Supporting information

Quantitative determination of the binding capabilities of individual large-ring cyclodextrins in complex mixtures

Dennis Larsen, Andreas Erichsen, Giorgia Masciotta, Sebastian Meier and Sophie R. Beeren*

Contents

S1 General experimental information 2
S1.1 Materials and chemicals 2
S1.2 Instruments 2
S2 Production and characterisation of LRCD mixture 3
S2.1 General procedure for production of LRCD mixture 3
S2.2 Characterization of LRCD mixture 3
S3 Assignment of LRCD anomeric peaks in HR-HSQC NMR spectra 6
S3.1 General procedure for isolation of individual LRCDs 6
S3.2 Acquisition of highly resolved ^{1}H-^{13}C HSQC NMR spectra 6
S3.3 Assignment of LRCD anomeric peaks in HR-HSQC NMR spectra 6
S4 Comparison of LRCD mixture to cycloamylose and linear α-1,4-glucans 8
S5 HSQC NMR spectroscopy titrations of LRCD mixture with KI_3 and SDS 9
S5.1 General procedure for HSQC titrations 9
S5.2 HSQC titration with KI_3 9
S5.3 HSQC titration with SDS 11

S1 General experimental information

S1.1 Materials and chemicals

Cycloamylose (Ezaki Glico Co., Ltd.), CD6–CD8, KI_3, SDS, NMR solvents, CGTase (Amano Enzymes, Inc.), α-amylase (Megazyme) and MALDI matrices were purchased from commercial suppliers, stored as specified by the supplier and used as received. HPLC or LCMS grade solvents from commercial suppliers were used for all chromatographic analyses. High purity water used for chromatography and to produce LRCD mixtures was obtained by filtering deionised water through a commercial water purifier (Merck Millipore Synergy UV). NMR samples were measured in standard 5 mm borosilicate glass NMR tubes with plastic caps.

S1.2 Instruments

HPLC chromatographic analysis was performed on a Thermo Scientific Dionex Ultimate 3000 HPLC system with a Waters Acquity UPLC BEH Amide column (2.1 × 150 mm, 1.7 µm particle size) maintained at 30 °C during all analyses. The system was connected to an Agilent 1260 Infinity ELSD operating at 90 °C and 70 °C for the evaporator and nebuliser temperatures, respectively, and an N_2 gas flow of 1.0 L/min. Gradient elution was applied using eluents with 0.1 % (vol.) formic acid. Typical gradients would go from 25 % water to 70 % water over the course of either 27 minutes with a flow rate of 0.4 mL/min for a slightly faster analysis, with some loss of peak separation, or over 36 minutes with a flow rate of 0.3 mL/min for better peak separation. The column was flushed with at least five column volumes of 100 % aqueous eluent, typically at a flow rate of 0.2 mL/min, after each run and equilibrated with at least five column volumes of the starting eluent mixture prior to each injection. 1D NMR spectra were acquired on a Bruker Avance III 400 MHz NMR spectrometer equipped with a Prodigy broadband observe (BBO) probe. HR-HSQC 2D NMR spectra were acquired on an 800 MHz Bruker Avance III instrument equipped with an 18.7 T Oxford magnet and a TCI cryoprobe. MALDI-TOF-MS analyses were performed on a Bruker Autoflex Speed instrument using 2,5-dihydroxybenzoic acid (DHB) as matrix on manufacturer-supplied ground steel target plates.

S2 Production and characterisation of LRCD mixture

S2.1 General procedure for production of LRCD mixture

Cycloamylose (1 g, Ezaki Glico Co., ltd. prod. no. 302-32406, avg. mol. weight 7300 g/mol) was dissolved in water (100 mL) into which was added CGTase stock solution (100 µL, *B. macerans*, used as supplied by Amano Enzyme, Inc). The resulting solution was stirred for 8 hours, after which it was heated to the boiling point where it was maintained for 15 minutes. The solution was allowed to cool to 60 °C, at which point α-glucosidase stock solution (100 µL, Megazyme, prod. no. E-TSAGS, 1500 units/mL) was added. The reaction mixture was maintained at 60 °C overnight. The reaction mixture was then heated to the boiling point where it was maintained for 30 minutes before the solvent was removed using a rotary evaporator. To the remaining solids were added water (50 mL) and the mixture was agitated by gentle shaking and sonication resulting in a slightly unclear solution that was filtered through a paper filter to remove the remaining solids. The filtrate was freeze-dried to yield the LRCD mixture as a water-soluble white solid (0.83 g).

S2.2 Characterization of LRCD mixture

The obtained LRCD mixture was characterised by 1D ^{1}H and high resolution ^{1}H-^{13}C HSQC NMR spectroscopy as well as MALDI-TOF-MS and HPLC-ELS. Integrals of the partially resolved anomeric peaks in the ^{1}H NMR spectrum (Figure S1) indicate that the mixture consists of primarily CD6–CD8 and a significant amount of glucose, although the integral values for particularly CD6–CD8 and the β-anomer of glucose are suppressed due to the proximity to the suppressed solvent residual signal. Thus, integral values observed in the HSQC spectrum (Figure S2) were deemed a better estimate of the relative concentrations in the mixture. This indicates that ca. 80% by weight of the LRCD mixture consists of CD6–CD8 and glucose, while the remaining 20% is made up of LRCDs. This was also supported by the HPLC-ELS data (see Figure 1b in the manuscript).

To verify that the LRCDs elute in order of increasing size, fractions were collected directly at the inlet to the ELS detector after injection of a large volume (20 µL) of a relatively concentrated sample (50 mg/mL). This was done after several repeated injections (of 2.0 µL each) had demonstrated that retention times were stable. The isolated fractions were analysed by MALDI-TOF-MS using DHB matrix, and the LRCDs were confirmed to elute in order of increasing DP by identification of their $[M+Na]^{+}$ adduct ions in the respective fractions. A representative set of MALDI-TOF-MS spectra are shown in Figure S3.

