Supplementary Informations

S1 Materials and Methods 2
S1.1 DNA sequences 2
S1.2 Buffer preparation 2
S1.3 Oligonucleotide preparation 2

S2 pKa determination of CD1 3

S3 Experimental variations monitored by Cryo-EM 6
S3.1 Sample preparation 6
S3.2 Fibres formation between **CD1** and dsDNA$_{18}$ 6
S3.3 Effect of the oligonucleotide length 7
S3.4 Fibres width 8
S3.5 Effect of the ionic strength 9
S3.6 Control experiments 10
S3.7 Changes in fibre forming conditions 12

S4 Synthesis and characterization 21
S4.1 General informations 21
S4.2 Structures of the CDs and their preparations 22
S4.3 Synthesis of the hydrophobic group precursors (**5** and **6**) 23
S4.4 Synthesis of the phenyl CD (**CD3**) 24
S4.5 Synthesis of the cyclohexyl CD (**CD4**) 32

S5 Surface Plasmon Resonance (SPR) 40
S5.1 Supplementary Text 40
S5.2 Material and methods 40
S5.3 Low density dsDNA 41
S5.4 Medium density dsDNA 42
S5.5 High density dsDNA 43

S6 cryo-EM for 3D reconstruction 44
S6.1 Sample preparation for data collection 44
S6.2 Cryo-EM data processing 44
S6.3 Model building and refinement 46

S7 Crystal Structure of CD1 49
S7.1 Methods 49
S7.2 Crystal structure and packing of **CD1** 49

S8 Estimation of volumes and packing 51

S9 Ionic strength for standard cryoEM conditions 53

S10 Molecular dynamics simulations 54
S10.1 Methods of simulation 54
S10.2 Bending of protofibrils 55
S10.3 Distances between neighboring **CD1** protofibrils 55
S10.4 Voids between dsDNA$_{18}$ and **CD1** protofibrils 56

S11 Statistical mechanical calculations 59
S11.1 Summary of findings 59
S11.2 Mass action model 59
S11.3 Bending fluctuations, Coulomb interactions and optimal tube radius. 64

S1 Materials and Methods

S1.1 DNA sequences

Sequences of ssDNAs of different lengths (Biomers):

5’-AATCGT-3’	$ssDNA_6$-1
5’-ACGATT-3’	$ssDNA_6$-2

5'-AATCCTGTACTGTCTACA -3'	$ssDNA_{18}$-1
5'-TGTAGACAGTACAGGATT-3'	$ssDNA_{18}$-2

5'- AATCCTGTACTGTCTACAAATCCTGTACTGTCTACAAATCCTGTACTGTCTACAATTCCTGTACTGTCTACAAATCCTGTACTGTCTACATATCCTGTACTGTCTACAAATCCTGTACTGTCTAATCCTGTACTGTCTAC-3'	$ssDNA_{140}$-1
5'- GTAGACAGTACAGGATTAGACAGTACAGGATTTGTAGACAGTACAGGATATGTAGACAGTACAGGATTTGTAGACAGTACAGGAATTGTAGACAGTACAGGATTTGTAGACAGTACAGGATTTGTAGACAGTACAGGATT-3’	$ssDNA_{140}$-2

S1.2 Buffer preparation

Tris(hydroxymethyl)aminomethane (TRIS) was purchased from Alfa Aesar and **T**RIS **B**oronic acid **E**DTA (TBE) running buffer 10X was purchased from Thermo Scientific. TBE 1X was prepared from TBE 10X and then adjusted to pH = 8 with HCl aq. 1 M and TRIS (200 mM) was prepared from TRIS and then adjusted to pH = 8 with HCl aq. 1 M.

S1.3 Oligonucleotide preparation

The same volume of $ssDNA_{18}$-1 (50 µM) and $ssDNA_{18}$-2 (50 µM) was annealed twice (95 °C for 20 minutes and then slowly cooled to 20 °C over 1.5 h) to give $dsDNA_{18}$ (25 µM).

The same volume of $ssDNA_6$-1 (50 µM) and $ssDNA_6$-2 (50 µM) was annealed twice (95 °C for 20 minutes and then slowly cooled to 20 °C over 1.5 h) to give $dsDNA_6$ (25 µM).

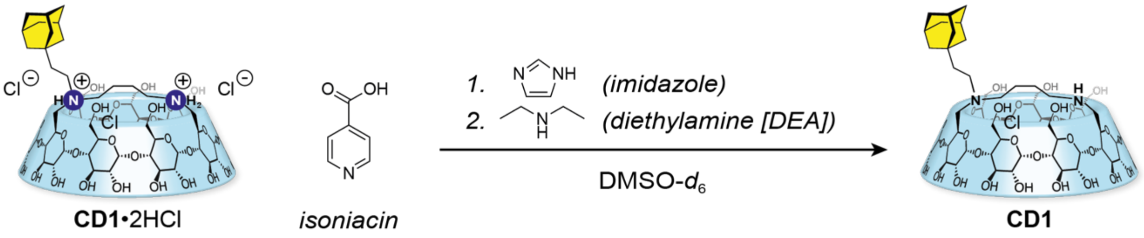
The same volume of $ssDNA_{140}$-1 (50 µM) and $ssDNA_{140}$-2 (50 µM) was annealed twice (95 °C for 20 minutes and then slowly cooled to 20 °C over 1.5 h) to give $dsDNA_{140}$ (25 µM).

S2 pKa determination of CD1

The determination of the two pKa values for **CD1** was performed using ^{1}H NMR spectroscopy (500 MHz, 300 K) in DMSO-d_6 in order to maintain **CD1** dissociated and in solution over the full range of pH explored.

The titration experiment was achieved using protonated **CD1** as a hydrochloric salt (**CD1**•2HCl, 1.9 mM), two pH probes (isoniacin and imidazole) and two different bases (imidazole and diethylamine [DEA]).[1]

The titration was carried out in two steps, *i.e*. from pH 5.5 to pH 8.3 and from pH 8.3 to pH 9.6. For the first part of the titration, isoniacin (5 mM) was used as pH NMR indicator and imidazole as a base (up to 52 eq, 104 mM). For the second part of the titration, the concentration of imidazole remained constant and played the role of a pH NMR indicator and DEA was used as a base (up to 74 eq.; 140 mM).



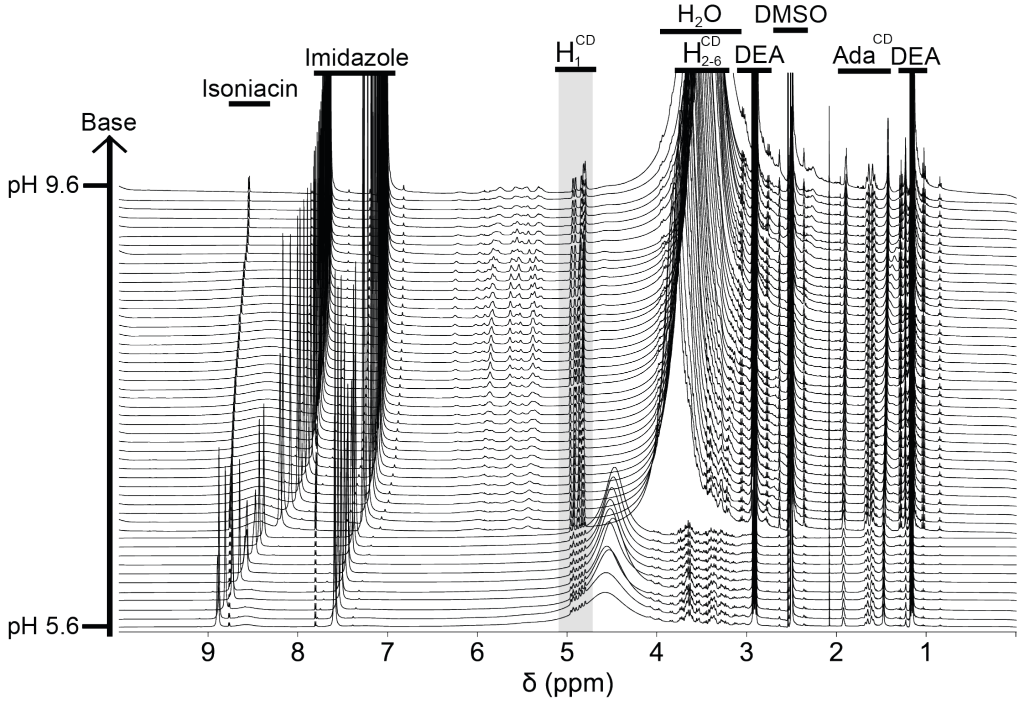


Figure S 1. ^{1}H NMR titration (500 MHz, 300 K) of **CD1**•2HCl in DMSO-d_6 with diethylamine (DEA). Isoniacin and imidazole were used as internal standard to estimate the pH during titration.

