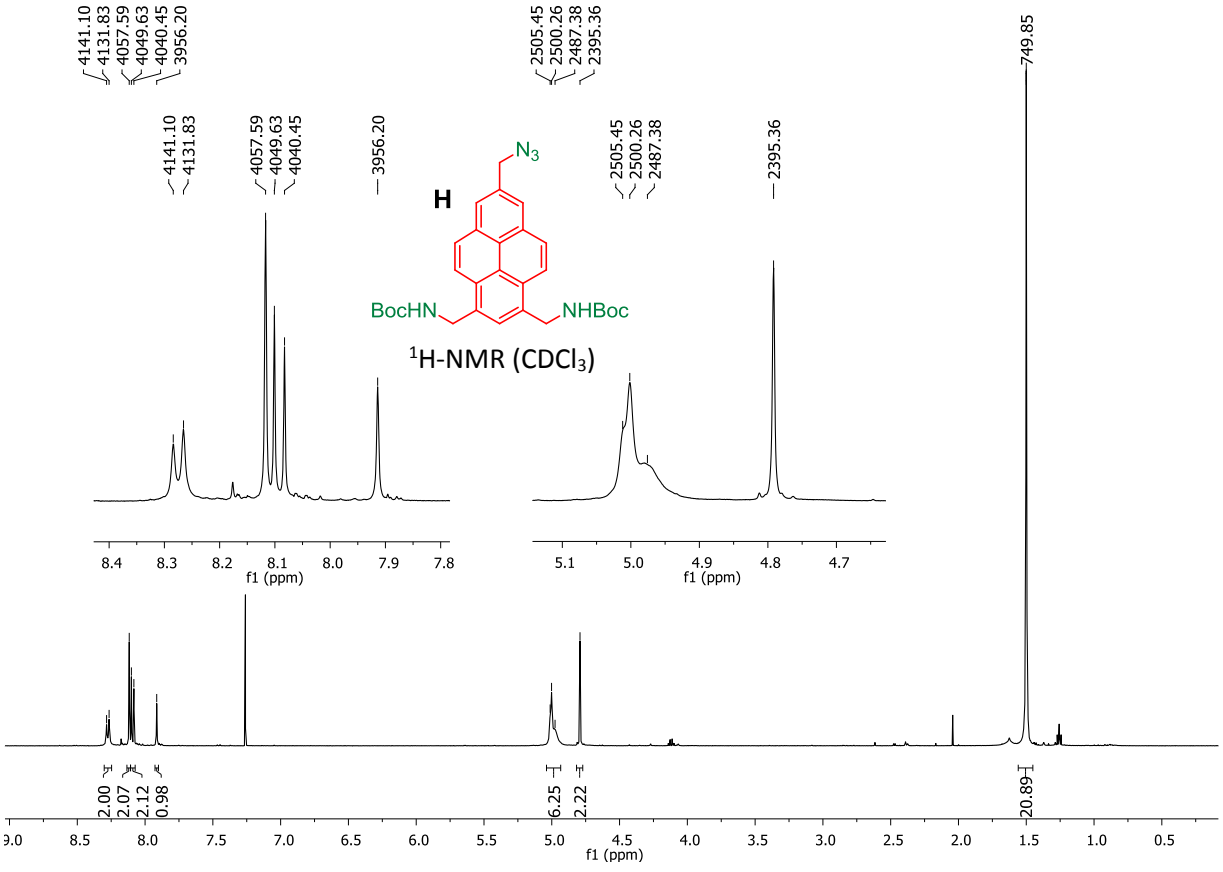

8.385
8.362
8.248
8.192
8.169
7.956
7.597
5.553
5.539
5.525
4.971
4.956
4.862
4.849
4.846
3.346
2.514
2.510
2.505
2.501
2.496
1.778
1.762
1.745
1.432
OH
G
BocHN
NHBoc
1H-NMR (DMSO-d6)
f1 (ppm)

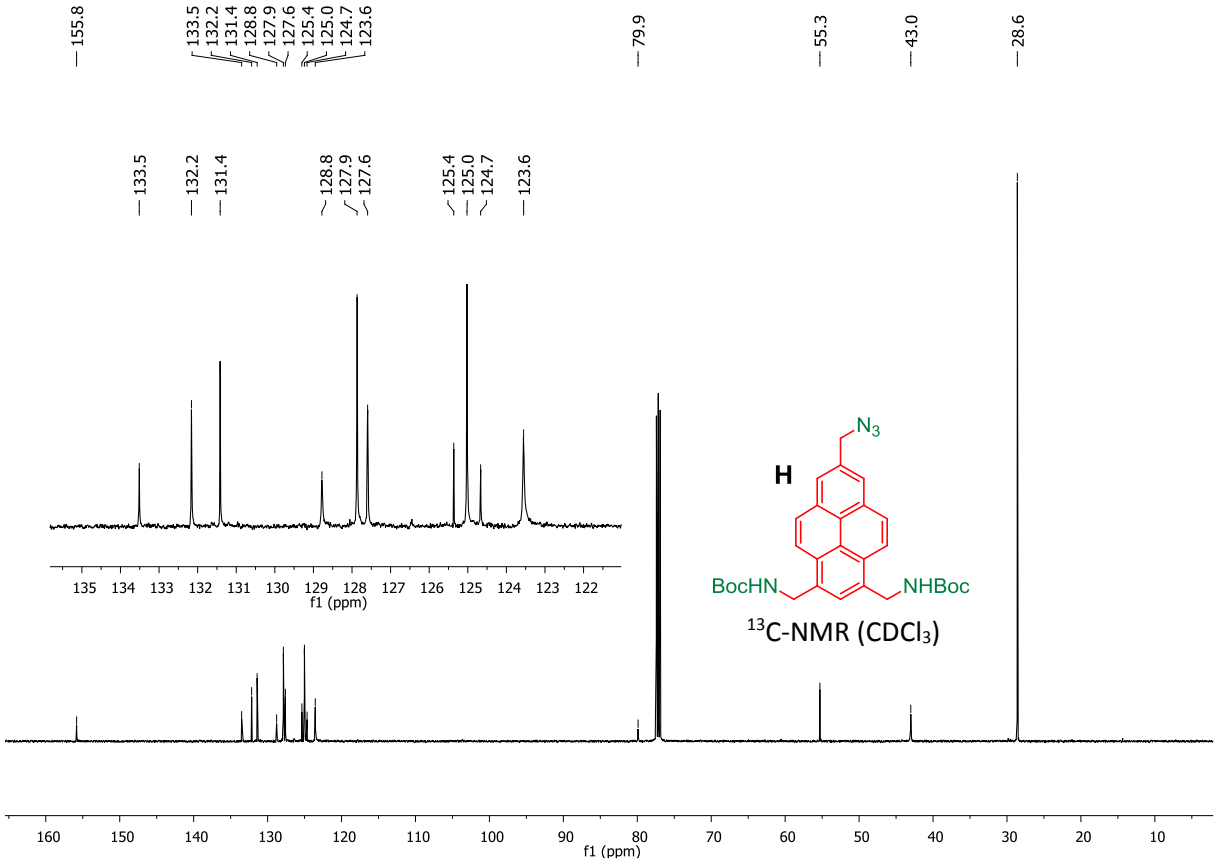

155.7710
140.6811
133.1090
130.3884
127.3011
127.2380
127.0056
126.9507
124.4495
124.3990
123.2904
123.1964
123.1637
77.8503
63.1812
41.7436
28.2840
f1 (ppm)

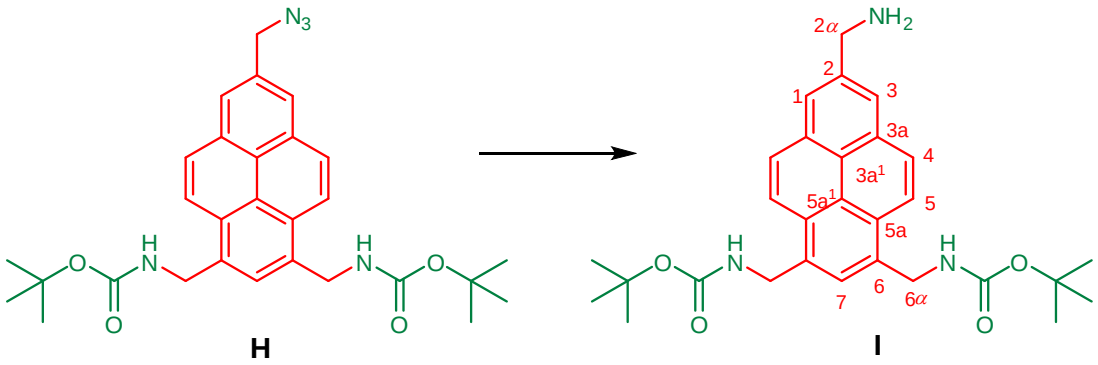

HSQC
(DMSO-d_6)
OH
G
BocHN
NHBoc
f2 (ppm)
HMBC
(DMSO-d_6)
OH
G
BocHN
NHBoc
f1 (ppm)
f2 (ppm)
f1 (ppm)



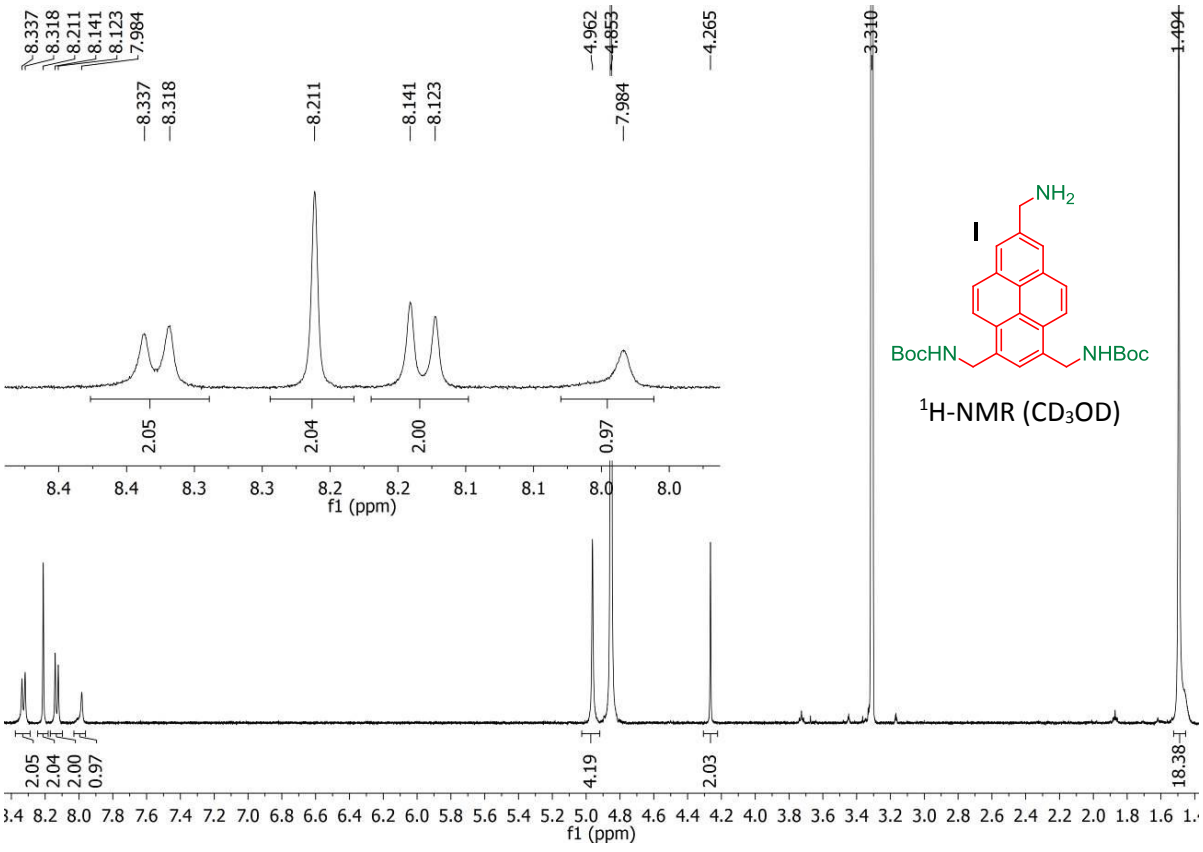
2-Azidomethyl-6,8-bis[(*N*-*tert*-butyloxycarbonyl)aminomethyl]pyrene (H). Alcohol **G** (0.25 g, 0.51 mmol) was placed under a nitrogen atmosphere and dry N_2-saturated DMF (2.5 mL) was added. The resulting suspension was cooled to 0 °C, then CBr_4 (0.34 g, 1.03 mmol) and PPh_3 (0.27 g, 1.03 mmol) were added under continuous stirring, resulting in a bright orange solution. Sodium azide (0.14 g, 2.15 mmol) was then added as one portion and the resulting solution was allowed to warm to room temperature and stirred for 16 h. Water (~25 mL) was added and the resulting precipitate was extracted with CH_2Cl_2 (4×25 mL). The collected organic phases were concentrated to dryness and the residue was purified by column chromatography (first eluted with CH_2Cl_2, then with 15:1 CH_2Cl_2/EtOAc) to afford azide **H** (0.13 g, 0.252 mmol, 50%) as a light yellow solid. R_f = 0.38 ($CHCl_3$/EtOAc 10:1); IR: n_{max} = 3356, 2982, 2935, 2879, 2095, 1686, 1675, 1519, 1270, 1249, 1162, 1049 and 860 cm^{-1}; ^{1}H NMR (500 MHz, $CDCl_3$): *d* = 8.27 (d, 2H, ^3J = 9.3 Hz, H-5), 8.12 (s, 2H, H-1,3), 8.09 (d, 2H, ^3J = 9.3 Hz, H-4), 7.91 (s, 1H, H-7), 5.01-4.97 (m, 6H, H-6*α*, N*H*), 4.79 (s, 2H, H-2*α*) and 1.50 (s, 18H, *t*-Bu-CH_3) p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3$): *d* = 155.8 (CO), 133.5 (C-2) 132.2 (C-6), 131.4 (C-3a), 128.8 (C-5a), 127.9 (C-4), 127.6 (C-7), 125.4 (C-5a^1), 125.0 (C-1,3), 124.7 (C-3a^1), 123.6 (C-5), 79.9 (*C*$(CH_3)_3$), 55.3 (C-2*α*), 43.0 (C-6*α*) and 28.6 (*t*-Bu-*C*H_3) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{29}H_{33}N_5O_4Na$ [M + Na]$^+$: 538.2430, found 538.2408.

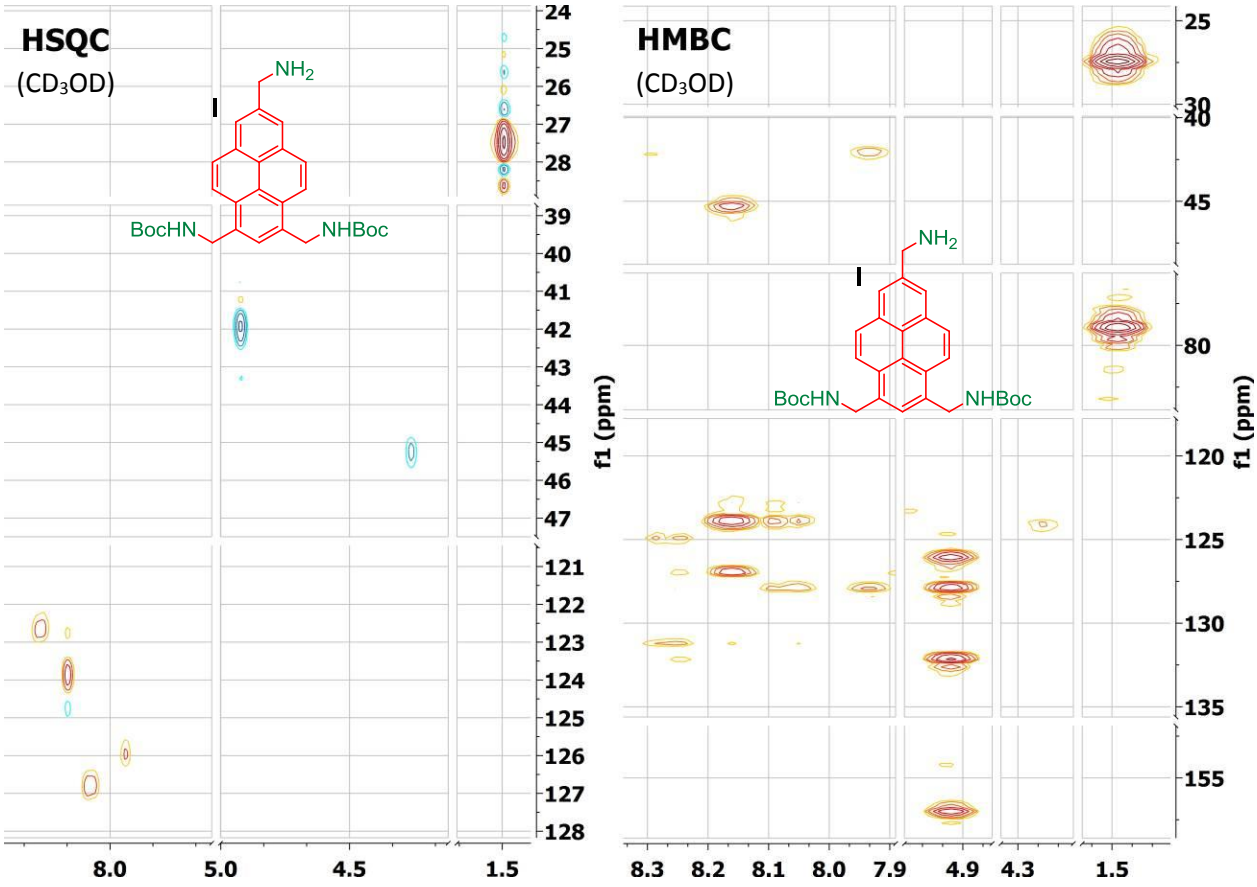


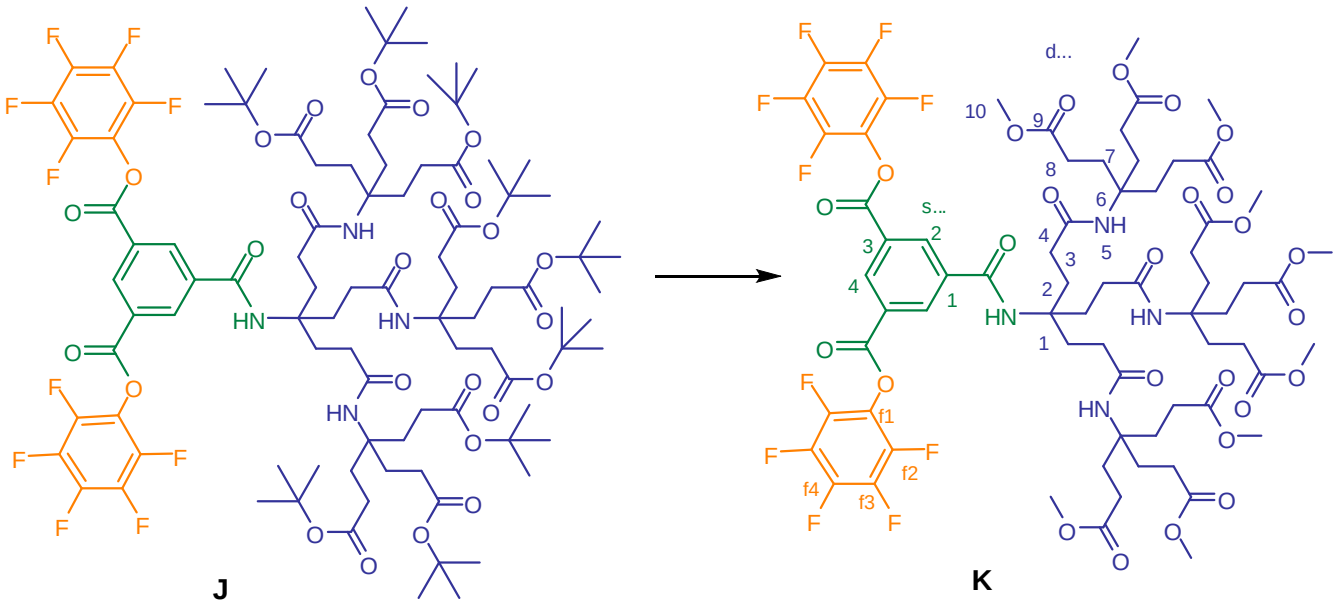

155.8
133.5
132.2
131.4
128.8
127.9
127.6
125.4
125.0
124.7
123.6
79.9
55.3
43.0
28.6
N3
H
BocHN
NHBoc
13C-NMR (CDCl3)
f1 (ppm)



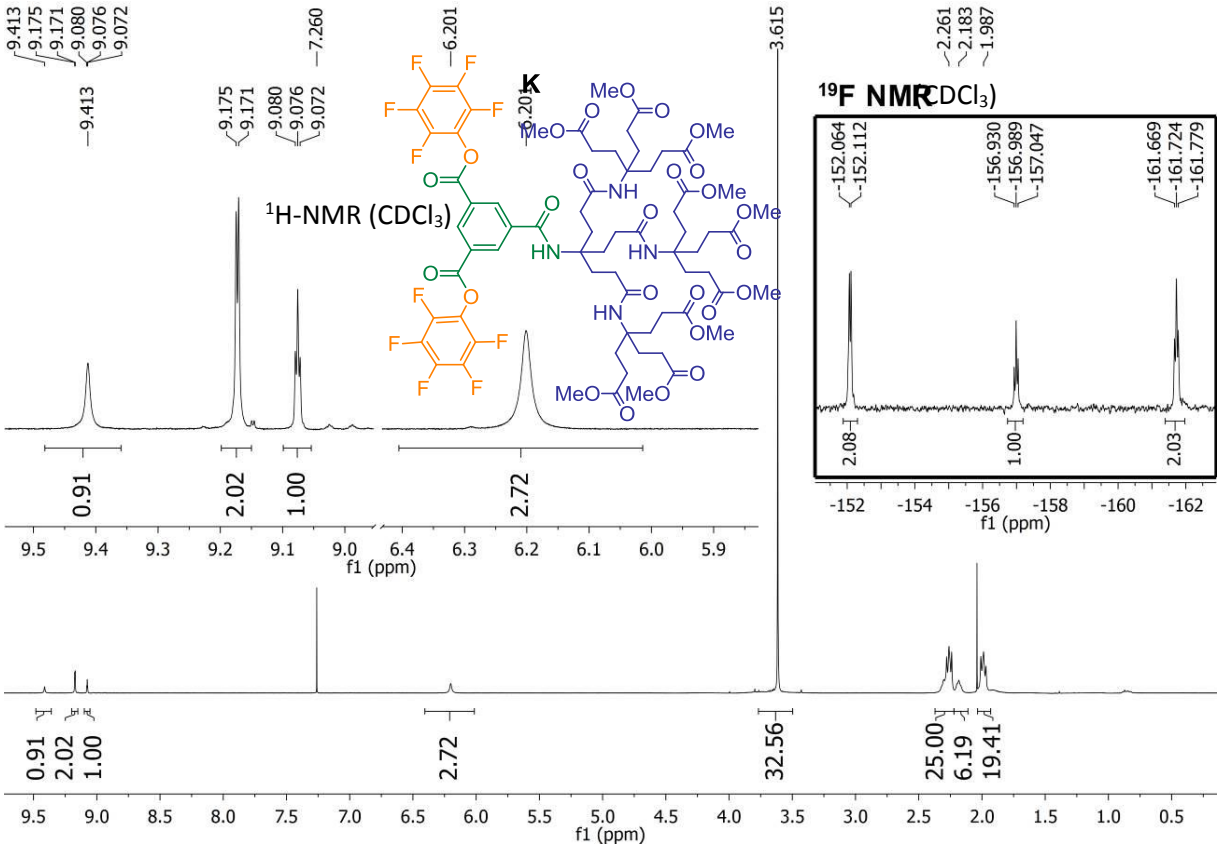
2-Aminomethyl-6,8-bis[(*N*-*tert*-butyloxycarbonyl)aminomethyl]pyrene (I). To a stirred solution of azide **H** (0.13 g, 0.252 mmol) in dry and degassed THF (30 mL), under a nitrogen atmosphere, was added a solution of $P(CH_3)_3$ in THF (1 M, 1.5 mL, 1.5 mmol). After stirring for 1 h, N_2-saturated water was added (2 mL) and the reaction mixture was stirred for an additional 1 h at 40 °C in an oil bath. Then, with the oil bath still at 40 °C, a gentle but steady stream of N_2 was passed over the solution until only a solid remained. DMSO (5 mL) was added and the solution was freeze-dried to yield amine **I** (123 mg, 252 mmol, 100%) as a light yellow solid. R_f = 0.42 (NH_4OH (aq.)/MeOH/CH_2Cl_2 1:9:90); ^{1}H NMR (500 MHz, CD_3OD): *d* = 8.33 (d, 2H, ^3J = 9.5 Hz, H-5), 8.21 (s, 2H, H-1,3), 8.14 (d, 2H, ^3J = 9.0 Hz, H-4), 7.98 (s, 1H, H-7), 4.96 (s, 4H, H-6*α*), 4.27 (s, 2H, H-2*α*) and 1.49 (s, 18H, *t*-Bu-C*H*$_3$) p.p.m.; {^{1}H-^{13}C}-HSQC (500 MHz, CD_3OD): *d* = 126.8 (C-4), 126.0 (C-7), 123.9 (C-1,3), 122.7 (C-5), 45.2 (C-2*α*), 41.9 (C-6*α*), 27.5 (*t*-Bu-*C*H$_3$) p.p.m.; {^{1}H-^{13}C}-HMBC (500 MHz, CD_3OD): 157.0 (*C*O), 132.1 (C-2), 131.2 (C-6), 127.9 (C3a), 126.9 (C-4), 126.1 (C-5a, C-7), 124.9 (C-3a^1, C-5a^1), 124.1 (C-5), 123.9 (C-1), 78.9 (*C*(CH$_3$)$_3$), 45.3 (C-2*α*), 42.0 (C-6*α*), 27.4 (*t*-Bu-*C*H$_3$) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{29}H_{35}N_3O_4Na$ [M + Na]$^+$: 512.2525, found 512.2541.

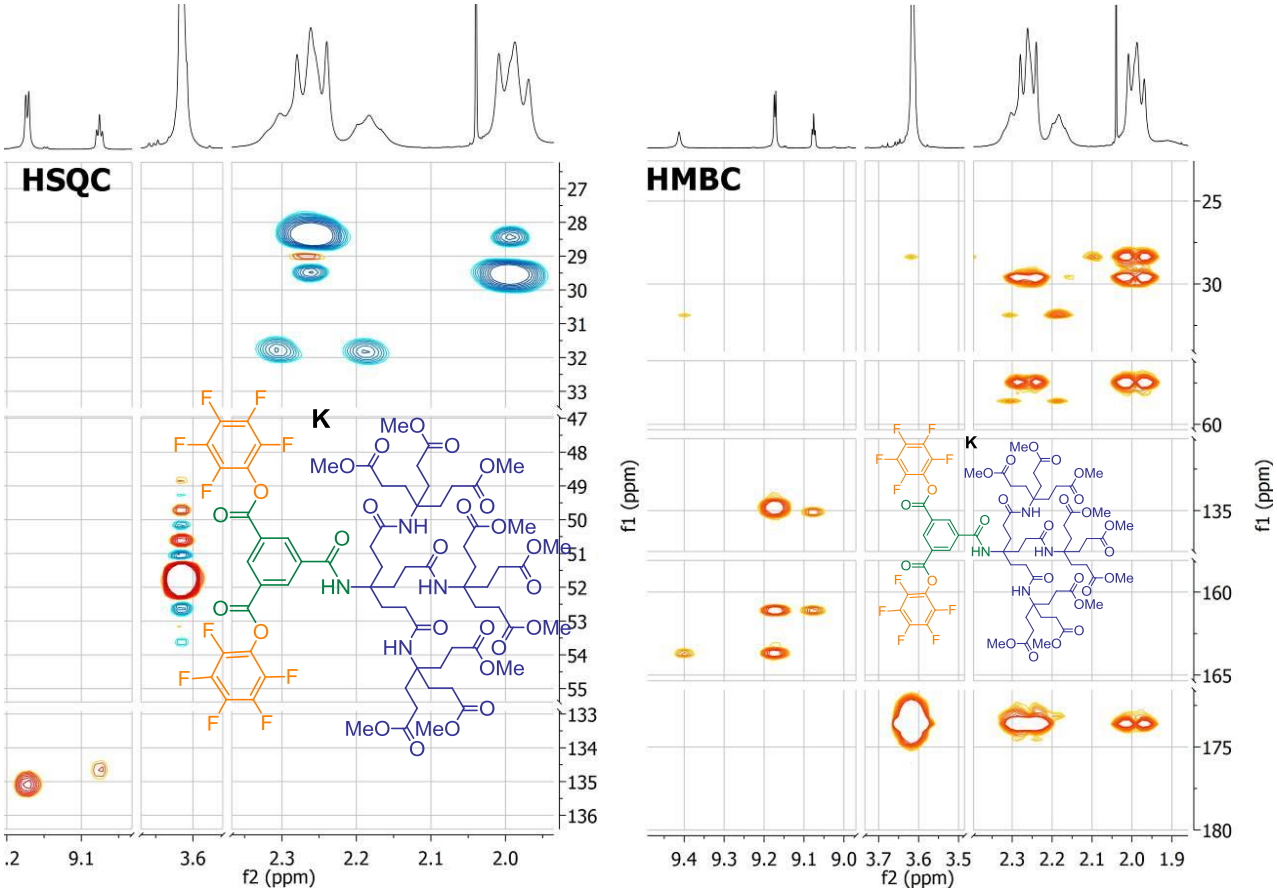


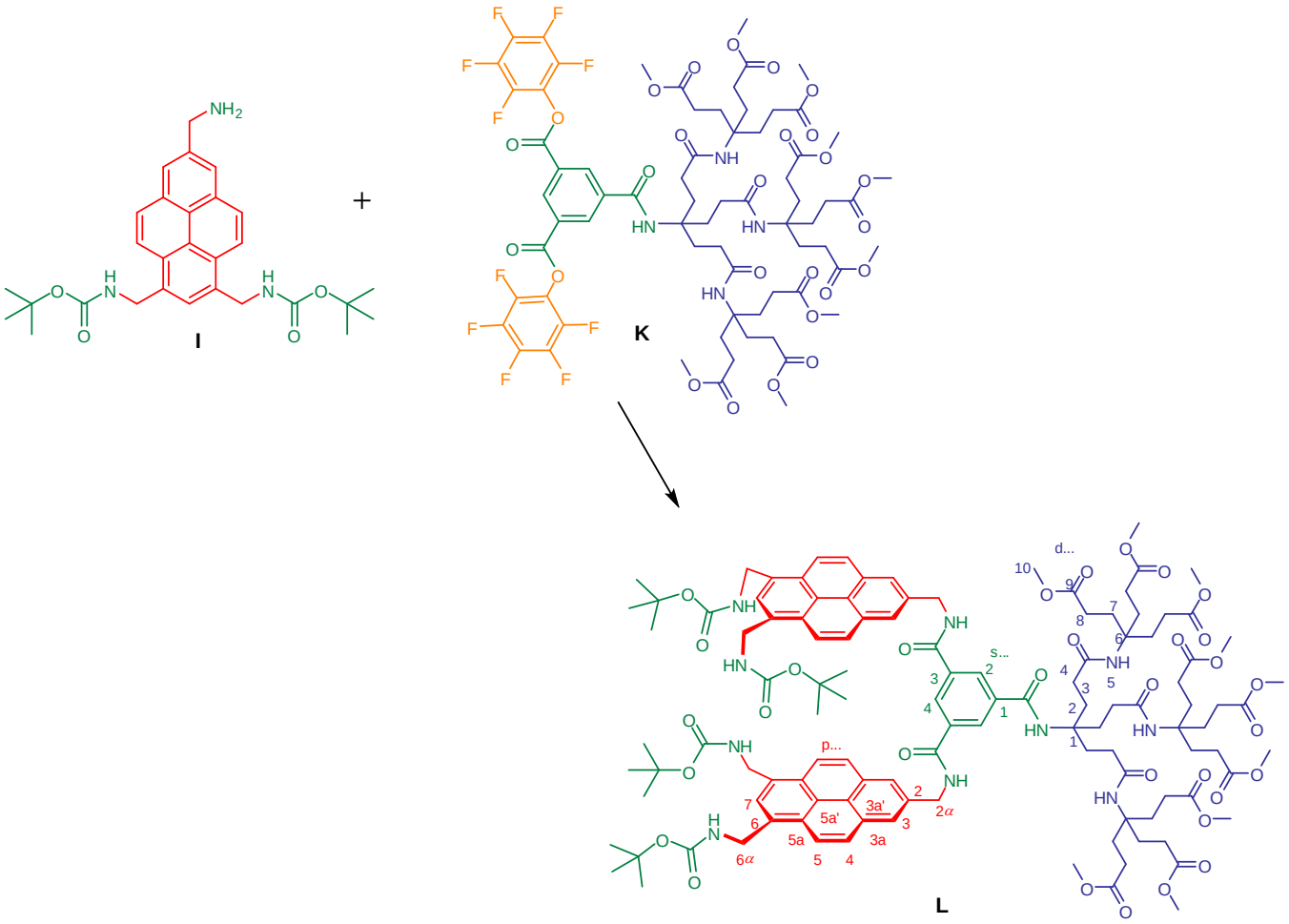

HSQC
(CD3OD)
NH2
BocHN
NHBoc
HMBC
(CD3OD)
NH2
BocHN
NHBoc
f1 (ppm)
f1 (ppm)



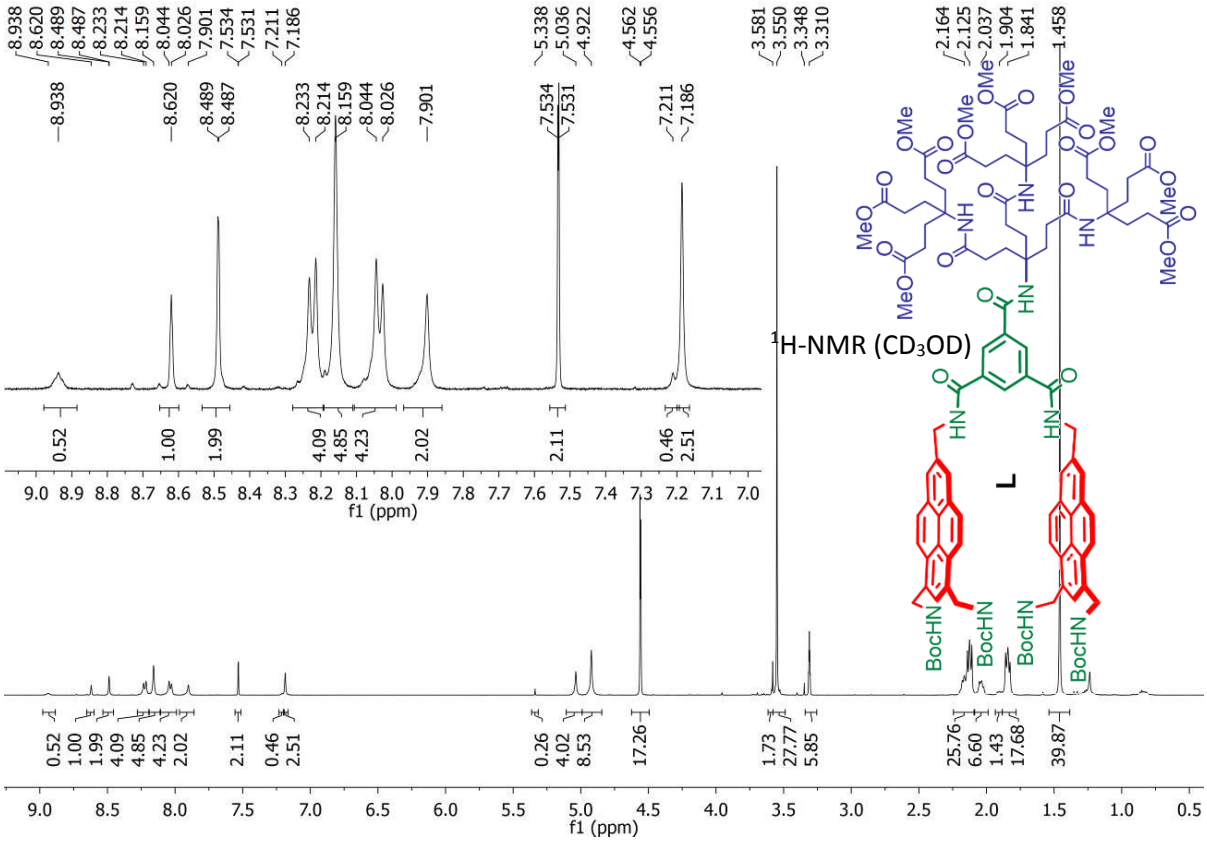
Bis(pentafluorophenyl) ester K. To a stirred solution of bis(pentafluoropheny)l ester **J** (*16*) (0.61 g, 0.31 mmol) in dry DCM (5 mL) was added TFA (5 mL) in a dropwise manner. The resulting solution was stirred for 16 h, after which the volatiles were removed by passing a N_2 stream over the stirred mixture. The remaining solid was dissolved in a 1:1 trimethyl orthoformate:MeOH mixture (15 mL), concentrated HCl was added (5 % v/v, 0.75 mL), and the mixture was stirred for 24 h. After removal of the volatiles, the solid was passed through a silica column (Ø = 2cm; ↕ = 10 cm) with EtOAc, and concentrated to dryness to yield bis(pentafluorophenyl) ester **K** as a white solid (0.46 g, 0.29 mmol, 94%). R_f = 0.55 (EtOAc). ^{1}H NMR (500 MHz, $CDCl_3$): *d* = 9.41 (s, 1H, s1-N*H*), 9.17 (d, 2H, ^4J = 1.0 Hz, H-s2), 9.08 (t, 1H, ^4J = 1.0 Hz, H-s4), 6.20 (s, 3H, d5-N*H*), 3.61 (s, 27H, H-d10), 2.26 (m, 24H, H-d3/H-d4), 2.18 (m, 6H, H-d2) and 1.99 (t, 18H, ^3J = 10.0 Hz, H-d7) p.p.m.; ^{19}F NMR (470.4 MHz, $CDCl_3$): *d* = -152.09 (d, 2F, ^3J = 11.3 Hz, F-f2), -156.99 (t, 1F, ^3J = 48.9 Hz, F-f4), -161.72 (t, 2F, ^3J = 25.9 Hz, F-f3) p.p.m.; {^{1}H-^{13}C}-HSQC (500 MHz, $CDCl_3$): *d* = 126.8 (C-4), 135.1 (C-s2), 134.6 (C-s4), 51.7 (C-d10), 31.8 (C-d3), 31.7 (C-d2), 29.5 (C-d7), 28.3 (C-d8) p.p.m.; {^{1}H-^{13}C}-HMBC (500 MHz, $CDCl_3$): 173.6 (C-d9), 173.5 (C-d4), 163.7 (C-s1*C*O), 161.1 (C-s3*C*O), 135.1 (C-s2), 134.8 (C-s4), 58.6 (C-d1), 57.5 (C-d6), 31.9 (C-d2/d3), 29.5 (C-d7), 28.4 (C-d8), p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{70}H_{82}F_{10}N_4O_{26}Na$ [M + Na]$^+$: 1607.4950, found 1607.4952.