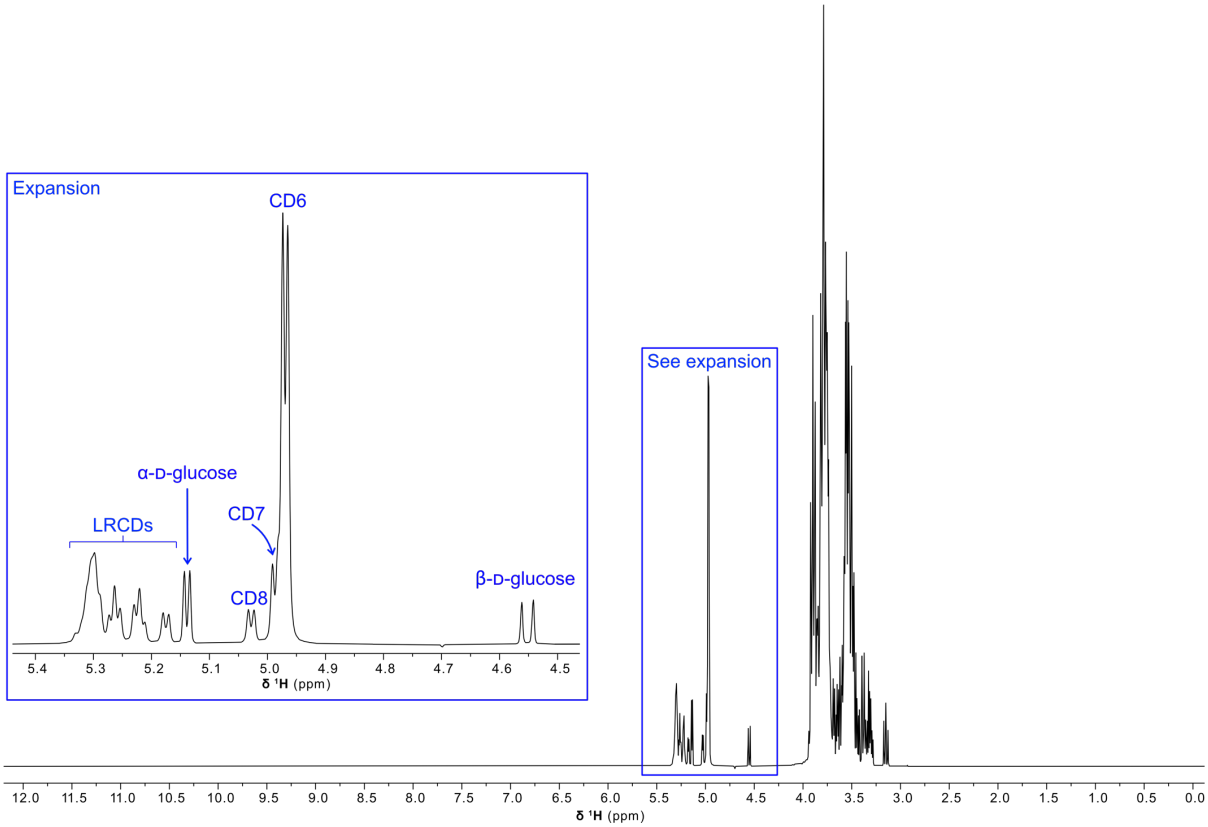

Figure S1: Solvent-suppressed ^{1}H NMR spectrum (400 MHz, 298 K) of isolated LRCD mixture (10 mg/mL) in D_2O. The expansion shows the partially resolved signals in the anomeric region.

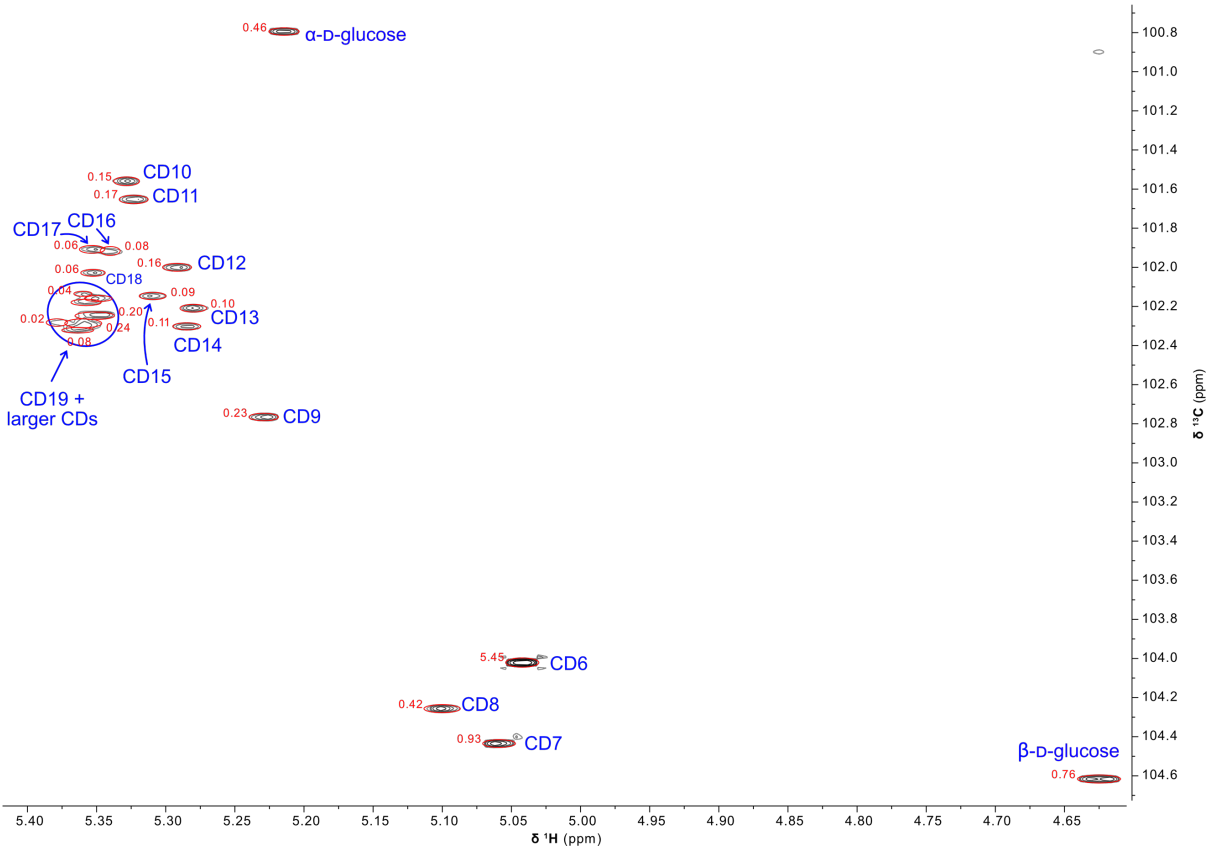

Figure S2: Anomeric region of the high-resolution ^{1}H-^{13}C HSQC NMR spectrum of the obtained LRCD mixture (800 MHz, 10 mg/mL in D_2O at 323K). Integrated areas are shown in red circles and the integral values (out of a total of 10) is written in red numerals next to each peak. Note that the peaks for both α- and β-D-glucose are aliased by one spectral width (6.00 ppm) in the ^{13}C dimension (i.e. their unaliased ^{13}C δ values are ca. 94.8 and 98.6, respectively). See section S3.2 for further acquisition details.