---

[1] G. Schenck, K. Baj, J. A. Iggo, M. Wallace, *Anal. Chem.* **94**, 8115–8119 (2022).

The monitoring of **CD1** signals clearly shows two distinct chemical shifts transitions during the pH titration. We thus tried to distinguish the nitrogens involved in these transitions (secondary and tertiary amino groups). This was made possible thanks to the following determination process (Figure S 2):

A NOE correlation allowed to identify the methylene connected to the adamantyl group in β position with respect to the tertiary amine ($CH_2β$). This later was further confirmed comparing its ^{13}C NMR chemical shift with previously synthesized adamantyl functionalized cyclodextrins,[2] and looking at the TOCSY spectrum showing a clear connection with another methylene group in α position to the tertiary amine ($CH_2α$).

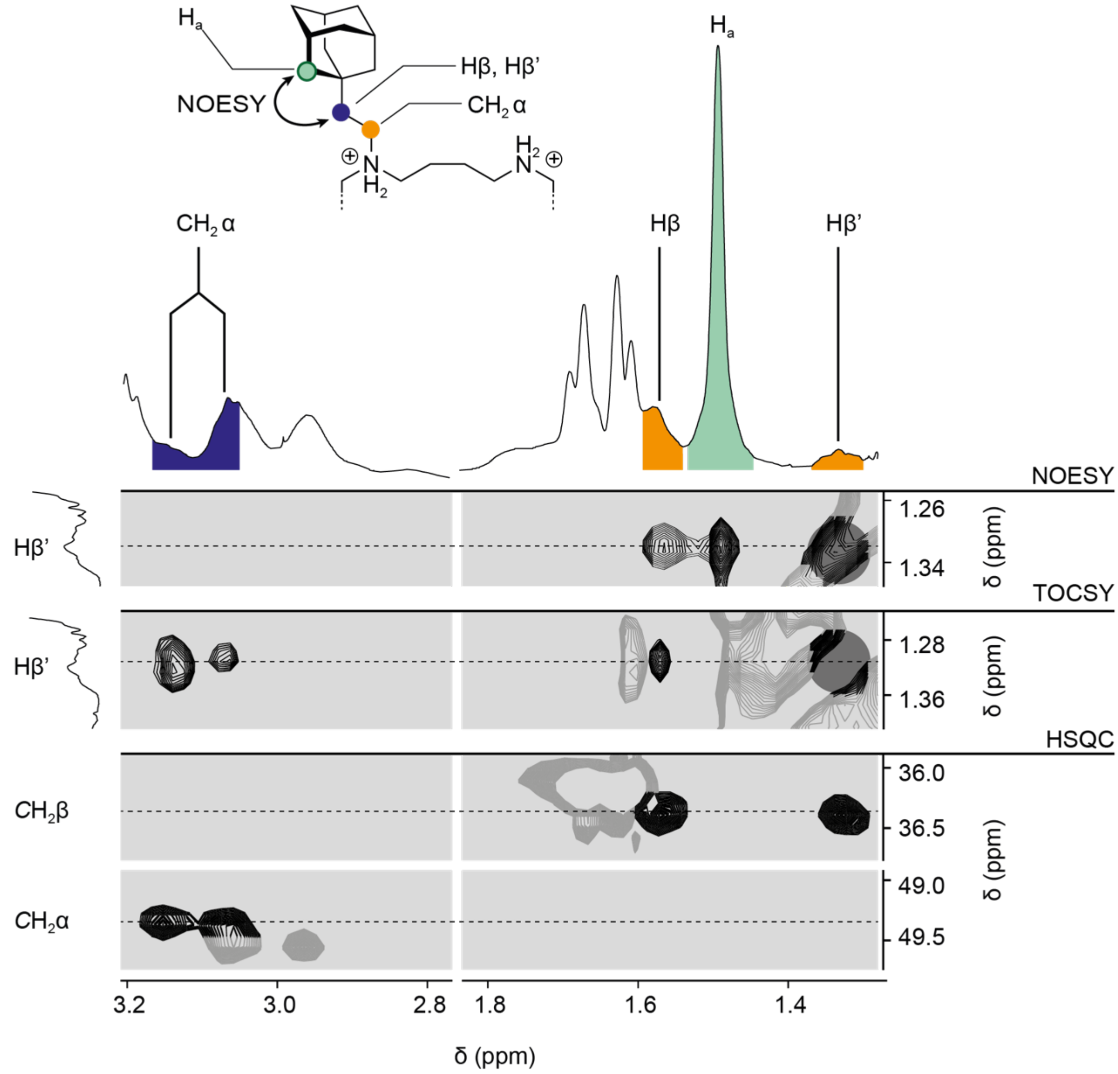


Figure S 2 Selected regions of NOESY, TOCSY, HSQC NMR (500 MHz, 300 K) of **CD1**•2HCl in DMSO-*d*$_6$.

---

[2] (a) P. Evenou, J. Rossignol, G. Pembouong, A. Gothland, D. Colesnic, R. Barbeyron, S. Rudiuk, A.-G. Marcelin, M. Ménand, D. Baigl, V. Calvez, L. Bouteiller, M. Sollogoub, *Angew. Chem. Int. Ed.* **57**, 7753–7758 (2018); (b) W. Hu, V. Libérioux, J. Rossignol, G. Pembouong, E. Derat, M. Ménand, L. Bouteiller, M. Sollogoub, *Angew. Chem. Int. Ed.* e202507069 (2025).

The protonation state was first monitored following the signals of one of the two diastereotopic proton (Figure S 3a, Hβ' – blue line). However, the monitoring of the signal along the titration was not possible due to severe overlapping. Nevertheless, the transition of Hβ' appeared to follow the same trend as one of the H_1^{CD} protons (Figure S 3b, blue line). Plotting the variation of its chemical shift against pH (Figure S 3c, blue curve) allowed to determine a pKa of 7.1 for the tertiary amine using the graphical tangent method. Similarly, a second transition was observed associated to another methylene group in β position to the secondary amine (Figure S 3a, Hβ' – red line). As previously observed, the transition of Hβ' suffered from overlapping while following the same trend as another isolated H_1^{CD} signal (Figure S 3b, red line). The corresponding titration plot allowed us to determine a pKa of 8.5 for the secondary amine using the graphical tangent method (Figure S 3c, red curve).

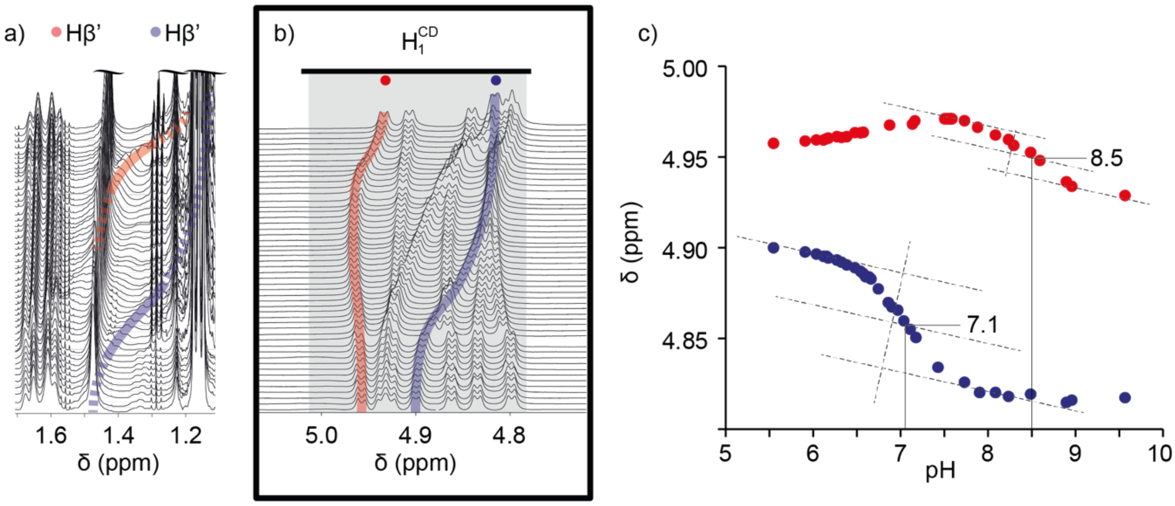


Figure S 3 a) b) Selected regions of ^{1}H NMR titration (500 MHz, 300 K) of **CD1**•2HCl in DMSO-*d*₆ with diethylamine (DEA). c) Corresponding titration plots.