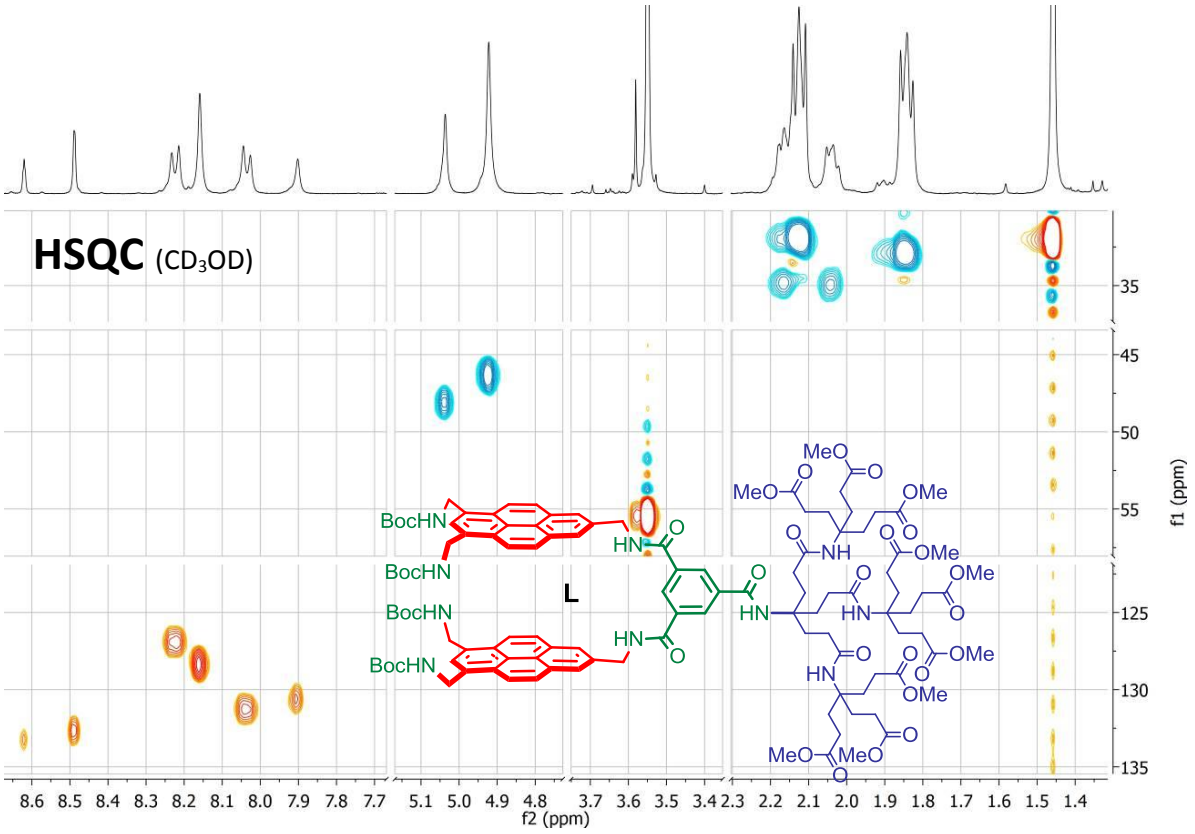

^{1}H-NMR ($CDCl_3$)
^{19}F NMR ($CDCl_3$)
K
f1 (ppm)

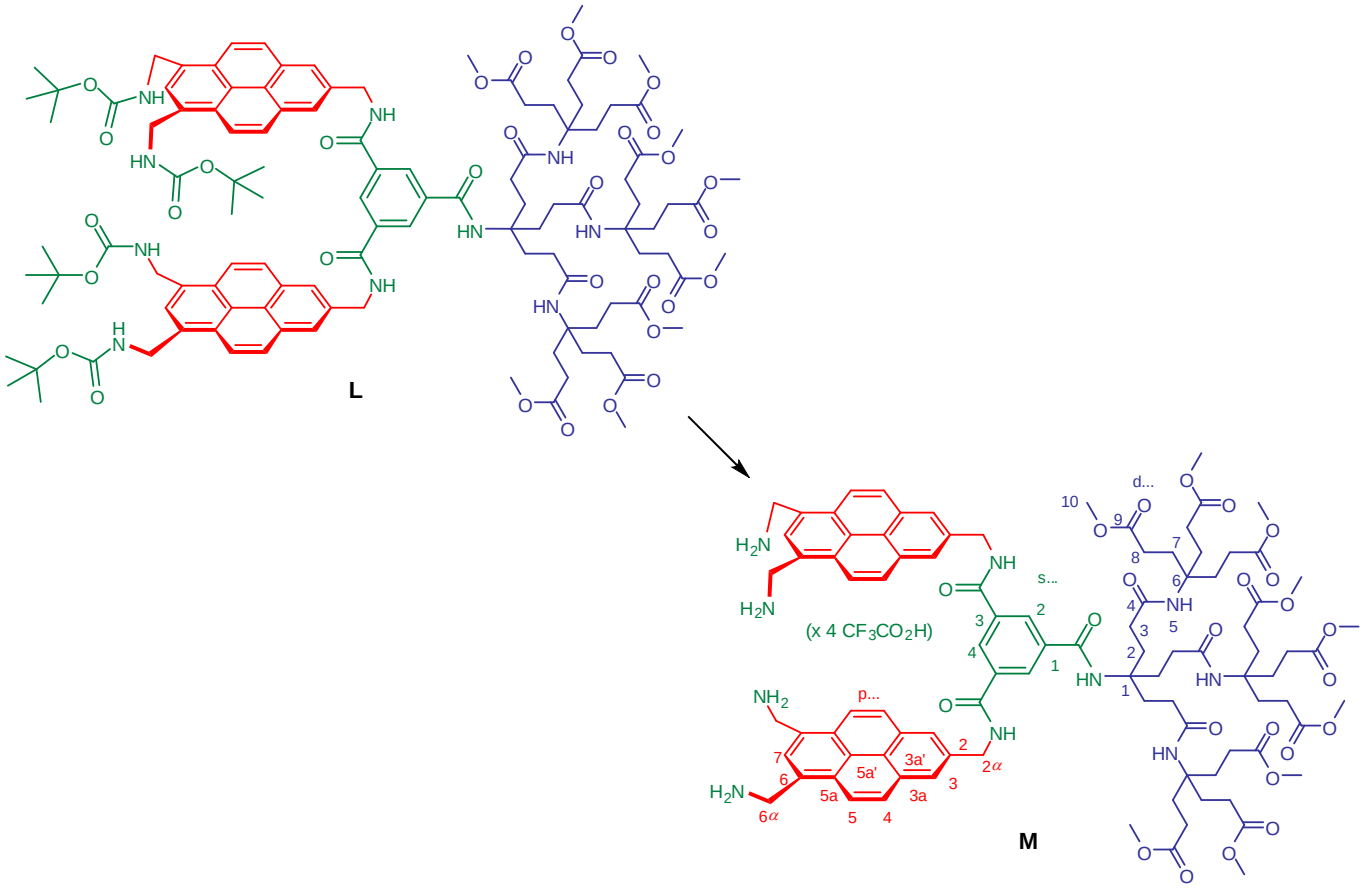

HSQC
HMBC
K
f2 (ppm)
f1 (ppm)



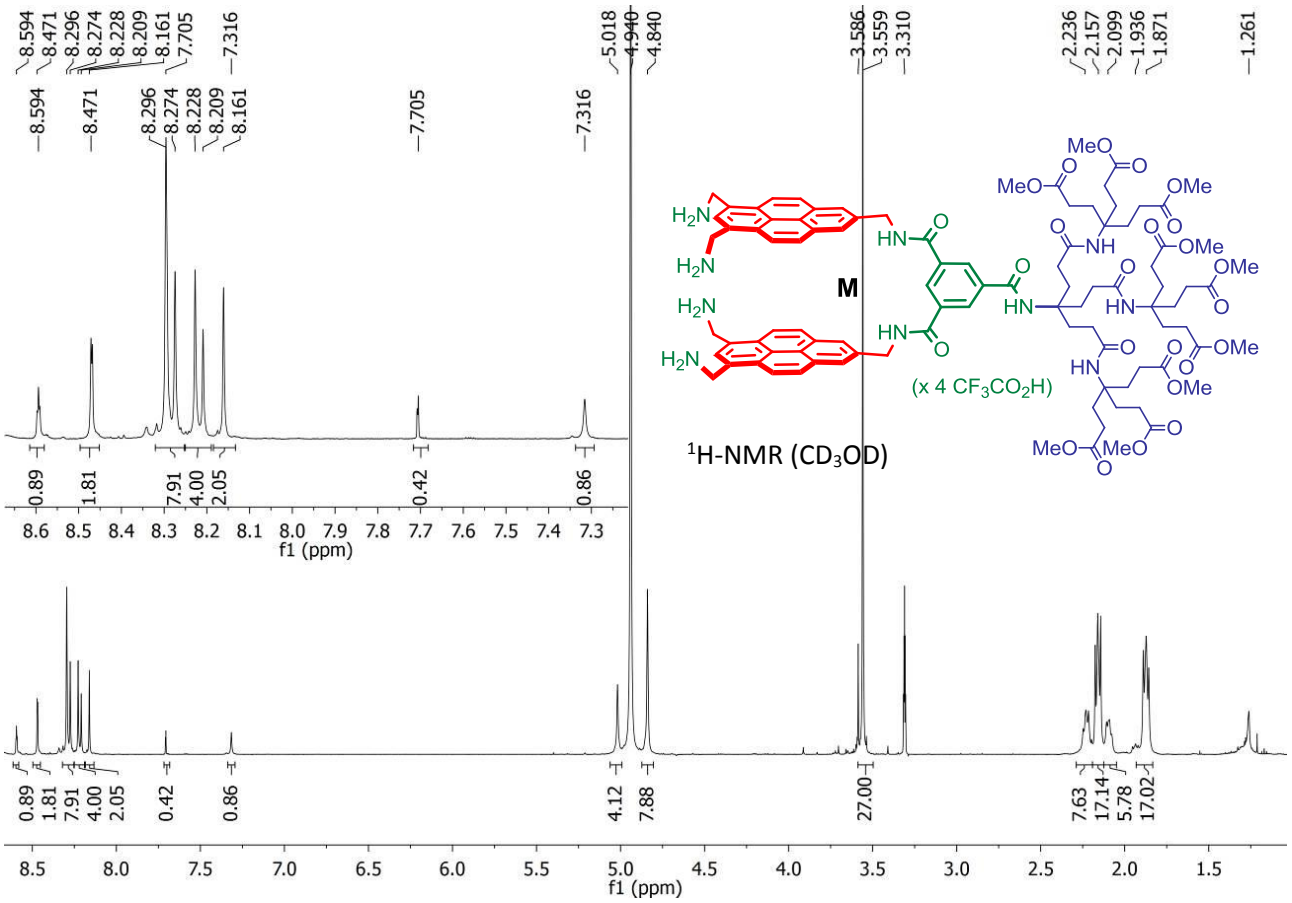
Boc-protected tetraamino-dipyrene L. To a stirred solution of bis-pentafluorophenyl ester **K** (200 mg, 0.126 mmol) in THF (1 mL) was added amine **I** (157 mg, 0.320 mmol) dissolved in THF (5×1 mL portions). DIPEA (0.44 mL, 2.52 mmol) was added dropwise, the solution was stirred for 12 h, and the volatiles removed under vacuum. Pure Boc-protected tetraamine **L** (261 mg, 0.126 mmol, 100%) was obtained by column chromatography using CH_2Cl_2/MeOH 9:1 as the eluent, while unreacted starting amine could be quantitatively recovered. R_f = 0.65 (CH_2Cl_2/MeOH 9:1); ^{1}H NMR (500 MHz, CD_3OD): *d* = 8.94 (t, 0.5H, ^4J = 1.0 Hz, s1-N*H*), 8.62 (t, 1H, ^4J = 6.5 Hz, H-s4), 8.49 (d, 2H, ^4J = 1.0 Hz, H-s2), 8.22 (d, 4H, ^3J = 9.5 Hz, H-p5), 8.16 (s, 5H, H-p3/p2*α*-N*H*), 8.03 (d, 4H, ^3J = 9.0 Hz, H-p4), 7.90 (s, 2H, H-p7), 7.53 (d, 2H, ^3J = 1.5 Hz, p6*α*-N*H*), 7.19 (s, 2.5H, H-d5), 5.04 (s, 4H, H-p2*α*), 4.92 (s, 8H, H-p6*α*), 3.55 (s, 27H, H-d10), 2.16-2.12 (m, 24H, H-d4/d8), 2.04 (m, 6H, H-d2), 1.84 (t, 18H, H-d7), 1.46 (s, 36H, *t*-Bu-C*H*$_3$); {^{1}H-^{13}C}-HSQC (500 MHz, CD_3OD): *d* = 135.1 (C-s4), 132.7 (C-s2), 131.3 (C-p4), 130.6 (C-p7), 128.4 (C-p3), 126.9 (C-p5), 55.5 (C-d10), 48.1 (C-p2*α*), 46.3 (C-p4*α*), 35.0 (C-d2), 34.9 (C-d3),32.9 (C-d7), 31.9 (*t*-Bu-*C*H$_3$), 31.8 (C-d8) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{116}H_{150}N_{10}O_{32}Na_2$ [M + 2Na]$^{2+}$: 1120.5101, found 1120.5063.

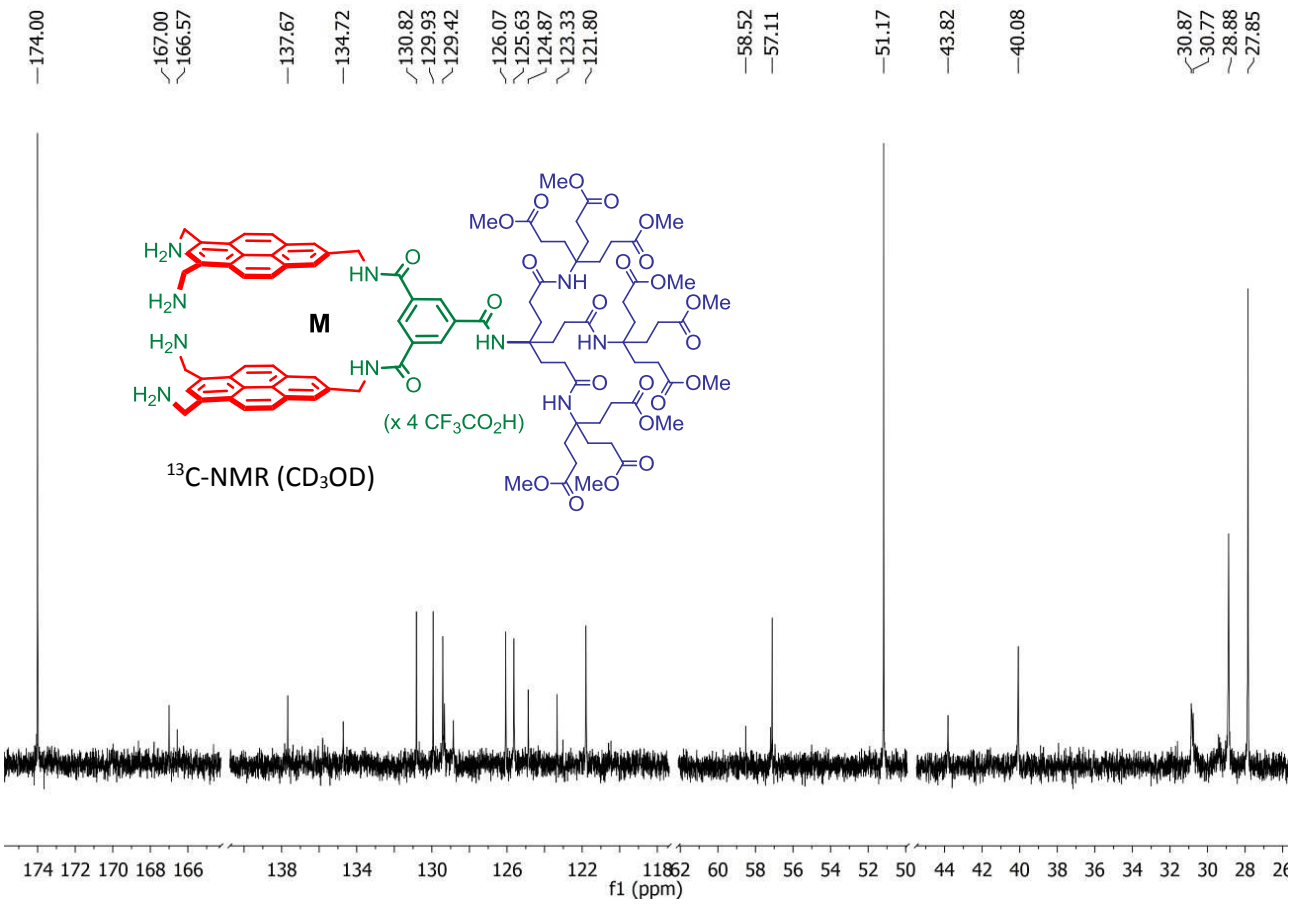

^{1}H-NMR (CD_3OD)
L
f1 (ppm)

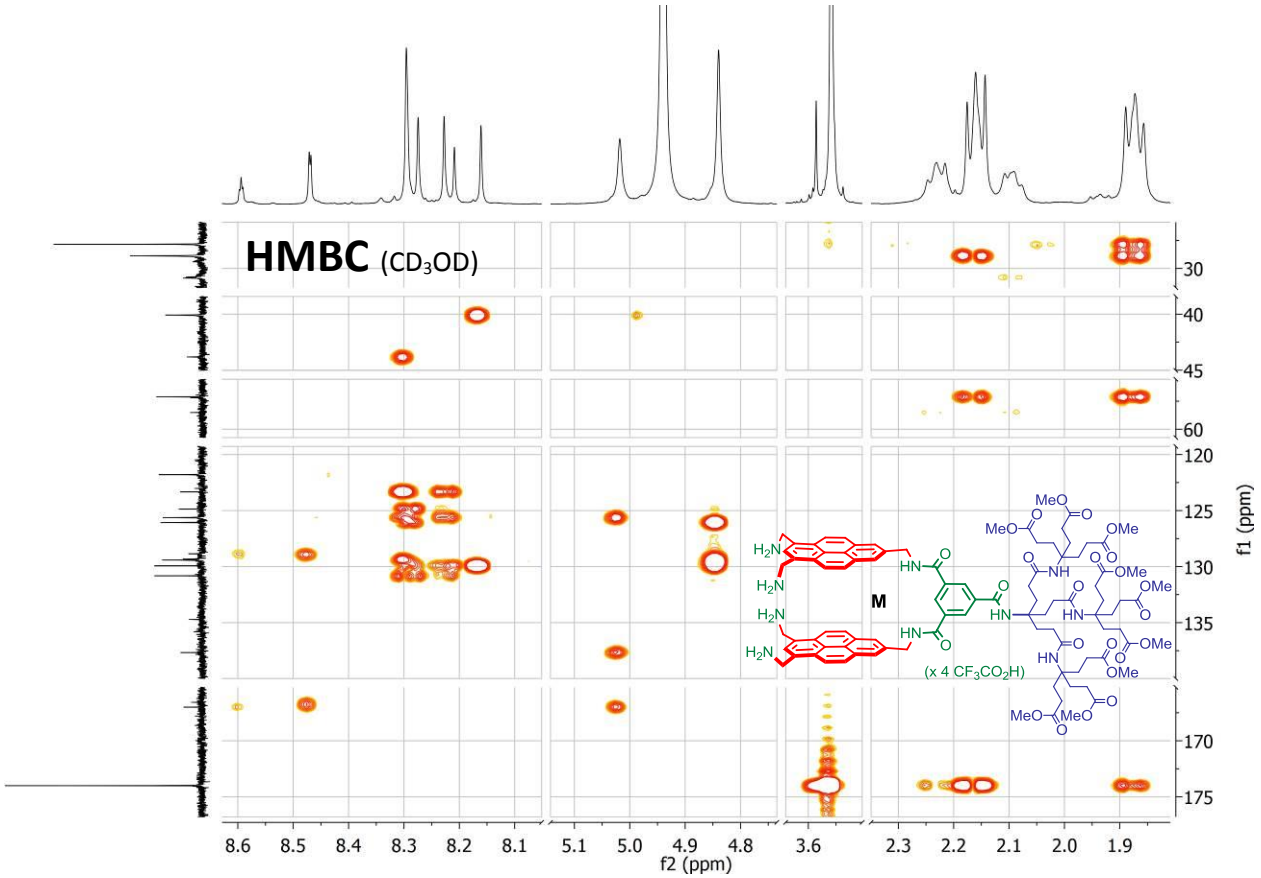

HSQC (CD_3OD)
L
f2 (ppm)
f1 (ppm)

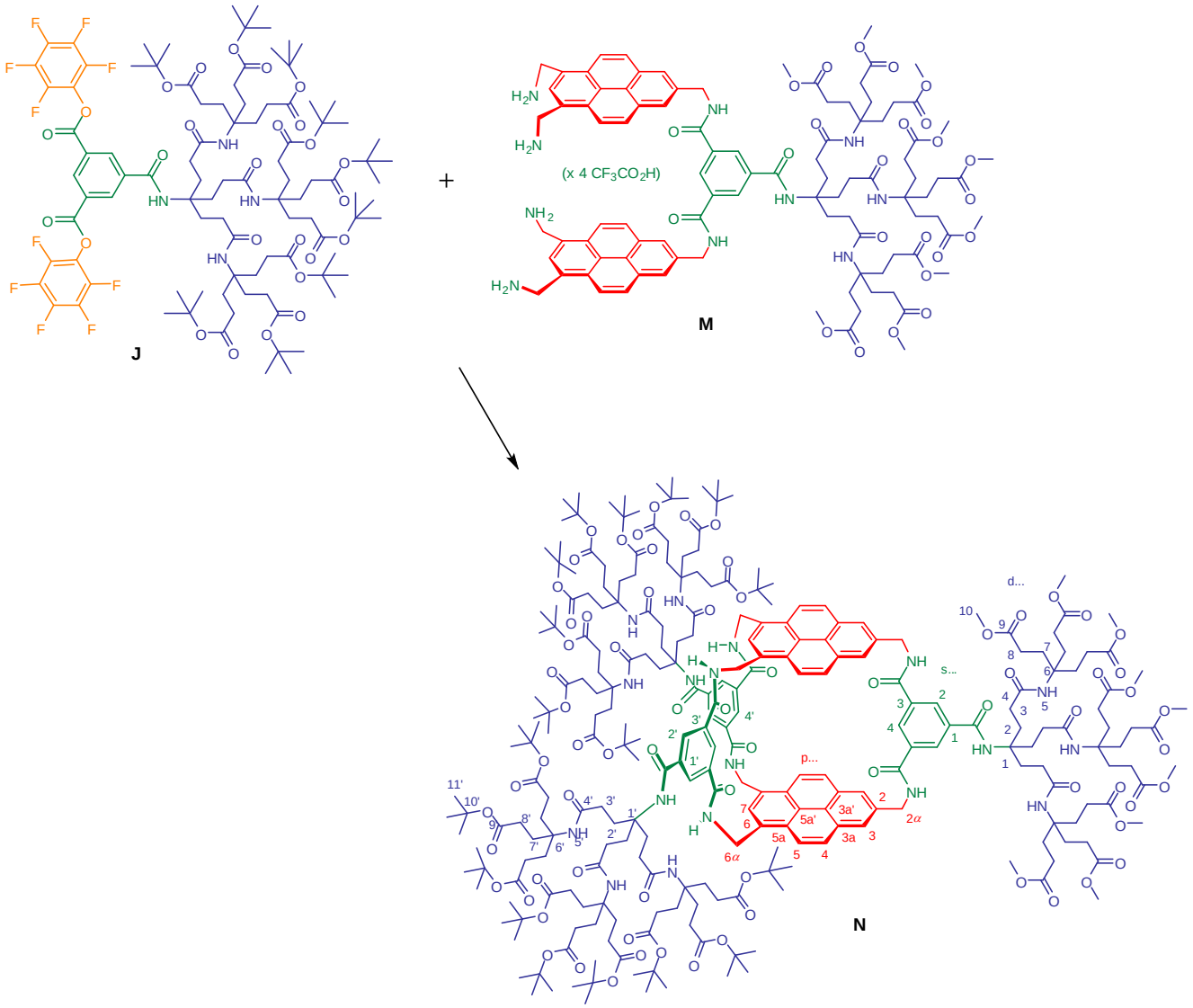


Tetraamine M (4·CF_3CO_2H) . To a stirred solution of Boc-protected tetraamine **L** (167 mg, 76 μmol) and $HSiEt_3$ (86 μL, 32 μmol) in DCM (5 mL) at 0 °C was added TFA (5 mL) in a dropwise manner. The resulting solution was stirred for 16 h, after which the volatiles were removed by passing a N_2 stream over the stirred mixture. Water (10 mL) was then added and the suspension was freeze dried to yield tetraamine **M**, presumably as the salt with 4 × CF_3CO_2H (quantitative by NMR). ^{1}H NMR (500 MHz, CD_3OD): *d* = 8.59 (t, 1H, ^4J = 1.5 Hz, H-s4), 8.47 (d, 2H, ^4J = 1.5 Hz, H-s2), 8.30 (s, 4H, H-p3), 8.28 (d, 4H, ^3J = 11.0 Hz, H-p5), 8.21 (d, 4H, ^3J = 9.5 Hz, H-p4), 8.16 (s, 2H, H-p7), 7.71 (d, 0.5H, ^3J = 1.5 Hz, p6α-N*H*), 7.32 (s, 1H, H-d5), 5.02 (s, 4H, H-p2α), 4.84 (s, 8H, H-p6α), 3.56 (s, 27H, H-d10), 2.24 (t, 6H, H-d4), 2.16 (t, 18H, H-d8), 2.10 (t, 6H, H-d2), 1.87 (t, 18H, H-d7) p.p.m.; ^{13}C NMR (125 MHz / {^{1}H-^{13}C}-HMBC 500 MHz, CD_3OD,): *d* = 174.0 (C-d9), 167.0 (C-s1*C*O), 166.6 (C-s3*C*O), 137.7 (C-s2), 134.7, (C-s4), 130.8 (C-p2), 129.9 (C-p6), 129.4 (C-p3a), 126.0 (C-p5a), 125.6 (C-p4), 124.9 (C-p7), 123.3 (C-p1,3), 121,8 (C-p5), 58.5 (C-d1), 57.1 (C-p6α), 51.2 (C-d10), 43.8 (C-d6), 40.1 (C-p6α), 30.9 (C-d3), 30.8 (C-d2), 28.9 (C-d7), 27.9 (C-d8) p.p.m.; HRMS (ESI$^+$): *m/z* calculated for $C_{96}H_{119}N_{10}O_{24}$ $[M + H]^+$: 1795.8393, found 1795.8429.


8.594
8.471
8.296
8.274
8.228
8.209
8.161
7.705
7.316
5.018
4.940
4.840
3.586
3.559
3.310
2.236
2.157
2.099
1.936
1.871
1.261
0.89
1.81
7.91
4.00
2.05
0.42
0.86
4.12
7.88
27.00
7.63
17.14
5.78
17.02
f1 (ppm)
M
(x 4 CF3CO2H)
1H-NMR (CD3OD)

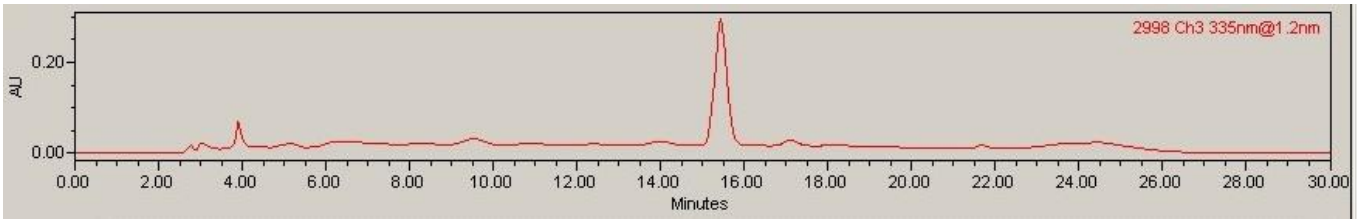

174.00
167.00
166.57
137.67
134.72
130.82
129.93
129.42
126.07
125.63
124.87
123.33
121.80
58.52
57.11
51.17
43.82
40.08
30.87
30.77
28.88
27.85
f1 (ppm)
M
(x 4 CF3CO2H)
13C-NMR (CD3OD)


HMBC (CD3OD)
f2 (ppm)
f1 (ppm)
M
(x 4 CF3CO2H)

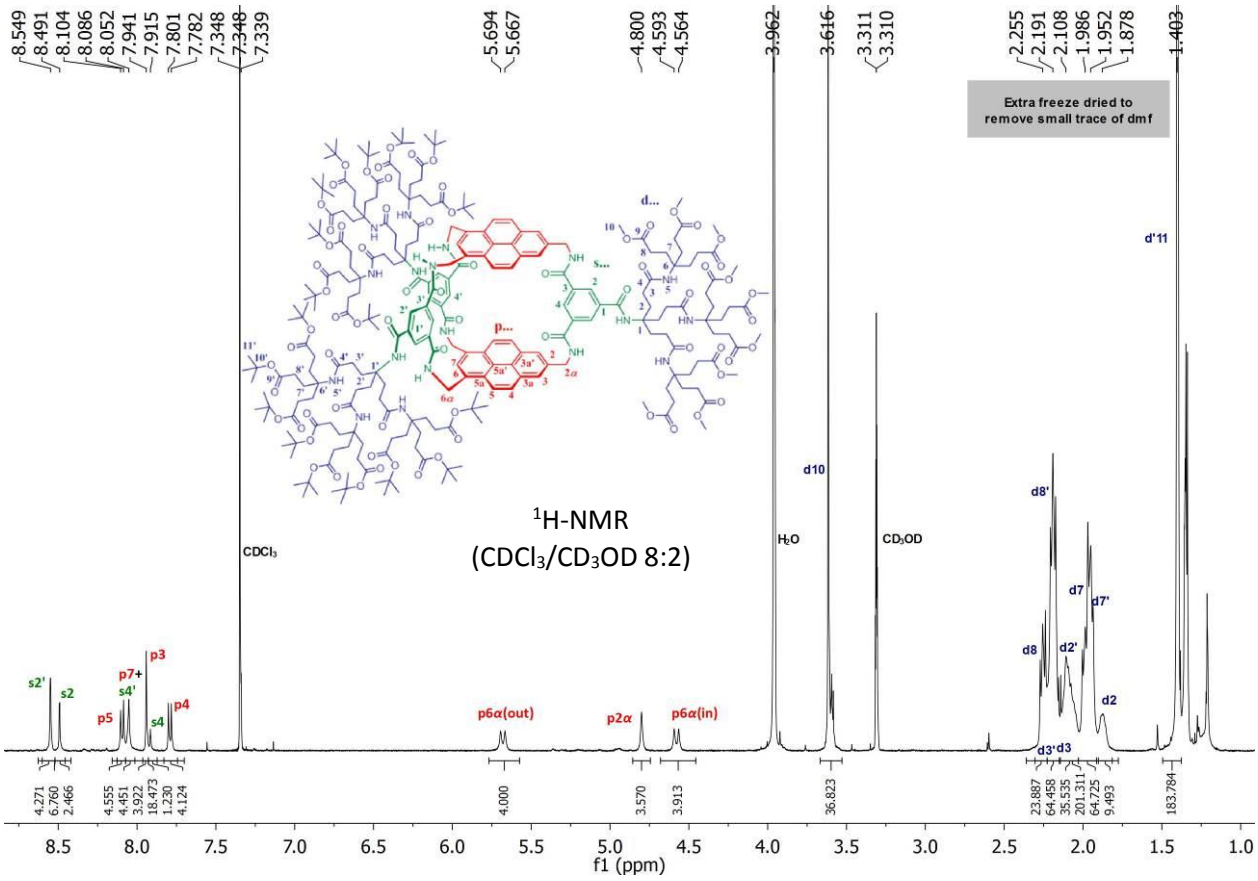


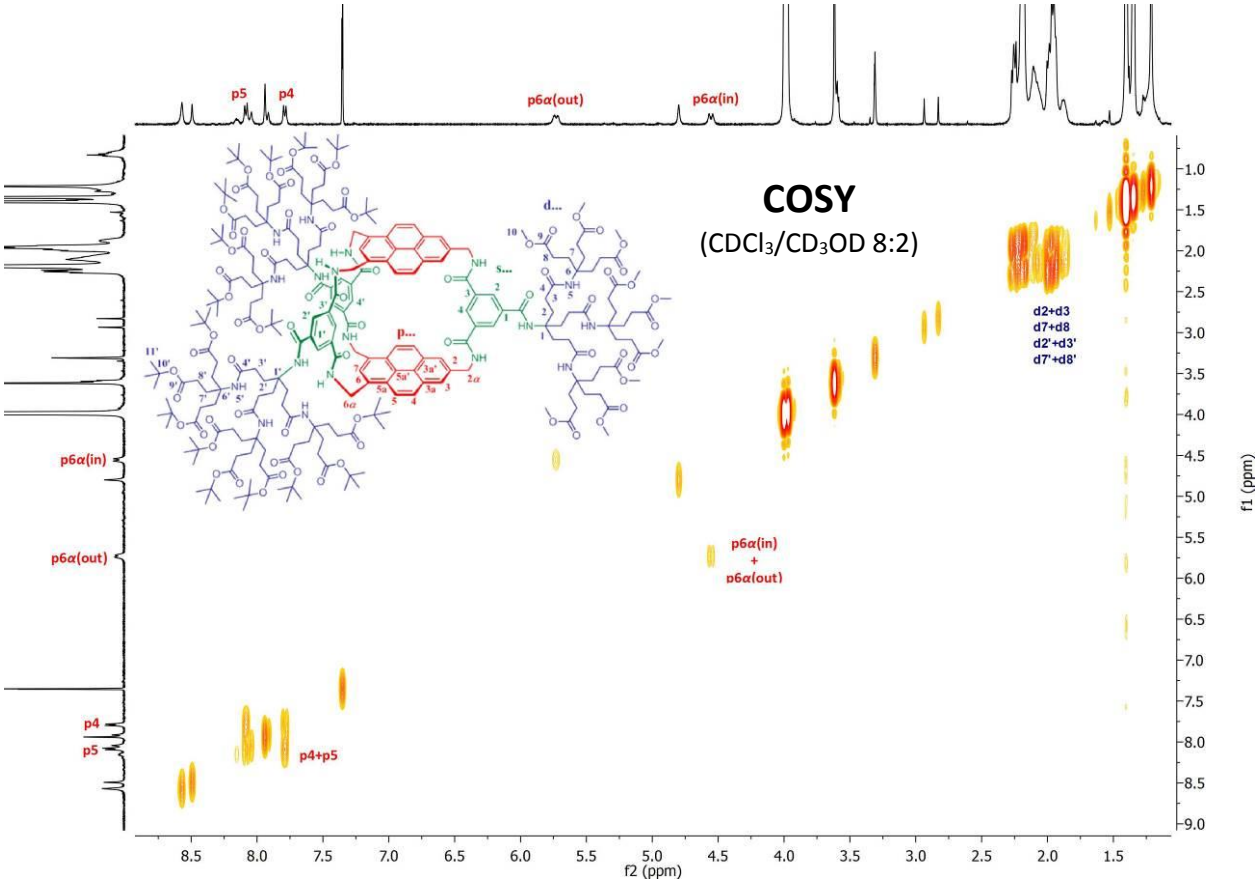
Protected receptor N. To a stirred solution of tetraamine **M** TFA salt (see above; 136 mg, 76 μmol) and DIPEA (2 mL, 11.5 mmol) in dry THF (750 mL) was added a solution of bis(pentafluorophenyl) ester **J** (299 mg, 152 μmol) in anhydrous THF (in 30 mL) over the course of 48 h. After the addition was complete, the reaction mixture was stirred for an additional 72 hours. The solvent was removed *in vacuo* and the resulting oil dissolved in CH_2Cl_2 (25 ml). This solution was extracted with saturated aq. NH_4Cl solution (25 ml), water (25 ml), and dried with $MgSO_4$. After removal of the solvent, the residue was dissolved in acetone/water 8:2 (30 mL) and purified with preparative HPLC (Waters CSH C18 5 μm 19 × 250 mm) by injecting 3 ml portions and eluting with acetone in water (from 80% to 92% in 18 min., then to 100% in 2 min., then back to 80% in 5 min.): R_t = 15.7 min (see next page for a typical preparative HPLC trace). Pure protected receptor **N** was thus obtained as a white solid (68 mg, 13.7 μmol, 18%). A full assignment could be made based on ^{1}H, ^{13}C-, {^{1}H-^{1}H}-COSY, {^{1}H-^{13}C}-HSQC and {^{1}H-^{13}C}-HMBC. ^{1}H NMR (500 MHz, $CDCl_3/CD_3OD$ 8:2): *d* = 8.58 (s, 4H, H-s2'), 8.50 (s, 2H, H-s2), 8.16 (s, 2H, H-s4'), 8.10 (d, 4H, ^3J = 9.4 Hz, H-p5), 8.05 (s, 2H, H-p7), 7.95 (s, 4H, H-p3), 7.93 (s, 1H, H-s4), 7.80 (d, 4H, ^3J = 9.2 Hz, H-p4), 5.74 (d, 4H, ^2J = 12.3 Hz, H-p6α(out)), 4.81 (s, 4H, H-p2α), 4.57 (d, 4H, ^2J = 14.7 Hz, H-p6α(in)), 3.63 (s, 27H, H-d10), 2.27 (t, 18H, H-d8), 2.22 (t, 12H, H-d3'), 2.20 (t, 36H, H-d8'), 2.13 (t, 6H, H-d3), 2.09 (t, 12H, H-d2'), 2.00 (t, 18H, H-d7), 1.97 (t, 36H, H-d7'), 1.90 (t, 6H, H-d2), 1.42 (s, 162H, H-d11') p.p.m.; ^{13}C NMR (125 MHz, $CDCl_3/CD_3OD$ 8:2): *d* = 175.19 (C-d9), 174.23 (C-d9'), 174.1 (C-d4), 174.08 (C-d4'), 168.54 (C-s3'CO), 168.12 (C-s1CO), 167.47 (C-s3CO), 167.14 (C-s1'CO), 135-137 (C-s1/s1'), 133.64 (C-p7), 133.62 (C-p2), 131.79 (C-s3'), 131.14 (C-s2'), 130.92 (C-s3), 130.86 (C-s2), 130.57 (C-p3a), 130.56 (C-p6), 128.71 (C-p4), 128.37 (C-s4'), 128.01 (C-s4), 126.94 (C-p3), 126.2 (C-p5a), 124.82 (C-p5), 124.81 (C-p5a'), 124.79 (C-p3a'), 82.04 (C-d10'), 58.67 (C-d6'), 58.47 (C-d6), 58.4 (C-d1'), 57.23 (C-d1), 53.04 (C-d10), 46.23 (C-p2α), 44.46