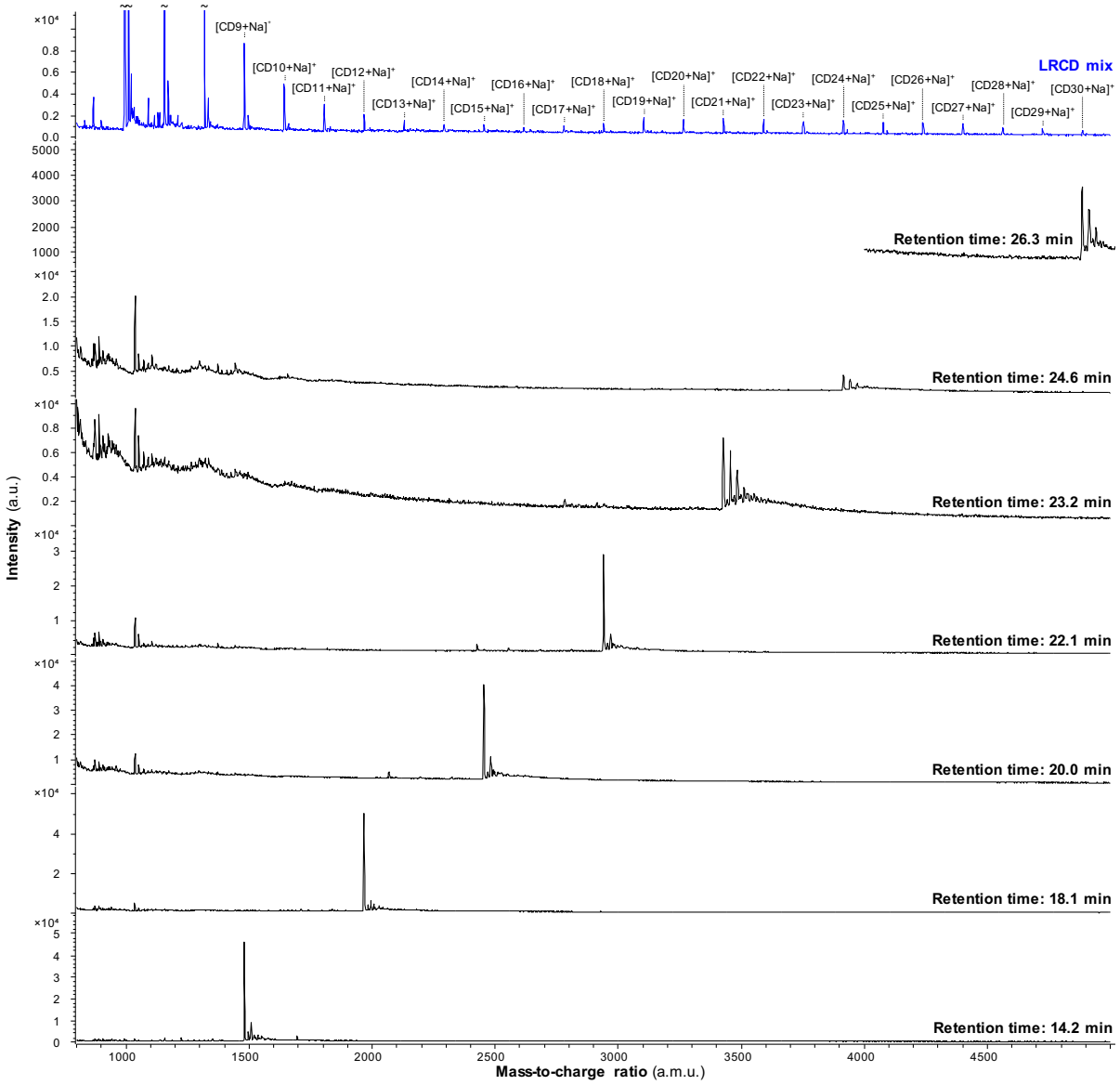

Figure S3: Representative MALDI-TOF-MS spectra of fractions isolated from analytical HPLC run. A spectrum of the entire LRCD mixture (blue) is shown at the top for reference. Note that the intensities of CD6–CD8 are much higher than shown here, where their peaks have been cut off to focus on the observed LRCD masses. Retention times refer to the peaks observed in the chromatogram shown in Figure 1b in the manuscript. For the later-eluting fractions, the mass range of the MALDI-TOF instrument was adjusted for optimal detection of the dilute samples, which is why only m/z ratios above 4000 are observed in the fraction isolated at 26.3 minutes.

S3 Assignment of LRCD anomeric peaks in HR-HSQC NMR spectra

S3.1 General procedure for isolation of individual LRCDs

To LRCD mixture (0.1 g) was added water/acetonitrile (1:1, 2 mL), leading to an almost clear solution that was filtered through a syringe filter and injected on a preparative HILIC type HPLC column (XBridge BEH Amide OBD Prep, Waters, 130 Å, 5 µm, 19 × 150 mm) installed on a preparative HPLC system (Büchi Reveleris) equipped with an ELS detector. Gradient elution was applied (from 25 % water in acetonitrile to 59 % water in acetonitrile over 94 minutes with a constant flow rate of 7.0 mL/min) and peaks corresponding to different (groups of) LRCDs were collected in separate fractions. 2–4 mg of each LRCD (CD9–CD18) was obtained in various purities as confirmed by analytical HPLC-ELS (Figure S4c).

S3.2 Acquisition of highly resolved ^{1}H-^{13}C HSQC NMR spectra

Sensitivity-enhanced ^{1}H-^{13}C HSQC NMR spectroscopy with optimised spectral widths, referred to as high-resolution HSQC NMR spectroscopy herein, was conducted at 323 K on a Bruker Avance III instrument operating at 800 MHz ^{1}H frequency, equipped with an 18.7 T Oxford magnet and a TCI cryoprobe. A Kel-F spinner turbine was used to reduce sample vibrations due to the higher weight compared to POM spinner turbines. All samples were locked on the D_2O solvent and were allowed to thermally equilibrate for 10 minutes prior to data acquisition. Sensitivity improved ^{1}H-^{13}C HSQC (1024×160 complex data points sampling 160 ms and 21.2 ms, respectively) spectra were recorded using a standard pulse sequence (hsqcetgpsisp.2) and bilevel adiabatic decoupling of ^{13}C during acquisition. The carrier offset in the ^{13}C dimension was placed near 101 ppm (typically 100.975 ppm or 100.987 ppm), and spectra were acquired using extensive aliasing in the ^{13}C dimension in spectra of 8 ppm (in the ^{1}H dimension) and 6 ppm (in the ^{13}C dimension, corresponding to a dwell time of 414.4 µs) spectral width. The resultant spectra sampled the FID for 160 and 133 ms in the ^{1}H and ^{13}C dimension. Nonuniform sampling of 50% of the data points was used and 12 transients were accumulated using an inter-scan relaxation delay of 1 s. This approach resulted in an acquisition time of 40 min per experiment. All spectra were acquired in Bruker Topspin 3.5 pl6. Spectra were processed and analysed in MestReNova v14.1.1-24571. All HSQC spectra were referenced to the anomeric signal of β-D-glucose (4.625 ppm and 98.620 ppm for the ^{1}H and ^{13}C dimensions, respectively, which was separately referenced to sodium trimethylsilylpropanesulfonate (DSS) in D_2O). Glucose was a convenient reference compound since it was present in all LRCD mixture samples and was not found to interact with the guest molecules used in this study. Figures of overlaid HSQC spectra were made using Bruker Topspin 4.3.0.