pKa values determined in DMSO are known to differ from those measured in water. However, a recent study has estimated the extent of these deviations, particularly for amines, and found that the variations are relatively minor.[3] Specifically, the pKa of secondary amines tends to increase slightly in DMSO (ΔpKa ≈ + 0.2 relative to water), while that of tertiary amines tends to decrease (ΔpKa ≈ – 0.4 relative to water).

Accordingly, the pKa values determined for the amines of **CD1** are estimated to vary as follows:

pKa (sec. amine) = 8.3 to 8.5
pKa (tert. amine) = 7.1 to 7.5

---

[3] E. Rossini, A. D. Bochevarov, E. W. Knapp, *ACS Omega* **3**, 1653–1662 (2018).

S3 Experimental variations monitored by Cryo-EM

S3.1 Sample preparation

For cryo-EM images, a 5 µL drop of the initial sample solution was deposited on "quantifoil"® carbon membrane grids (Quantifoil Micro Tools GmbH, Germany). The excess of liquid on the grid was absorbed with a filter paper and the grid was flash-frozen in liquid ethane to form a thin vitreous ice film using a homemade mechanical cryo plunger. Once placed in a Gatan 626 cryo-holder cooled with liquid nitrogen, the samples were transferred in the microscope and observed at low temperature (-180 °C). Cryo-EM images were recorded on ultrascan 1000, 2k x 2k pixels CCD camera (Gatan, USA), using a LaB6 JEOL JEM2100 (JEOL, Japan) cryo-microscope operating at 200 kV with a JEOL low dose system (Minimum Dose System, MDS, JEOL, Japan) to protect the thin ice film from any irradiation before imaging and reduce the irradiation during the image capture.

Xray Energy Dispersive Spectra (XEDS) characterizing the elemental composition of the sample have been taken with JEOL XEDS detector with 140 eV resolution.

S3.2 Fibres formation between **CD1** and dsDNA$_{18}$

CD1 (1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (88.5 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (33.7 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{18}$] = 6 µM, [**CD1**]/[dsDNA$_{18}$] = 117) that was further incubated for 12 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

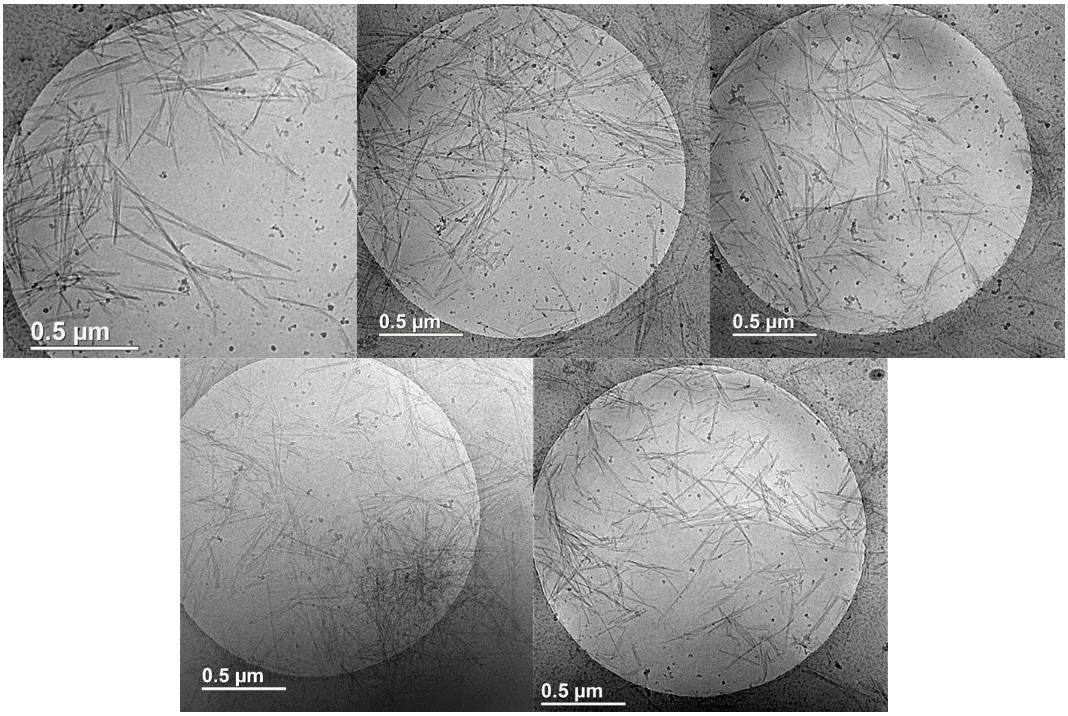


Figure S 4

S3.3 Effect of the oligonucleotide length

S3.3.1 dsDNA$_6$

CD1(1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.9 µL) of this solution was diluted with TBE 3X (15,3 µL, pH = 8). To the corresponding solution, was added dsDNA$_6$ (72 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_6$] = 18 µM, [**CD1**]/[dsDNA$_6$] = 39) that was further incubated for 16 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

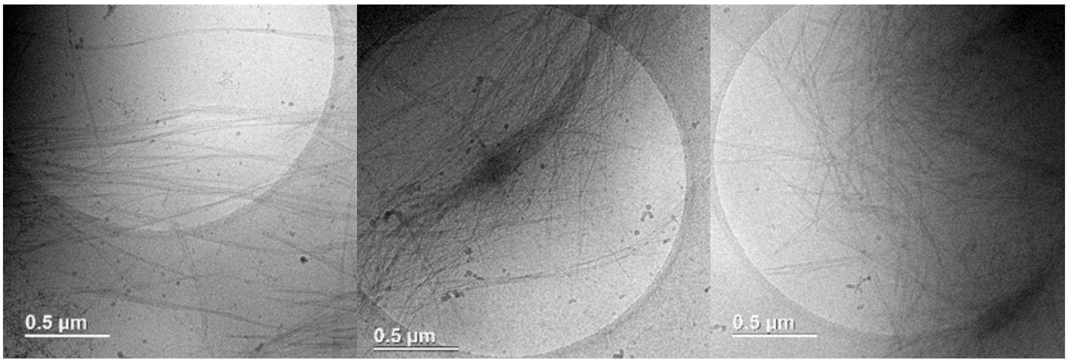


Figure S 5

S3.3.2 dsDNA$_{140}$

CD1 (1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (118.0 µL, pH = 8). To the corresponding solution, was added dsDNA$_{140}$ (4.2 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{140}$] = 0.75 µM, [**CD1**]/[dsDNA$_{140}$] = 933) that was further incubated for 1 h 20 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