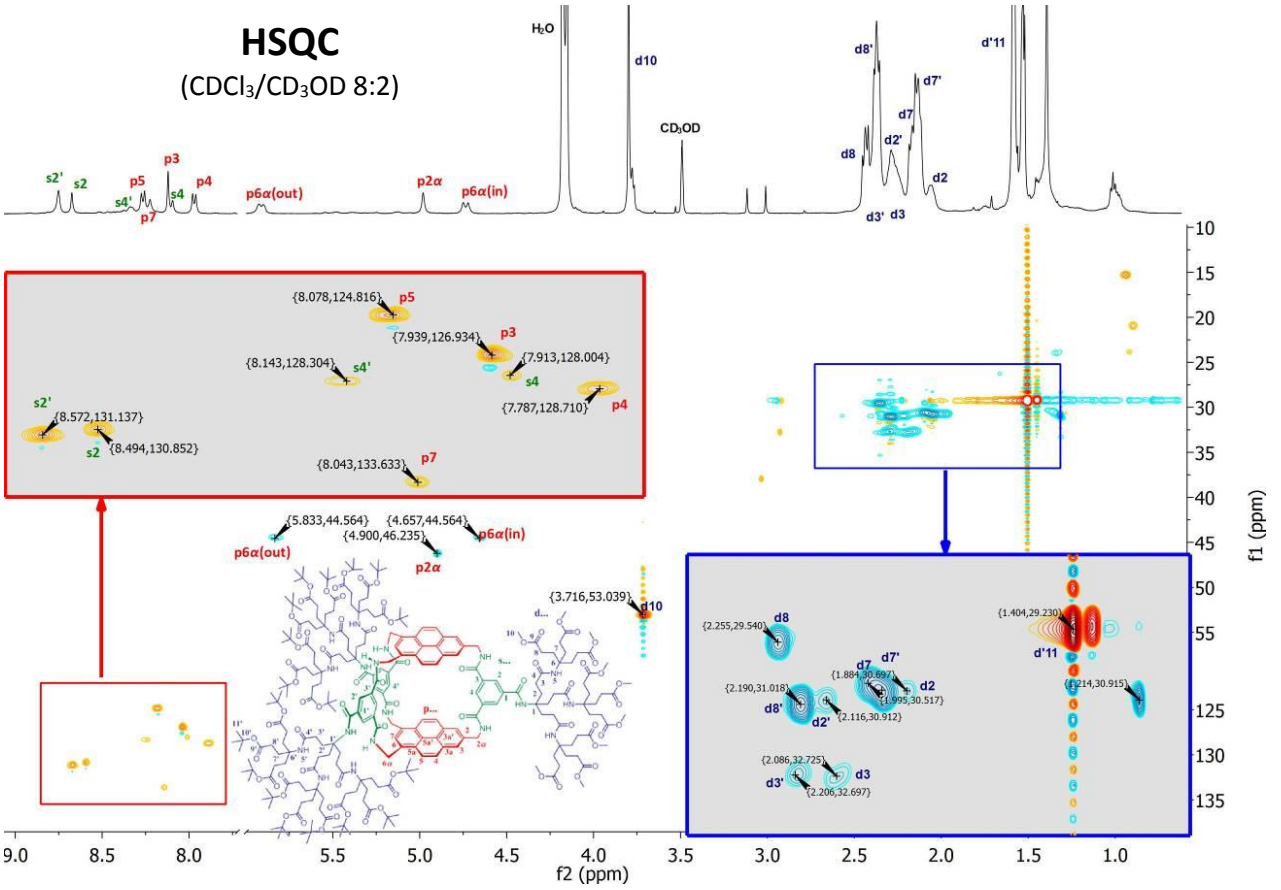
(C-p6α(in)), 44.46 (C-p6α(out)), 32.82 (C-d2'), 32.71 (C-d3'), 31.03 (C-d8'), 30.93 (C-d3), 30.71 (C-d2), 30.69 (C-d7'), 30.54 (C-d7), 29.55 (C-d8), 29.23 (C-d11'); MALDI-TOF: *m/z* calculated for $C_{266}H_{386}N_{18}O_{72}Na$ $[M+Na]^+$; 5007.7, found 5007.9; ESQUIRE-MS (ESI^+): *m/z* calculated for $C_{266}H_{386}N_{18}O_{72}Na_2$ $[M+2Na]^{2+}$: 2517.0, found 2516.4; $C_{266}H_{386}N_{18}O_{72}Na_3$ $[M+3Na]^{3+}$: 1685.7, found 1686.0; $C_{266}H_{386}N_{18}O_{72}Na_4$ $[M+4Na]^{4+}$: 1270.0, found 1270.7.

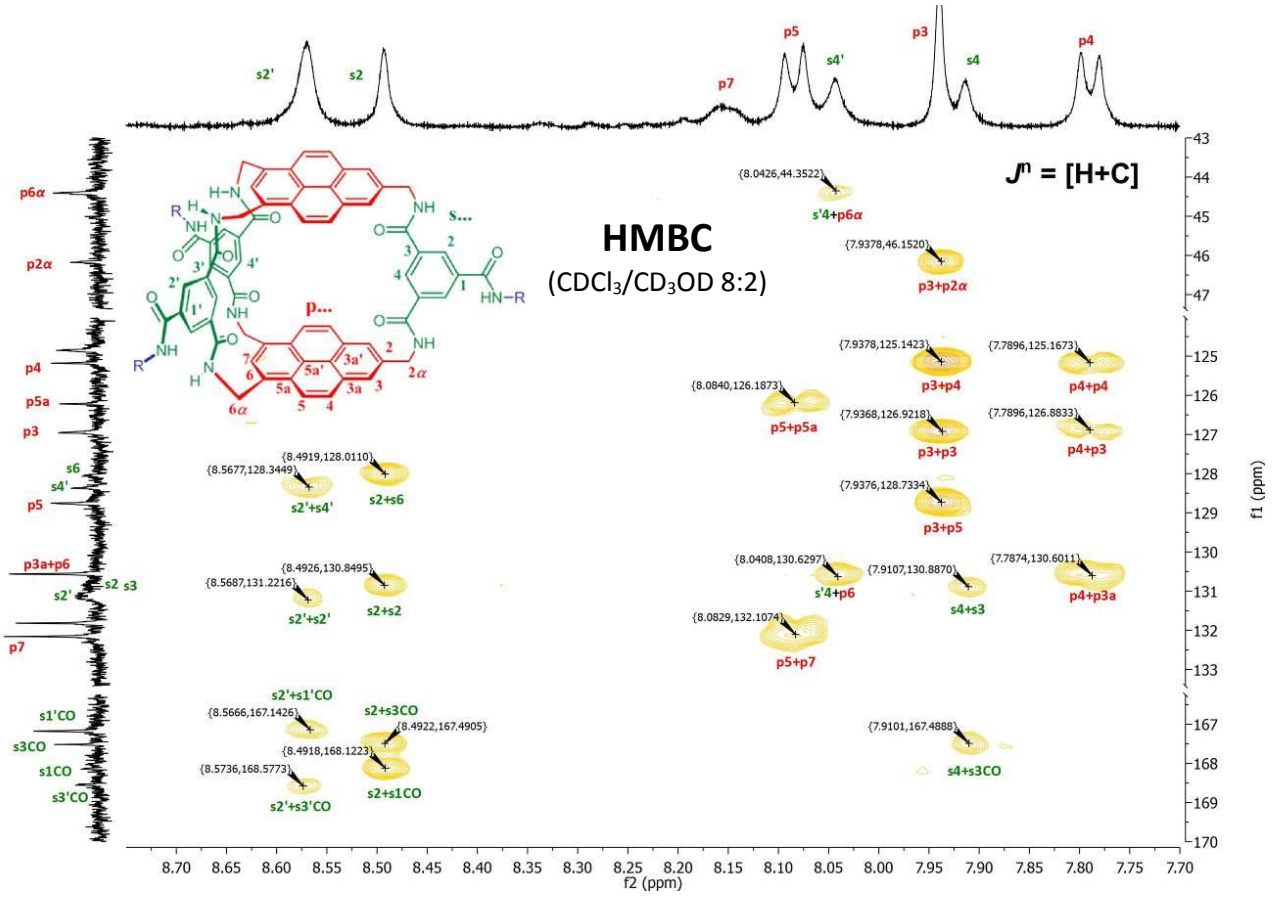


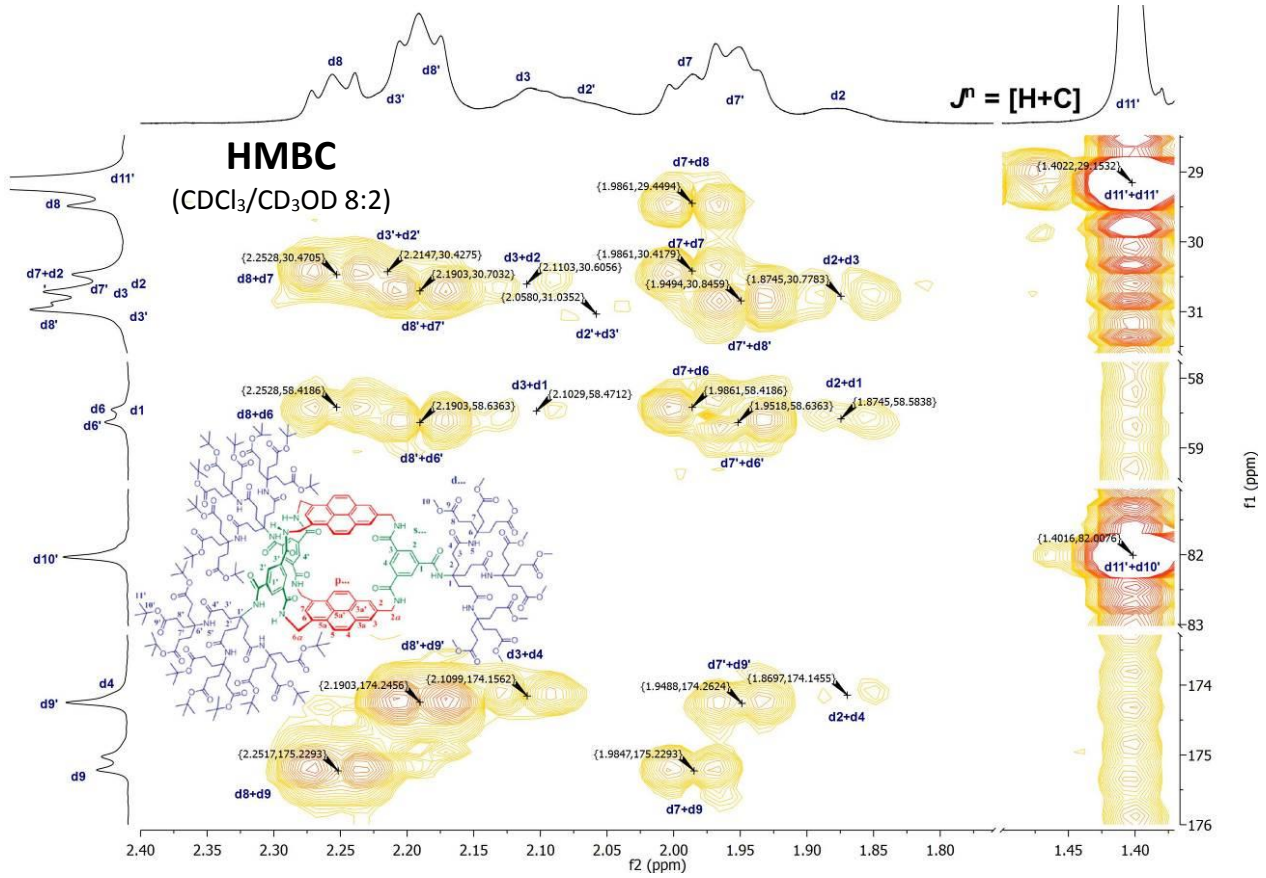
Typical preparative HPLC trace for the purification of protected receptor **N**, monitored at 335 nm.

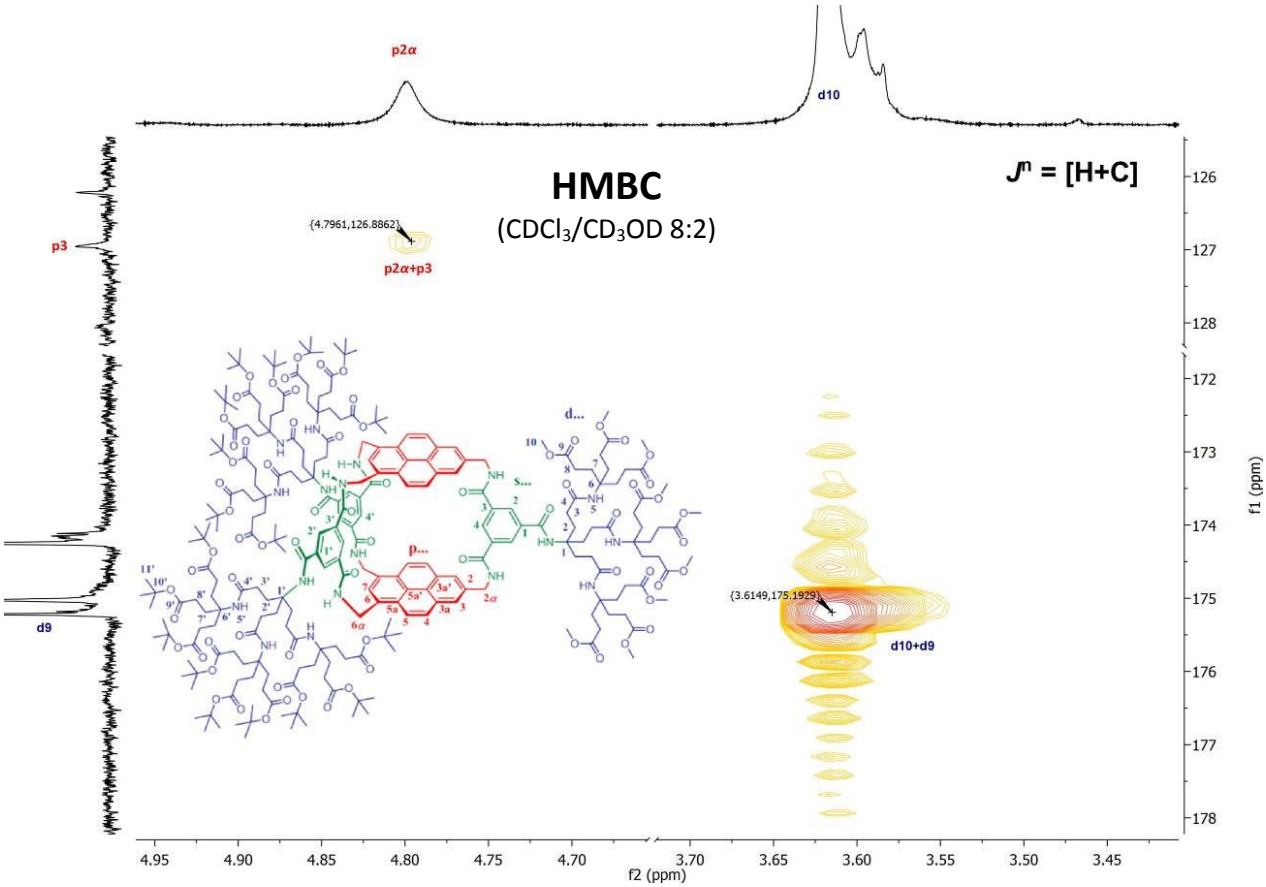

Extra freeze dried to remove small trace of dmf
^{1}H-NMR
($CDCl_3$/CD_3OD 8:2)
CDCl3
H2O
CD3OD
d10
d'11
d8'
d7
d7'
d8
d2'
d2
d3'd3
p3
p7+
s4'
s2'
s2
p5
s4
p4
p6α(out)
p2α
p6α(in)
f1 (ppm)

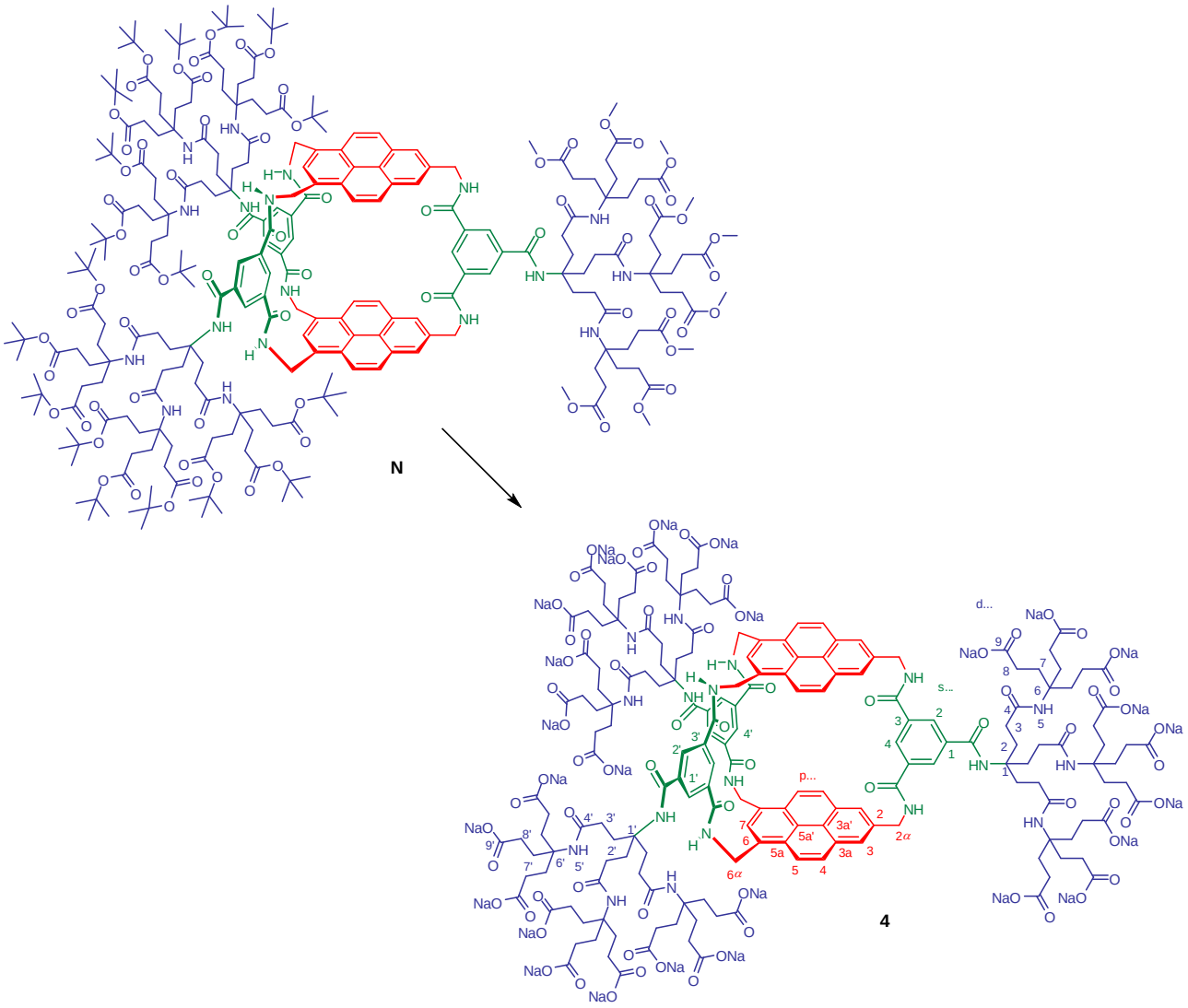

COSY
($CDCl_3$/CD_3OD 8:2)
p5
p4
p6α(out)
p6α(in)
d2+d3
d7+d8
d2'+d3'
d7'+d8'
p6α(in)
+
p6α(out)
p4
p5
p4+p5
f2 (ppm)
f1 (ppm)


HSQC
(CDCl3/CD3OD 8:2)
H2O
d10
CD3OD
d8'
d'11
d7'
d7
d2'
d8
d2
d3' d3
s2' s2 s5 s4' p3 s4 p7 p4
p6α(out)
p2α
p6α(in)
{8.078,124.816} p5
{7.939,126.934} p3
{8.143,128.304} s4'
{7.913,128.004} s4
{7.787,128.710} p4
s2' {8.572,131.137}
s2 {8.494,130.852}
{8.043,133.633} p7
{5.833,44.564} {4.657,44.564}
{4.900,46.235}
p6α(out) p6α(in)
p2α
{3.716,53.039} d10
{2.255,29.540} d8
{2.190,31.018} d8'
{1.884,30.697} d7
d7'
{1.995,30.517} d2
{2.116,30.912} d2'
{2.086,32.725} d3'
{2.206,32.697} d3
{1.404,29.230}
d'11
{1.214,30.915}
f2 (ppm)
f1 (ppm)

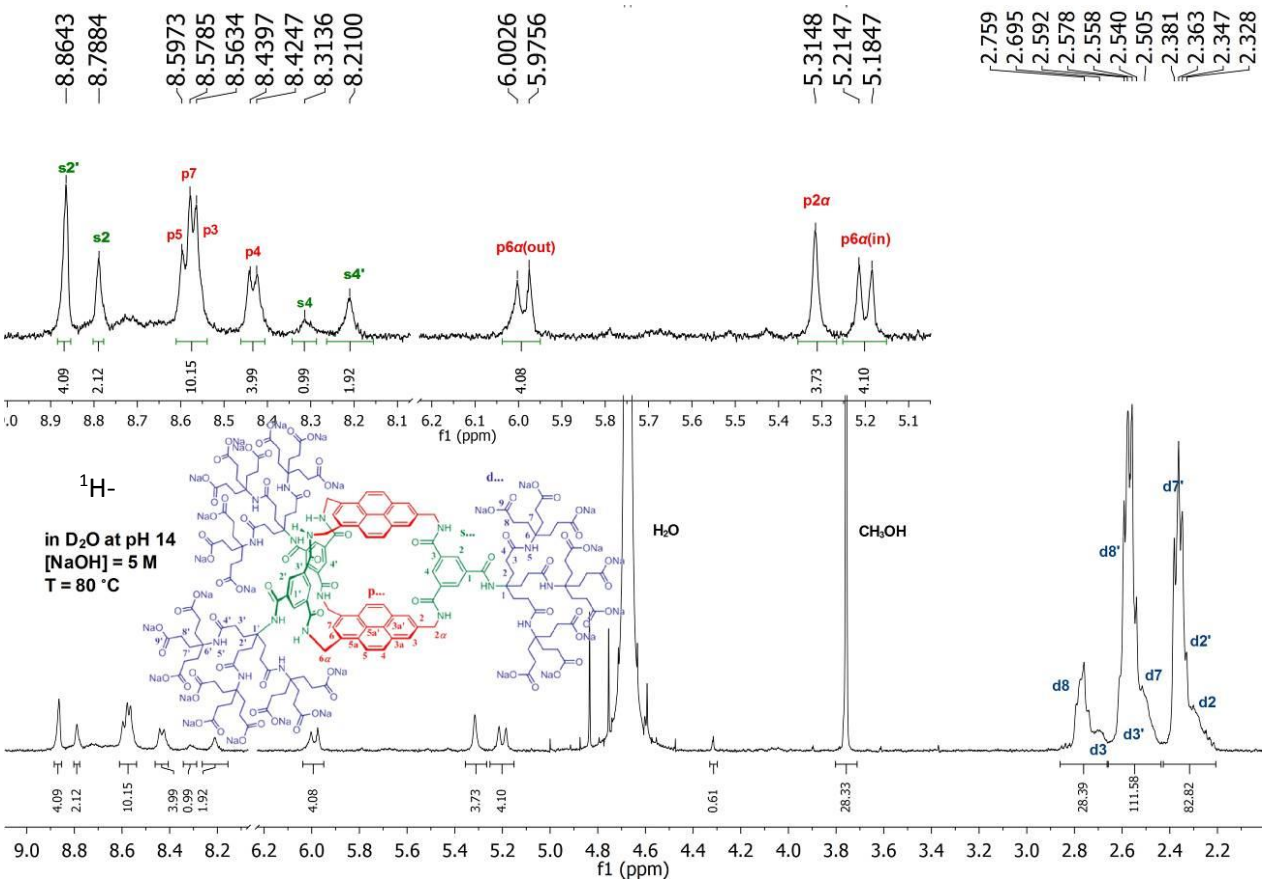

HMBC
(CDCl3/CD3OD 8:2)
Jⁿ = [H+C]
s2' s2 p7 p5 s4' p3 s4 p4
p6α p2α p4 p5a p3 s6 s4' p5 p3a+p6 s2' s2 s3 p7 s1'CO s3CO s1CO s3'CO
{8.0426,44.3522} s'4+p6α
{7.9378,46.1520} p3+p2α
{7.9378,125.1423} p3+p4
{7.7896,125.1673} p4+p4
{8.0840,126.1873} p5+p5a
{7.9368,126.9218} p3+p3
{7.7896,126.8833} p4+p3
{8.4919,128.0110} {8.5677,128.3449} s2'+s4' s2+s6
{7.9376,128.7334} p3+p5
{8.0408,130.6297} s'4+p6
{7.9107,130.8870} s4+s3
{7.7874,130.6011} p4+p3a
{8.4926,130.8495} {8.5687,131.2216} s2'+s2' s2+s2
{8.0829,132.1074} p5+p7
s2'+s1'CO s2+s3CO
{8.5666,167.1426} {8.4922,167.4905}
{8.4918,168.1223} {8.5736,168.5773} s2'+s3'CO s2+s1CO
{7.9101,167.4888} s4+s3CO
f2 (ppm)
f1 (ppm)

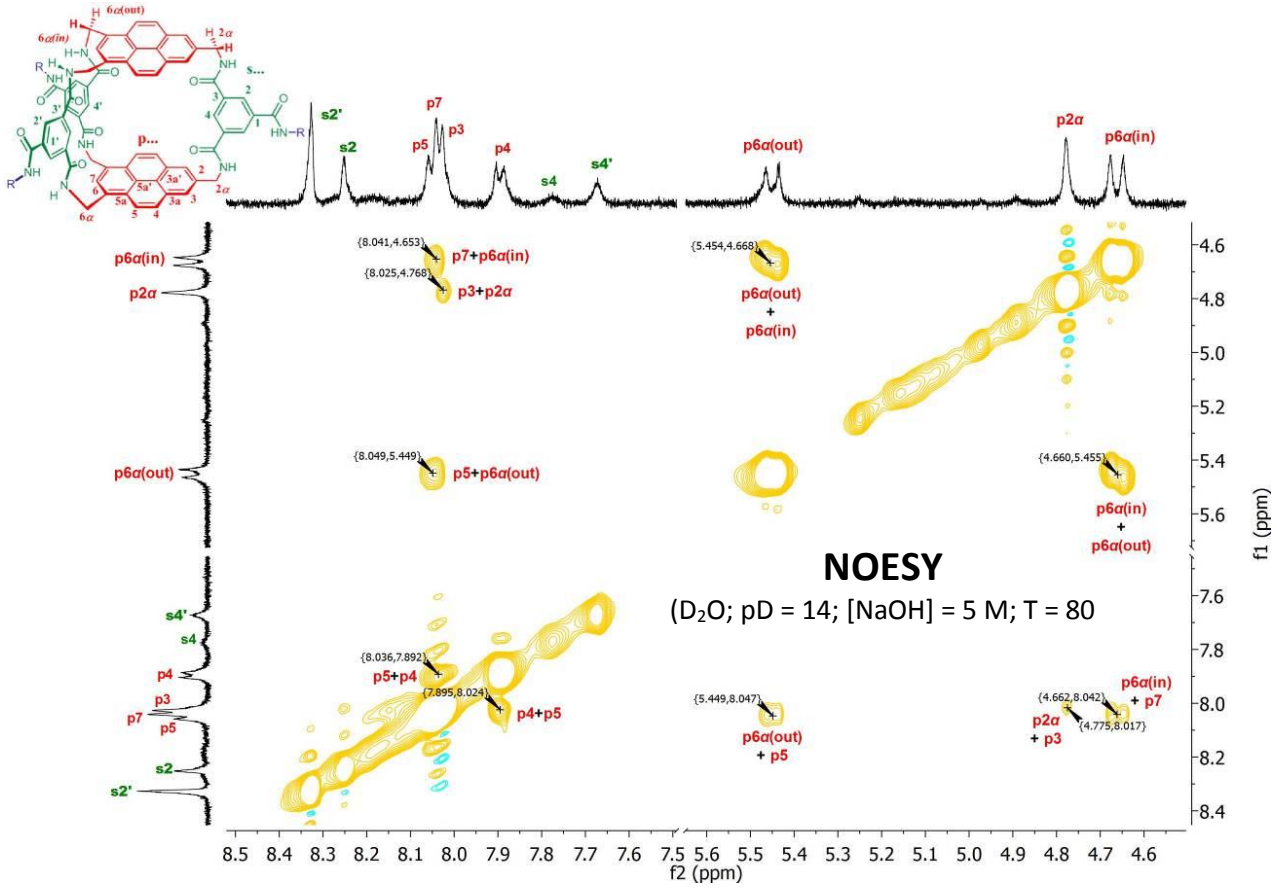

HMBC
(CDCl3/CD3OD 8:2)
Jn = [H+C]
f2 (ppm)
f1 (ppm)

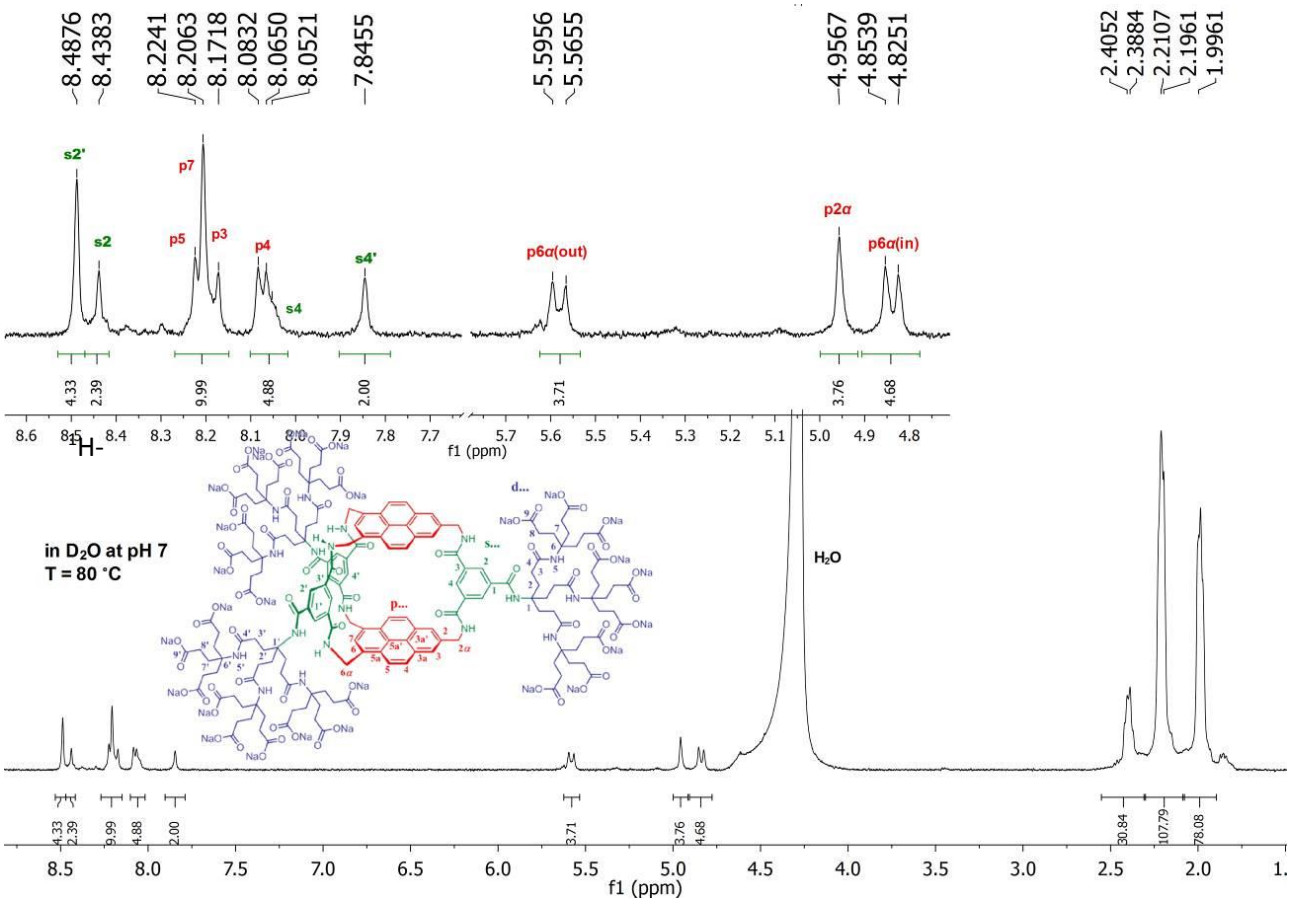

HMBC
(CDCl3/CD3OD 8:2)
Jn = [H+C]
f2 (ppm)
f1 (ppm)

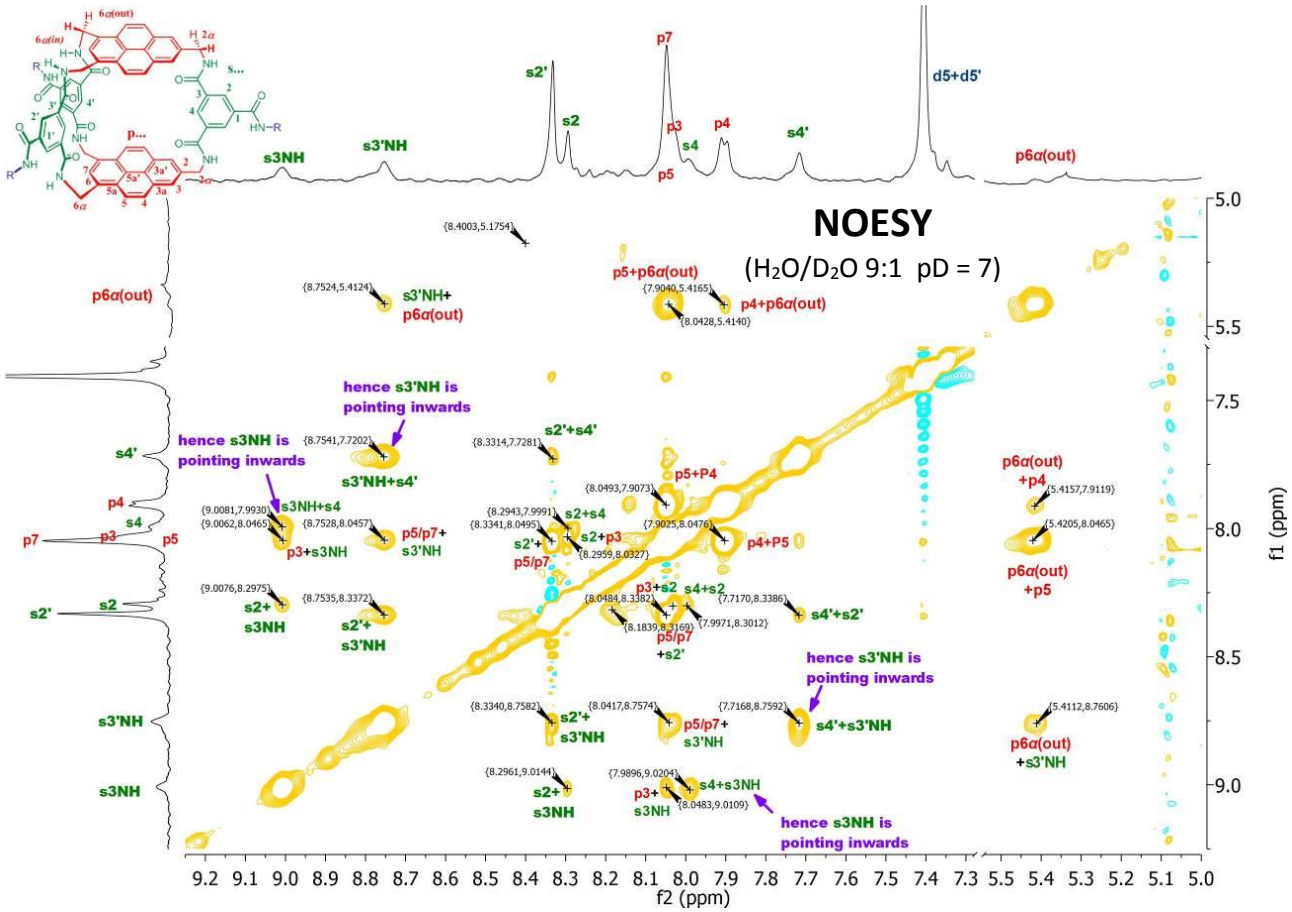


Receptor 4. A 25 mL round bottom flask was charged with N_2, a magnetic stirrer, protected receptor **N** (19.2 mg, 3.85 μmol), and dry DCM (5 mL). The solution was cooled to 0 °C, triethylsilane was added (0.11 mL, 0.689 mmol), followed by a dropwise addition of TFA (5 mL). The resulting solution was allowed to slowly warm to laboratory temperature (ca. 25 °C), and stirred for about 12 h. The volatiles were removed by passing a N_2 stream over the solution and the remaining solid was suspended in water (10 mL) and freeze-dried (twice). The resulting white solid was then suspended in water (3 mL), NaOH was added (0.2 mL, 5 M), and the resulting solution was stirred for 15 minutes. Ion exchange resin (Amberlyst 15, hydrogen form; thoroughly washed with water) was then added until the pH < 3. The resulting suspension was then neutralised using NaOH (5 M, 1 M, and 0.1 M solutions) to pH ≈ 7.1. After passing the resulting solution through a 45 and a 20 μm filter, it was freeze-dried to yield the receptor **4** in its sodium form as an off-white solid (17.1 mg, 3.85 μmol, 100%). A full assignment could be made based on ^{1}H, {^{1}H-^{1}H}-COSY, and {^{1}H-^{1}H}-NOESY NMR spectroscopy. ^{1}H NMR (500 MHz, D_2O): *d* = 9.01 (s, 2H, H-s3NH), 8.75 (s, 2H, H-s3'NH), 8.33 (s, 1H, H-s1NH), 8.035 (d, 4H, ^3J = 9.7 Hz, H-p5), 8.033 (s, 2H, H-p7), 8.012 (s, 4H, H-p3), 8 (s, 1H, H-s1'NH), 7.89 (d, 4H, ^3J = 9.2 Hz, H-p4), 5.441 (d, 4H, ^2J = 13.9 Hz, H-p6α(out)), 4.766 (s, 4H, H-p2α), 4.649 (d, 4H, ^2J = 15 Hz, H-p6α(in)), 2.25 (s, 6H, H-d3), 2.25 (s, 12H, H-d3'), 2.04 (s, 6H, H-d2), 2.04 (s, 18H, H-d8), 2.04 (s, 12H, H-d2'), 2.04 (s, 36H, H-d8'), 1.84 (s, 18H, H-d7), 1.84 (s, 36H, H-d7') p.p.m.. MALDI-TOF (from 5% v/v formic acid in water): *m/z* calculated for $C_{185}H_{224}N_{18}O_{72}Na$ $[M+Na]^+$; 3872.4, found 3872.0; $C_{185}H_{226}N_{18}O_{73}Na$ $[M+Na+H_2O]^+$; 3890.4, found 3890.7; $C_{185}H_{229}N_{18}O_{74}$ $[M+H+2H_2O]^+$; 3904.5, found 3904.8; $C_{186}H_{226}N_{18}O_{74}Na$ $[M+Na+HCOOH]^+$; 3918.4, found 3918.8.