S3.3 Assignment of LRCD anomeric peaks in HR-HSQC NMR spectra

Isolated (groups of) LRCDs were dissolved in D_2O (ca. 0.2–2 mg/mL) and analysed using the high-resolution HSQC NMR method described in Section S3.2. A lock tube containing glucose (0.1 mg/mL) in D_2O was added to each NMR tube to allow referencing of the obtained spectra to the anomeric signal of glucose. Figure S4a shows the anomeric region of the HR-HSQC NMR spectra obtained from individual (groups of) LRCDs and the commercially sourced cycloamylose that was used as the starting material for the synthesis of the LRCD mixture. The exact same region of the HR-HSQC NMR spectrum of the LRCD mixture is shown in Figure S4b for comparison. These spectra allowed the assignment of the anomeric signal of each LRCD in the range from CD9 to CD18. The anomeric signals of α, β and γ-CD as well as D-glucose (not visible in the region shown in Figure S4) were identified by comparison to commercially available reference compounds.

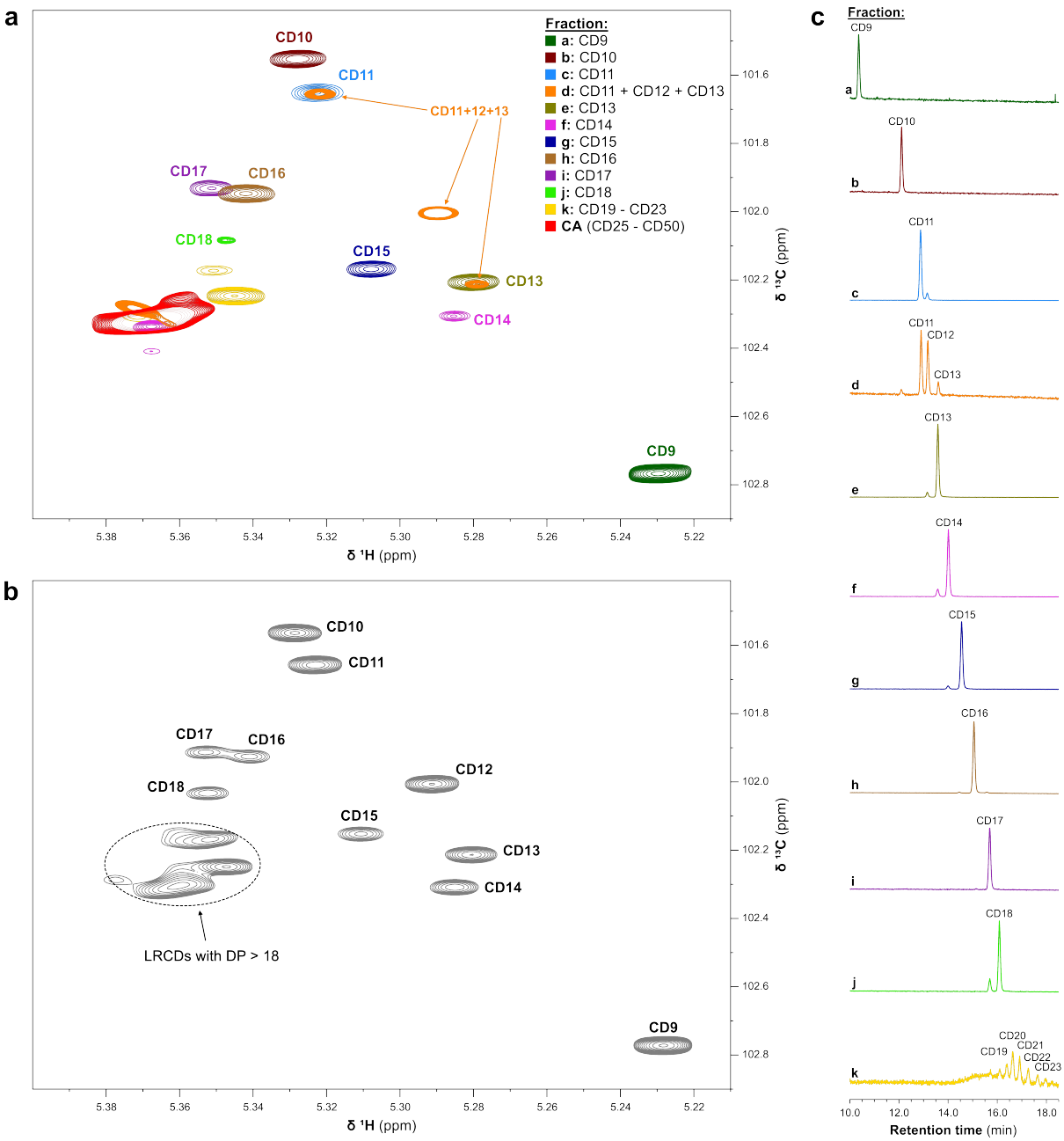

Figure S4. (a) Anomeric region of the HSQC spectrum of (partially) purified LRCDs. Spectra of 11 fractions (a–k) and cycloamylose (CA) are shown with 12 overlaid spectra in different colors as per the legend (323 K, 800 MHz). (b) Same region of the HSQC spectrum of the mixture of LRCDs (323 K, 800 MHz) with labels showing the resulting assignment of each peak for CD9–CD18. (c) Chromatograms (ELS detection) of the (partially) purified LRCDs.

S4 Comparison of LRCD mixture to cycloamylose and linear α-1,4-glucans

The high-resolution ^{1}H-^{13}C HSQC NMR spectra of cycloamylose (Ezaki Glico Co., ltd. prod. no. 302-32406, avg. mol. weight 7300 g/mol) and a mixture of maltooligosaccharides DP1–DP20[22] (10 mg/ml) were recorded and overlaid with the high-resolution ^{1}H-^{13}C HSQC NMR spectrum of the obtained LRCD mixture. The anomeric region of the spectra is shown in Figure S5. It was noted that the anomeric signal for the smaller LRCDs (from CD9 up to ca. CD18) are well-resolved and at significantly different chemical shift values than both the larger LRCDs (i.e. cycloamylose, from ca. CD25 and up to ca. CD50) and the anomeric signals from linear α-glucans.