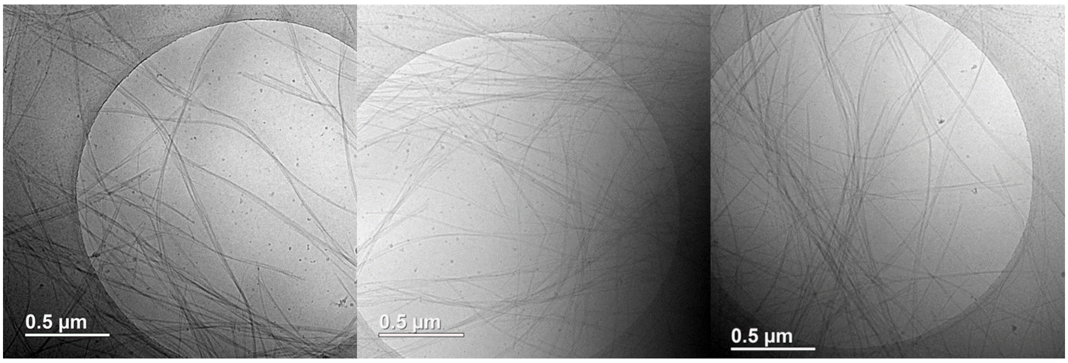


Figure S 6

S3.4 Fibres width

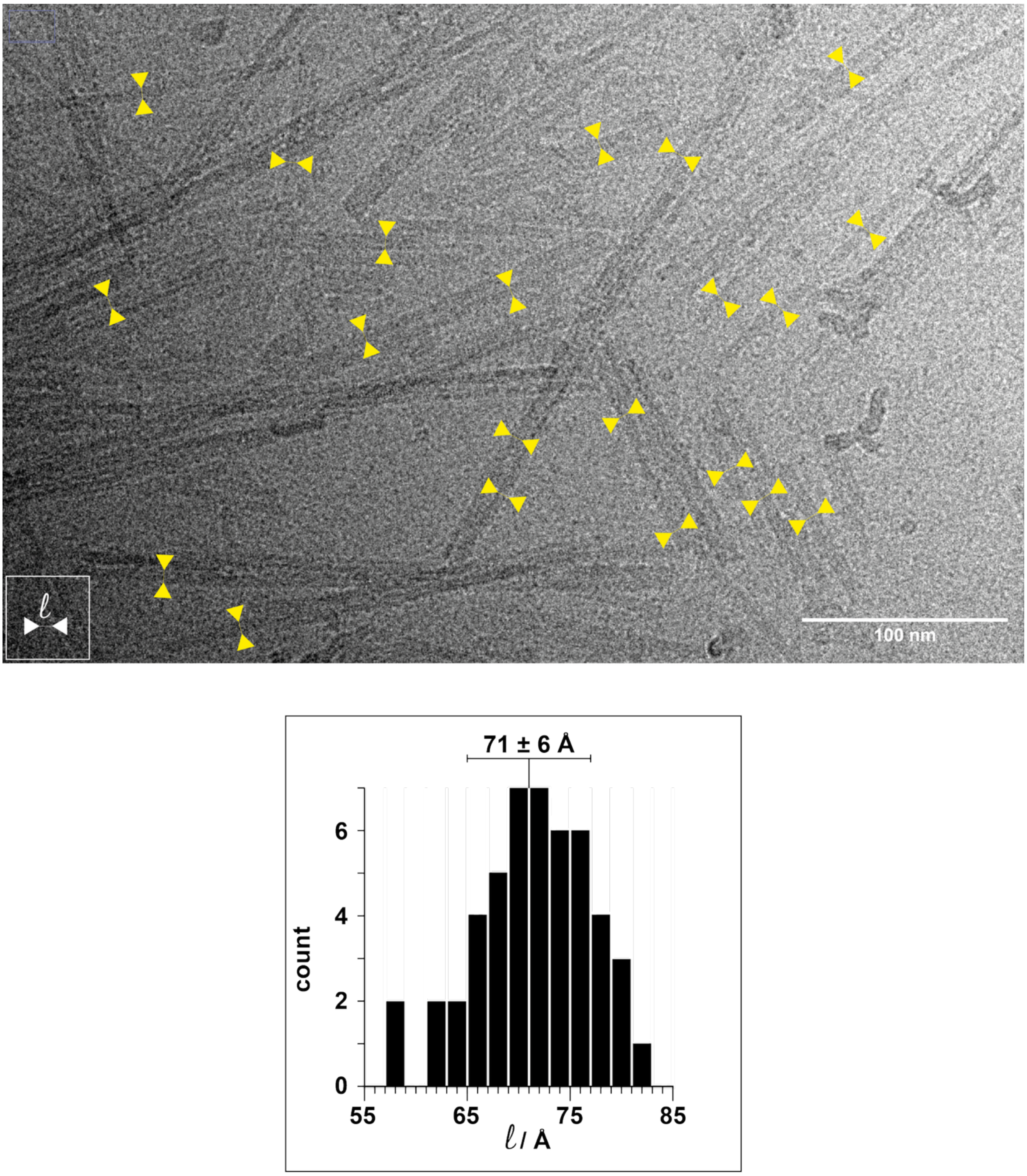


Figure S 7. High resolution micrograph of **CD1**-dsDNA$_{18}$ fibres and their width distribution (50 measurements using Image J).

S3.5 Effect of the ionic strength

CD1 (1.0 mg, 0.74 *µ*moles) was dissolved in 135 µL of distilled H_2O (5.5 mM). 17.9 µL of this solution was diluted with 88.6 µL of buffer (TBE 1X, pH = 8, [KCl] = 3 M]). 33.6 µL of $dsDNA_{18}$ (25 µM) was then added to the solution. The mixture was incubated for 21 minutes and then vitrified in liquid ethane. Cryo-TEM images were then acquired.

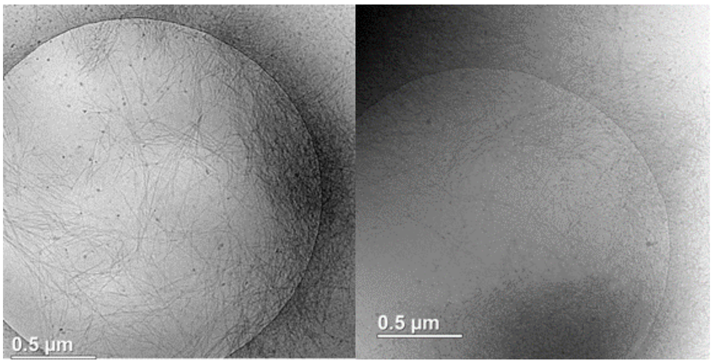


Figure S 8

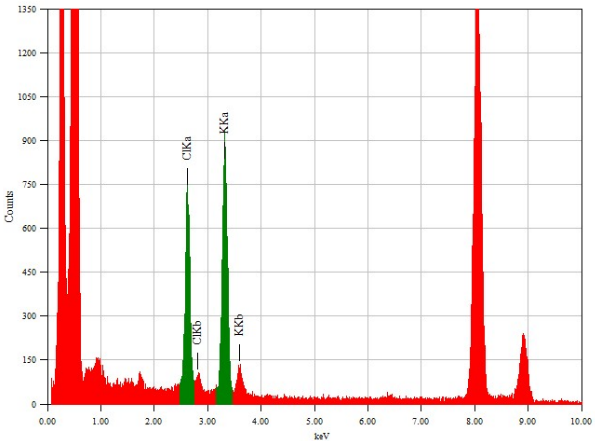


Figure S 9 Energy Dispersive X-ray Spectroscopy (EDS) analysis of a micrograph recorded with the same cryo-EM grid as experiment

S3.6 Control experiments

S3.6.1 ***CD1** alone*

CD1(1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (122.2 µL, pH = 8) to afford the final solution ([**CD1**] = 0.7 mM) that was further incubated for 7 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

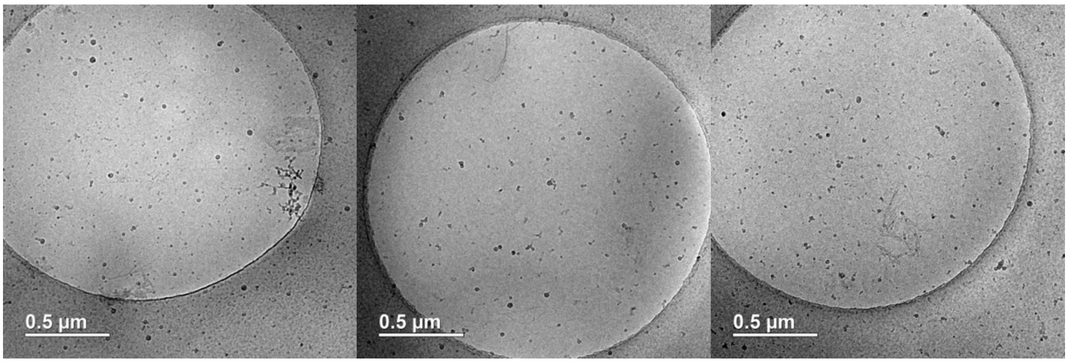


Figure S 10

S3.6.2 dsDNA$_{18}$ alone

dsDNA$_{18}$ (33.7 µL at 25 µM in water) was diluted with TBE 1X (88.5 µL, pH = 8.4) to afford the final solution ([dsDNA$_{18}$] = 6.9 µM) that was further incubated for 5 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