8.8643
8.7884
8.5973
8.5785
8.5634
8.4397
8.4247
8.3136
8.2100
6.0026
5.9756
5.3148
5.2147
5.1847
2.759
2.695
2.592
2.578
2.558
2.540
2.505
2.381
2.363
2.347
2.328
s2'
s2
p7
p5
p3
p4
s4
s4'
p6α(out)
p2α
p6α(in)
f1 (ppm)
1H-
in D2O at pH 14
[NaOH] = 5 M
T = 80 °C
H2O
CH3OH
d8
d8'
d7
d7'
d3
d3'
d2
d2'
f1 (ppm)


p6α(in)
p2α
p6α(out)
p7+p6α(in)
p3+p2α
p5+p6α(out)
p5+p4
p4+p5
p6α(out) + p6α(in)
p6α(in) + p2α
p6α(out) + p5
p6α(in) + p7
p2α + p3
NOESY
(D2O; pD = 14; [NaOH] = 5 M; T = 80
f2 (ppm)
f1 (ppm)

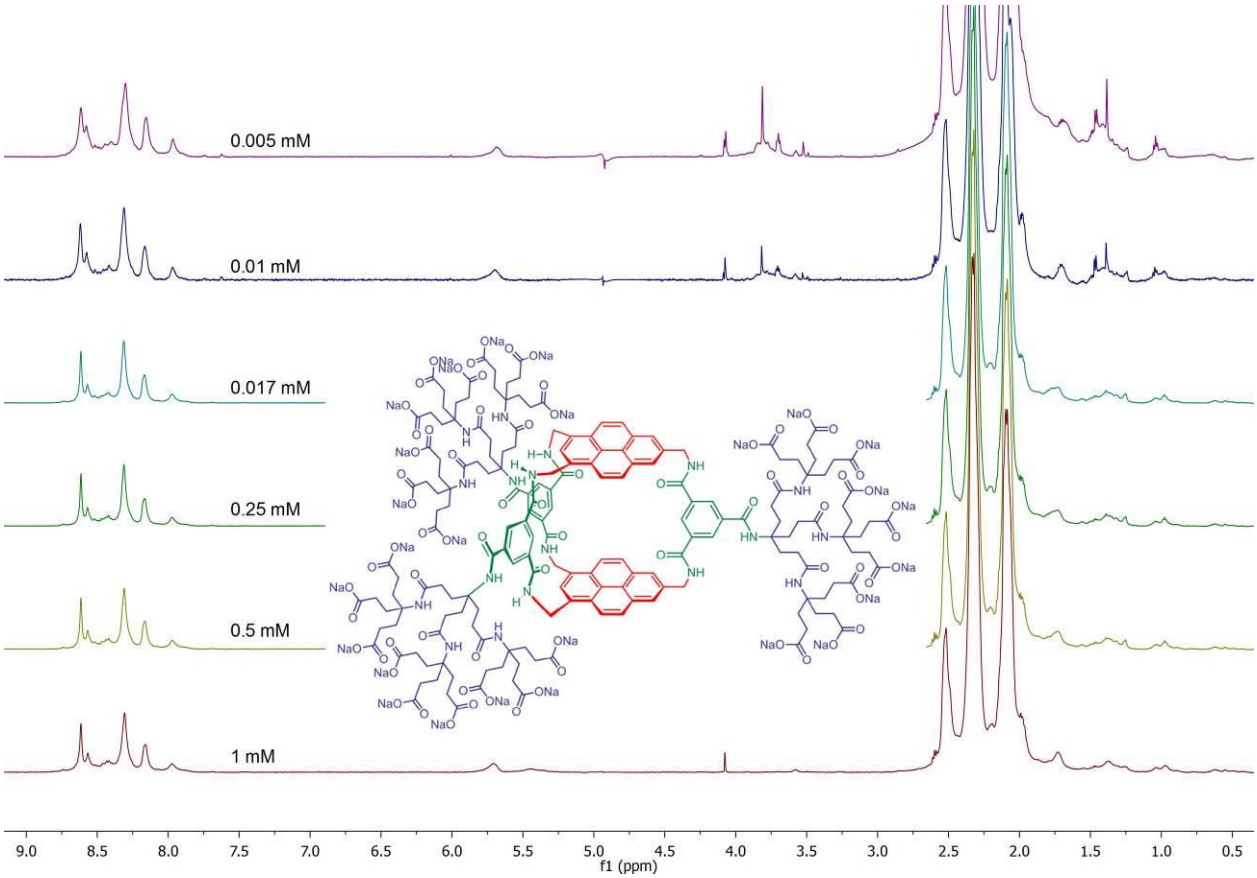




NOESY

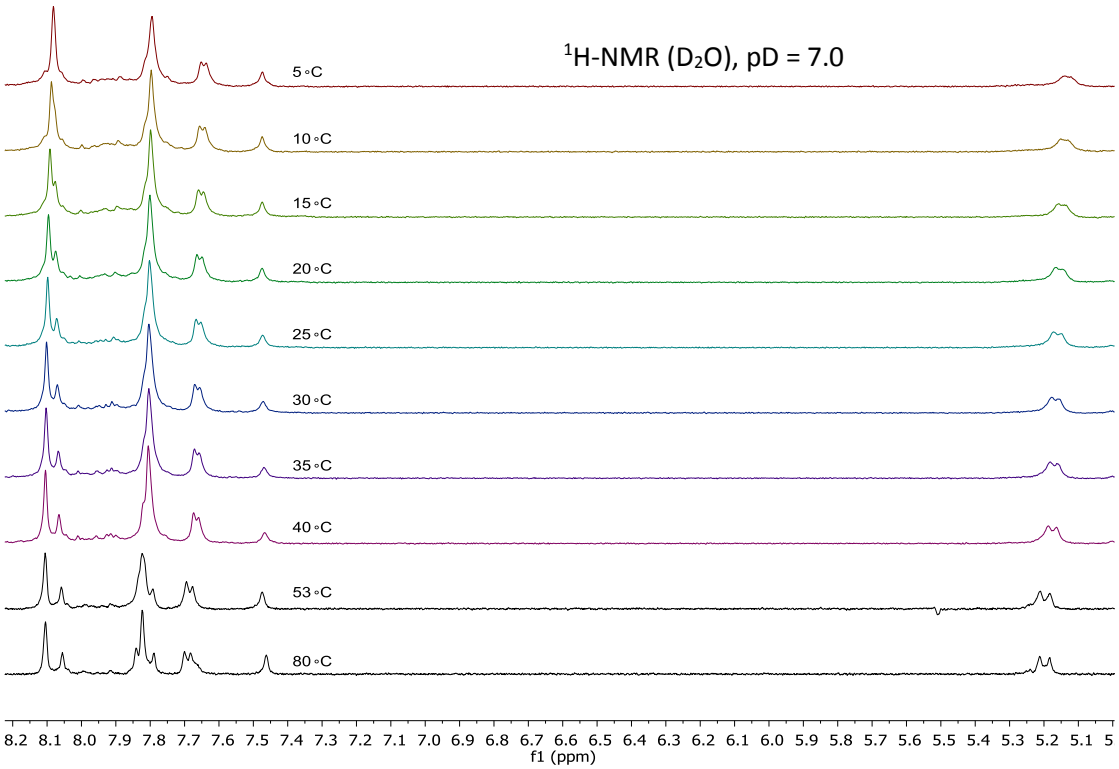
(H_2O/D_2O 9:1 pD = 7)

^{1}H NMR Dilution Study of 4 in D_2O.

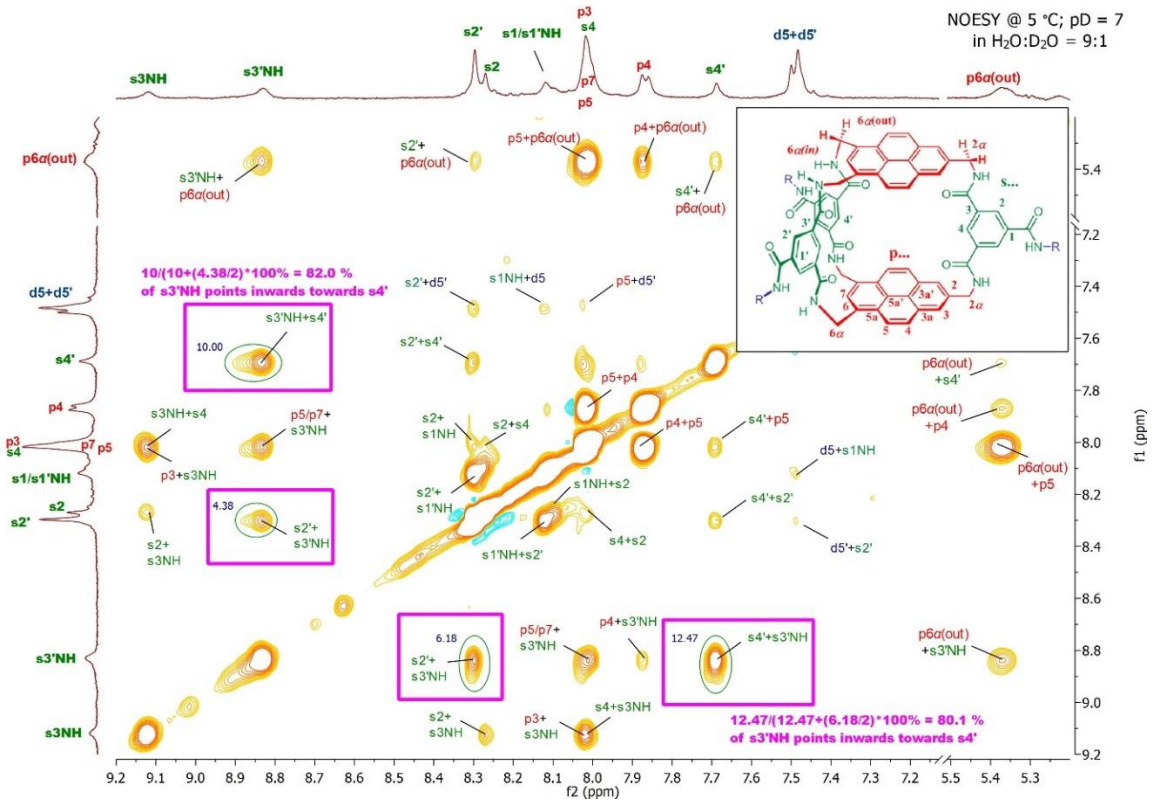


Supplementary Figure 2 | ^{1}H NMR dilution study of receptor **4** in D_2O at pD = 7.0 and T = 25 °C.

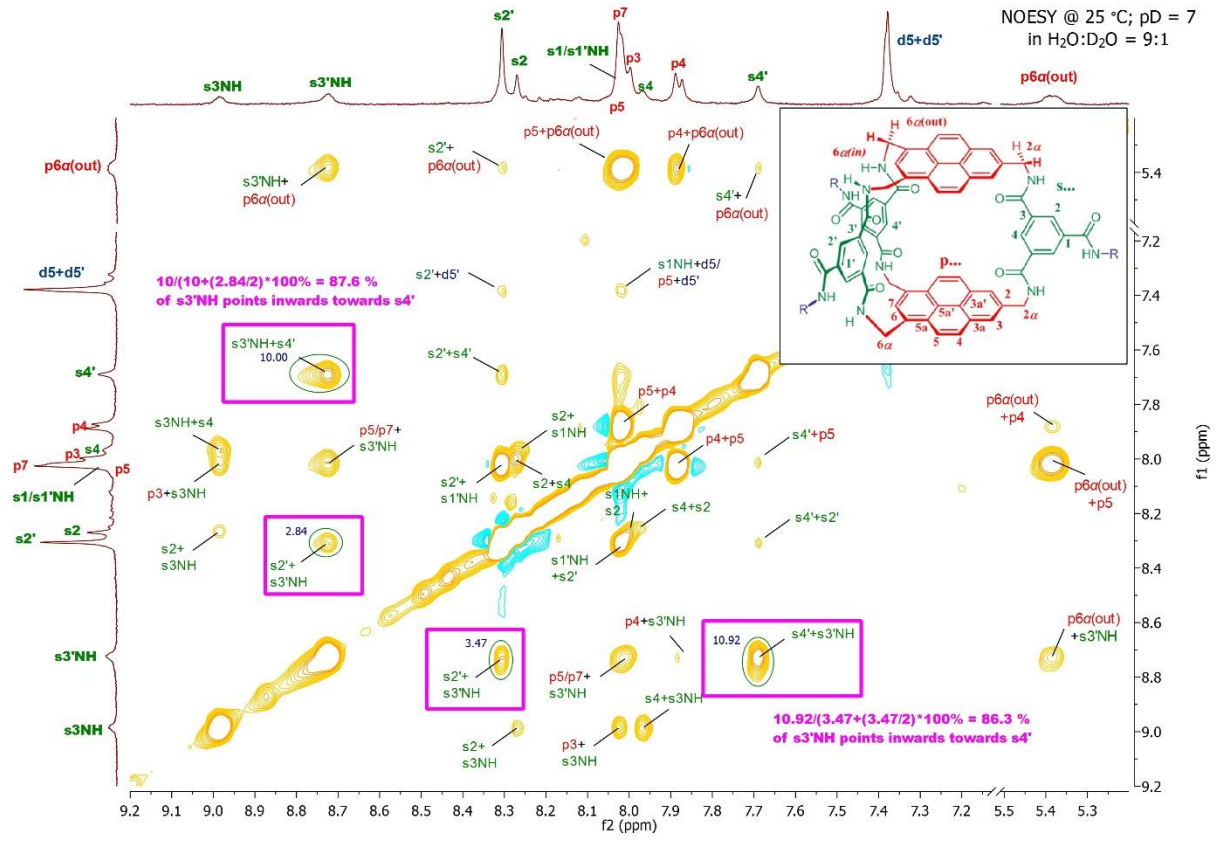
Variable Temperature Study of 4 in H_2O/D_2O.



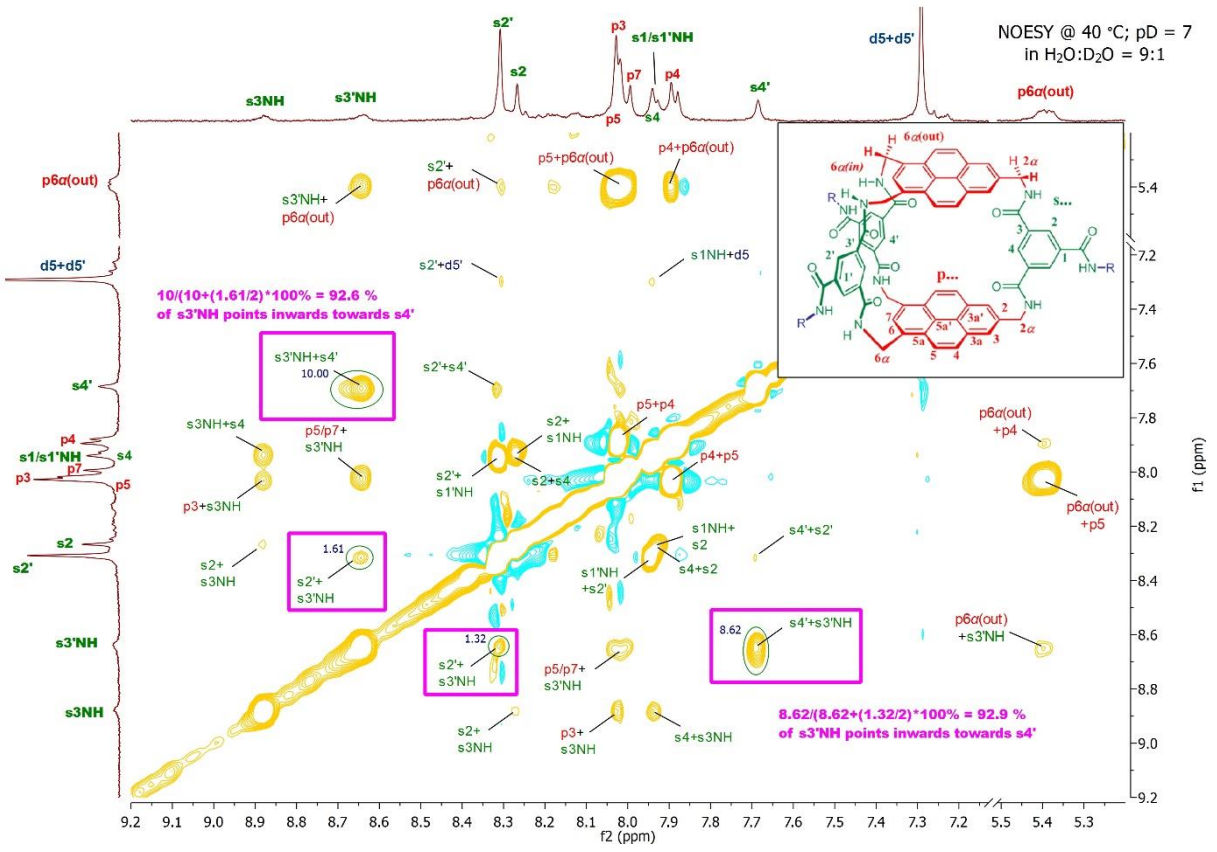
Supplementary Figure 3 | ^{1}H NMR variable temperature study of receptor **4** in D_2O at pD = 7.0 and T = 25 °C.



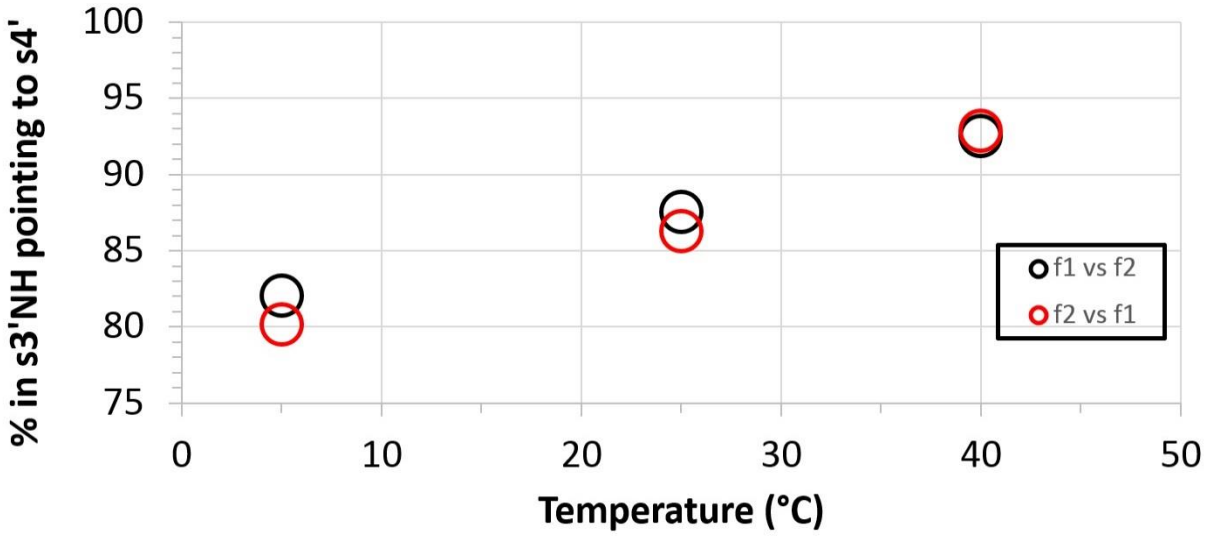
Supplementary Figure 4 | NOESY spectrum of receptor **4** in H_2O/D_2O 9:1 at pH = 7.0 and T = 5 °C.



Supplementary Figure 5 | NOESY spectrum of receptor **4** in H_2O/D_2O 9:1 at pH = 7.0 and T = 25 °C.



Supplementary Figure 6 | NOESY spectra of receptor **4** in H_2O/D_2O 9:1 at pH = 7.0 and T = 40 °C.



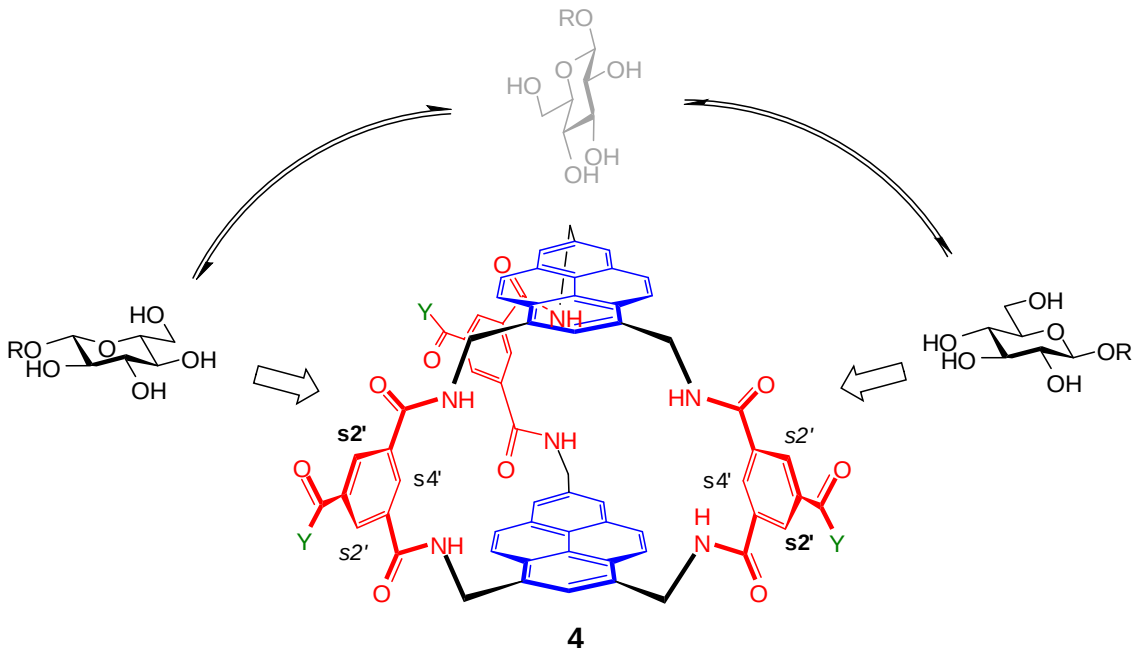
Supplementary Figure 7 | Percentage of inward-directed s3'NH as a function of temperature, as determined from the NOESY spectra in Supplementary Figures 4-6.

2. Binding studies

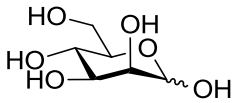
2.1. ^{1}H NMR titrations

NMR spectra were recorded at 298 K and 600 MHz on a Varian INOVA 600, or Varian VNMR S600 Cryo. Chemical shifts (δ) are reported in parts per million (p.p.m.). A solution of receptor **4** in D_2O (99.9% -D, typically 0.15 – 0.25 mM) was prepared and typically 500 μL was transferred into a new or thoroughly cleaned and dried NMR tube. This same receptor solution was then used to prepare a stock solution of carbohydrate. In the case of reducing sugars, the stock solution was left overnight so that α and β forms could equilibrate. Aliquots of increasing volume were added from the stock solution into the NMR tube and, after thorough mixing, the ^{1}H-NMR spectrum was recorded. The first addition typically was 0.5 μL; subsequent additions were made using double the volume of the previous addition (i.e. 1.0, 2.0, 4.0, 8.0 μL, etc.). Signals due to receptor protons were observed to move, as expected for binding with fast exchange on the ^{1}H NMR timescale. Some signals were found to split into two, for reasons explained below in Supplementary Figure 8. The changes in chemical shift ($\Delta\delta$) were first fitted to a 1:1 binding model using a non-linear least squares curve-fitting programme implemented within Excel. The programme yields binding constants K_a and limiting Δδ as output. An estimated error for K_a could be obtained from individual data points by assuming the determined K_a and δ_{HG}. For monosaccharides and chitobiose **3** (n = 1), the fits obtained were good to reasonable (mostly with $r \geq 0.999$), and the thus obtained K_a and Δδ were considered trustworthy. Spectra and analysis curves are shown in Supplementary Figures 9-13 and 16. For cellodextrins **1** the signals continued moving at high substrate concentrations in a manner which suggests weak binding of a second (and possibly a third) substrate molecule. In these cases, the data was reprocessed using the program 'HypNMR'[6] assuming a combination of 1:1 and 1:2 receptor **4**:carbohydrate stoichiometries. This procedure is less reliable than fitting to a simple 1:1 model, as several parameters are varied simultaneously. However, the results were generally consistent with those obtained by ITC (see later). Examples of spectra and fitting curves (both 1:1 and 1:1+1:2 binding models) for cellobiose **1** (n = 1) and cellotriose **1** (n = 2) are shown in Supplementary Figures 14 and 15 respectively.

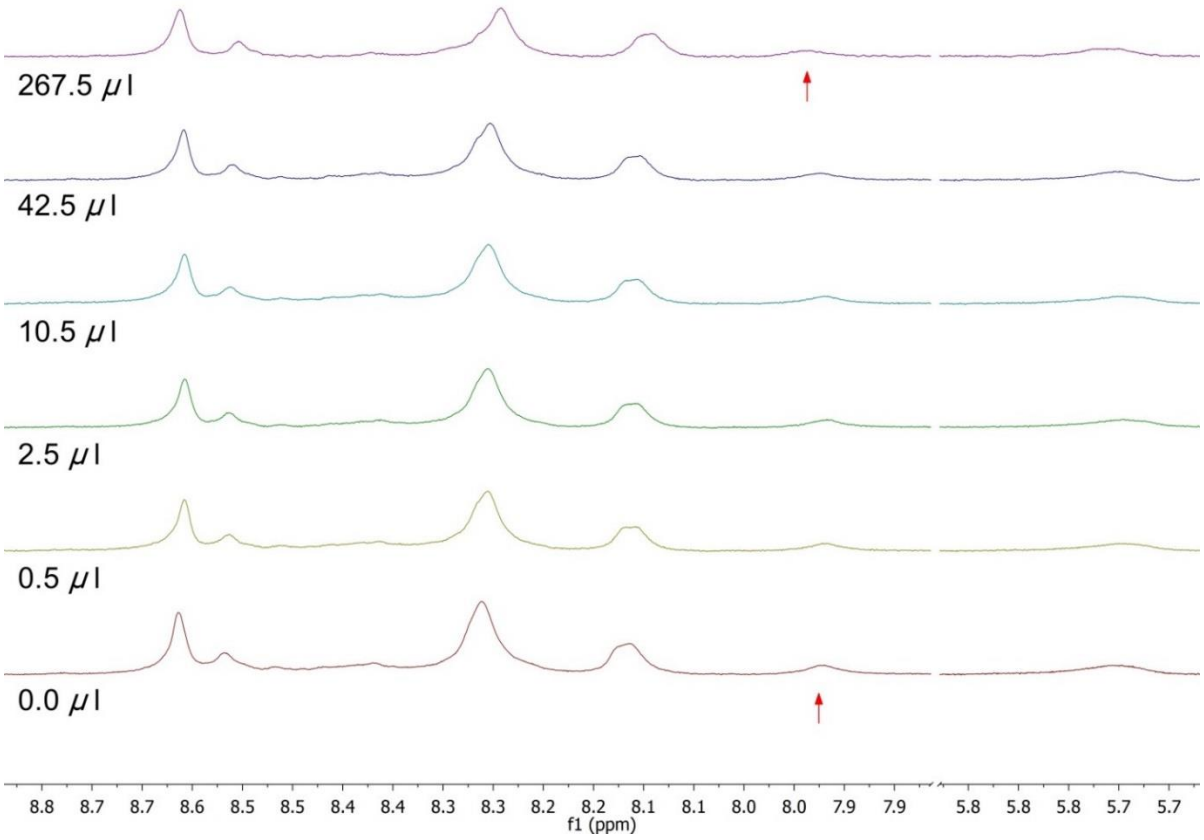
[6] http://www.hyperquad.co.uk/hypnmr.htm (April 2015). The programmes is able to fit multiple stoichiometries, also uses non non-linear least squares curve-fitting, and has K_a and limiting Δδ as output.

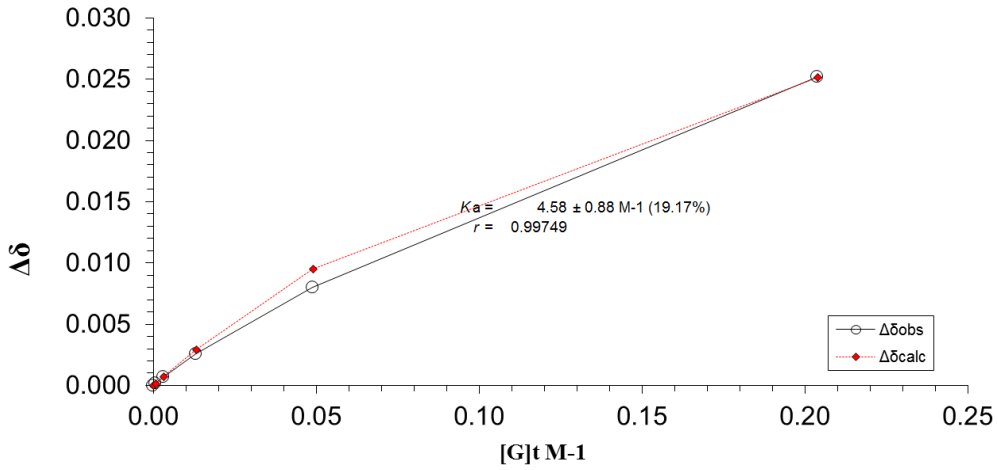


Supplementary Figure 8 | Desymmetrisation of proton environments in 4 on binding to carbohydrates. In the unbound state receptor **4** features two planes of symmetry so that, for example, the four protons s2' give just one NMR signal. On entering the cavity, the carbohydrate causes complete loss of symmetry, so that protons s4' and all four protons s2' now occupy different environments. For binding with slow exchange on the ^{1}H NMR timescale, the two signals in the free receptor thus become six in the complex. However, if exchange is now speeded up, the substrate will move rapidly between two equivalent orientations (as illustrated), causing pairs of signals to coalesce. The two s4' signals become one, so these protons appear as a single peak in the complex (as in the receptor). Thus addition of carbohydrate to **4** causes s4' to move but not to split (cf. Fig. 2a). For protons s2', exchange between the two substrate orientations again causes pairs of protons to coalesce. However, as there are four s2' protons, this results in two separate signals (one for **s2'** and one for *s2'*). Addition of carbohydrate to **4** therefore causes s2' to split (cf. Fig. 2a).

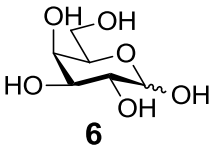


5

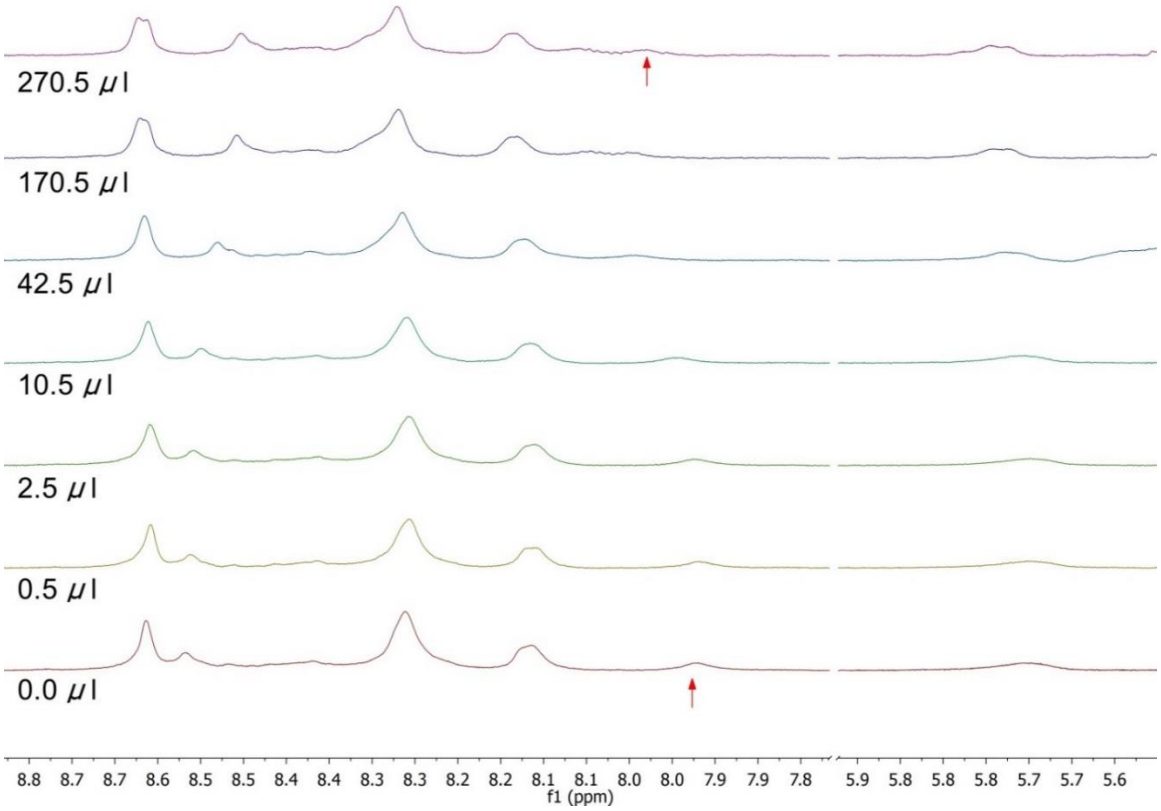


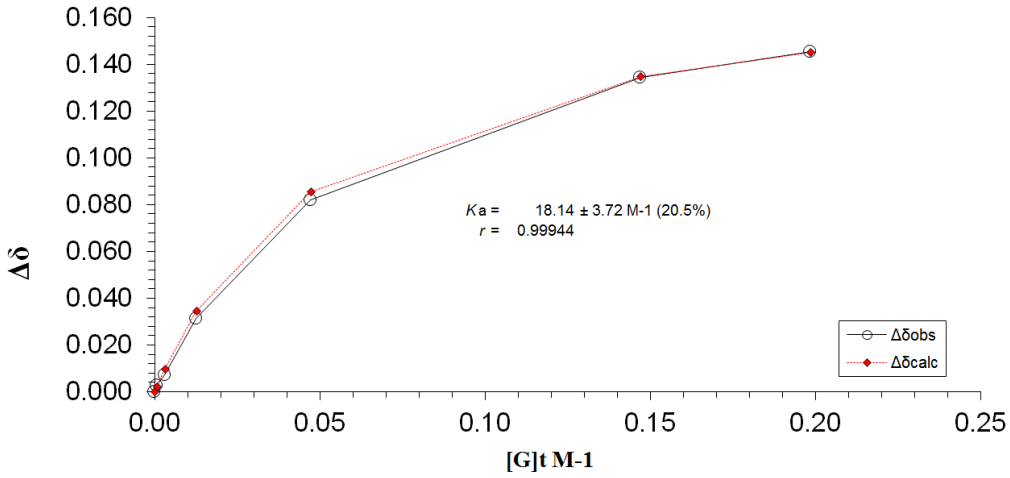


Supplementary Figure 9 | ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **4** (0.15 mM, 500 μL) titrated with mannose (**5**) (509 mM) in D_2O, also containing **4** (0.15 mM). Limiting $\Delta\delta$ = 0.052 p.p.m.