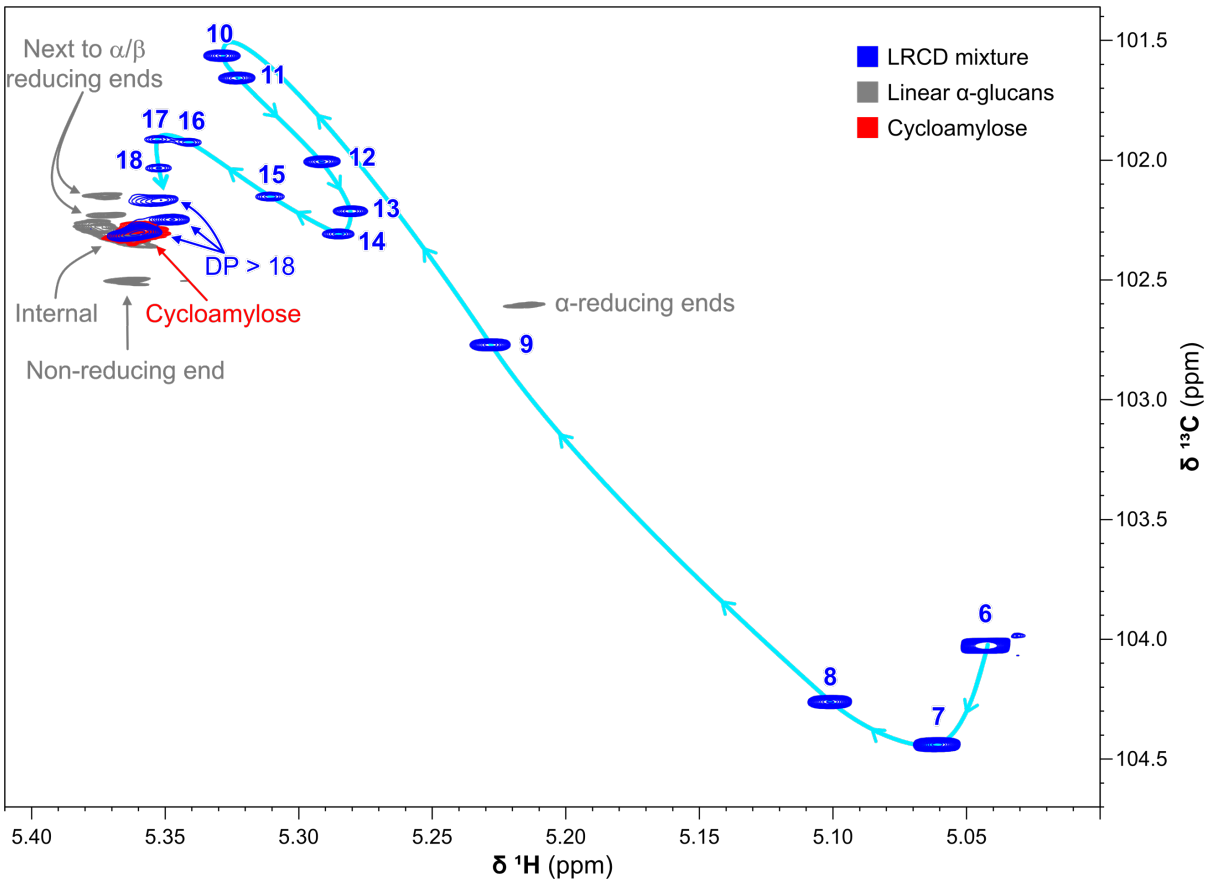

Figure S5. Anomeric region of high-resolution ^{1}H-^{13}C HSQC spectra of the obtained LRCD mixture (blue), commercially available cycloamylose (red) and a mixture of linear α-glucans (grey), all in D_2O (323 K, 800 MHz). The light blue line indicates the systematic nature of how the position of the peak for the anomeric signals change as a function of DP.

S5 HSQC NMR spectroscopy titrations of LRCD mixture with KI_3 and SDS

S5.1 General procedure for HSQC titrations

The titrant (SDS or KI_3) was dissolved (20 mM) in a stock solution of LRCD mixture (10 mg/mL) in D_2O. 500 µL of the LRCD mixture stock solution was transferred to an NMR tube. The NMR tube was loaded into the instrument where it was allowed to thermally equilibrate to the probe temperature for 10 minutes or until a steady lock signal indicated thermal equilibration had been reached. The data were acquired according to the procedure described in Section S3.2. Manual shimming was conducted during 32 dummy scans to ensure that the sample (including the spin systems) is in a steady state prior to data acquisition and that the magnetic field is as homogeneous as possible during data acquisition. Symmetric line shapes were obtained in this manner. After acquisition, the tube was taken out of the instrument, an appropriate amount of titrant stock solution was added, the contents were thoroughly mixed by turning the capped tube upside down several times (avoiding shaking motions during the SDS titration to prevent formation of bubbles). The tube was re-installed, allowed to thermally equilibrate and the same acquisition procedure was followed once again. This procedure was followed up to 10 mM of titrant. Solutions with higher amounts of titrant were made by a reverse titration procedure, starting with 500 µL of the 20 mM titrant solution that was sequentially diluted with LRCD mixture stock solution.

S5.2 HSQC titration with KI_3

A brown stock solution of KI_3 (20 mM) and LRCD mixture (10 mg/mL) was made by dissolving I_2 (20 mM) in a stock solution of KI (20 mM) and LRCD mixture (10 mg/mL). The titration was performed according to the general procedure described above. The anomeric regions of the resulting HSQC spectra are overlaid in Figure S6.