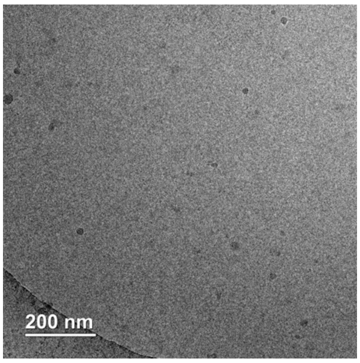


Figure S 11

S3.6.3 *CD lacking the hydrophobic group – **CD2** + dsDNA$_{18}$*

CD2 (1.0 mg, 0.84 µmoles) was first dissolved in distilled water (153 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (88.6 µL, pH = 8.4). To the corresponding solution, was added dsDNA$_{18}$ (33.6 µL at 25 µM in water) to afford the final solution ([**CD2**] = 0.7 mM, [dsDNA$_{18}$] = 6 µM) that was further incubated for 8 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

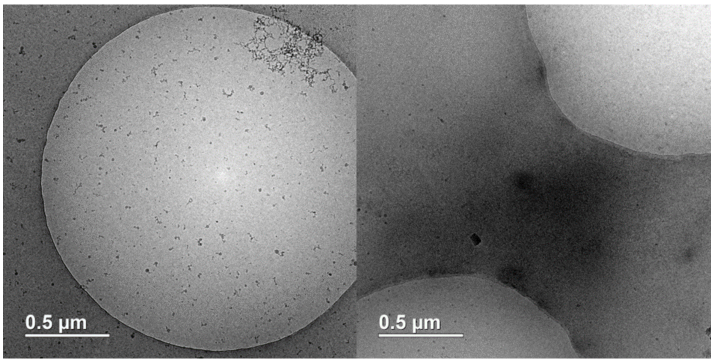


Figure S 12

S3.7 Changes in fibre forming conditions

S3.7.1 Effect of the global concentration

S3.7.1.1 Supplementary text

Diluting the **CD1**/DNA solution tenfold does not produce any fibres (Figure S 13), while a tenfold increase in concentration produces fibre aggregates and bundles (Figure S 14), suggesting that long rod-like particles are quite sensitive to bundle formation induced by even quite weak attractive interactions.[4] This kind of behaviour is not that of a simple isodesmic type, where large variations in concentrations have mild effects on the length and number of assemblies.

S3.7.1.2 Tenfold dilution

CD1 (1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (1.2 µL) of this solution was diluted with TBE 1X (88.6 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (2.2 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.07 mM, [dsDNA$_{18}$] = 0.6 µM) that was further incubated for 20 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

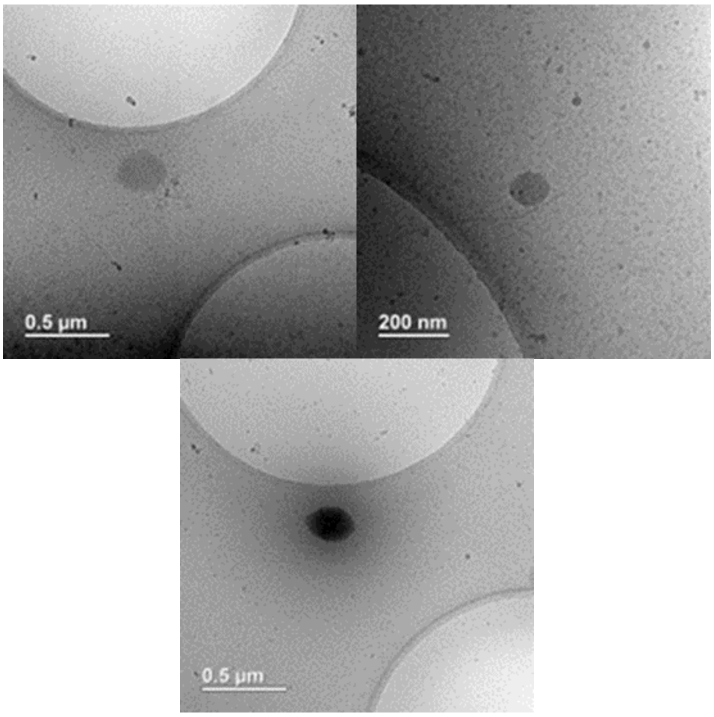


Figure S 13

---

[4] A. V Kyrylyuk, P. van der Schoot, Continuum percolation of carbon nanotubes in polymeric and colloidal media. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8221-8226 (2008).

S3.7.1.3 Tenfold concentration

CD1 (1.0 mg, 0.74 µmoles) was dissolved in TBE 3X (40.1 µL, pH = 8). dsDNA$_{18}$ (60.6 µL at 100 µM in water) was added to afford the final solution ([**CD1**] = 7 mM, [dsDNA$_{18}$] = 60 µM) that was further incubated for 15 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

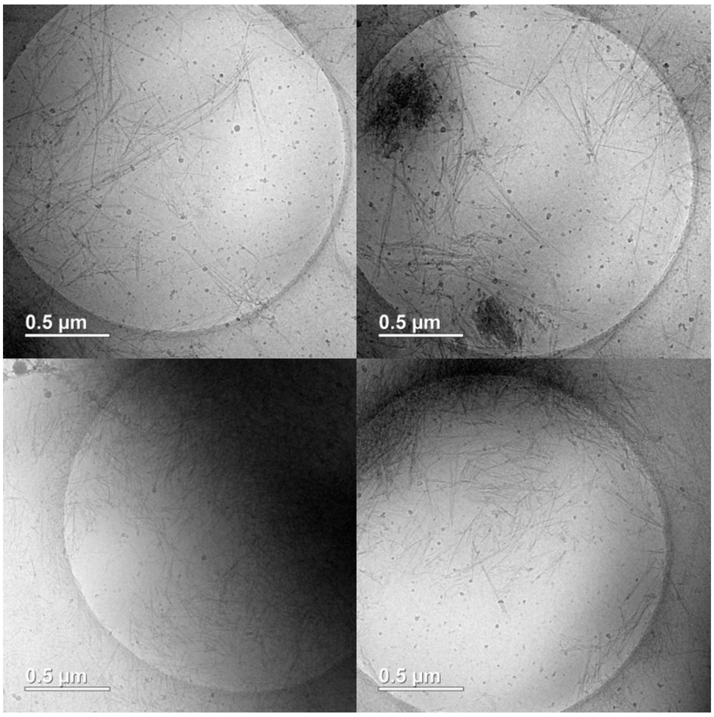


Figure S 14

S3.7.2 Effect of the relative concentrations, N/P ratio

S3.7.2.1 Supplementary text

Changes in the ratio of **CD1** to DNA, which can be translated in terms of changes in the charge balance (N/P ratio – nitrogen on the CD to phosphorus ratio), either leads to spherical particles at N/P=1 (Figure S 15) or fibrillar assemblies at higher values of N/P (of 6.4 and 159) (Figure S 4 and Figure S 16), indicating that a large excess of CD relative to DNA is necessary to build the assemblies.