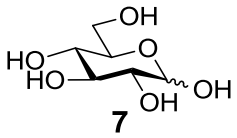


6

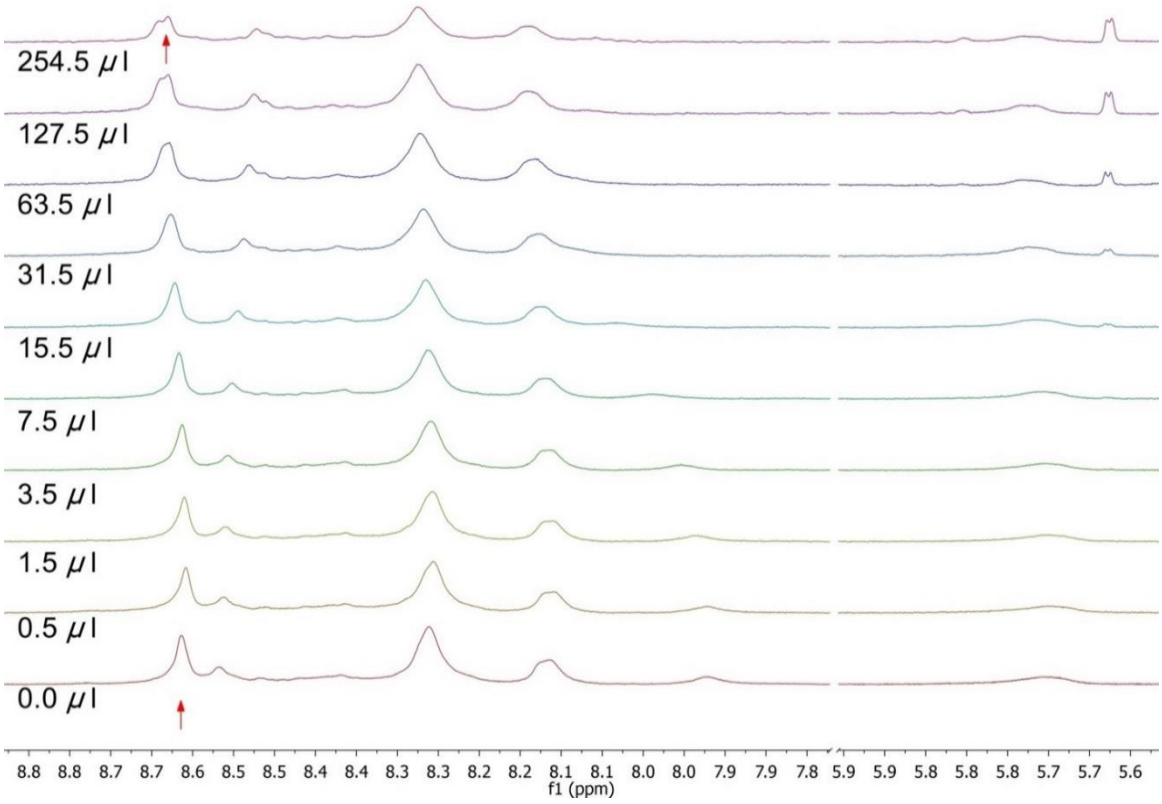


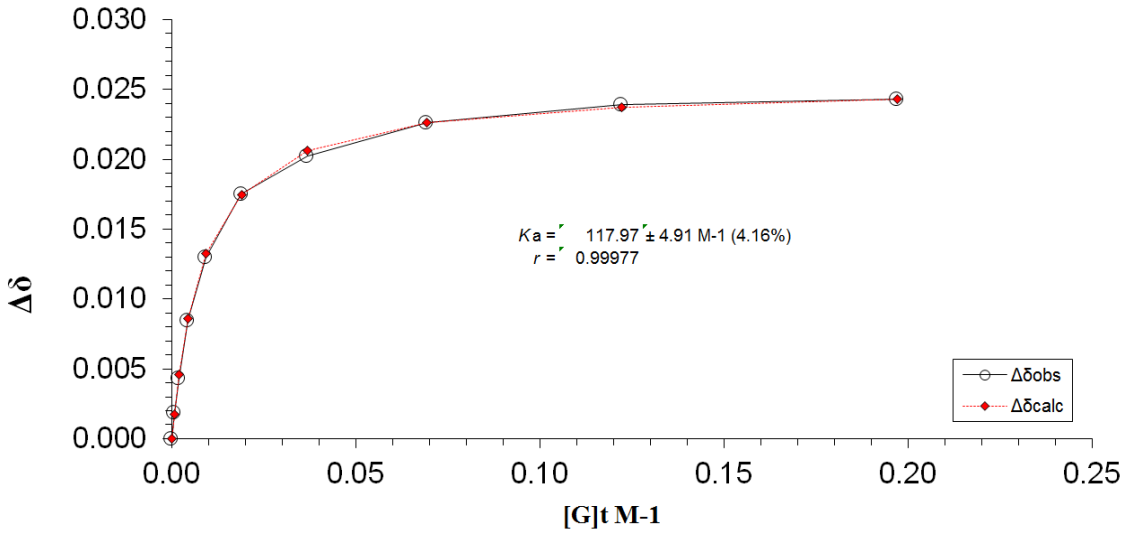


Supplementary Figure 10 | ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **4** (0.15 mM, 500 µL) titrated with galactose (**6**) (492 mM) in D_2O, also containing **4** (0.15 mM). Limiting Δδ = 0.185 p.p.m.

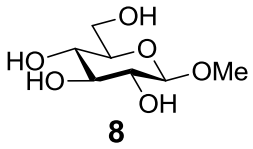


7

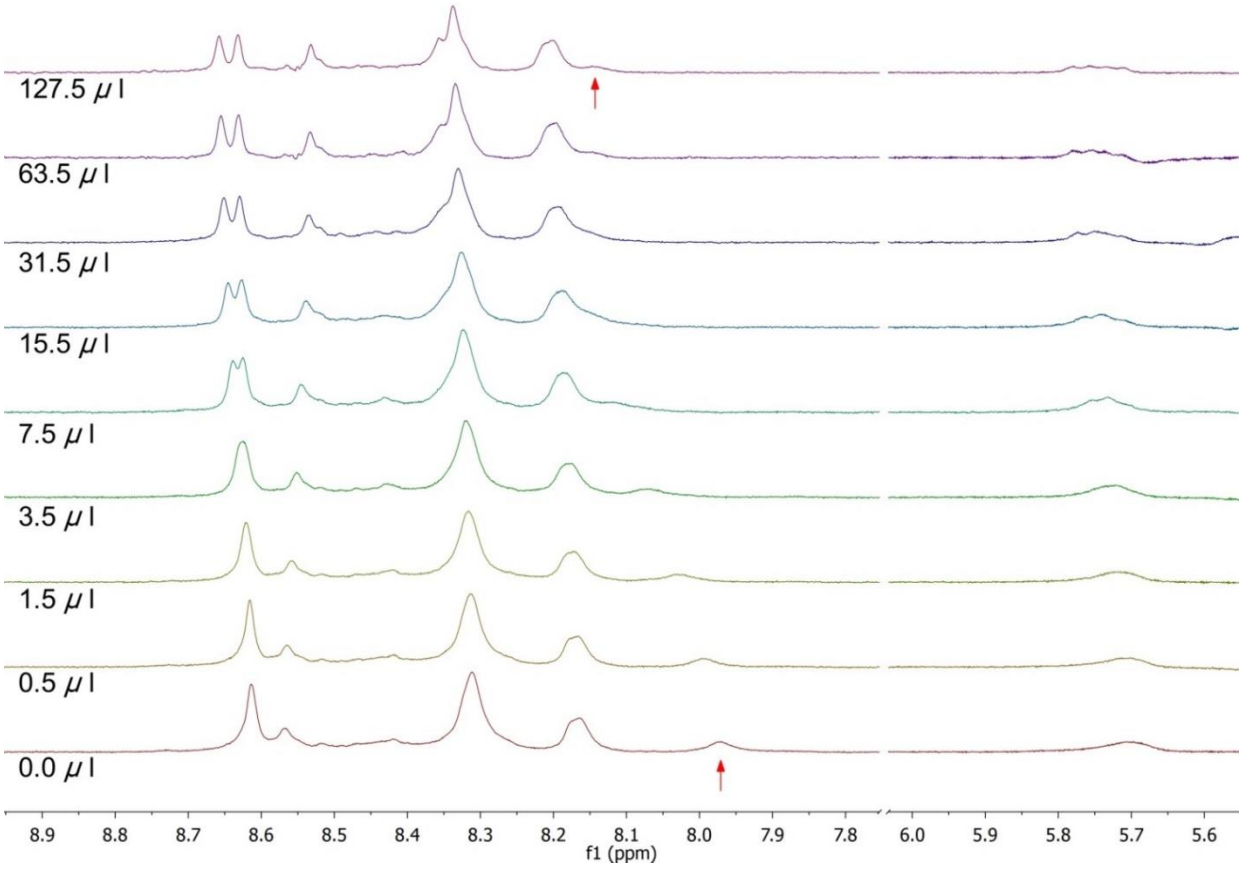


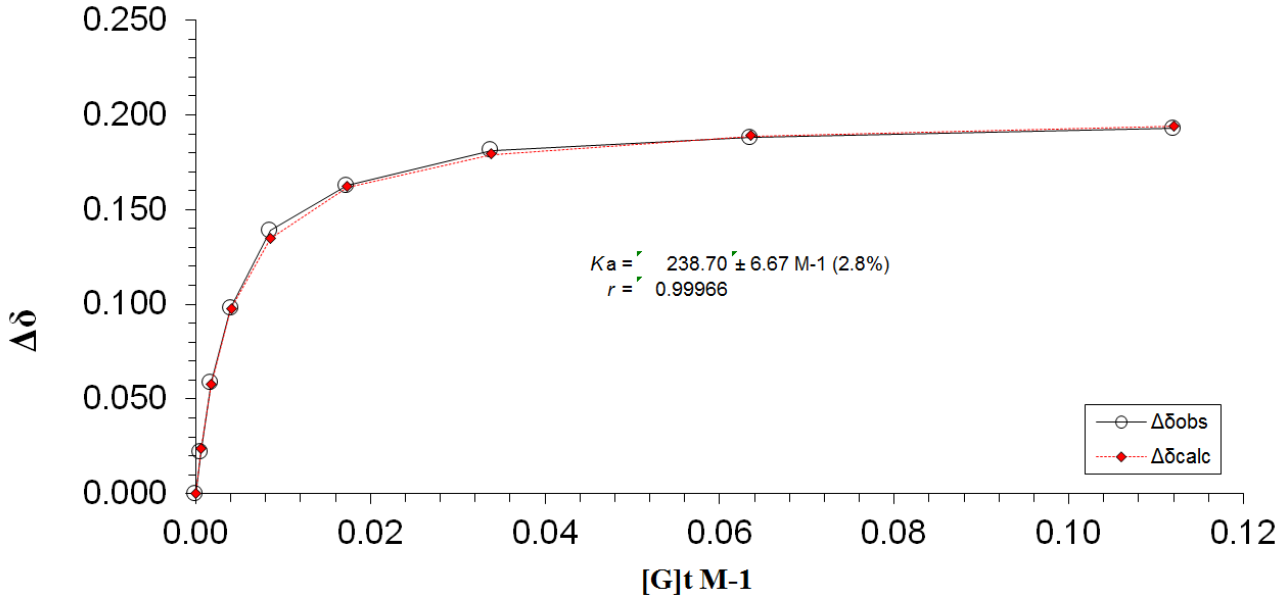


Supplementary Figure 11 | ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **4** (0.15 mM, 500 μL) titrated with glucose (**7**) (505 mM) in D_2O, also containing **4** (0.15 mM). Limiting Δδ = 0.025 p.p.m.

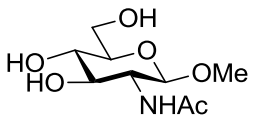


8

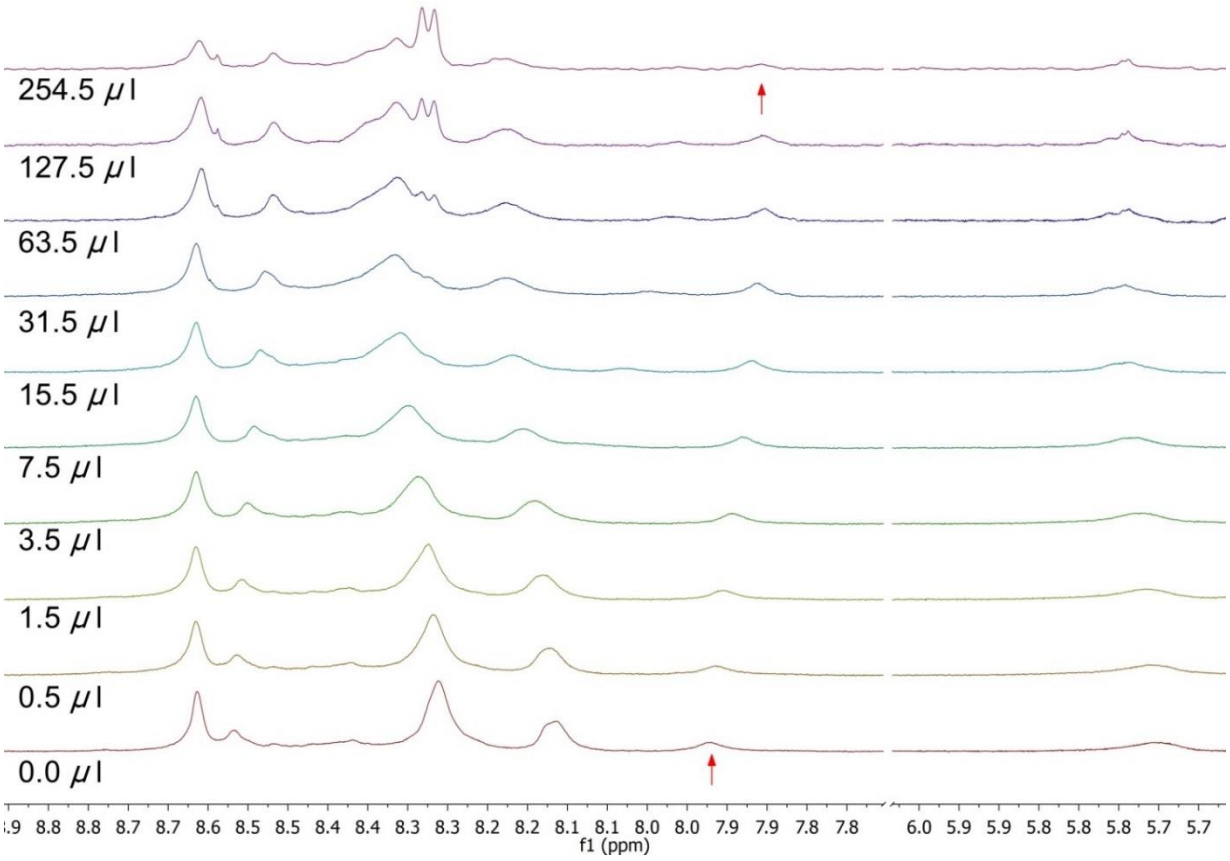


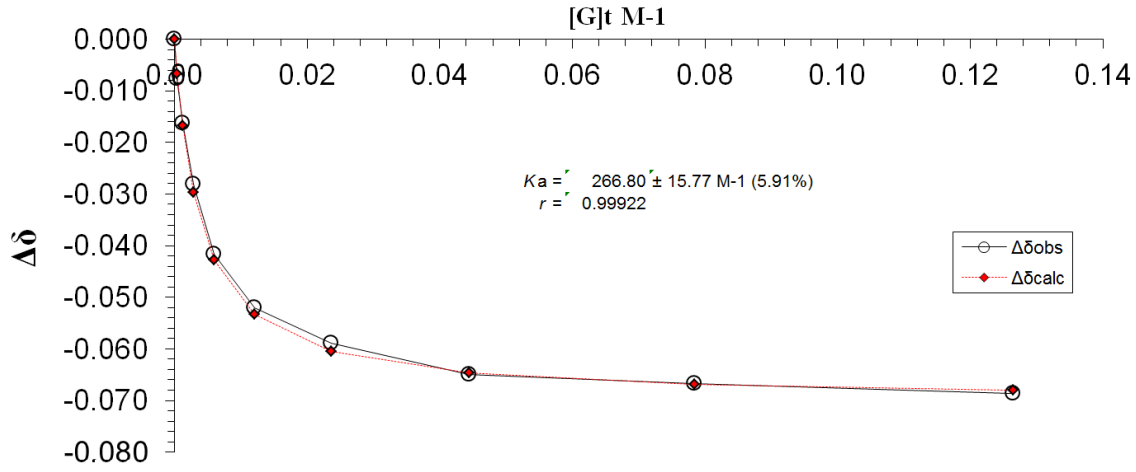


Supplementary Figure 12 | ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **4** (0.15 mM, 500 μL) titrated with methyl β-D-glucoside (**8**) (505 mM) in D_2O, also containing **4** (0.15 mM). Limiting Δδ = 0.201 p.p.m.

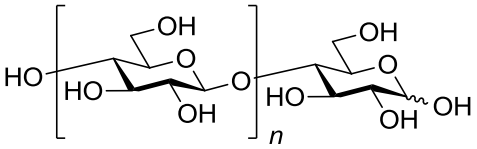


9

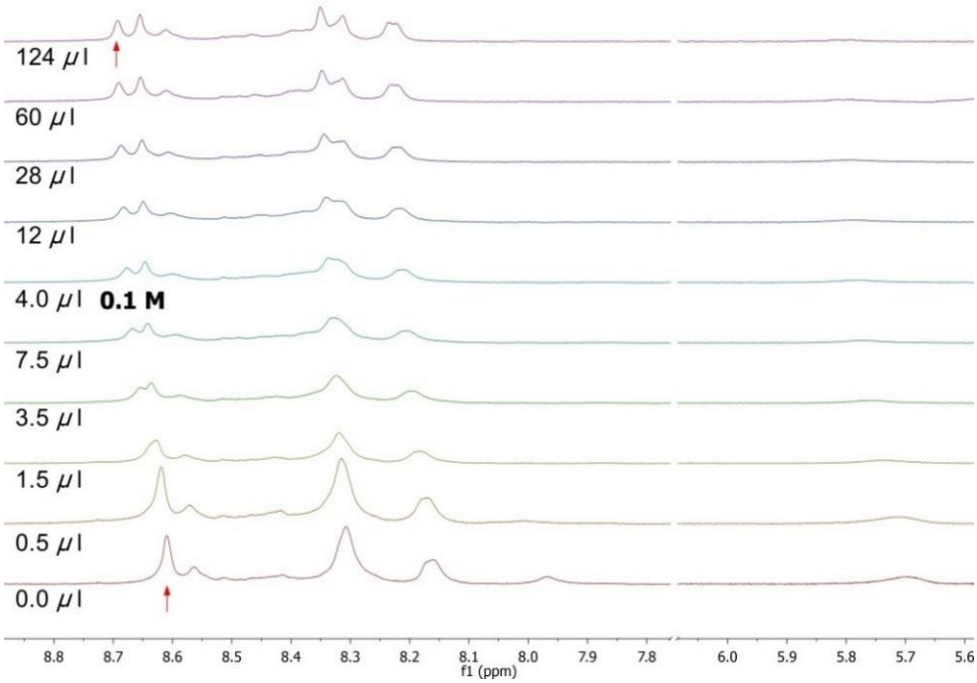


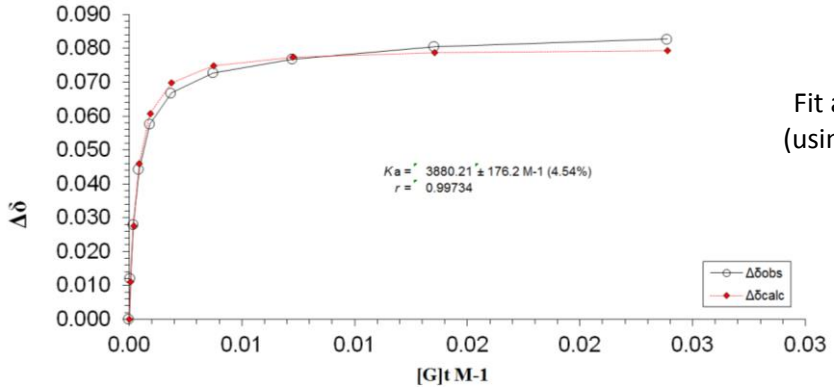


Supplementary Figure 13 | ^{1}H NMR spectra (top) and binding analysis curve (bottom) for receptor **4** (0.15 mM, 500 μL) titrated with methyl *N*-acetyl-*β*-D-glucosaminide (**9**) (325 mM) in D_2O, also containing **4** (0.15 mM). Limiting Δδ = –0.070 p.p.m.

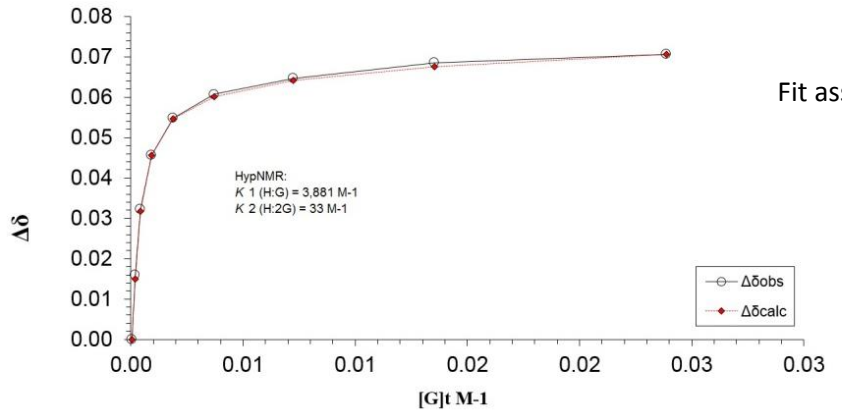


1, *n* = 1



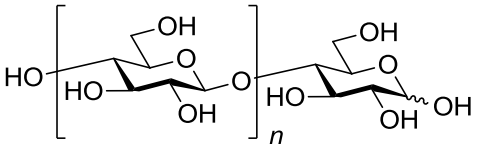


Fit assuming a 1:1 binding model
(using in-house Excel spreadsheet)

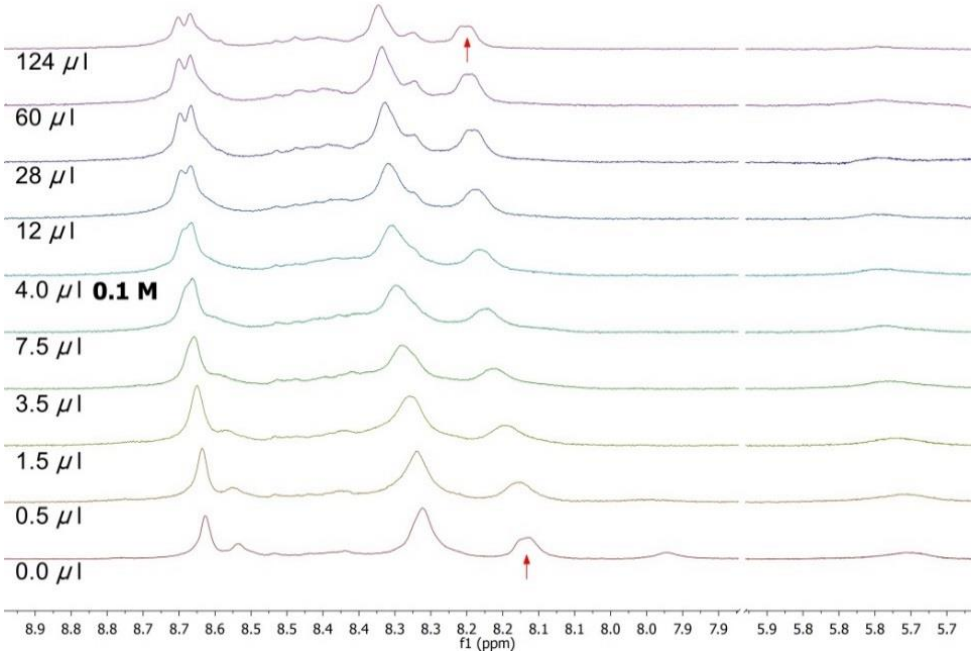


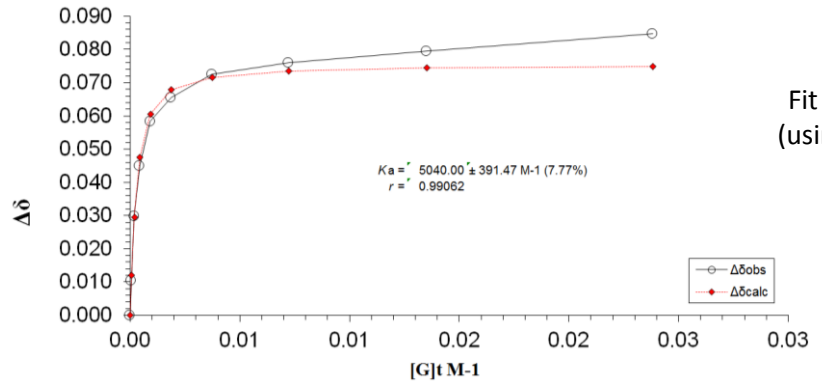
Fit assuming a 1:1 + 1:2 binding model
(using 'HypNMR')

Supplementary Figure 14 | ^{1}H NMR spectra (top) and binding analysis curves (bottom) for receptor **4** (0.15 mM, 500 µL) titrated with cellobiose (**1**, *n* = 1) (49.7 mM, then 99.4 mM) in D_2O, also containing **4** (0.15 mM). Limiting Δδ = 0.080 p.p.m.

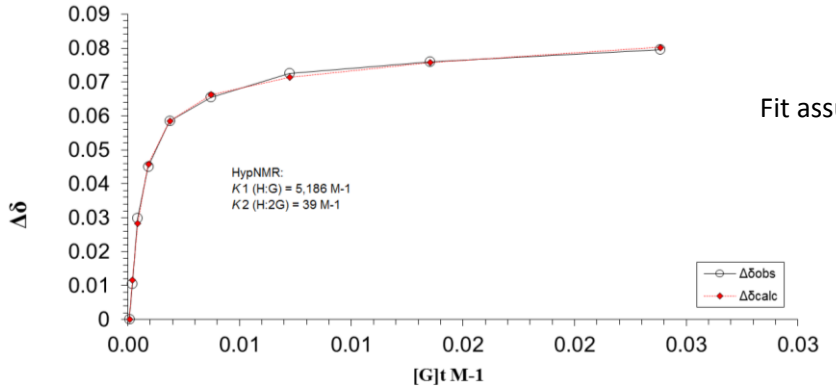


1, *n* = 2



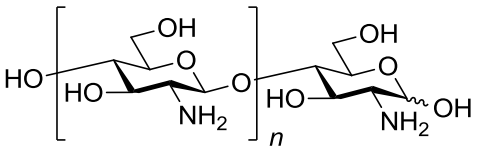


Fit assuming a 1:1 binding model
(using in-house Excel spreadsheet)

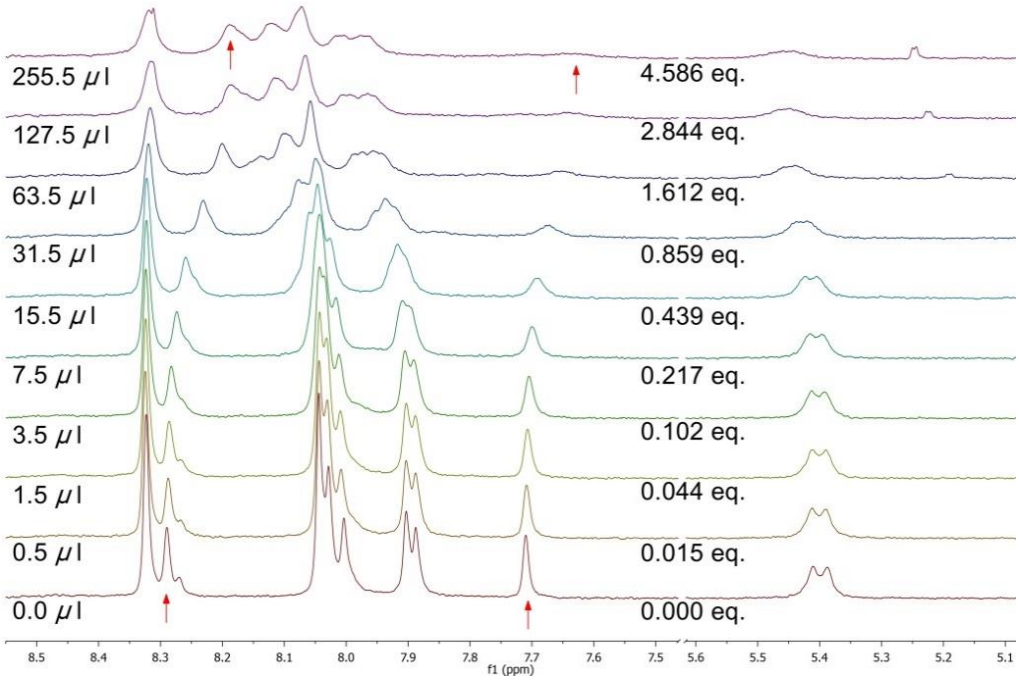


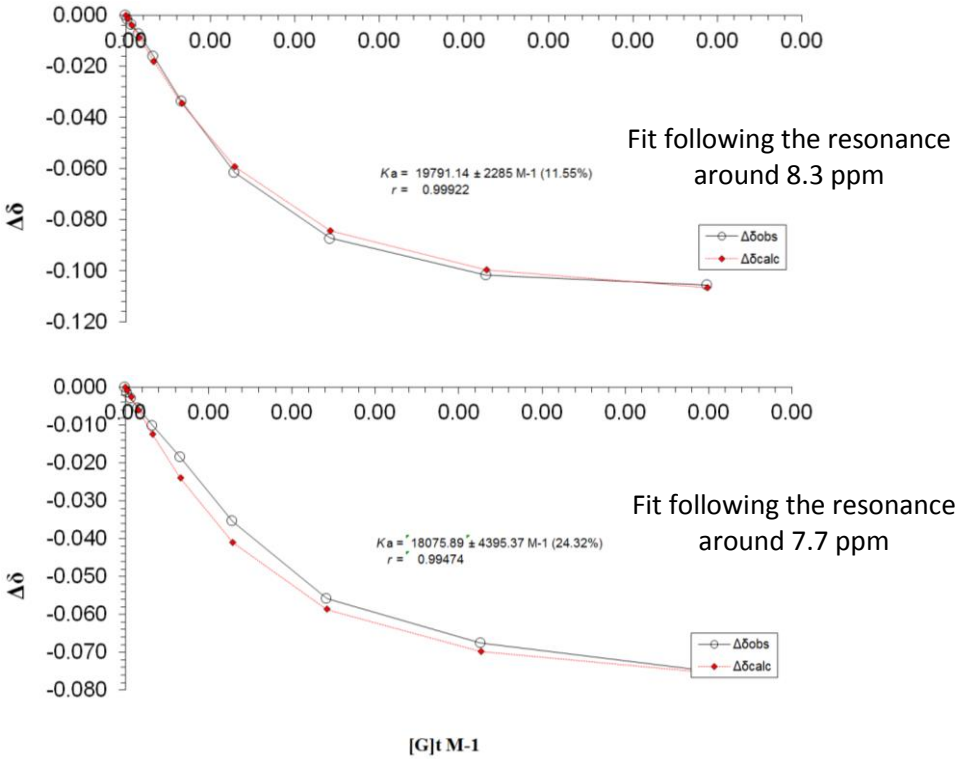
Fit assuming a 1:1 + 1:2 binding model
(using 'HypNMR')

Supplementary Figure 15 | ^{1}H NMR spectra (top) and binding analysis curves (bottom) for receptor **4** (0.15 mM, 500 µL) titrated with cellotriose (**1**, *n* = 2) (49.6 mM, then 99.1 mM) in D_2O, also containing **4** (0.15 mM). K_a (1:1) taken from corresponding ITC experiment.



3, *n* = 1

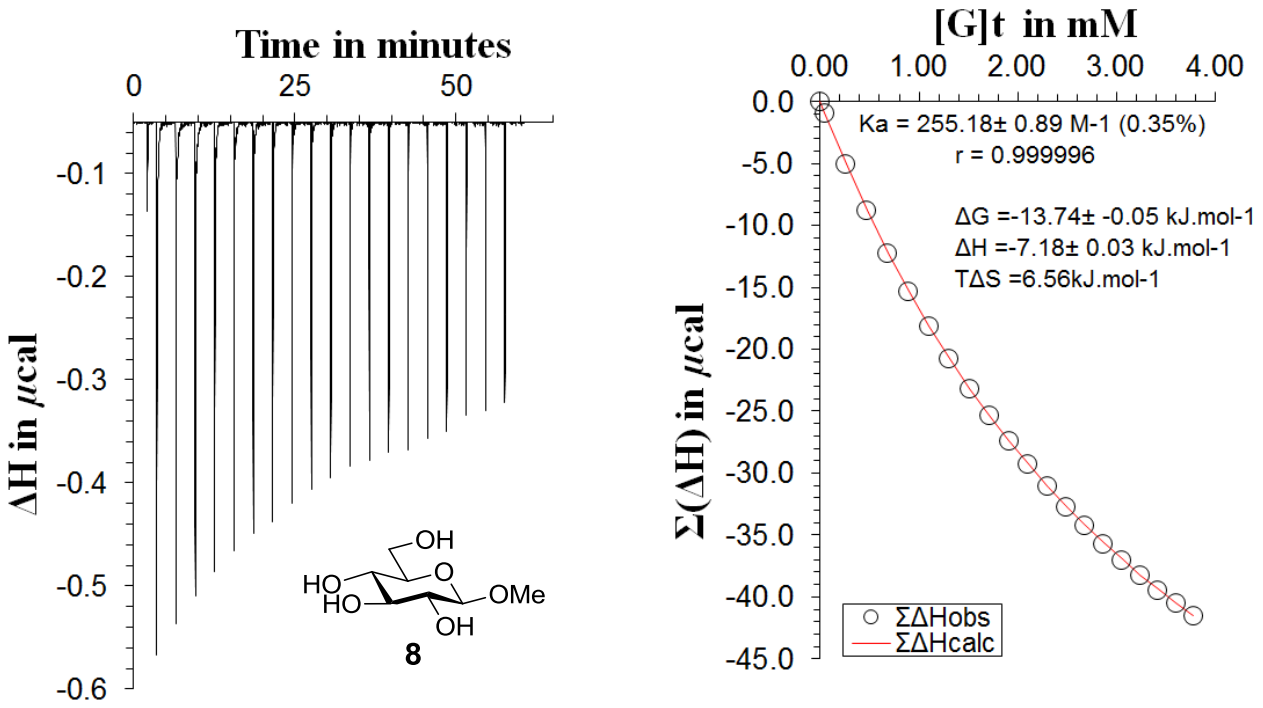




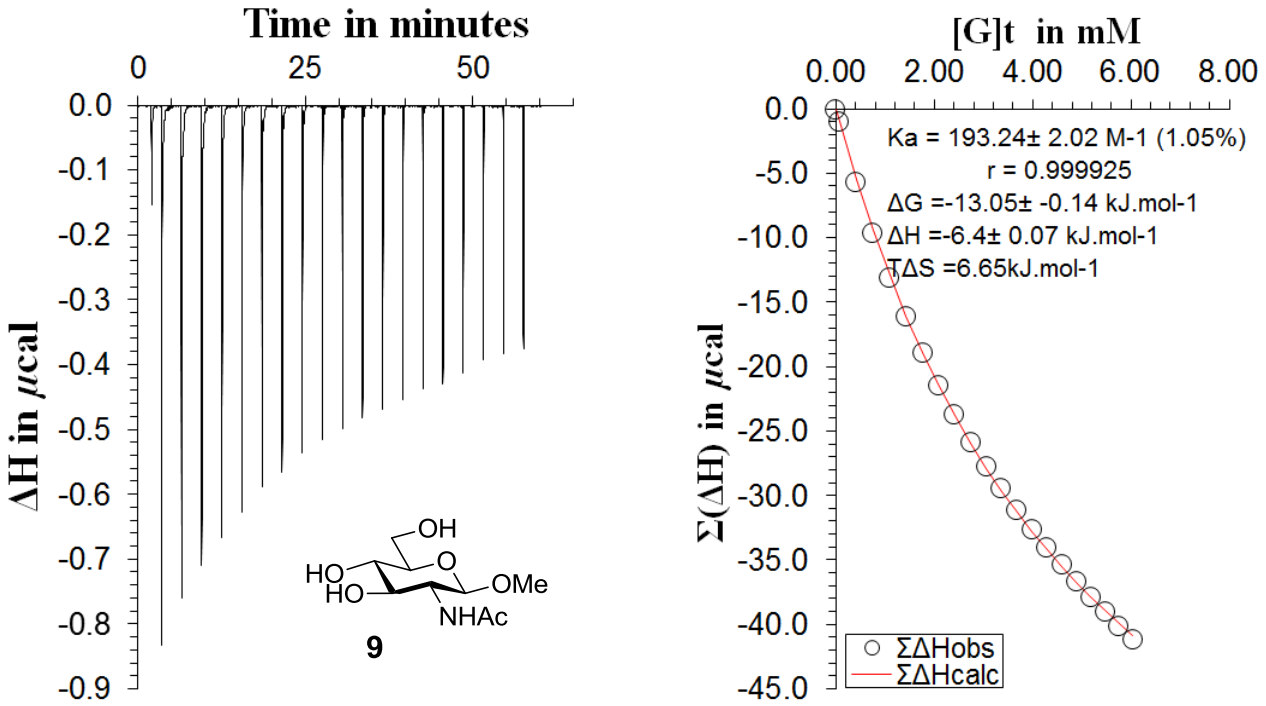
Supplementary Figure 16 | ^{1}H NMR spectra (top) and binding analysis curves (bottom) for receptor **4** (0.15 mM, 500 µL) titrated with chitobiose (**3**, n = 1) (49.6 mM, then 99.1 mM) in D_2O, also containing **4** (0.15 mM). pD = 7. Limiting Δδ = –0.120 p.p.m. (following the resonance ~ 8.3 p.p.m.); –0.08 p.p.m. (following the resonance at ~ 7.7 p.p.m.).