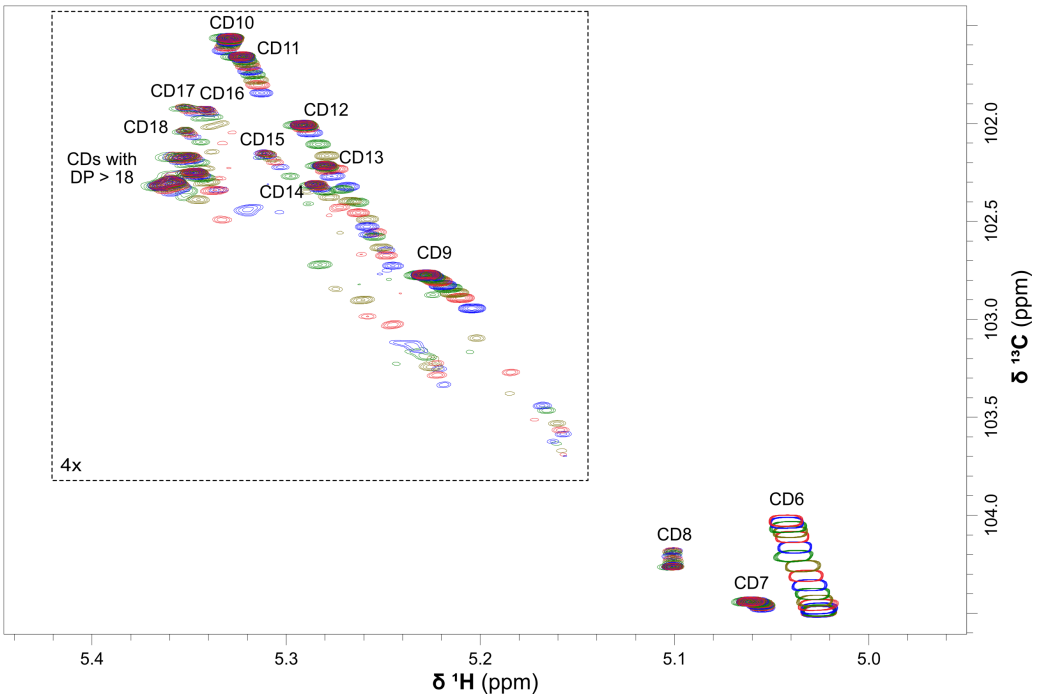

Figure S6. Overlay of the anomeric region of the high-resolution ^1H-^{13}C HSQC spectra obtained in the KI_3 titration.

Signals from individual CDs were monitored and the changes in their ^{13}C NMR chemical shifts (Δδ) were plotted as a function of the concentration of added KI_3. To determine the concentrations of KI_3 at half saturation ($[KI_3]_{1/2}$), the best fit of the obtained data to a logistic dose-response curve was found by non-linear regression (Figure S7). The chosen logistic curve equation is shown below:

$$\Delta\delta = \frac{A_1 - \Delta\delta_{\max}}{1 + \left(\frac{[KI_3]}{[KI_3]_{1/2}}\right)^p} + \Delta\delta_{\max}$$

where A_1 is the value of Δδ at the starting point of the titration (i.e. at 0 mM KI_3 added) and $\Delta\delta_{max}$ is the value at full saturation. A_1 is ideally zero, but to allow for error on the determination of the chemical shift in the starting point of the titration, this

value was fitted too. As can be seen in Table S1, the obtained values are all near zero, as expected. The exponent *p* defines the slope (i.e. steepness) of the curve going from A_1 to $\Delta\delta_{max}$ through $[KI_3]_{1/2}$. All the obtained values are listed in Table S1.

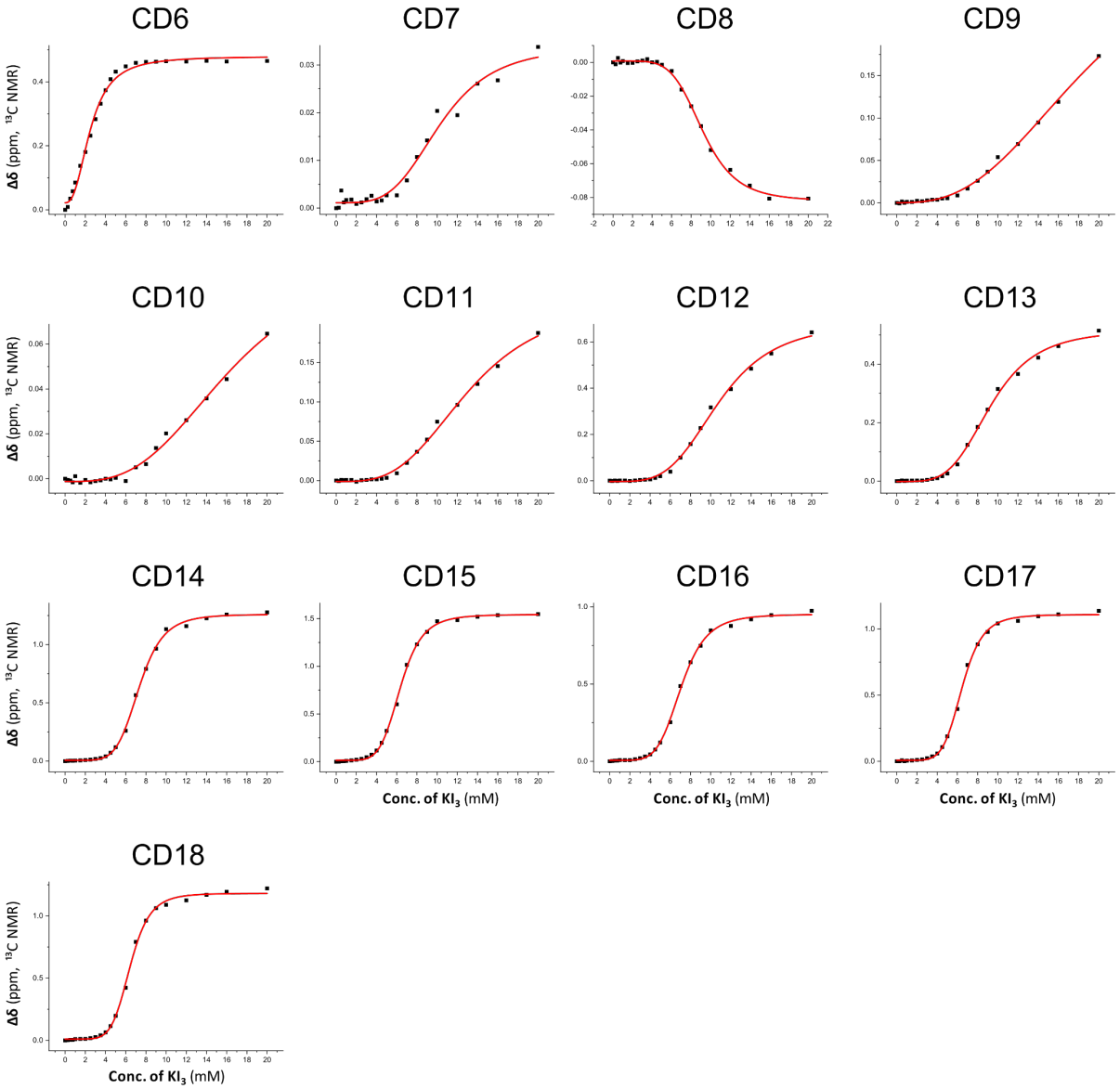

Figure S7. Observed changes in the ^{13}C NMR chemical shifts of CD6-CD18 as a function of the concentration of KI_3 (data points represented by black squares) and best fit to a logistic dose response curve (red line).