S3.7.2.2 N/P = 1

CD1(1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (67.7 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (54.5 µL at 100 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{18}$] = 38.9 µM) that was further incubated for 11 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

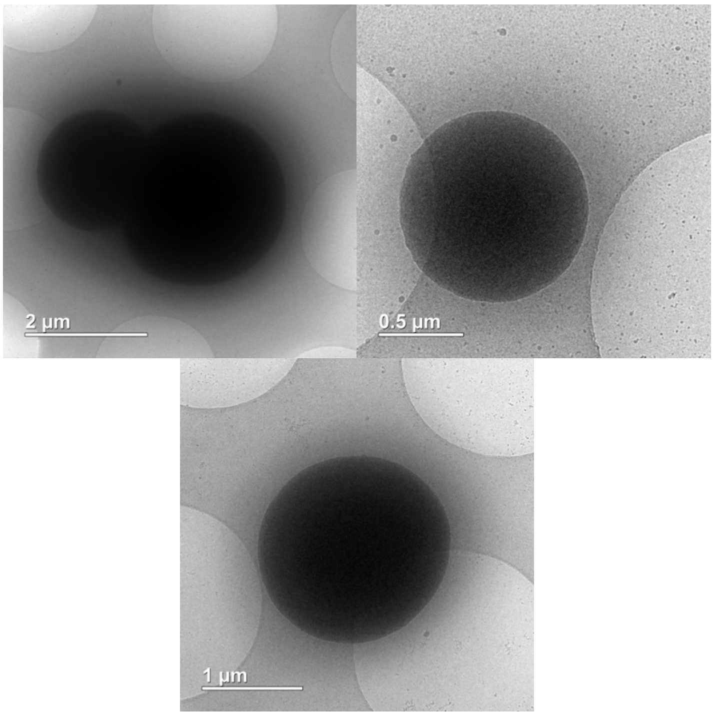


Figure S 15

S3.7.2.3 N/P = 159

CD1(1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (120.8 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (1.37 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{18}$] = 0.25 µM) that was further incubated for 13 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

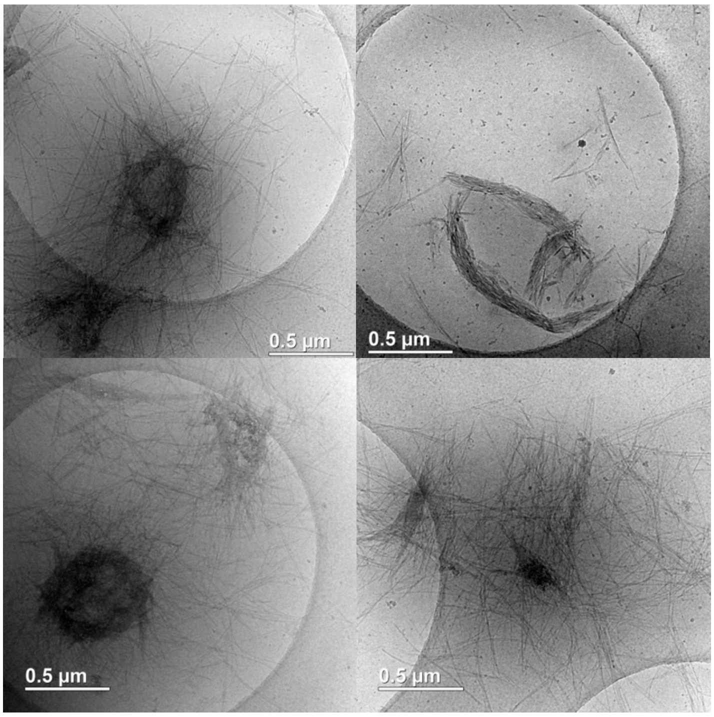


Figure S 16

S3.7.3 Effect of the hydrophobic group affinity

S3.7.3.1 Supplementary text

Decreasing the affinity between the CDs in the SMP, by replacing the adamantyl group in **CD1** (with a binding constant of ca. 10^4 M^{-1}) by either a phenyl (**CD3**) or a cyclohexyl (**CD4**) group (Figure 1g and S4), of which the binding constants are respectively 1 and 2 orders smaller in magnitude, prevents the production of fibres under the same conditions (Figure S 17 and Figure S 18). This all-or-nothing behaviour is typical of cooperative assembly needing a critical concentration to trigger fibre growth. This conclusion is supported by our model calculations (see text and S11), showing the importance of having an excess of CD as well, promoting the CDs to form SMPs.

*S3.7.3.2 **CD3** – phenyl group*

CD3 (1.0 mg, 0.77 µmoles) was first dissolved in distilled water (139 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (88.6 µL, pH = 8). To the corresponding solution, was added $dsDNA_{18}$ (33.6 µL at 25 µM in water) to afford the final solution ([**CD3**] = 0.7 mM, [$dsDNA_{18}$] = 6 µM) that was further incubated for 12 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

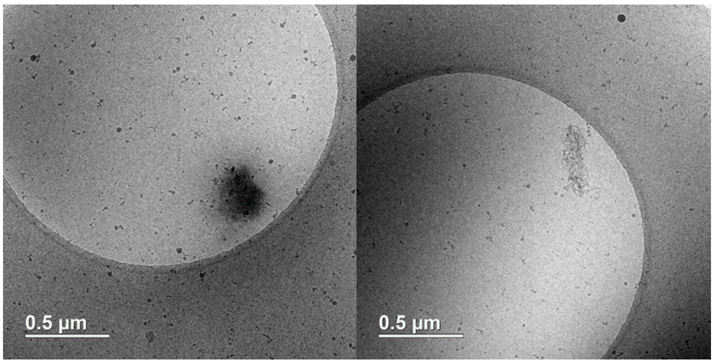


Figure S 17

*S3.7.3.3 **CD4** – cyclohexyl group*

CD4 (1.0 mg, 0.77 µmoles) was first dissolved in distilled water (140 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (88.6 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (33.6 µL at 25 µM in water) to afford the final solution ([**CD4**] = 0.7 mM, [dsDNA$_{18}$] = 6 µM) that was further incubated for 12 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

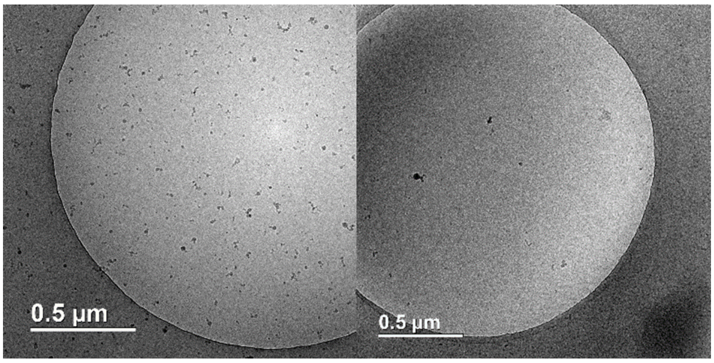


Figure S 18

S3.7.4 Effect of the pH

S3.7.4.1 Supplementary text

The cationic charge of **CD1** plays a pivotal role, as evidenced by pH-dependent variations in fibre formation, where at pH 7 and 7.5 only seemingly large (Figure S 19 and Figure S 20), structureless spheres are observed, while at pH 8 and 8.5 fibres are formed (Figure S 21 and Figure S 22) and at pH 9 and over the assemblies precipitate. The pKa values of **CD1** support this pH-related behaviour with a tertiary amine (pKa < 7.1, see S2) remaining largely neutral, while the secondary amine (pKa 8.5) is protonated in the pH range of the study.

S3.7.4.2 Samples preparation

CD1 (1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with TBE 1X (88.5 µL, pH adjusted to the desired value with 1 M HCl solution). To the corresponding solution, was added dsDNA$_{18}$ (33.7 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{18}$] = 6 µM, [**CD1**]/[dsDNA$_{18}$] = 117) that was further incubated for 12 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