2.2. Isothermal Titration Microcalorimetry (ITC)

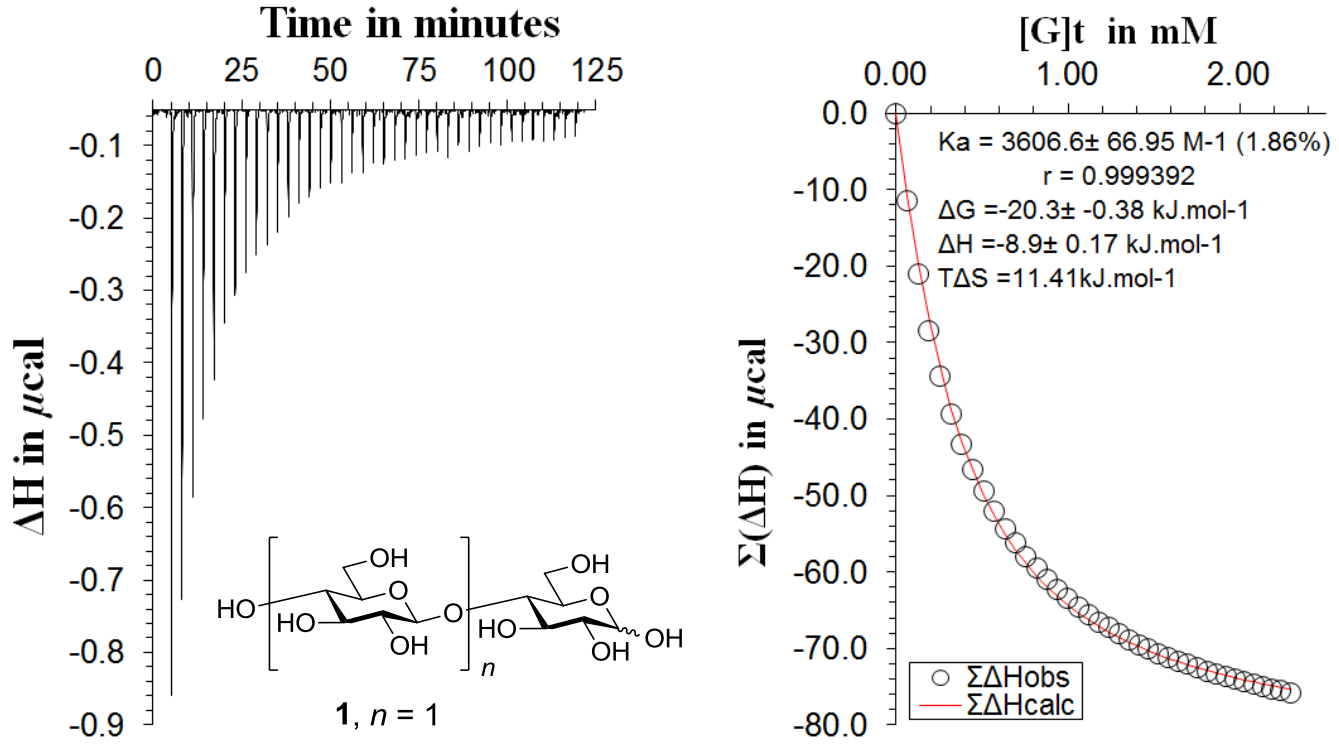
Isothermal Titration MicroCalorimetry (ITC) experiments were performed on a MicroCal iTC200 microcalorimeter. ITC experiments were carried out at 298 K. Saccharide solutions were prepared in HPLC-grade water and allowed to equilibrate overnight if necessary (see ^{1}H-NMR titrations above). The sample cell was charged with 200 μL of a receptor solution in HPLC-grade water (typically 0.2 – 0.5 mM). Then, aliquots (typically 1.0 μL) of carbohydrate solution were added and the evolution of heat was followed as a function of time. Heats of dilution were measured by injecting the same carbohydrate solution into HPLC-grade water, using identical conditions. For every addition, the heat of dilution was subtracted from the heat of binding using a MicroCal software programme implemented in ORIGIN 5.0. This gave an XY matrix of heat *vs*. total guest concentration. This matrix was then imported into a specially written Excel programme to fit the data to a 1:1 binding model, using similar mathematics as employed in the ^{1}H-NMR titrations, giving K_a. ΔG can be derived from K_a, and ΔS can be derived from ΔH and ΔG using common thermodynamic equations. This method of analysis consistently produced more accurate fits than the MicroCal software. Raw data and analysis curves are shown in Supplementary Figures 17-25.



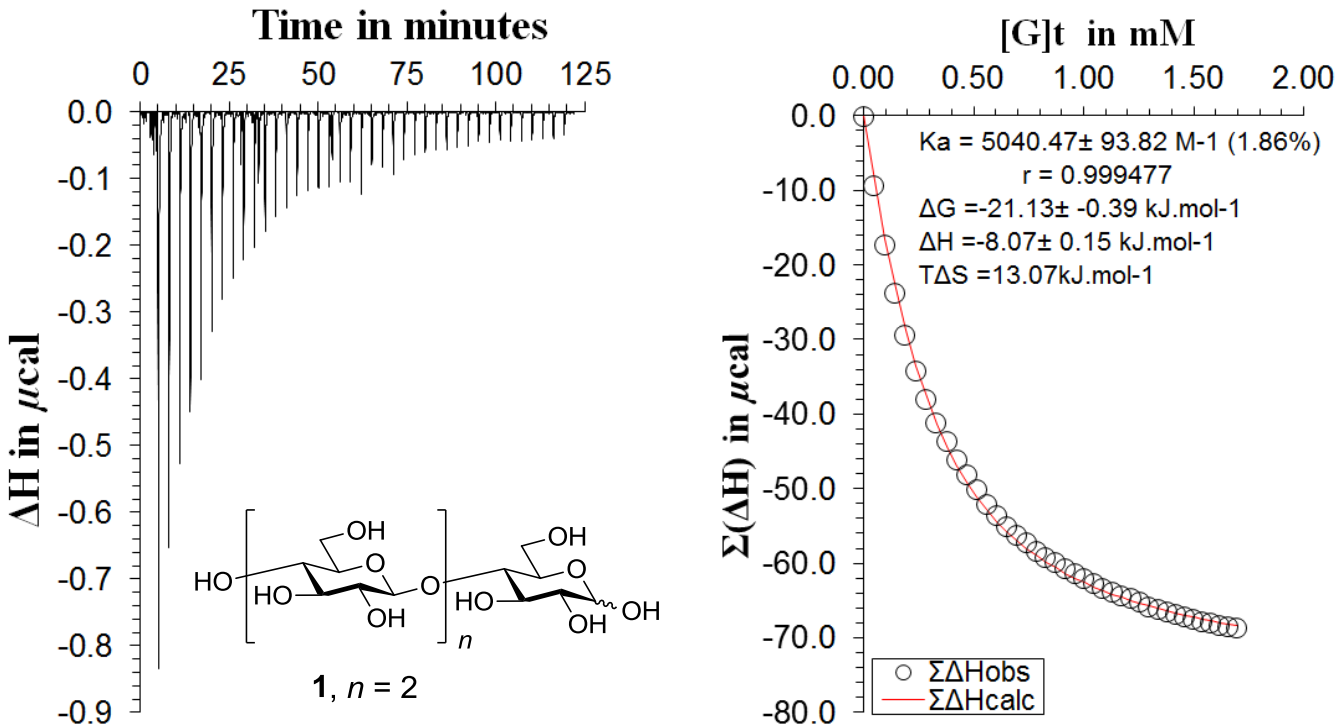
Supplementary Figure 17 | ITC data from addition of methyl β-D-glucoside (**8**) (21.9 mM) to **4** (0.25 mM, 200 μL). The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 255 ± 1 M^{-1} (0.4%) and r = 1.000 (thermodynamic data also given in the figure).



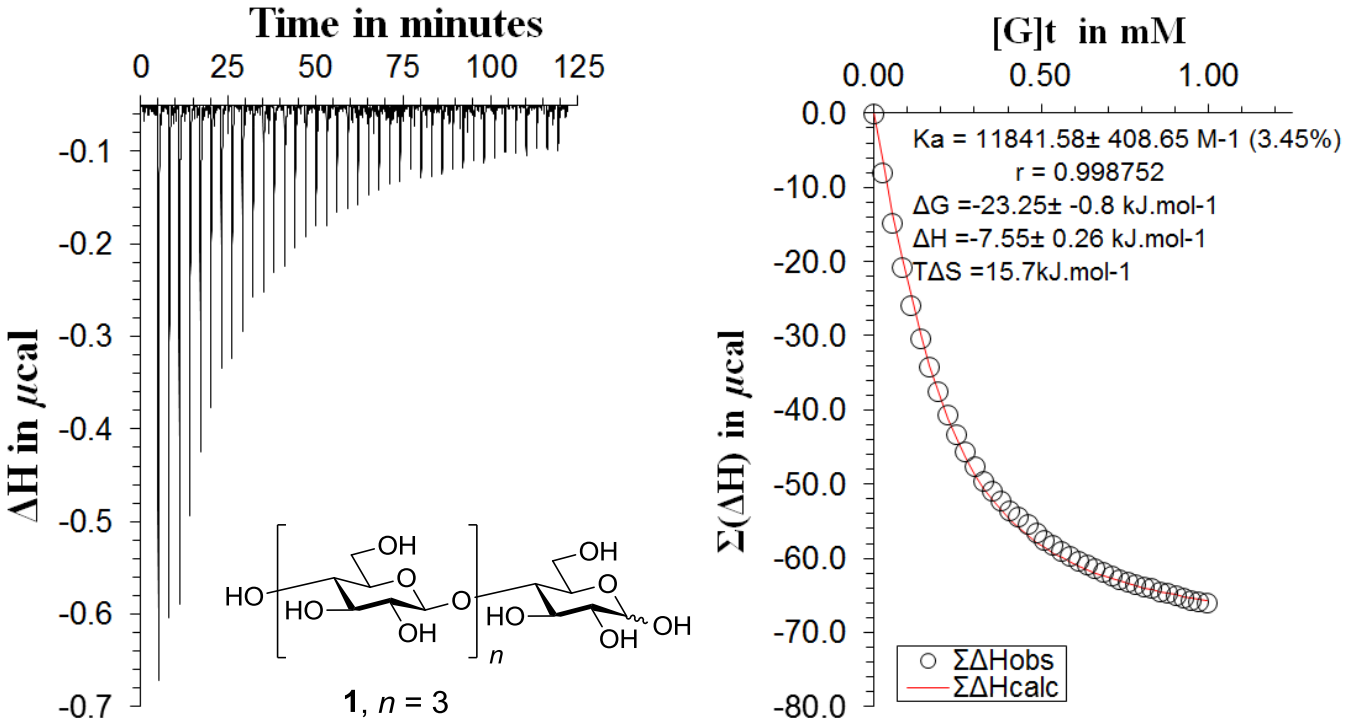
Supplementary Figure 18 | ITC data from addition of methyl *N*-acetyl-*β*-D-glucosaminide (**9**) (34.8 mM) to **4** (0.25 mM, 200 μL). The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 193 ± 2.0 M^{-1} (1.1%) and r = 0.9999 (thermodynamic data also given in the figure).



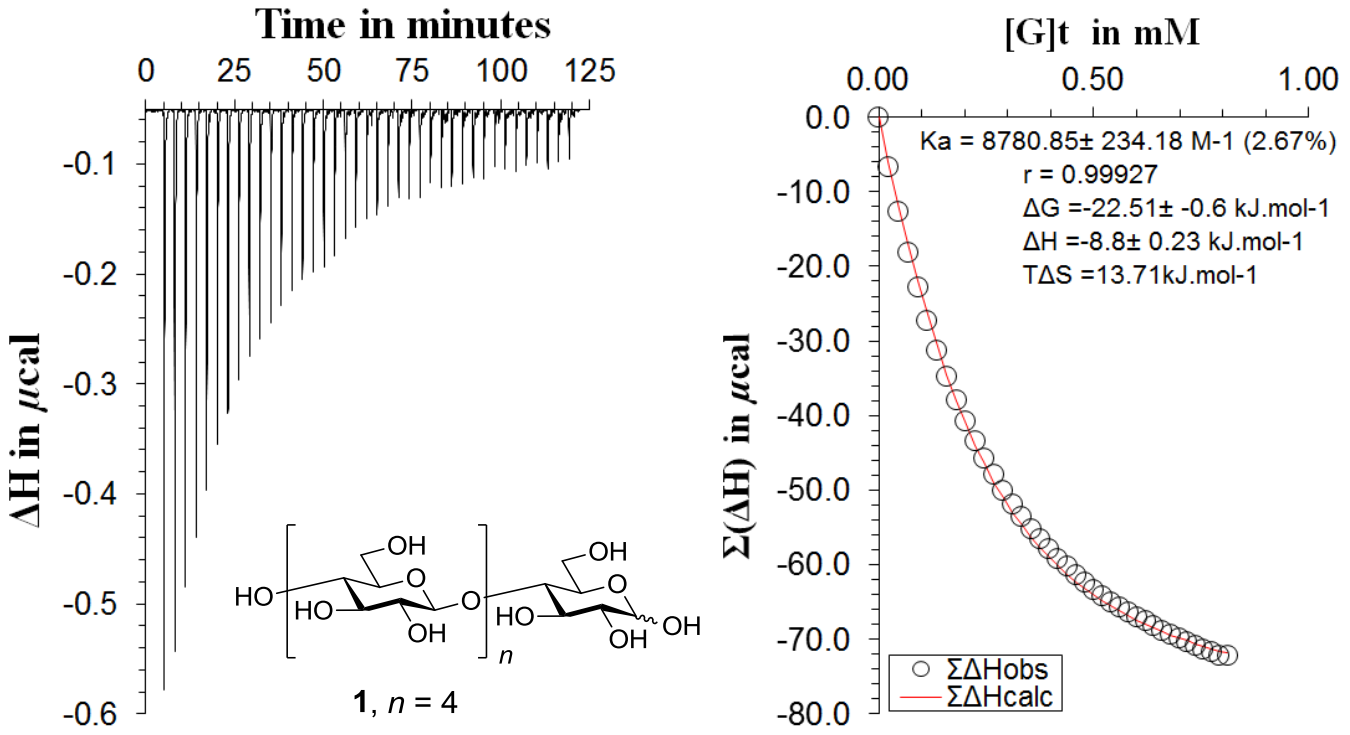
Supplementary Figure 19 | ITC data from addition of cellobiose (**1**, n = 1) (13.05 mM) to **4** (0.2 mM, 200 μL). The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 3,607 ± 67 M^{-1} (1.9%) and r = 0.9994 (thermodynamic data also given in the figure).



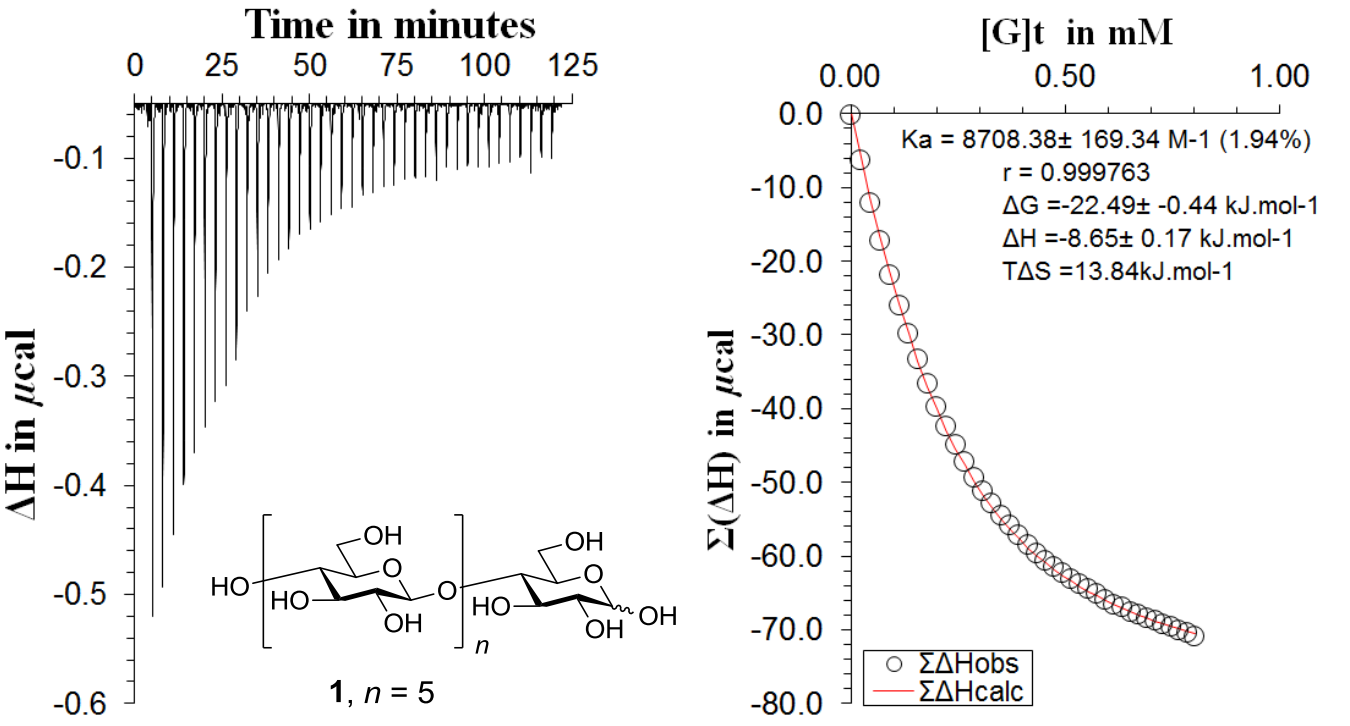
Supplementary Figure 20 | ITC binding study of receptor **4** (0.2 mM, 200 μL) titrated with D-cellotriose (**1**, n = 2) (13.05 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 5,040 ± 94 M^{-1} (1.9%) and r = 0.9995 (thermodynamic data also given in the figure).



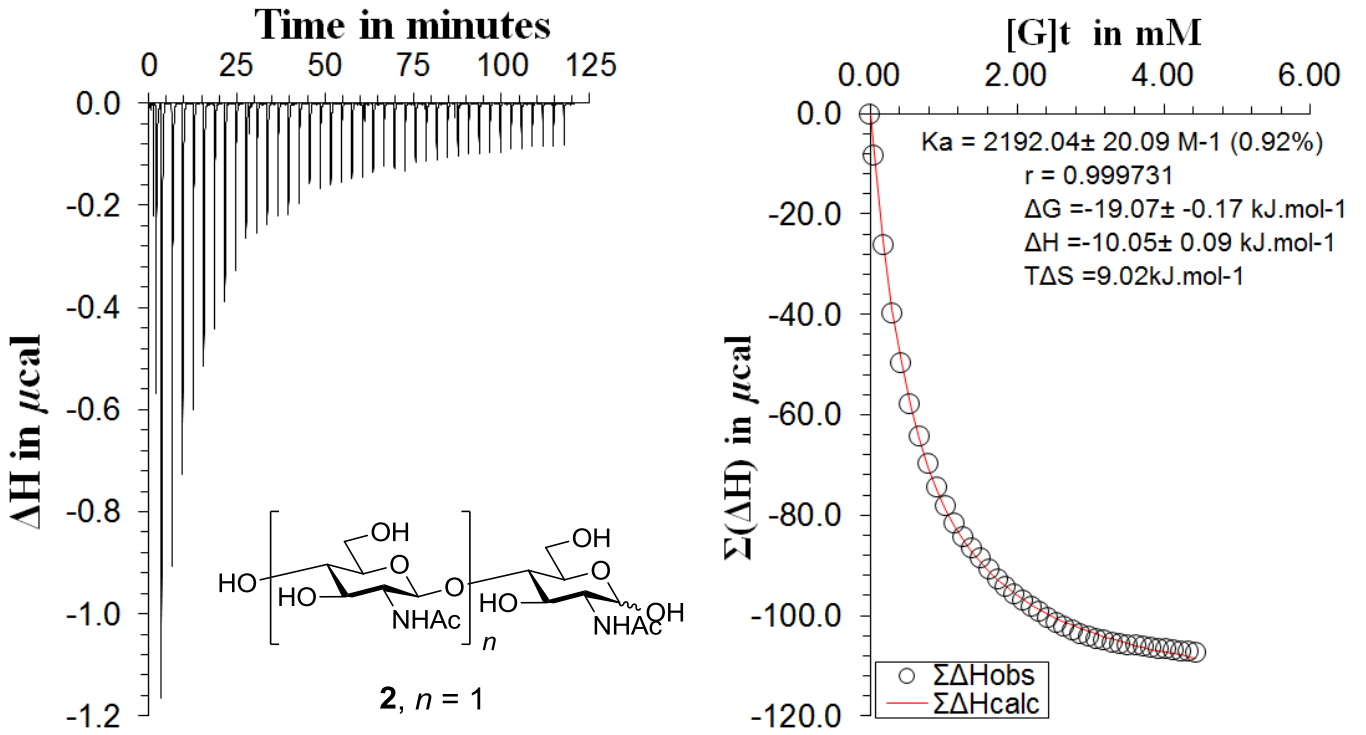
Supplementary Figure 21 | ITC binding study of receptor **4** (0.2 mM, 200 μL) titrated with D-cellotetraose (**1**, n = 3) (5.66 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 11,842 ± 409 M^{-1} (3.5%) and r = 0.9988 (thermodynamic data also given in the figure).



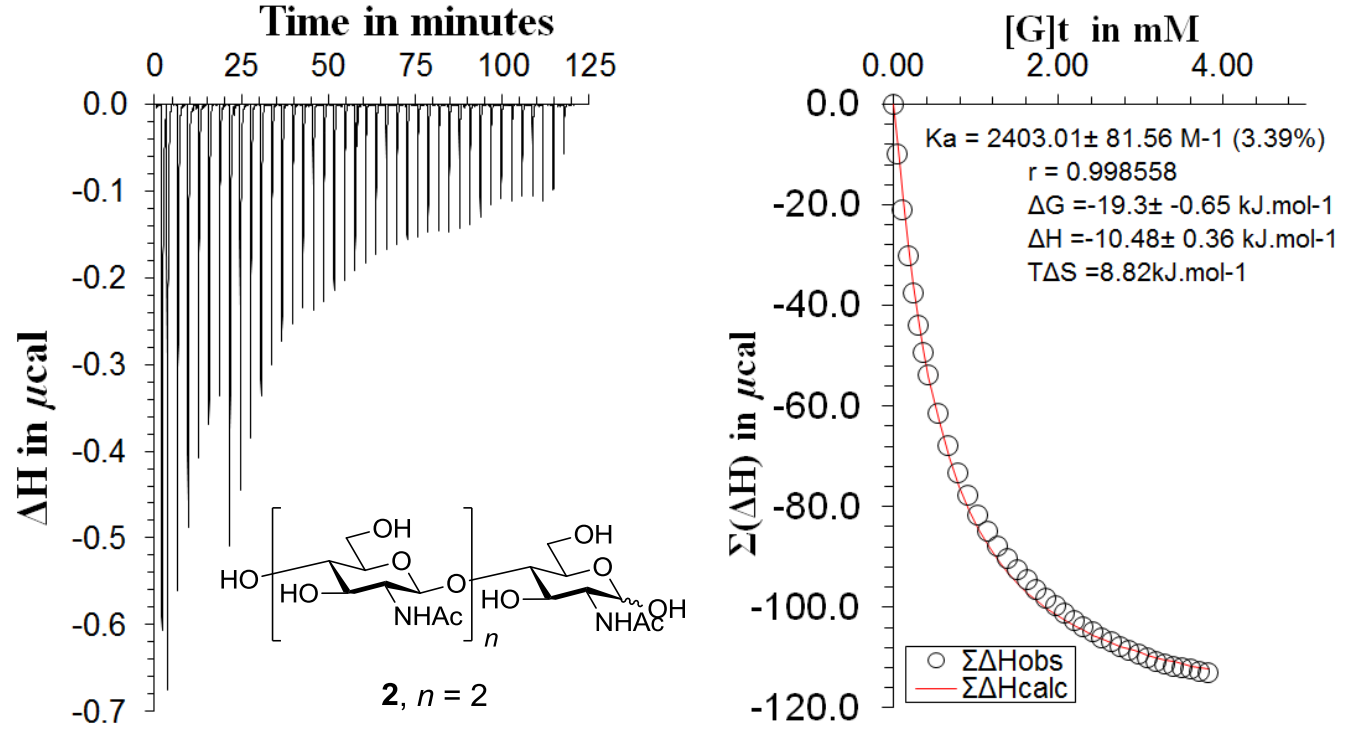
Supplementary Figure 22 | ITC binding study of receptor **4** (0.2 mM, 200 μL) titrated with D-cellopentaose (**1**, n = 4) (4.61 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 8,781 ± 234 M^{-1} (2.7%) and r = 0.9993 (thermodynamic data also given in the figure).



Supplementary Figure 23 | ITC binding study of receptor **4** (0.2 mM, 200 μL) titrated with D-cellohexaose (**1**, n = 5) (4.56 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 8,708 ± 169 M^{-1} (1.9%) and r = 0.9998 (thermodynamic data also given in the figure).



Supplementary Figure 24 | ITC binding study of receptor **4** (0.25 mM, 200 μL) titrated with N,N'-diacetyl-D-chitobiose (**2**, n = 1) (25.3 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating K_a = 2,192 ± 20 M^{-1} (0.9%) and r = 0.9997 (thermodynamic data also given in the figure).



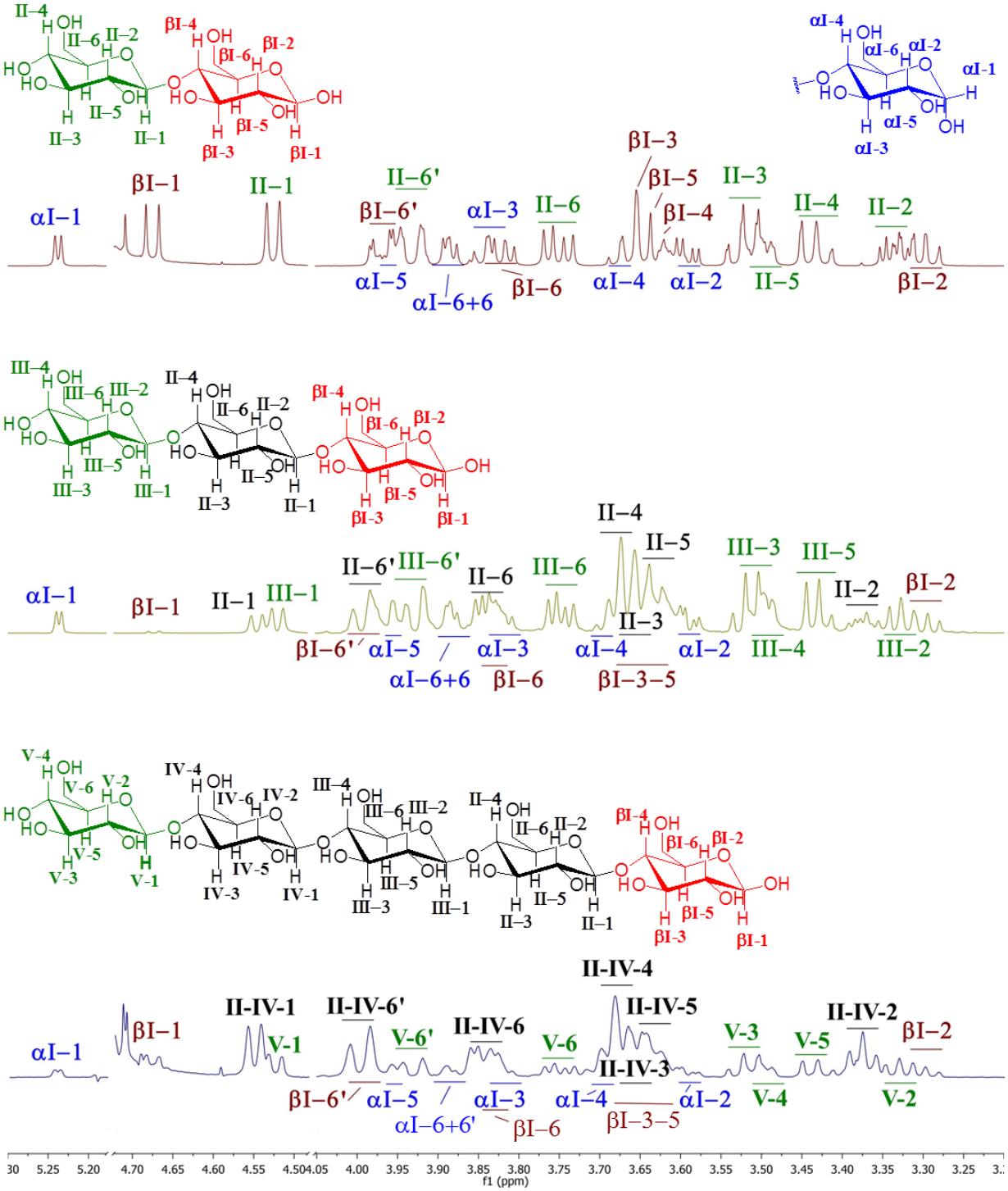
Supplementary Figure 25 | ITC binding study of receptor **4** (0.25 mM, 200 μL) titrated with N,N',N"-triacetyl-D-chitotriose (**2**, n = 2) (25.2 mM) in H_2O. The sum of heat evolution was plotted as a function of the concentration of carbohydrate and fitted to a 1:1 binding model indicating 2,403 ± 82 M^{-1} (3.4%) and r = 0.9986 (thermodynamic data also given in the figure).

3. NMR structural studies

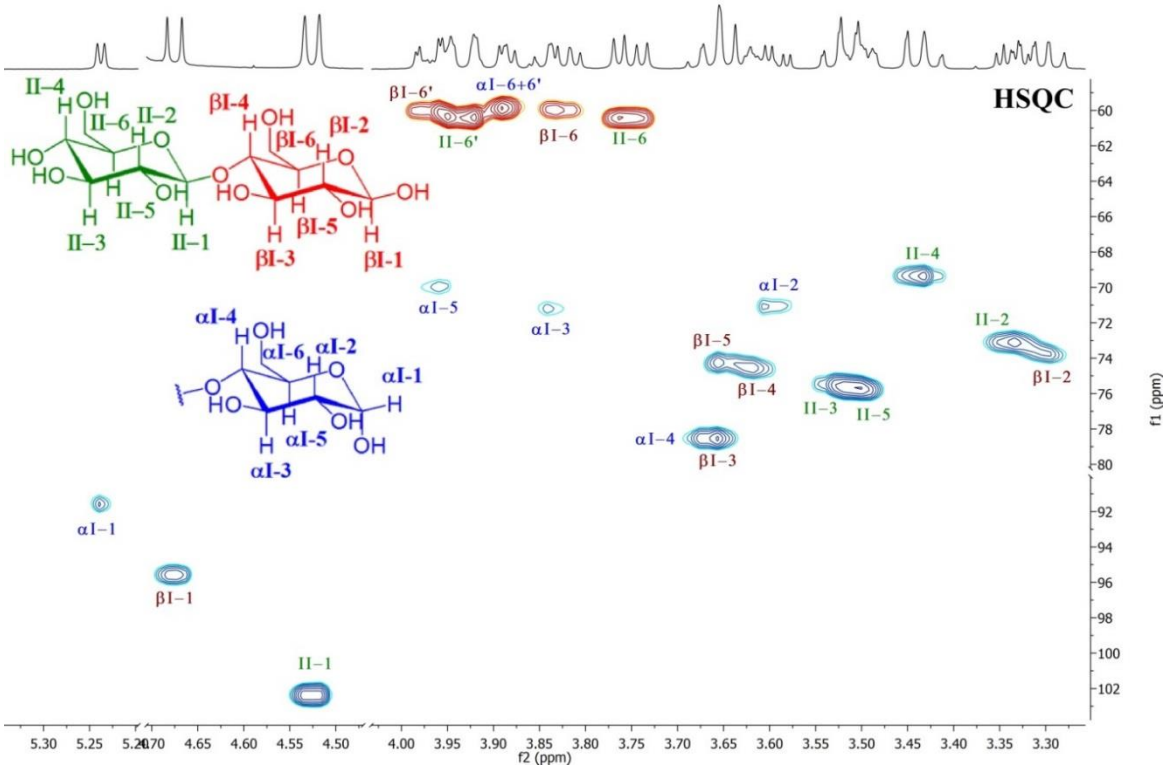
3.1. ^{1}H-NMR assignments of cellodextrins

The NMR spectra were acquired at 298 K at 600 MHz using a Varian VNMR S600 Cryo. Chemical shifts (δ) are reported in parts per million (p.p.m.). The ^{1}H NMR spectra were recorded at 298 K at 600 MHz on a Varian VNMR s600 Cryo. These spectra of the cellodextrins in D_2O were first assigned on the basis of 2-D NMR spectra with guidance from the literature[7]. The assignments are summarised in Supplementary Figure 26, and the 2D spectra are shown in Supplementary Figures 27-33. Note that the intensity of resonances due to αI-1 and βI-1 is weakened by solvent suppression.

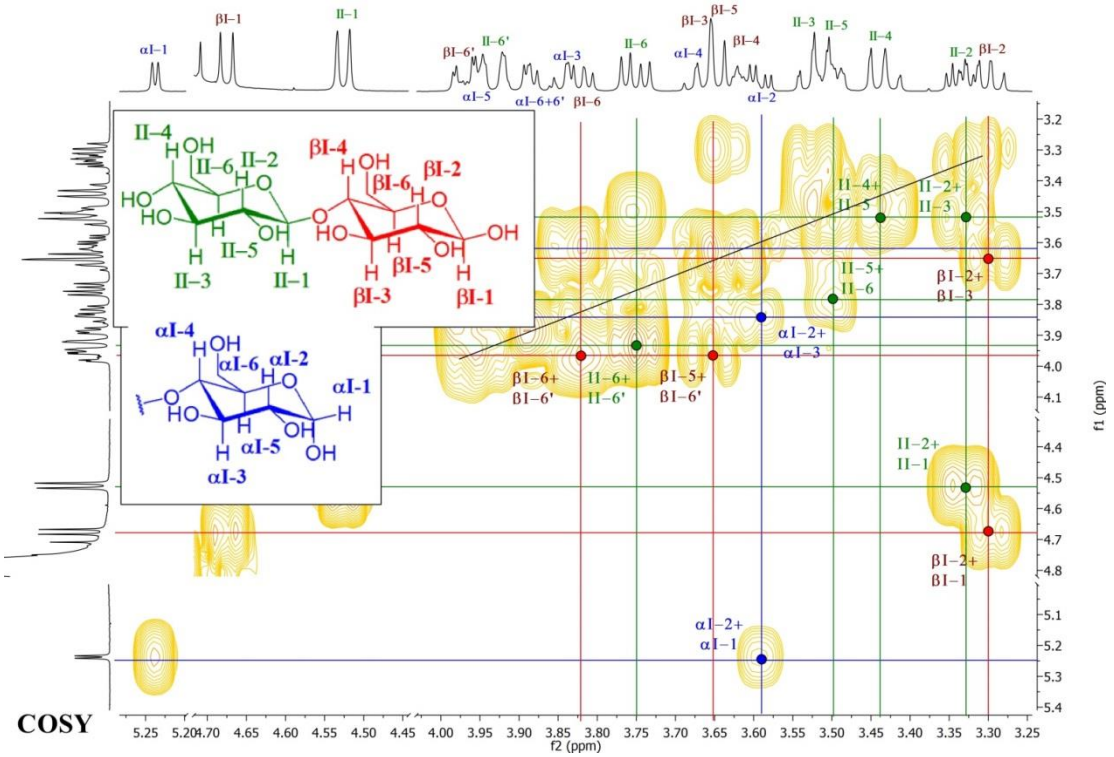
[7] Flugge, L. A., Blank, J. T. & Petillo, P. A. Isolation, modification, and NMR assignments of a series of cellulose oligomers. *J. Am. Chem. Soc.* **121**, 7228-7238 (1999); Roslund, M. U., Tahtinen, P., Niemitz, M. & Sjoholm, R. Complete assignments of the H-1 and C-13 chemical shifts and J(H,H) coupling constants in NMR spectra of D-glucopyranose and all D-glucopyranosyl-D-glucopyranosides. *Carbohydr. Res.* **343**, 101-112 (2008); Shintate, K., Kitaoka, M., Kim, Y. K. & Hayashi, K. Enzymatic synthesis of a library of beta-(1 -> 4) hetero-D-glucose and D-xylose-based oligosaccharides employing cellodextrin phosphorylase. *Carbohydr. Res.* **338**, 1981-1990 (2003); Sugiyama, H., Hisamichi, K., Usui, T., Sakai, K. & Ishiyama, J. A study of the conformation of beta-1,4-linked glucose oligomers, cellobiose to cellohexaose, in solution. *J. Mol. Struct.* **556**, 173-177 (2000).