Table S1. Values obtained from logistical fit to the data from the titration with KI_3[a]

	A1 (ppm)		**$\Delta\delta_{max}$** (ppm)		$[KI_3]_{1/2}$ (mM)		***p***	
	Value	Std. error	Value	Std. error	Value	Std. error	Value	Std. error
CD6	0.02191	0.00830	0.47987	0.00725	2.50636	0.08239	2.50538	0.19397
CD7	0.00112	0.00053	0.03403	0.00269	10.30895	0.62175	3.85605	0.62376
CD8	0.00084	0.00044	-0.08218	0.00134	9.19802	0.10153	5.35986	0.27534
CD9	0.00018	0.00083	0.33323	0.04233	19.55712	1.73228	2.72828	0.16689
CD10	-0.00127	0.00055	0.10035	0.01510	16.62263	1.68045	3.08898	0.32178
CD11	-0.00118	0.00115	0.23092	0.01331	13.19548	0.54937	3.28849	0.21515
CD12	-0.00364	0.00351	0.68416	0.02009	10.83677	0.22762	3.88311	0.20850
CD13	-0.00142	0.00307	0.51780	0.01143	9.30494	0.15387	4.13461	0.22424
CD14	0.00636	0.00549	1.25966	0.01050	7.33698	0.04910	6.14351	0.22108
CD15	0.01416	0.00586	1.54353	0.00946	6.36287	0.03672	5.90393	0.16370
CD16	0.00671	0.00416	0.95179	0.00779	7.07116	0.04965	5.62953	0.19042
CD17	0.00736	0.00495	1.10907	0.00797	6.47126	0.04204	6.41655	0.21983
CD18	0.01065	0.00661	1.18121	0.01050	6.44107	0.05169	6.62195	0.28901

[a]Values are reported as they were calculated in the regression analysis, rounded off to 1/10,000[th] of a base unit. The number of significant digits is not meant to convey that the values are known with this accuracy.

S5.3 HSQC titration with SDS

Stock solutions were made as described in section S5.1. The titration was performed according to the general procedure described in section S5.1. The anomeric regions of the resulting HSQC spectra are overlaid in Fig. 3a of the manuscript. Signals from individual CDs were monitored and the changes in their ^{13}C NMR chemical shifts (Δδ) were plotted as a function of the concentration of added SDS (Fig. 3b).

To obtain binding constants, the changes in the ^{13}C NMR chemical shifts for each CD was plotted against the changes in ^{13}C NMR chemical shift of CD8, which was used as a reference in this study due to prior knowledge of binding affinity (see ref. 27 in the manuscript). The data was fitted by non-linear regression to equation 12 (derived below) according to the method described in ref. 22 of the manuscript. The data and resulting fits are shown in Figure S8, and the values obtained for K_a and $\Delta\delta_{max}$ from the fit are listed in Table S2.

Derivation of equation for determination of relative binding constants in a competition NMR titration.

For a 1:1 binding interaction between a host H_x and a guest G, the association constant $K_{a,x}$ can be defined according to equations 1 and 2, where [H], [G] and [HG] are the equilibrium concentrations of free host, free guest, and host/guest complex, respectively.

$$\mathrm{H} + \mathrm{G} \rightleftharpoons \mathrm{HG} \qquad (1)$$

$$K_a = \frac{[\mathrm{HG}]}{[\mathrm{H}][G]} \qquad (2)$$

For two different hosts (H_x and H_y), the ratio of the association constants ($K_{a,x}/K_{a,y}$) can then be defined by equation (3.

$$\frac{K_{a,x}}{K_{a,y}} = \frac{[H_xG]}{[H_x][\mathrm{G}]} \Big/ \frac{[H_yG]}{[H_y][\mathrm{G}]}$$

$$\frac{K_{a,x}}{K_{a,y}} = \frac{[H_xG][H_y]}{[H_yG][H_x]} \qquad (3)$$

Equation (3) can be rearranged to give equation (4).

$$\frac{K_{a,x}}{K_{a,y}} = \frac{[H_y]}{[H_yG]} \Big/ \frac{[H_x]}{[H_xG]} \qquad (4)$$

In an NMR titration in fast exchange, we observe the formation of the complex (HG) from a host (H) by the change in chemical shift of the peak corresponding to these two species in fast exchange. The observed chemical shift (δ_{obs}) is determined by

the weighted average of the chemicals shifts of the free host (δ_H) and the host-guest complex (δ_{HG}) according to equation (5).

$$\delta_{obs} = \delta_H \frac{[H]}{[H]+[HG]} + \delta_{HG} \frac{[HG]}{[H]+[HG]} \qquad (5)$$

Rearranging equation (5) gives equation (6), where $\Delta\delta_{HG}$, often referred to as $\Delta\delta_{max}$, is the difference in chemical shift between δ_{HG} and δ_H, and $\Delta\delta$ is the observed chemical shift change during the titration upon aliquot addition of the guest (G).

$$\delta_{obs} = \left(1 - \frac{[HG]}{[H]+[HG]}\right)\delta_H + \frac{[HG]}{[H]+[HG]}\delta_{HG}$$

$$\delta_{obs} - \delta_H = (\delta_{HG} - \delta_H)\frac{[HG]}{[H]+[HG]}$$

$$\Delta\delta = \frac{[HG]}{[H]+[HG]}\Delta\delta_{max} \qquad (6)$$

Equation (6) can be rearranged to give equation (7) as follows:

$$\frac{\Delta\delta_{max}}{\Delta\delta} = \frac{[H]+[HG]}{[HG]}$$

$$\frac{\Delta\delta_{max}}{\Delta\delta} = \frac{[H]}{[HG]} + 1$$

$$\frac{[H]}{[HG]} = \frac{\Delta\delta_{max}+\Delta\delta}{\Delta\delta} \qquad (7)$$

Therefore, in the case where we have competing hosts (H_x and H_y), we have two equations (8) and (9) that relate the ratio of the concentrations of the free and bound hosts to the chemical shift changes in the titration.

$$\frac{[H_x]}{[H_xG]} = \frac{\Delta\delta_{max,x}-\Delta\delta_x}{\Delta\delta_x} \qquad (8)$$

$$\frac{[H_y]}{[H_yG]} = \frac{\Delta\delta_{max,y}-\Delta\delta_y}{\Delta\delta_y} \qquad (9)$$

Substitution of equations (8) and (9) into equation (4) and rearrangement gives equation (10), which relates the ratio of association constants for ($K_{a,x}$ and $K_{a,y}$) to the chemical shift changes observed in the titrations. It is noteworthy that the concentration of the hosts and guests are absent from equation (10).

$$\frac{K_{a,x}}{K_{a,y}} = \frac{\Delta\delta_{max,y}-\Delta\delta_y}{\Delta\delta_y} \Big/ \frac{\Delta\delta_{max,x}-\Delta\delta_x}{\Delta\delta_x}$$

$$\frac{K_{a,x}}{K_{a,y}} = \frac{\Delta\delta_x(\Delta\delta_{max,y}-\Delta\delta_y)}{\Delta\delta_y(\Delta\delta_{max,x}-\Delta\delta_x)} \qquad (10)$$

By further rearrangement we obtain an expression of $\Delta\delta_x$ as a function of $\Delta\delta_y$, $\Delta\delta_{\Delta\delta_{max,x}}$, $\Delta\delta_{max,y}$, $K_{a,x}$ and $K_{a,y}$ (Equation 11).