S3.7.4.3 pH 7

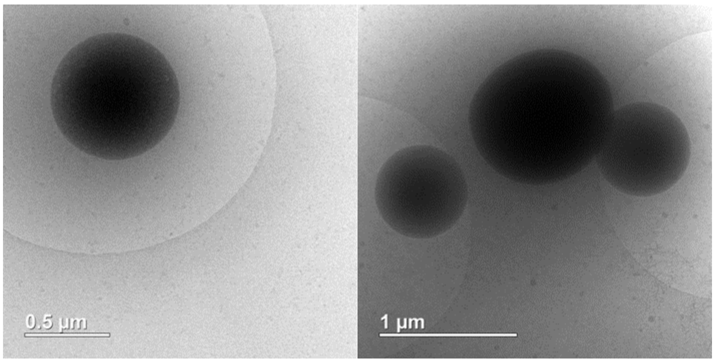


Figure S 19

S3.7.4.4 pH 7.5

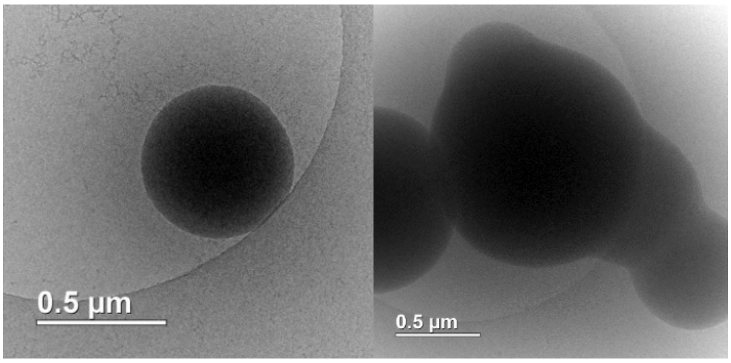


Figure S 20

S3.7.4.5 pH 8

reference experiment

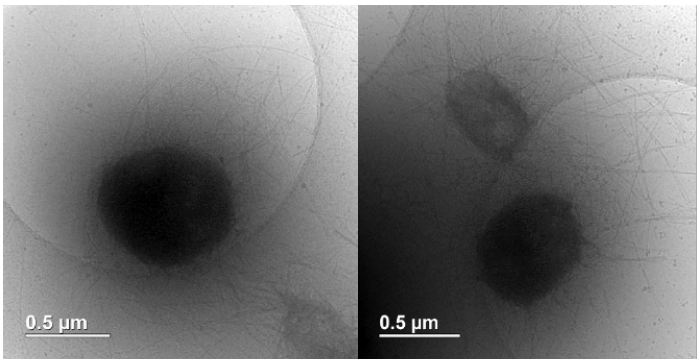


Figure S 21

S3.7.4.6 pH 8.5

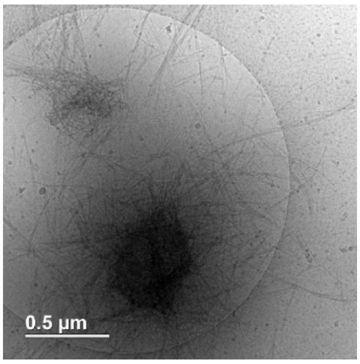


Figure S 22

S3.7.5 Effect of the buffer composition

Changing the buffer composition from TBE to Tris (Figure S 23) has no discernible impact on fibre formation.

S3.7.5.1 TRIS

CD1(1.0 mg, 0.74 µmoles) was first dissolved in distilled water (135 µL, 5.5 mM) and an aliquot (17.8 µL) of this solution was diluted with Tris 200 mM (88.6 µL, pH = 8). To the corresponding solution, was added dsDNA$_{18}$ (33.6 µL at 25 µM in water) to afford the final solution ([**CD1**] = 0.7 mM, [dsDNA$_{18}$] = 6 µM) that was further incubated for 11 min at R.T. The sample was then deposited on a carbon membrane grid following the sample preparation process mentioned above (see section S3.1).

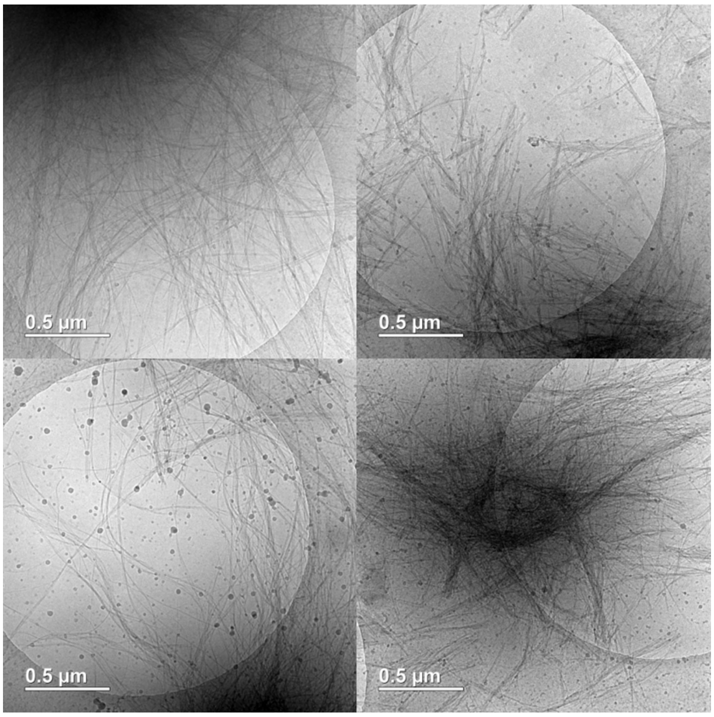


Figure S 23

S4 Synthesis and characterization

S4.1 General informations

The reactants were purchased from commercial sources and used without further purification. Tetrahydrofurane (THF) was freshly distilled over sodium/benzophenone. DMF was dried over molecular sieves. NMR spectra were recorded on a Bruker Avance I - 400 MHz or a Bruker Avance III - 600 MHz using $CDCl_3$, DMSO-d_6, and D_2O as solvents. All the spectra were calibrated with solvent signals: $CHCl_3$, H_2O, DMSO, and MeOH for carbon spectra in D_2O.

Assignments of the signals were done using COrrelation SpectroscopY (COSY), Nuclear Overhauser Effect SpectroscopY (NOESY), Heteronuclear Single Quantum Coherence Spectroscopy (HSQC), Heteronuclear Multiple Bound Correlation (HMBC), TOtal Correlation SpectroscopY (TOCSY), Transverse Rotating-frame Overhauser Enhancement Spectroscopy (T-ROESY). If total attribution is described, all sugar units are sort from the unit with the less deshielded H1 (α) to the more deshielded one (η). All units were then renamed from unit A to unit G in the counterclockwise (first rim).

High Resolution Mass Spectrometry (HRMS) was performed on a Bruker microTOF spectrometer using Agilent ESI-L Low Concentration Tuning-Mix, Sodium formiate or Lithium formiate as reference.

Optical rotations were measured on a Perkin-Elmer 341 digital polarimeter with a path length of 1 dm.

CD1, **CD2** and **CD7** were synthesized according to the literature.[2a]