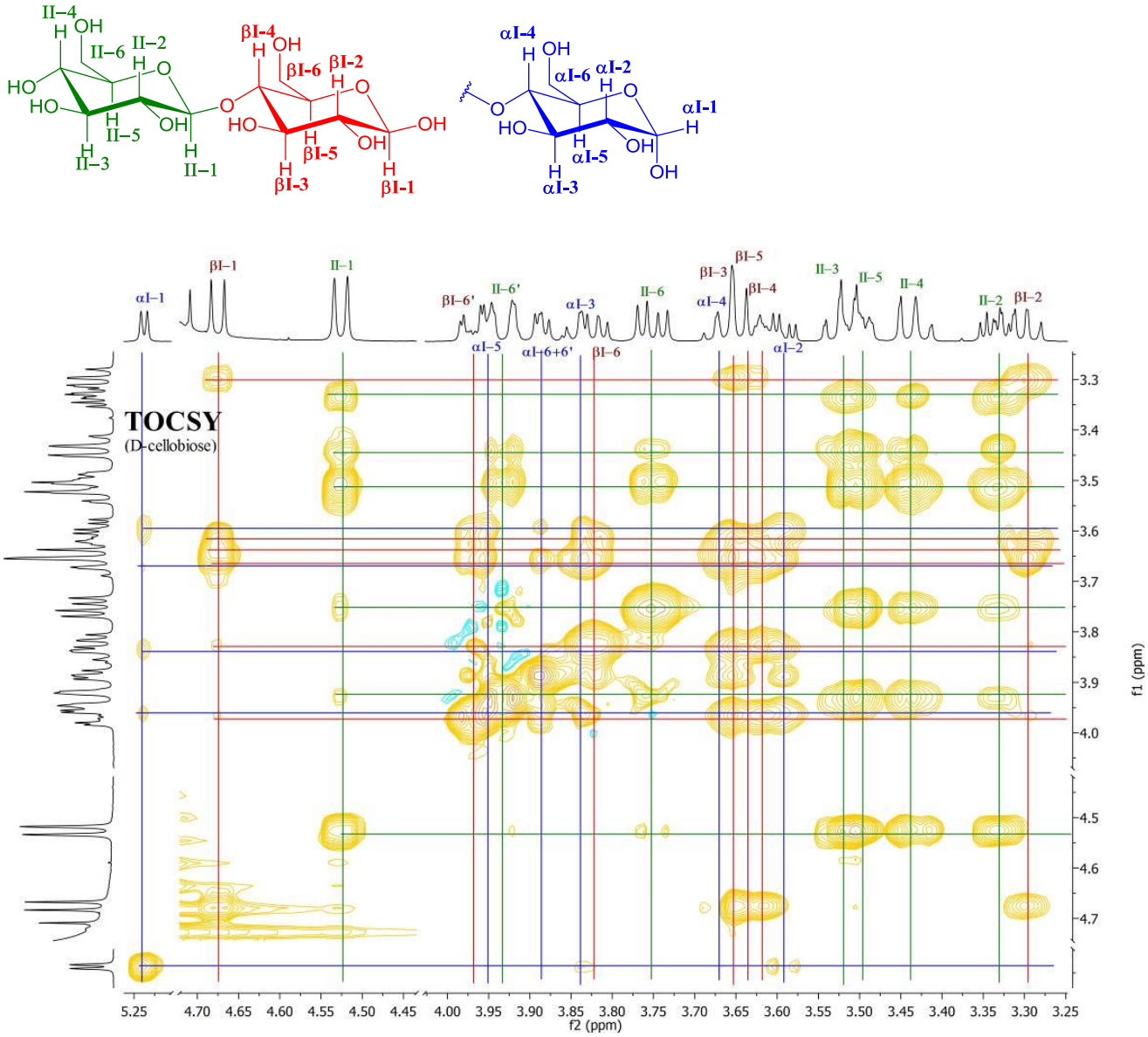
Supplementary Figure 26 | Assigned ^{1}H-NMR spectra recorded at 600 MHz of D-cellobiose (**1**, n = 1; top), D-cellotriose (**1**, n = 2; middle), and D-cellopentaose (**1**, n = 4; bottom) in D_2O. Note that the intensity of resonances due to *α*I-1 and *β*I-1 is weakened by solvent suppression.



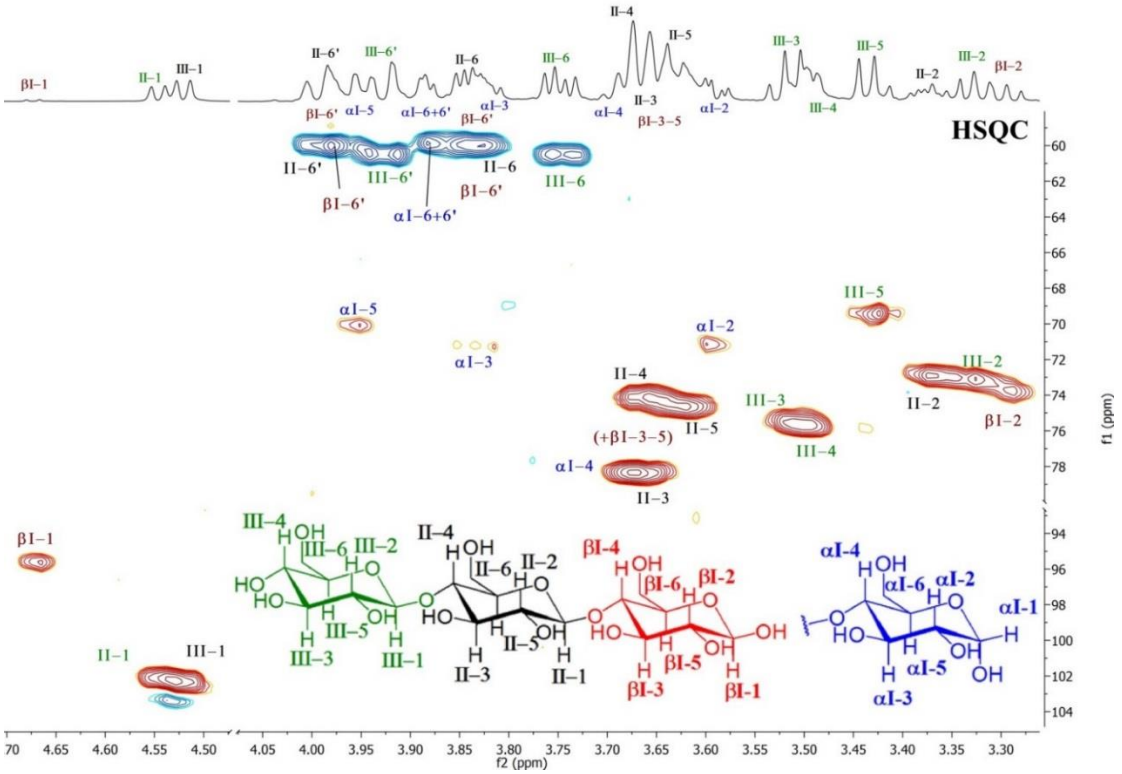
Supplementary Figure 27 |{^{1}H-^{13}C}-HSQC spectrum recorded at 600 MHz of D-cellobiose (**1**, n = 1) in D_2O, together with an assignment of individual resonances.



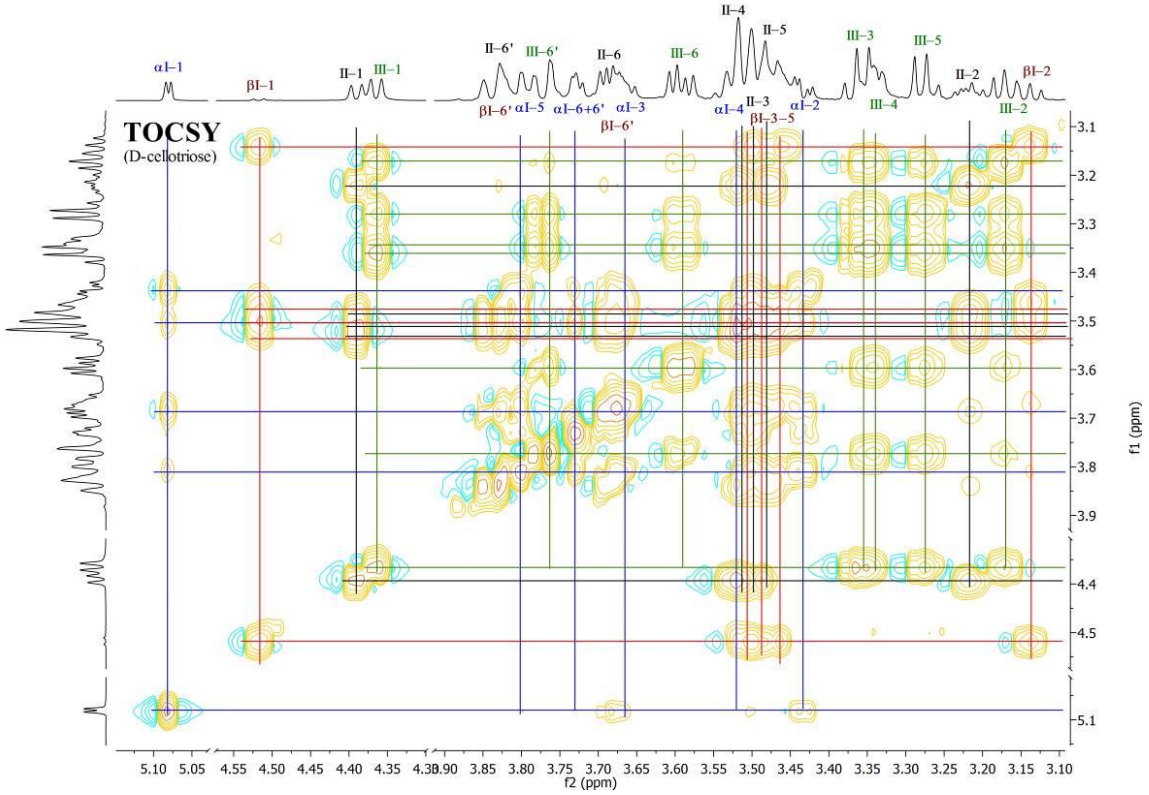
Supplementary Figure 28 |{^{1}H-^{1}H}-COSY spectrum recorded at 600 MHz of D-cellobiose (**1**, n = 1) in D_2O, together with an assignment of individual resonances.



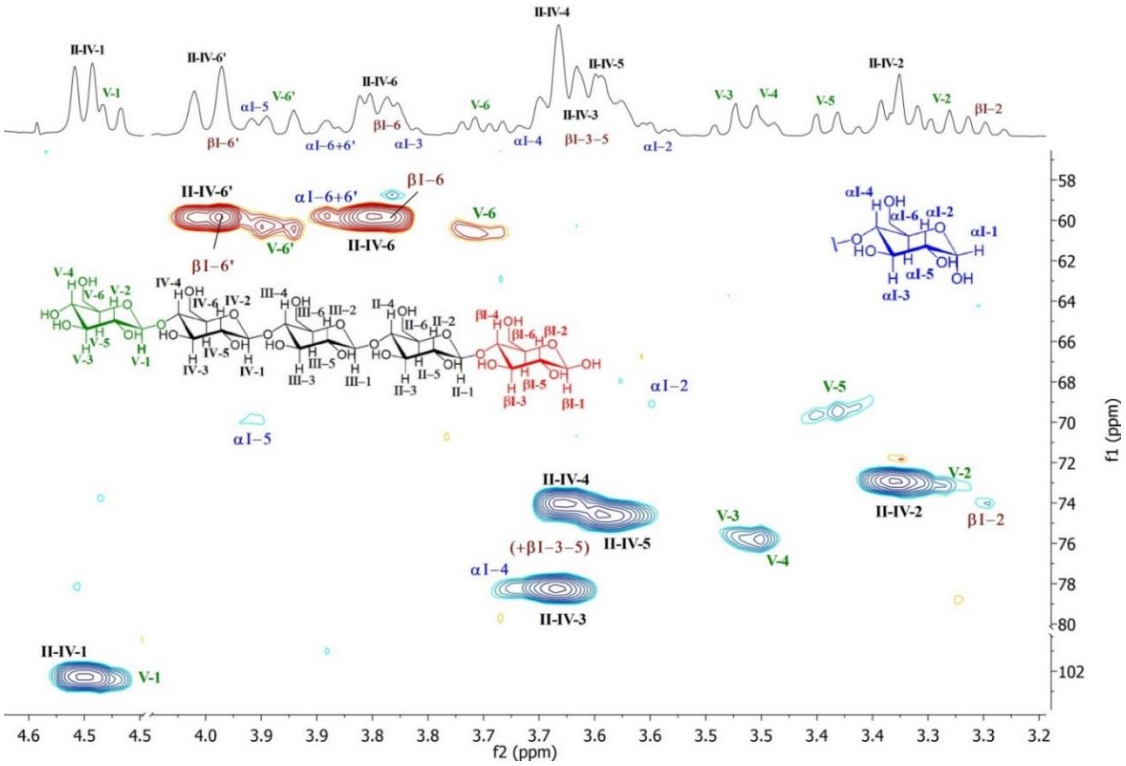
Supplementary Figure 29 | {^{1}H-^{1}H}-TOCSY spectrum recorded at 600 MHz of D-cellobiose (**1**, n = 1) in D_2O, together with an assignment of individual resonances. The colored horizontal and vertical lines are guides for identifying cross peaks (blue = *α*I, red = *β*I, green = II).



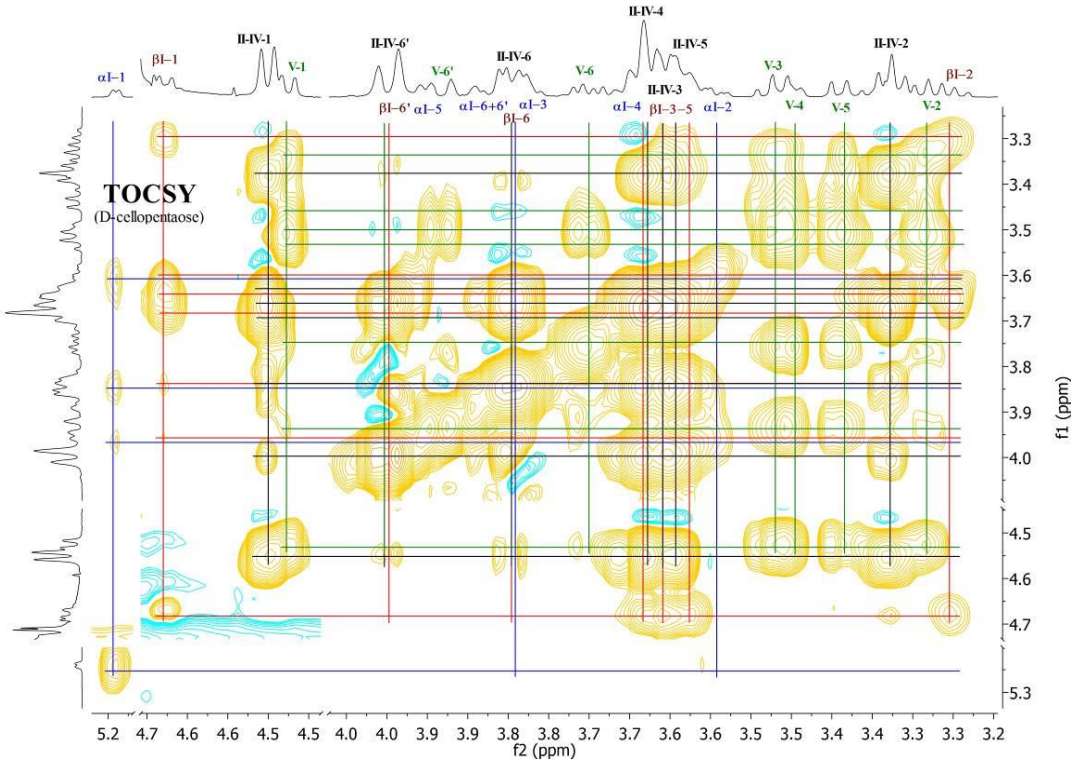
Supplementary Figure 30 |{^{1}H-^{13}C}-HSQC spectrum recorded at 600 MHz of D-cellotriose (**1**, n = 2) in D_2O, together with an assignment of individual resonances.



Supplementary Figure 31 |{^{1}H-^{1}H}-TOCSY spectrum recorded at 600 MHz of D-cellotriose (**1**, n = 2) in D_2O, together with an assignment of individual resonances. The colored horizontal and vertical lines are guides for identifying cross peaks (blue = *α*I, red = *β*I, black = II, green = III).



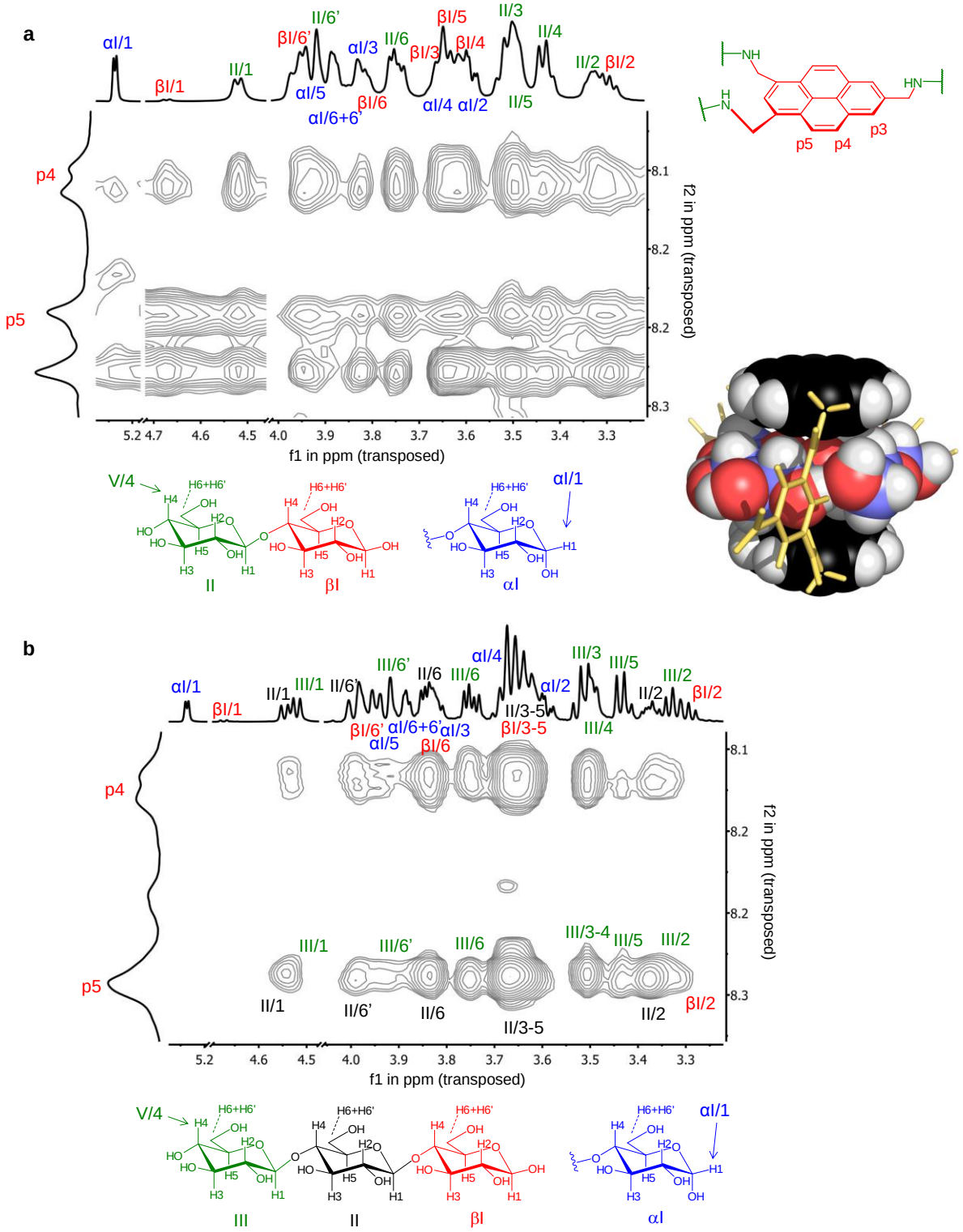
Supplementary Figure 32 | {^{1}H-^{13}C}-HSQC spectrum recorded at 600 MHz of D-cellopentaose (**1**, n = 4) in D_2O, together with an assignment of individual resonances.



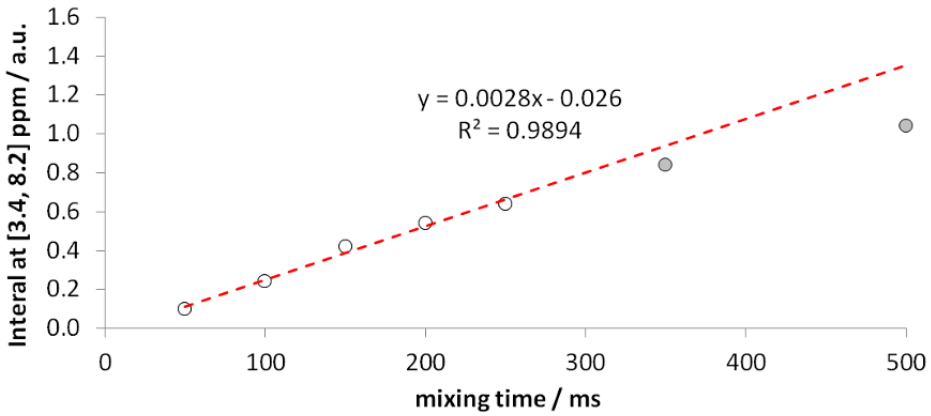
Supplementary Figure 33 | ^{1}H NMR assignment of cellopentaose 1 (n = 4). {^{1}H-^{1}H}-TOCSY spectrum of cellopentaose in D_2O, recorded at 600 MHz. The connections in this spectrum were used to assign the resonances as indicated. The coloured horizontal and vertical lines are guides for identifying cross peaks (blue = *α*I, red = *β*I, black = II-IV, green = V).

3.2. {^{1}H-^{1}H}-NOESY studies of [receptor 4···carbohydrate] complexes in D_2O

Nuclear Overhauser effect spectroscopy (NOESY) was employed to elucidate the structures of the complexes between receptor **4** and the cellodextrins cellobiose (**1**, n = 1), cellotriose (**1**, n = 2) and cellopentaose (**1**, n = 4). NOESY spectra (recorded at 600 MHz in D_2O with a mixing time of 150 ms) were acquired of **4** in the presence of excess cellobiose, cellotriose and cellopentaose as shown in Figure 2b and Supplementary Figure 34. f1 and f2 were transposed to enable viewing the carbohydrate region on the wider axis while displaying the highest resolution NOEs. The assignment of p4 and p5 was inferred from following the movement of p4/p5 in the corresponding titration experiment and the strong NOESY cross-peak present between p4 and p5.

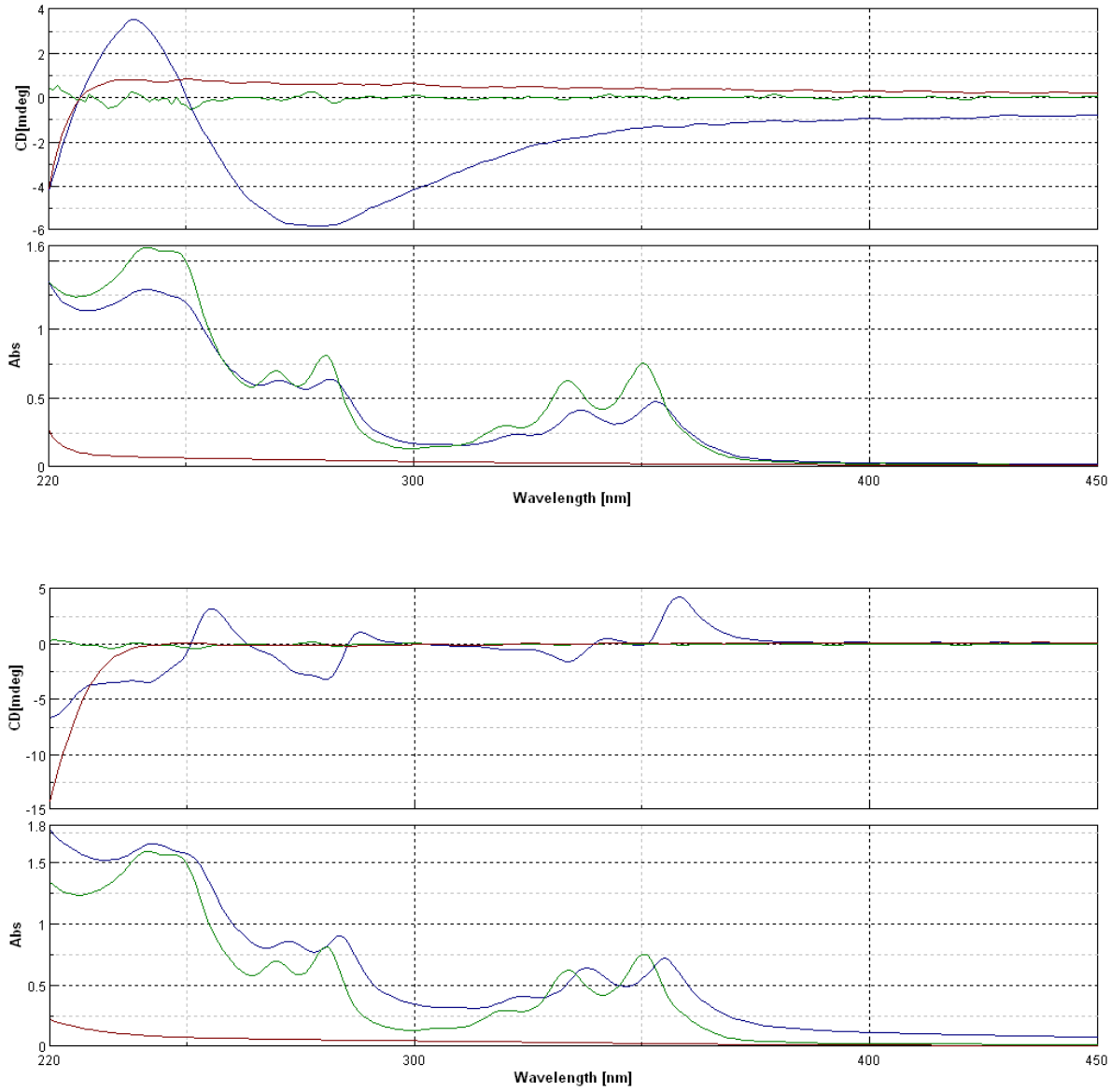


Supplementary Figure 34 | NOESY spectra of 4 with cellobiose and cellotriose. **a,** Partial NOESY spectrum (D_2O, 600 MHz) of **4** (0.25 mM) + D-cellobiose **1** (n = 4) (24 mM), with assignments. Mixing time (t_m) = 150 ms. Cross peaks are observed from receptor p4/p5 to all carbohydrate CH's. Inset: Computational model of the complex, consistent with the spectrum. Colour code as for Fig. 2c, except that side-chains are omitted for clarity. **b,** Corresponding spectrum for **4** (0.15 mM)+ cellotriose (24 mM). Connections to the reducing terminus (I) appear weak, but cross-peaks to the two non-reducing residues (II and III) are strong. Taken together, the spectra in **a** and **b** imply (i) that residues within the binding site exhibit strong cross-peaks to p4/p5, and (ii) that the reducing terminus will avoid the binding site if possible.



Supplementary Figure 35 | Integral of NOE cross peak at [3.35, 8.28] p.p.m. as a function of {^{1}H-^{1}H}-NOESY mixing time (t_m, at 600 MHz) of D-cellopentaose (**1**, n = 4; 4.7 mM) and receptor **4** (0.25 mM) in D_2O. The dashed red trend line reveals a linear relationship (R^2 = 0.9894) between t_m = 50–250 ms (open circles). The integral values of t_m = 350 and 500 ms (grey circles) were not included in the trend line.

4. Induced Circular Dichroism (ICD) studies



Supplementary Figure 36 | CD spectroscopy of receptor 4 + polysaccharides. Top: Receptor **4** (0.25 mM) + cellulose (25 mM glucosyl), alone and mixed, in 1 M NaOD. Bottom: Receptor **4** (0.25 mM) + chitosan **3** (n = 375-750) (25 mM glucosaminyl), alone and mixed in D_2O at pH = 6.4. In each case the top panel shows the CD spectra and the bottom panel the UV/visible absorbance spectra. Green = receptor alone, brown = polysaccharide alone, blue = receptor + polysaccharide.

5. Atomic Force Microscopy (AFM) studies

5.1. Equipment and sample preparation

Equipment: Atomic force microscopic analysis was performed using a Multi-mode VIII (Bruker, CA, USA) atomic force microscope in ambient conditions. Non-resonant PeakForce control was utilised to maintain a constant level of force interaction between the cantilever tip and the sample, providing accurate height values for measured sample topography. Cantilevers used had nominal values of 2 nm tip radius and 0.4 N/m spring constant (SCANASYST-AIR-HR, Bruker, CA, USA) enabling low force interaction in both fast scanning, using Bruker's fast scan HR unit, and in slower high resolution imaging.

Cellulose samples: Cellulose (Sigmacell 101, 2.027 mg) was suspended in a solution of receptor **4** (0.25 mM, 0.5 mL) in aq. NaOH (1 M), then freeze-thawed three times with liquid nitrogen and warm water.[8] The resulting clear solution thus contained 0.25 mM receptor **4** and 25 mM of glucosyl monomer (contained within cellulose chains). This solution was diluted $2.5 \cdot 10^5$ times with HPLC-grade water to give [**4**] = 1 nM and [monomer] = 100 nM. About 10 μL of this solution was deposited on freshly cleaved atomically flat Mica ($H_2KAl_3(SiO_4)_3$, 5 × 5 mm). While holding the mica at an angle of about 45°, HPLC-grade water (2 × 1 mL) was passed over the surface. The surface was allowed to dry before analysing with atomic force microscopy. A control sample containing only cellulose was obtained using the same procedure.

Chitosan samples: Chitosan (Sigma-Aldrich, M_v 60,000-120,000, [β-(1→4)-D-glucosamine]$_{375\text{-}750}$, 2.015 mg) was suspended in a solution of receptor **4** (0.25 mM, 0.5 mL) which was then acidified to pH = 1.0 with concentrated HCl (followed with pH probe). The resulting clear solution thus contained 0.25 mM receptor **4** and 25 mM of glucosaminyl monomer. The solution was diluted $2.5 \cdot 10^7$ times with HPLC-grade water to give [**4**] = 0.01 nM and [glucosaminyl monomer] = 1 nM. About 10 μL of this solution was deposited on freshly cleaved atomically flat mica ($H_2KAl_3(SiO_4)_3$, 5 × 5 mm) and allowed to dry before analysing with atomic force microscopy. Control samples containing only D-chitosan and only receptor **4** were obtained using the same procedure.

[8] Isogai, A. & Atalla, R. H. Dissolution of cellulose in aqueous NaOH solutions. *Cellulose* **5**, 309-319 (1998).

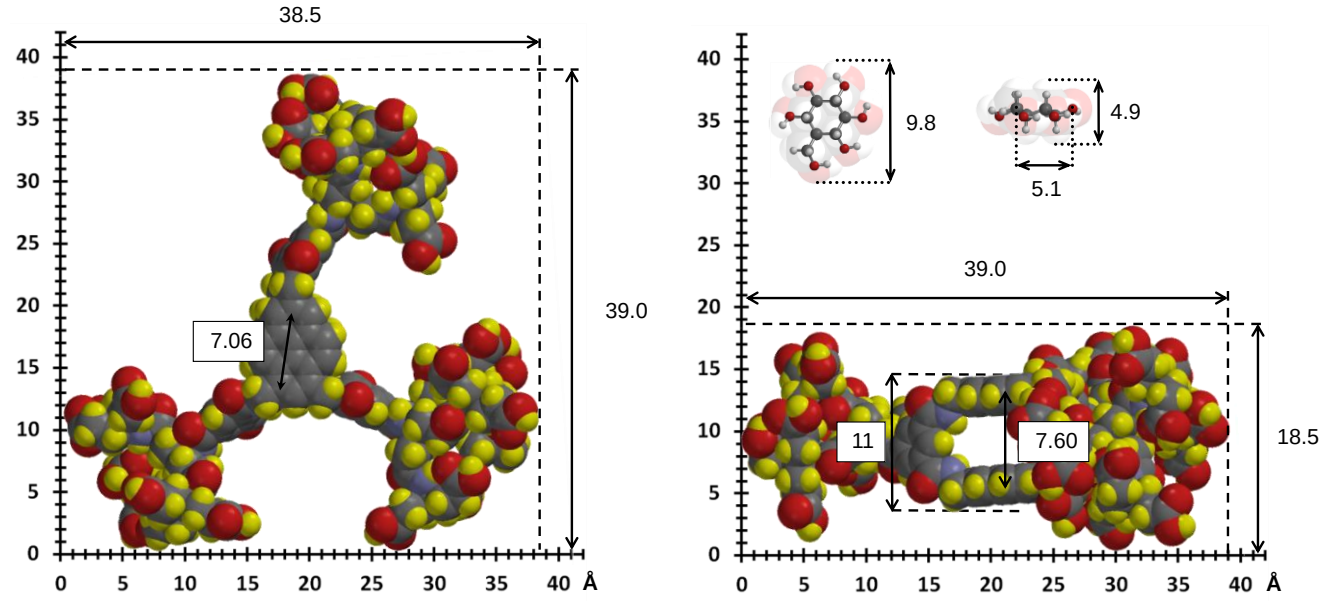
5.2. Approximation of cellulose, chitosan, and receptor dimensions

As illustrated in Supplementary Figure 37 (top right), the height, width, and length of a cellulose monomer can be approximated by considering the dimensions of a three dimensional *β*-glucose model. The dimensions shown are derived from interatomic distances and the van der Waals radii of H (1.09) and O (1.55). The volume of a model of $(Glc)_{50}$ was estimated as 6813 Å^3, meaning that the volume of a single glucose unit within a cellulose strand is about 136.26 Å^3. As glucose and glucosamine are similar in size, these measurements can be applied to both cellulose and chitosan monomers.

A three dimensional computer model of the receptor was subjected to a conformational analysis using molecular mechanics (Spartan 2007, MMFFaq force field). In Supplementary Figure 37 are shown top and side-views of the computed ground state conformation in space filling mode. Using easily measurable interatomic distances of the pyrene cage (indicated in white boxes), an approximate scale was drawn with graphical software, which in turn allowed easy estimation of the maximal length, width, and height of the receptor. The minimal height was estimated assuming the pyrene core is rigid and open by measuring the pyrene-pyrene distance and adding 2 × the van der Waals radius of carbon (1.70 Å).

The above approach thus led to the following dimensions:

	$[Glc]_n/[GlcNH_2]_n$	receptor
Height:	5 Å	11 Å (core), 18.5 Å (side-chain)
Width:	10 Å	38.5 Å
Length:	5 × *n* Å	39 Å

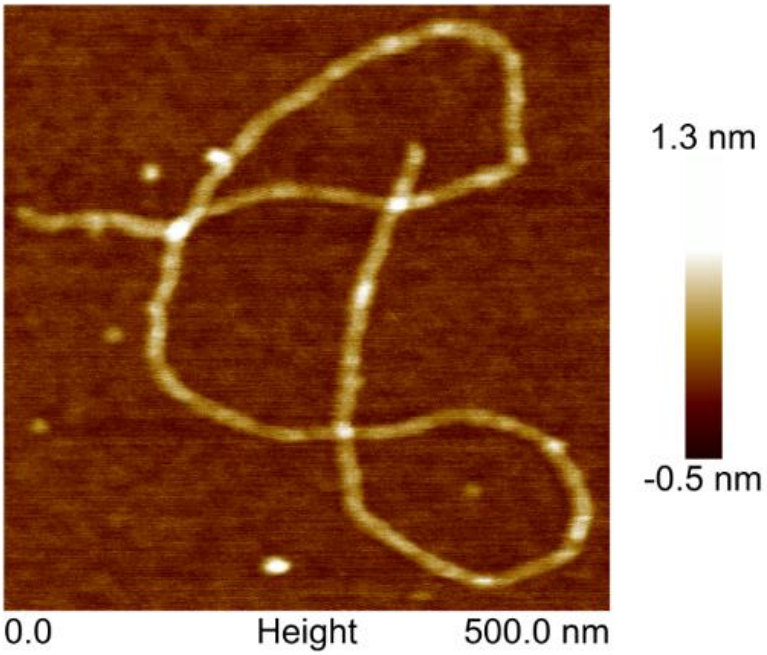


Supplementary Figure 37 | AFM measurements on 4 + cellulose. Estimated dimensions (Å) of **4** and glucose monomer, based on interatomic distances with addition of van der Waals atomic radii.