$$\Delta\delta_y\Delta\delta_{max,x}K_{a,x} - \Delta\delta_y\Delta\delta_xK_{a,x} = \Delta\delta_x\Delta\delta_{max,y}K_{a,y} - \Delta\delta_x\Delta\delta_yK_{a,y}$$

$$\Delta\delta_x = \frac{\Delta\delta_{max,x}\Delta\delta_yK_{a,x}}{\Delta\delta_{max,y}K_{a,y}-\Delta\delta_yK_{a,y}+\Delta\delta_yK_{a,x}} \qquad (11)$$

If $K_{a,x}$ and $\Delta\delta_{max,x}$ are known for one host (H_x), $K_{a,y}$ and $\Delta\delta_{max,y}$ can be determined for H_y and all other hosts in the mixture using equation (11). Observing the changes of chemical shift for the hosts in a competition titration, $\Delta\delta_y$ is plotted against $\Delta\delta_x$ and the data is fitted using non-linear regression to equation (11) to obtain the two unknowns, $K_{a,y}$ and $\Delta\delta_{max,y}$.

In this system, CD8 is used as the reference host (H_y), and K_a values for CD9–CD18 are determined relative to K_a(CD8). Therefore, equation (11) can be rewritten as equation (12) (given as equation (2) in the manuscript).

$$\Delta\delta_{(CDx)} = \frac{\Delta\delta_{max\,(CDx)} \cdot \Delta\delta_{(CD8)} \cdot K_{a(CDx)}}{\Delta\delta_{max(CD8)} \cdot K_{a(CD8)} - \Delta\delta_{(CD8)} \cdot K_{a(CD8)} + \Delta\delta_{(CD8)} \cdot K_{a(CDx)}}$$

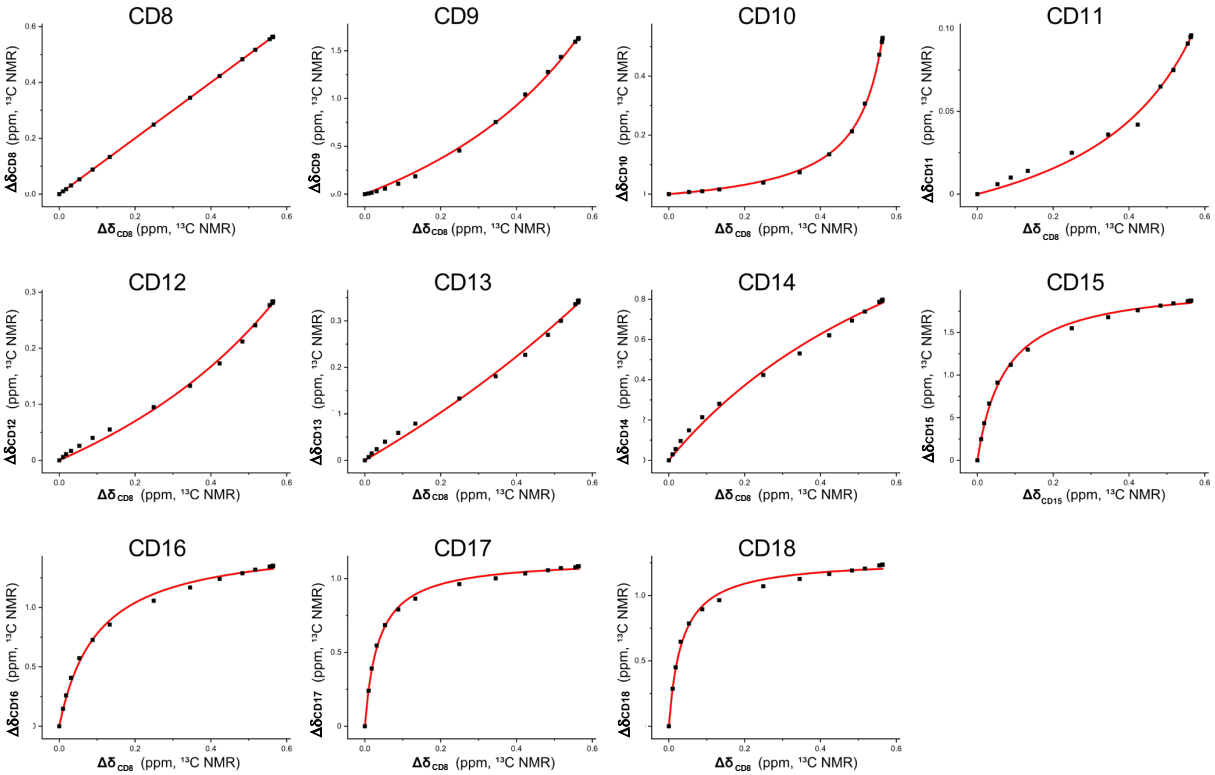

Figure S8. Observed changes in the ^{13}C NMR chemical shifts of CD8-CD18 as a function of the concentration of SDS (data points represented by black squares) and best fit to the equation in the manuscript (red line).

Table S2. Values obtained from fit of the data from the SDS titration to the equation in the manuscript[a]

	K_a (M^{-1})			$\Delta\delta_{max}$ (ppm)		
	Value	Std. error (fit)	Compounded error[b]	Value	Std. error (fit)	Compounded error[b]
CD8[c]	1125.88	57.18		0.563[e]	0.01[d]	
CD9	599.33	22.56	63.65	1.6420	0.0122	0.1248
CD10	131.33	3.00	12.00	0.5259	0.0025	0.0385
CD11	403.73	33.84	61.51	0.0947	0.0016	0.0081
CD12	681.80	40.16	86.89	0.2802	0.0032	0.0224
CD13	897.09	57.44	118.93	0.3381	0.0041	0.0273
CD14	1865.69	127.22	255.11	0.7812	0.0095	0.0630
CD15	9790.59	385.19	1056.29	1.8471	0.0120	0.1386
CD16	7466.07	362.41	874.18	1.3256	0.0108	0.1017
CD17	18820.64	705.57	1995.65	1.0663	0.0062	0.0793
CD18	19600.64	1108.52	2452.06	1.2044	0.0105	0.0930

[a]Values here are reported as they were calculated in the regression analysis, rounded off to 1/100th (for K_a) and 1/10,000th (for $\Delta\delta_{max}$) of a base unit. [b]Compounded errors are the sum of the relative errors for the values determined previously for CD8 and the errors on the values obtained in this titration. [c]Based on fit to a 1:1 model with ^{1}H NMR spectroscopic data reported earlier (see ref. 27 in the manuscript). There is no compounded error to report for CD8 as it is the reference compound. [d]$\Delta\delta_{max}$ for CD8 based on the value in the final data point (i.e. with [SDS] = 20 mM) with an estimated error of 0.01 ppm.