5.3. Analysis of AFM data

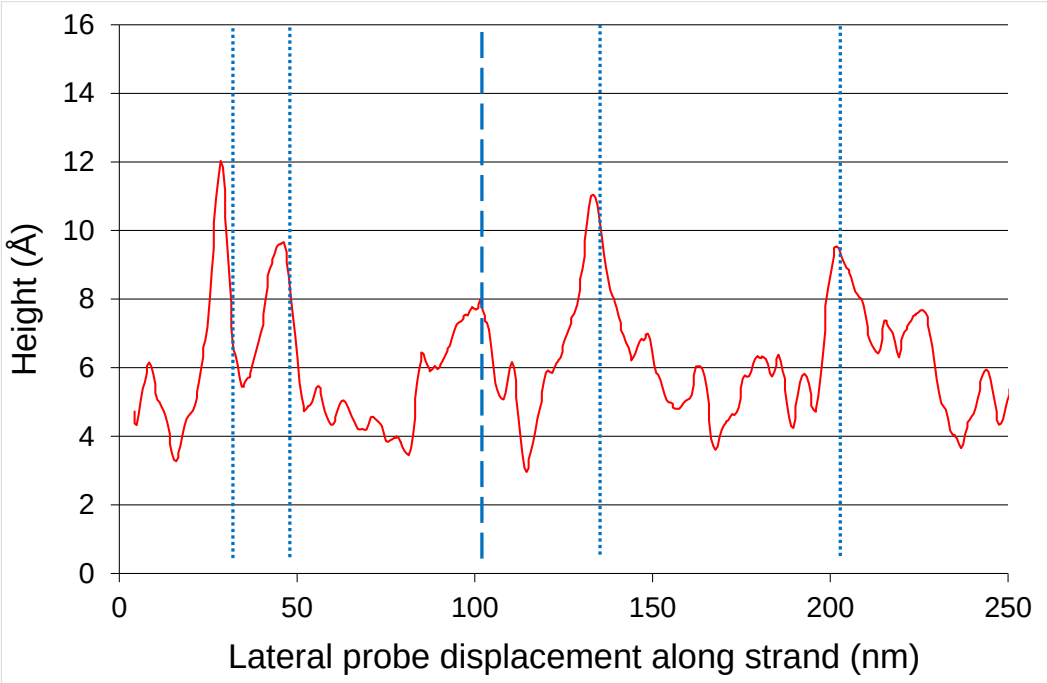
5.3.1. Cellulose + receptor **4**

Images were recorded of two samples of cellulose·**4**, prepared separately, as well as the control sample prepared from cellulose without receptor. An image from cellulose·**4** 'sample 1' is shown in Figure 4a. An image from cellulose·**4** 'sample 2' is shown in Supplementary Figure 38 and the control sample is shown in Figure 4b.

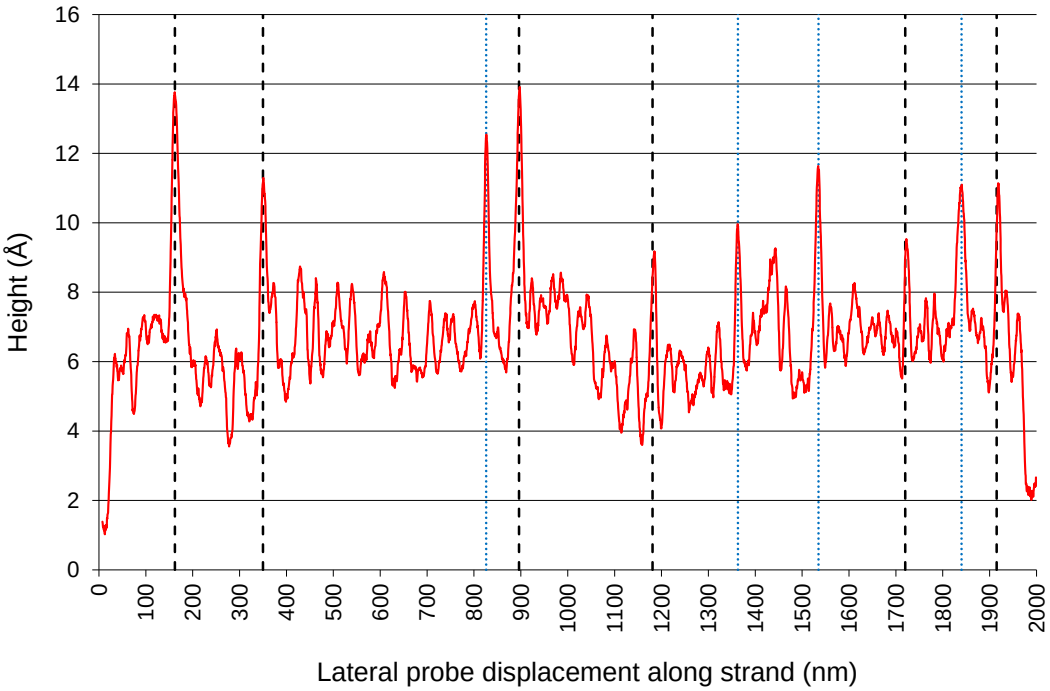


Supplementary Figure 38 | AFM picture of cellulose·**4** 'sample 2', deposited from diluted aqueous NaOH solutions, following the procedure discussed in section 5.1 above.

Both images of cellulose·**4** appear to show long strands, consistent with cellulose molecules, but with variable height profiles suggestive of threaded receptor molecules. A plots of the height-profile as a function of probe displacement along the polymer for a portion of 'sample 1' (Fig. 4a) is shown in Supplementary Figure 39. A similar plot for 'sample 2' is shoen in Supplementary Figure 40. In both plots, proposed positions for threaded receptors have been marked with blue dotted lines. The dashed blue line represents what might also be a receptor. Points where the polymer crosses itself have been marked with black dashed lines.



Supplementary Figure 39 | Height profile along polymer of cellulose·**4** 'sample 1' (red solid line). The profile is restricted to the section magnified in Fig. 4a. Proposed positions for threaded receptors have been marked with blue dotted lines. The dashed blue line represents what might also be a receptor.



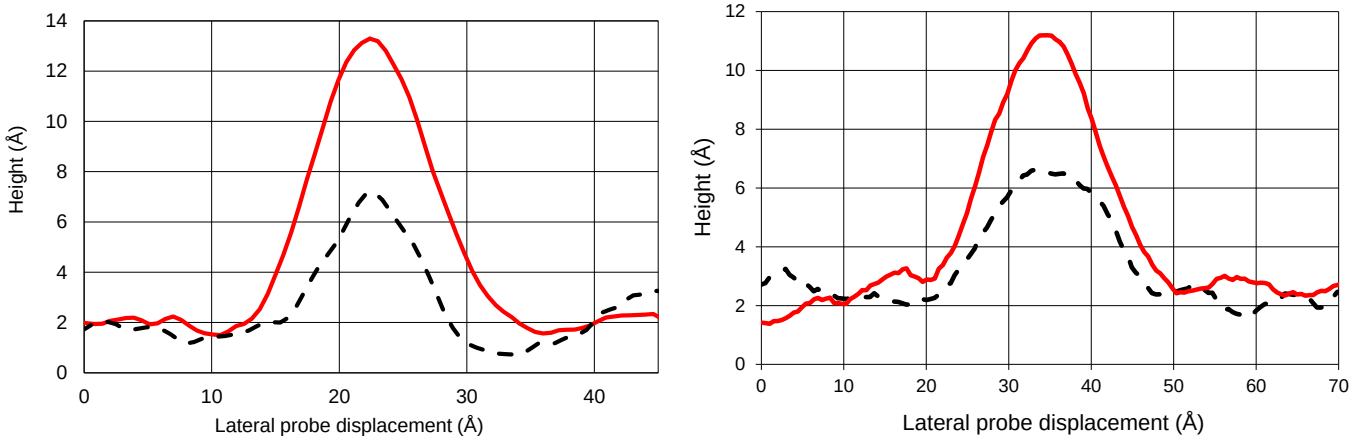
Supplementary Figure 40 | Height profile along polymer of cellulose·**4** sample 2 (red solid line). Proposed positions for threaded receptors have been marked with blue dotted lines. Points where the polymer crosses itself have been marked with black dashed lines.

Although height profiles measured by AFM are relatively reliable, resolution in the plane of measurement is impaired by tip-broadening effects. To obtain information about the width of the protrusions found on the polymer strands, it was necessary to correct for these. As the width of the polymer (W_P) is approximately known, the

measured polymer width (W_{Pm}) was used to derive the protrusion width (W_R) from the measured width (W_{Rm}) according to:

$$W_R = \left(\frac{W_{Rm}}{4 \times \frac{W_{Pm}}{4 \times \sqrt{W_P}}} \right)^2$$

For this purpose, cross-sections of both the bare polymer chains and the individual protrusions on the polymer chains were measured in various places. The averaged cross-sections of polymer (black dashed lines) and protrusions on polymer (red solid lines) are shown in Supplementary Figure 41 for 'sample 1' (left) 'sample 2' (right).



Supplementary Figure 41 | Averaged cross sections of cellulose·**4** sample 1 (left) and sample 2 (right), showing the measures height and width (not corrected for tip broadening) of the cellulose strands (black dashed lines) and the protrusions on the cellulose strands (red solid lines).

These data were used to compute the height and the width of the cellulose chains and protrusions as found on the polymer chains. The results are summarised in Supplementary Table 1 below.

Supplementary Table 1 | Overview of calculated and measured dimensions of cellulose polymer and receptor molecules (on the polymer).

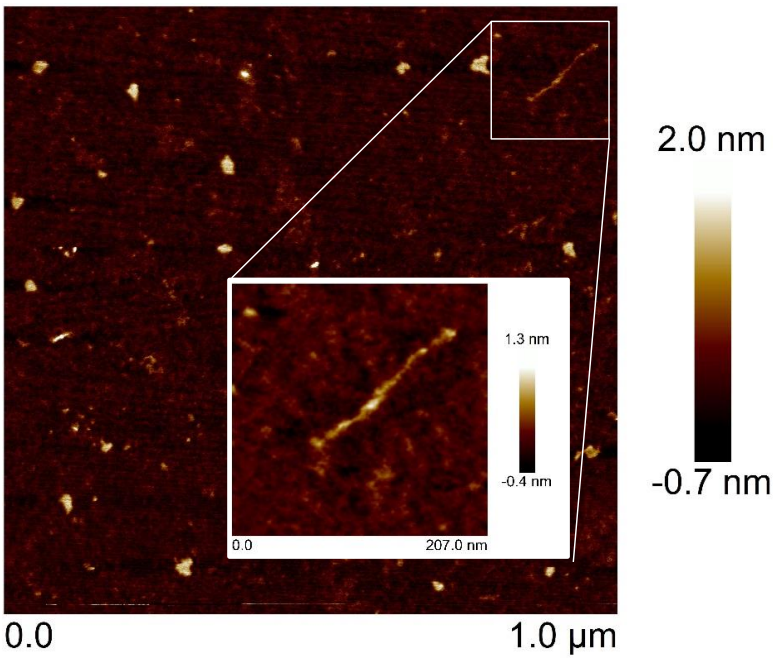
Description	**Calculated**	**Measured**	
		cellulose·4 sample 1	**cellulose·4 sample 2**
Background roughness	–	0.81 Å	0.98 Å
Height of cellulose molecules	5 Å	4.7 ± 1.0 Å	4.6 ± 1.0 Å
Width of cellulose molecules	10 Å	10 ± 4.2 Å	11 ± 4.0 Å
Length of cellulose molecules		~ 22,100 Å	~ 20,000 Å
n in [monomer]$_n$		~ 4,333	~ 4,000
Height of receptor threaded on cellulose	**11-18.5 Å**	**11.4 ± 1.7 Å**	**9.2 ± 1.0 Å**
Height of receptor overlaying cellulose	**16-23.5 Å**	" "	" "
Width of threaded receptor	38.5 Å (max.)	29 ± 9 Å	36 ± 4 Å

The measured dimensions of both cellulose strands correspond well with those estimated by molecular modelling. Also, the chain lengths are within reasonable bounds[9] and very similar for both samples. Importantly, the heights measured for protrusions along the polymer strands (9.2 – 11.4 ± ~1.5 Å) are consistent with the minimum value expected for the threaded receptor molecules (i.e. the height of the central core, 11 Å; see Supplementary Table 1). This implies that the cellulose is threaded through the cavity of **4**. In the alternative mode of binding, where the receptor molecule lies on top of the chain, the calculated minimum height is 11 + 5 = 16 Å which is inconsistent with the data. The measured widths of protrusions (29 – 36 Å; obtained after correction for tip-broadening effects) of the receptor are also consistent with the estimated values for the receptor. These measured widths might seem a little low, but one must bear in mind that the dendrimeric side chains are flexible; they are thus not likely to be stretched to their full length under solvent free conditions, and may also retreat (to some degree) from the AFM probe.

[9] Klemm, D., Heublein, B., Fink, H. P. & Bohn, A. Cellulose: Fascinating biopolymer and sustainable raw material. *Angew. Chem., Int. Ed.* **44**, 3358-3393 (2005).

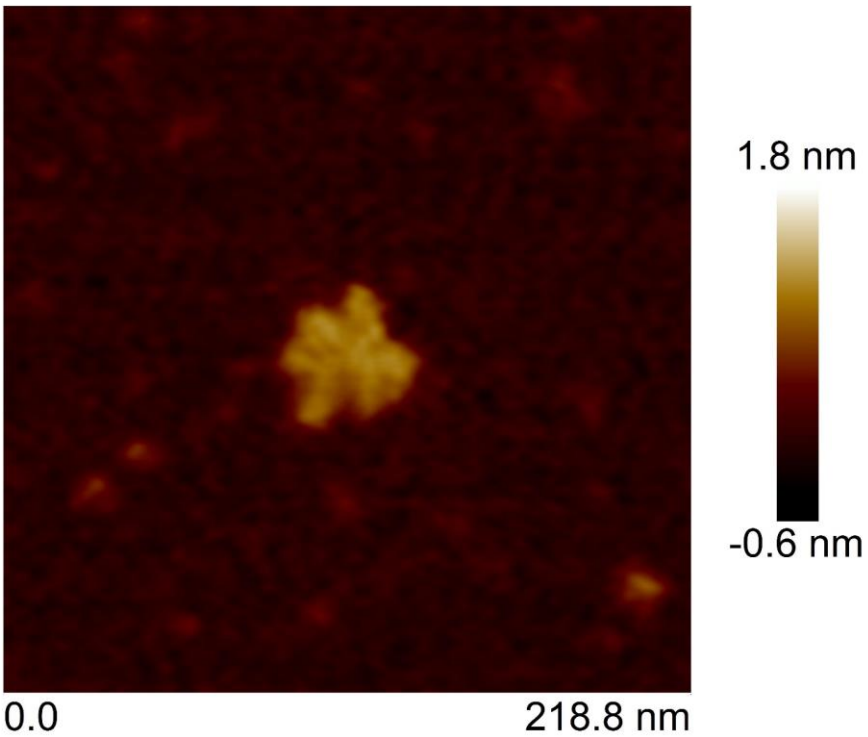
5.3.2. Chitosan + receptor **4**

An image of chitosan deposited on mica from an aqueous solution diluted from pH = 1 (control sample) is shown in Supplementary Figure 42. A selected region is also shown in Figure 5c. The images reveal that individual strands are typically coiled, and possibly aggregated. One strand in an extended conformation is also observable, as highlighted in Supplementary Figure 42. The measured length of this strand observed (~150 nm) is close to the expected range of 188 – 375 nm (0.5 nm × *n* for [*β*-(1→4)-D-glucosamine]$_n$, where *n* = 375-750). The coiling/aggregation is unsurprising as the HCl added to protonate the amino groups in chitosan might be partially lost during dilution and evaporation of the sample. This should resulting in an amino group/HCl ratio of less than unity, leading to the formation of -NH_3^+···H_2N- hydrogen bonds and thus coiled conformations and/or aggregates. The extended strand highlighted in the inset might represent an exceptional case where all amino groups remain protonated.



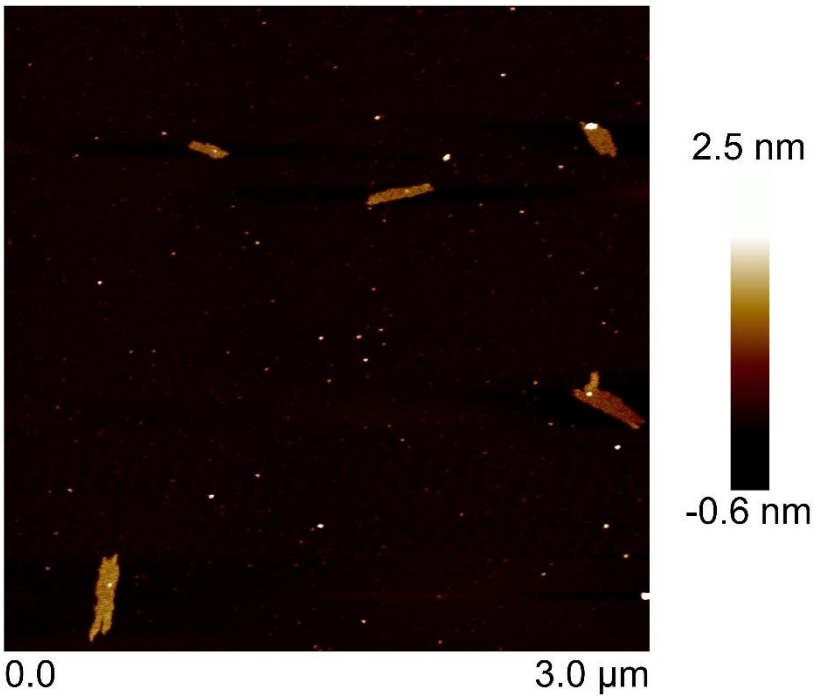
Supplementary Figure 42 | AFM image of chitosan deposited from acid. Chitosan deposited from aqueous acid, following the procedure used for **4** + chitosan. In the absence of receptor, the polysaccharide seems to form mostly coiled structures, although an isolated linear strand can also be observed. The differences may reflect differing protonation states.

As described earlier, a control sample of receptor **4** was prepared similarly. As illustrated in Supplementary Figure 43 (see also Figure 5d), AFM images showed irregular objects of various shapes and sizes, up to ~40 nm wide and 8 nm in height. These were presumed to be aggregates of **4** held together by hydrogen bonding between carboxyl groups. The shapes did not possess the uniform height, elongated morphology and/or fine structure observed for chitosan·**4** as discussed below.



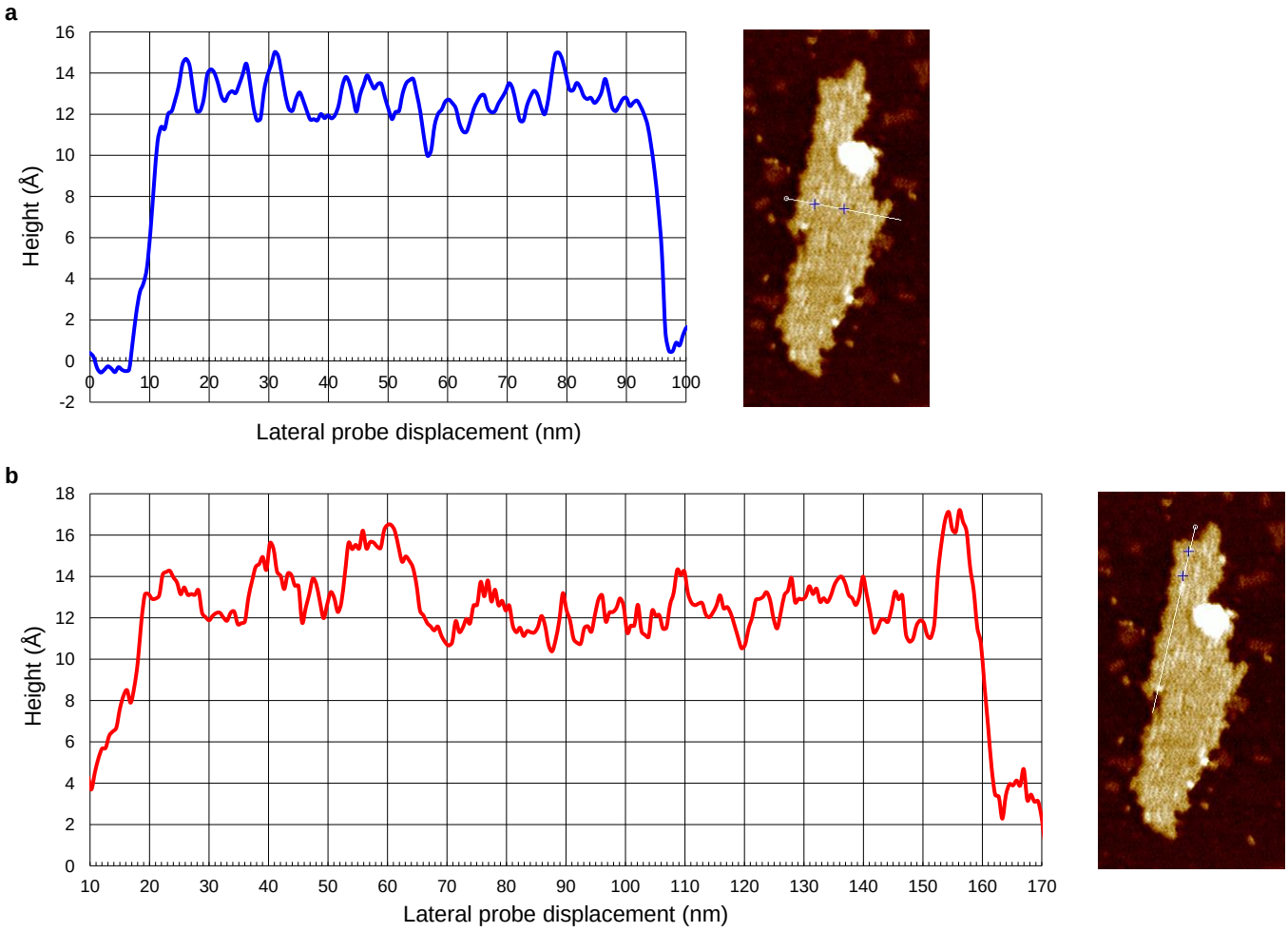
Supplementary Figure 43 | AFM image of receptor 4 deposited from acid. Receptor **4** deposited from aqueous acid, following the procedure used for **4** + chitosan. Most aggregates observed were several molecules thick (up to 8 nm in height). Even the few which could be monolayers (as here) lack the organised structures shown in Fig. 5a and Supplementary Figure 44.

Images of the chitosan·**4** sample are shown in Figure 5a and Supplementary Figure 44. Aggregation of the chitosan·**4** sample was found to be extensive; no individual strands could be found after multiple attempts. The morphology of the aggregates was rod-like with lengths in the range 200 – 500 nm, close to the estimated length of the polymer (188 – 375 nm, see above). The height of these features (see below for details) was typically ~ 13-14 Å. As with the cellulose·**4** sample, this height is only consistent with the height of threaded receptor molecules, i.e. 11 – 18.5 Å, and not consistent with a scenario in which chitosan and receptor molecules are stacked, i.e. 16 – 23.5 Å. This height is also inconsistent with aggregates of only chitosan, which should be a multiple of 5 Å (the height of chitosan in 5 Å, the width 10 Å).



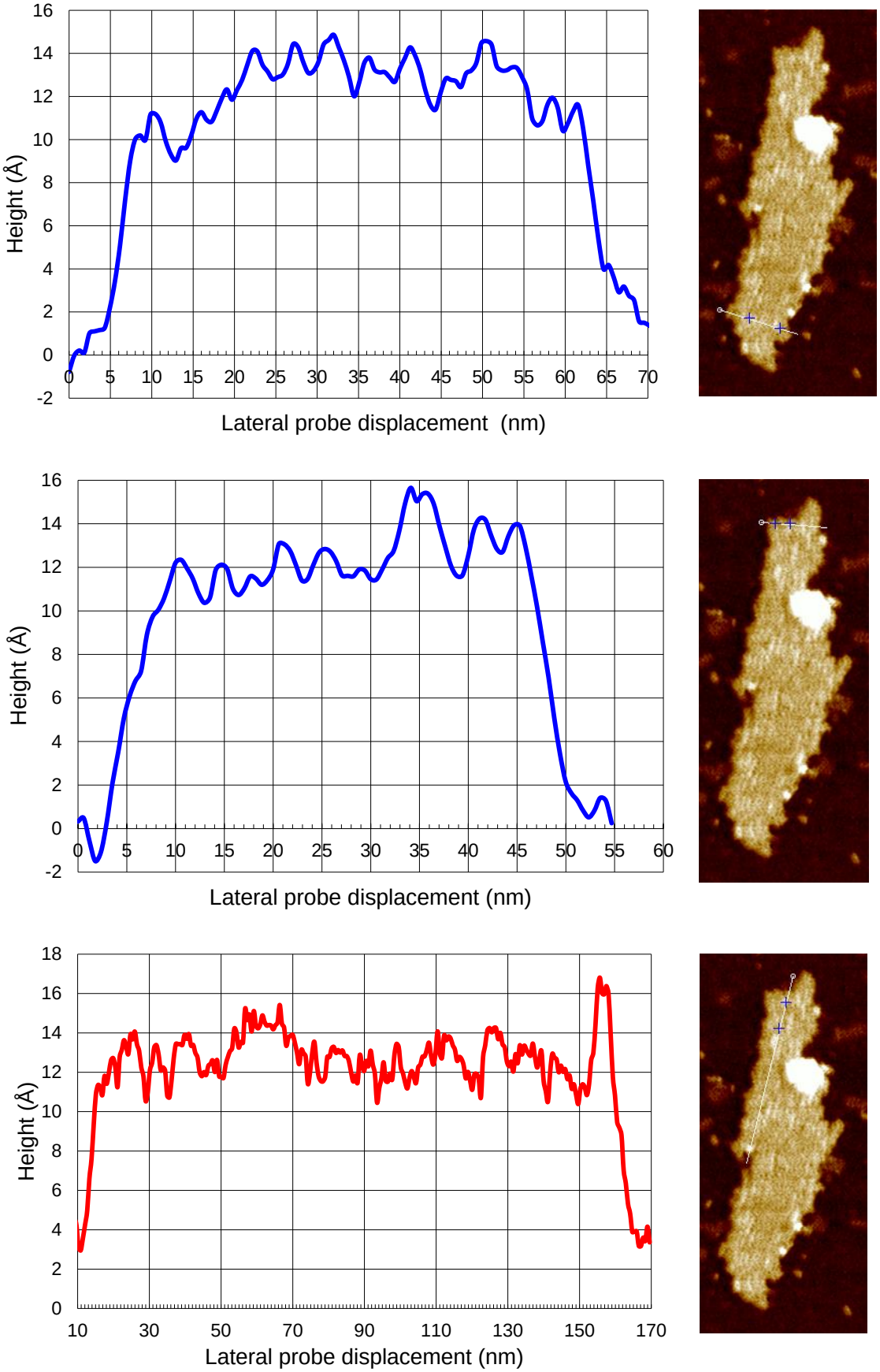
Supplementary Figure 44 | AFM image of 4 + chitosan. The image was obtained in the same way as that in Fig. 5a, except that a less sharp AFM tip was employed.

The aggregate recorded at maximal resolution (Fig. 5a) was subjected to a detailed analysis. Height profiles measured along the short axis (blue line) and along the long axis (red line) of the aggregate are shown on the same scale for direct comparison in Supplementary Figure 45 (see also Fig. 5b). Further examples are given in Supplementary Figure 46.[10]



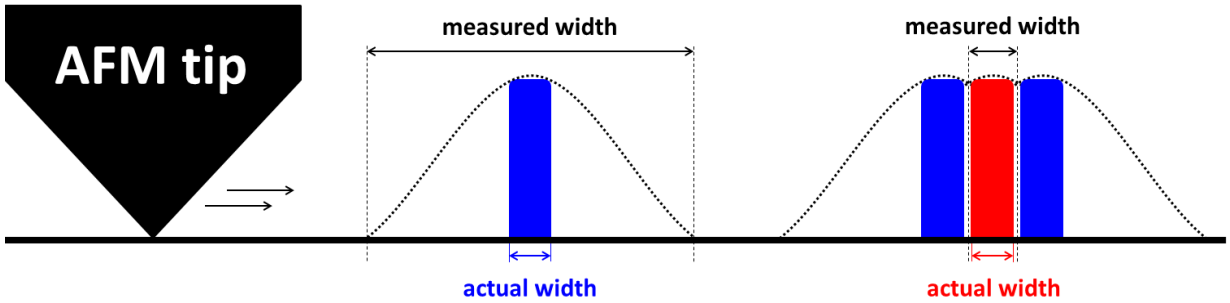
Supplementary Figure 45 | Height profiles for receptor 4 + chitosan (high resolution image, Fig. 5a). **a,** Height profile taken across the short axis of the aggregate, along the line indicated in the image. Variations are seen to be fairly regular and consistent with long-range order, as expected for molecular chains ~3.4 nm in width lying parallel to each other. The same periodicity was observed throughout the aggregate. The average height of 13 Å is consistent with threaded receptor **4**. **b,** Height profile taken along the long axis of the aggregate, along the line indicated in the image. Variations are less regular as expected for polypseudorotaxane chains in which receptor orientations and occupation levels may vary. The length is consistent with a chitosan chain containing 280 monomers. The profiles in **a** and **b** are best explained by a monolayer of polypseudorotaxane chains lying along the long axis of the aggregate, held together by hydrogen bonds between receptor carboxylic acid substituents.

[10] The measurements taken along the long axis of the aggregate all involve the somewhat smaller top left part of the aggregate, rather than the larger bottom right part. This was done because the section at the bottom right showed variation over longer horizontal distances. This “waviness” would have compromised attempts to obtain an accurate representation of the aggregate's fine-structure.



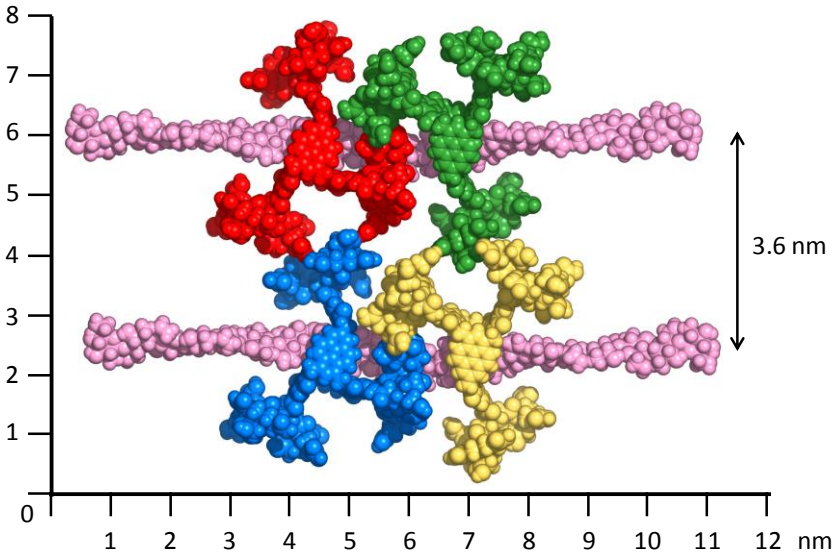
Supplementary Figure 46 | Additional height profiles taken along the short axis (blue, top and middle) and long axis (red, bottom) of the chitosan·**4** aggregate. Details and commentry as for Supplementary Figure 45.

It should be noted that tip-broadening has little effect when measuring the fine-structure of the aggregate, except at the periphery. The reason is illustrated in Supplementary Figure 47. Tip broadening is sensitive to the height of the object that is being measured. When measuring a feature from valley-to-valley, the tip only dips in slightly (about 2 nm) and does not reach all the way back to the mica surface (about 13-14 nm). Moreover, due to the tight packing, tip-broadening effects are largely cancelled out due to overlapping broadening of neighbouring features.



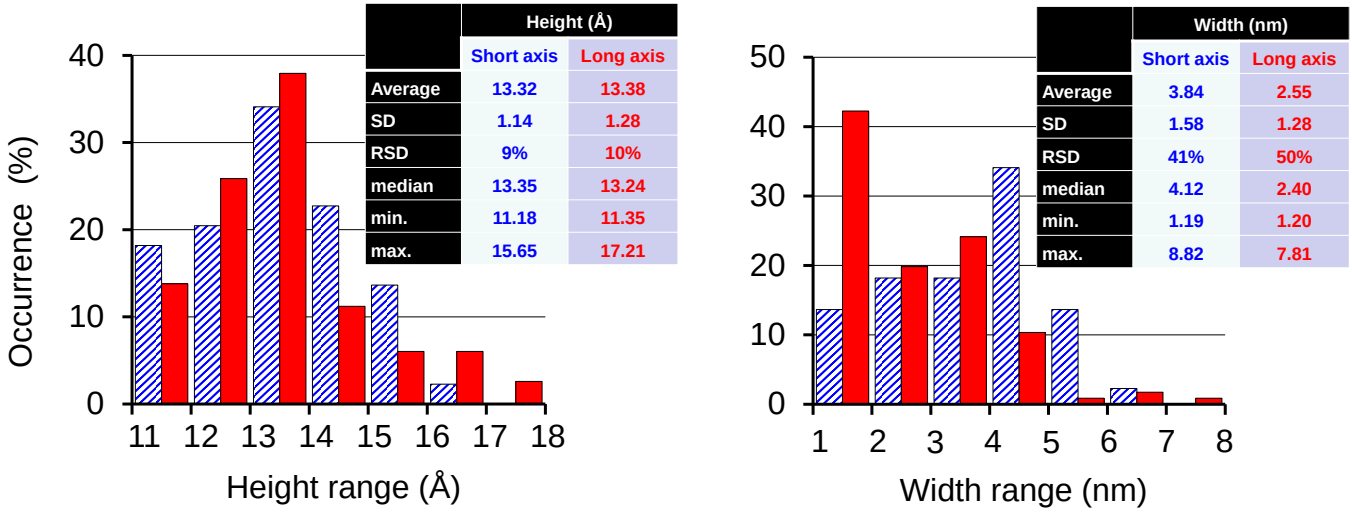
Supplementary Figure 47 | Illustration of why the measured dimensions of the fine-structure of the aggregate approximate the actual dimensions; the magnitude of tip-broadening effects depend on the height of a feature and are largely cancelled out when neighbouring features are packed closely.

From these measurements it is clear that the aggregate has a fine-structure with fairly constant height (~12-14 Å) consistent with that expected for threaded receptor molecules. It is also notable that the profiles across the short axis show a degree of regularity, implying a structure which is not perfectly organised but with underlying periodicity. The pattern is strongly suggestive of parallel chains oriented along the long axis of the aggregate. The average spacing between the peaks varies between 3.4 and 3.8 nm within the three trajectories measured. Modelling suggests that this distance is nicely consistent with parallel chains of polyrotaxanes. Under the acidic conditions used for deposition, the receptor side-chains should be protonated. This should allow the receptors to pack quite densely on the chitosan, assisted by hydrogen bonding between adjacent dendrimers. The chains can then align to form aggregates, again held together by hydrogen bonding between carboxyl groups (well-known to act as both hydrogen bond donors and acceptors). Supplementary Figure 48 shows a magnification of the model used in Figure 5a. Carboxyl and amino groups are assumed to be fully protonated (i.e. $-CO_2H$, $-NH_3^+$). The structure was built and minimised using MacroModel 10.3 (Maestro 9.7 interface) with the Merck Molecular Force Field (static) and Generalised Born/Surface Area (water) continuum solvation. The image thus represents an energy minimisation of a small fragment of the proposed polyrotaxane structure. This Model predicts a chitosan-chitosan repeat distance of ~3.6 nm, within the range measured.



Supplementary Figure 48 | Energy-minimised structure of four molecules of receptor **4** (red, green, blue, yellow) on two adjacent chitosan chains (pink; [*β*-(1→4)-D-glucosamine]$_{20}$) (see also Figure 5a). Carboxyl and amino groups are assumed to be fully protonated (i.e. -CO_2H, -NH_3^+). The structure was built and minimised using MacroModel 10.3 (Maestro 9.7 interface) with the Merck Molecular Force Field (static) and Generalized Born/Surface Area (water) continuum solvation.

Numerical analysis of the height profiles provides further support for the model shown above. The data are summarised in Supplementary Figure 49 for measurements taken along the short axis (striped blue bars, *N* = 45) and along the long axis (solid red bars, *N* = 116) of the chitosan·**4** aggregate. As shown on the left-hand side, the average height of features along both the short and long axes of the aggregate is ~ 13.3 (± 1.2) Å, consistent with a threaded mode of binding. Across the short axis, the most common width is 4-5 nm (right-hand side, blue columns), consistent with the span of a single receptor molecule. Along the long axis of the aggregate, the most common width is 1-2 Å (right-hand side, red columns). This is consistent with a pyrene unit overlaying a chitosan chain, or with an individual side-chain.



Supplementary Figure 49 | Bar diagrams and numerical data for measurements taken along the short axis (striped blue bars, *N* = 45) and along the long axis (solid red bars, *N* = 116) of the chitosan·**4** aggregate (see Figure 5a). The left-hand diagram shown the distribution of the height of the fine-structure of the aggregate, while the right-hand diagram represents the distributions of the width of the aggregate's fine-structure. The trajectories used for these measurements are shown in Supplementary Figures 45 and 46.