S4.2 Structures of the CDs and their preparations

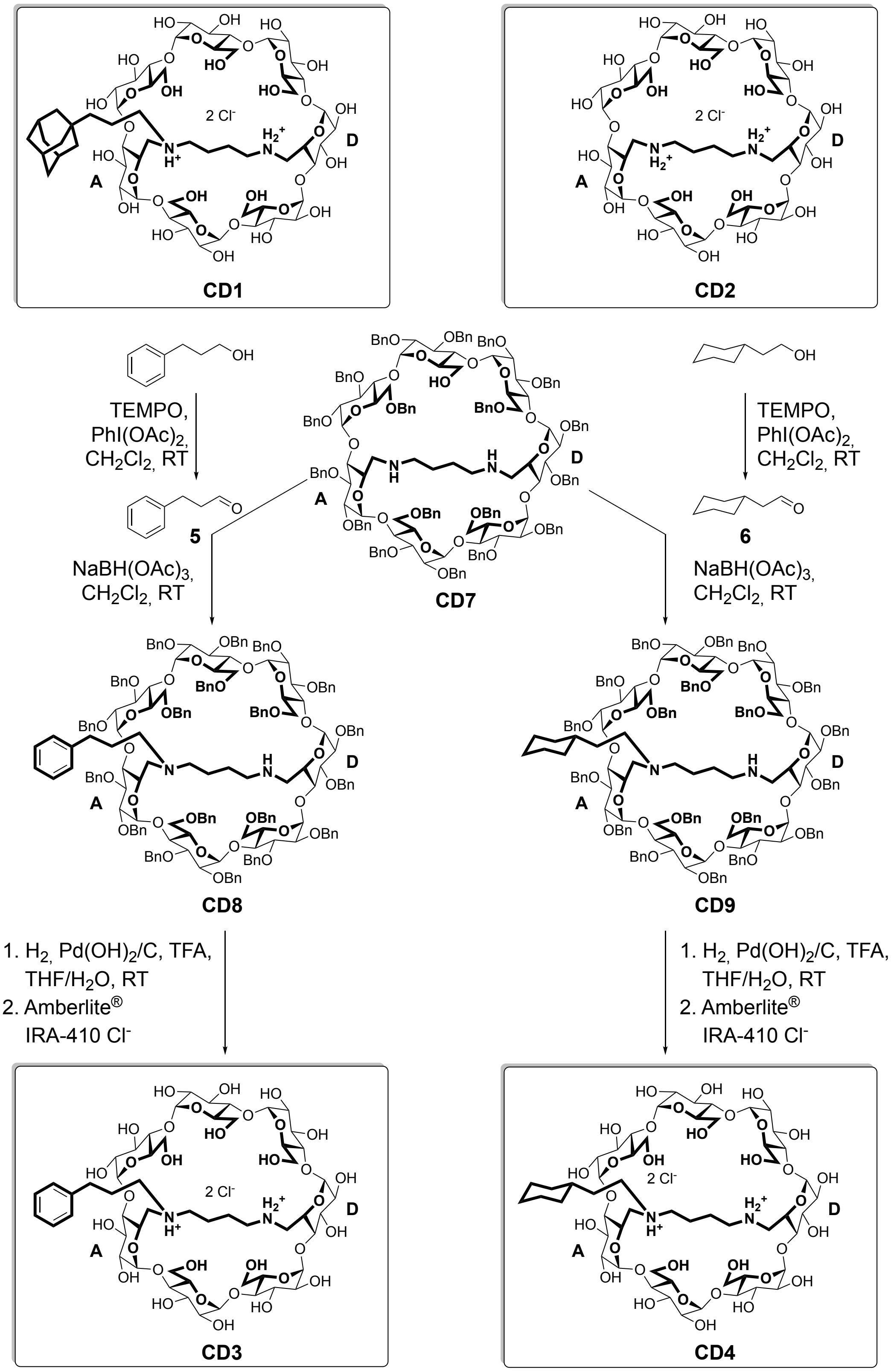
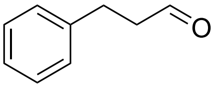


Figure S 24

S4.3 Synthesis of the hydrophobic group precursors (**5** and **6**)

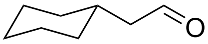
*S4.3.1 3-phenylpropanal (**5**)*



To a flask containing 3-phenylpropanol (763 mg, 5.60 mmol) in DCM (30 mL) under argon TEMPO (88 mg, 0.56 mmol, 0.1 eq.) then iodobenzene I,I-diacetate (2.0 g, 6.21 mmol, 1.1 eq) were added. The reaction was stirred for 2 h at RT. The reaction was followed by TLC (Cyclohexane/EtOAc 3:1, R_f aldehyde = 0.50, revelation CAM). Upon completion, the reaction mixture was diluted with DCM (50 mL). Saturated aqueous solution of $Na_2S_2O_3$ (100 mL) was added and extracted with DCM (2 x 50 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ (50 mL), brine (50 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The crude brown oil was purified by silica gel chromatography, eluted with ether/pentane (1:9) to provide 3-phenylpropanal **5** as a light red oil (749 mg, 100%).

The structure was confirmed by comparison with literature data.[5]

*S4.3.2 2-cyclohexylacetaldehyde (**6**)*



To a flask containing 2-cyclohexylethanol (1.1 g, 8.5 mmol) in DCM (40 mL) under argon, TEMPO (125 mg, 0.80 mmol, 0.1 eq.) then iodobenzene I,I-diacetate (2.84 g, 8.8 mmol, 1.1 eq) were added. The reaction was stirred for 3 h at RT. The reaction was followed by TLC (Cyclohexane/EtOAc 85:15, R_f aldehyde = 0.75, revelation $KMnO_4$). Upon completion, the reaction mixture was diluted with DCM (50 mL). Saturated aqueous solution of $Na_2S_2O_3$ (100 mL) was added and extracted with DCM (2 x 50 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ (50 mL), brine (50 mL), dried over $MgSO_4$ and concentrated under reduced pressure. The crude orange oil was purified by silica gel column chromatography and eluted with ether/pentane (5:95) to provide 2-cyclohexylacetaldehyde **6** as colorless oil (1.1 g, 100%).

The structure was confirmed by comparison with literature data.[6]

---

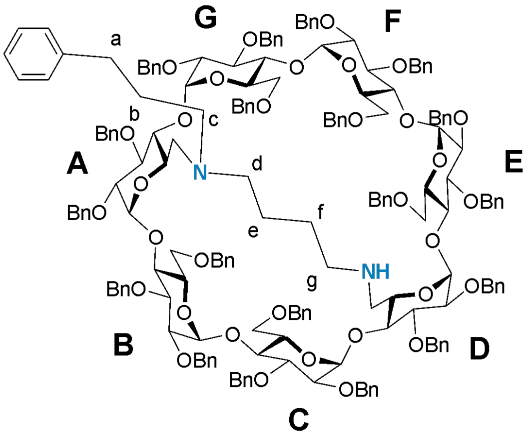
[5] **Selective Reduction of Carboxylic Acids to Aldehydes Catalyzed by $B(C_6F_5)_3$**, D. Bézier, S. Park, M. Brookhart, *Org. Lett.* **2013**, *15*, 496-499.

[6] **Assessing Synthetic Strategies: Total Syntheses of (±)-Neodolabellane-Type Diterpenoids**, C. Valente, M. G. Organ, *Chem. Eur. J.* **2008**, *14*, 8239-8245.

S4.4 Synthesis of the phenyl CD (**CD3**)

S4.4.1 Synthesis of ***CD8***

6^{A},6^{D}-Dideoxy-6^{A},6^{D}-diamino-N,N'-butyl-N-(propylphenyl)-N'-amine-$2^{A\text{-}G}$,$3^{A\text{-}G}$,$6^{B\text{-}C;\ E\text{-}G}$-nonadeca-O-benzyl-β-cyclodextrin



To a solution of **CD7** (1.34 g, 0.46 mmol) in DCM (200 mL) under an argon atmosphere, sodium triacetoxyborohydride (586 mg, 2.77 mmol, 5.1 eq) was added. Then 3-phenylpropanal **5** (68 mg, 0.51 mmol, 1.1 eq.) was added at 0 °C dropwise over 1 h. The reaction was stirred overnight at RT. The reaction was monitored by ESI-TOF MS. Upon completion, water (40 mL) was added, and layers were separated. The aqueous layer was extracted with DCM (2 x 30 mL). Organic layers were combined, washed with saturated $NaHCO_3$ aqueous solution (30 mL), brine (30 mL), dried with $MgSO_4$, filtered, and concentrated under reduced pressure. The resulting crude product was purified on silica gel (120 g) by flash chromatography, eluted with cyclohexane/EtOAc (gradient from 95:5 to 1:1 with 1% Et_3N at 1:1) to afford **CD8** AD-isomer (278 mg, 20%)[7,8] as a white foam.

Only the AD-isomer of **CD8** was fully analyzed and characterized.

R_f = 0.54 (Cyclohexane/EtOAc, 7:3, 5% H_2SO_4 in EtOH).

NMR ^{1}H (600 MHz, $CDCl_3$, 300 K): δ =7.36-6.94 (m, 100 H, 95 x H-Ar_{Bn}, 5 x H-Ar_{Ph}), 5.76 (d, ^3J = 3.9 Hz, 1 H, H-1C), 5.51 (d, ^2J = 10.4 Hz, 1 H, CH*H*-Ph), 5.36 (d, ^3J = 3.2 Hz, 1 H, H-1G), 5.28-5.23 (m, 3 H, 2 x CH*H*-Ph, H-1F), 5.20 (d, ^2J = 11.0 Hz, 1 H, CH*H*-Ph), 5.17 (d, ^2J = 10.9 Hz, 1 H, CH*H*-Ph), 4.98 (d, ^3J = 3.0 Hz, 1 H, H-1E), 4.95 (d, ^3J = 3.0 Hz, 1 H, H-1B), 4.89 (d, ^3J = 2.2 Hz, 1 H, H-1A), 4.87-4.83 (m, 2 H, 2 x CH*H*-Ph), 4.80-4.76 (m, 3 H, 3 x CH*H*-Ph), 4.74-4.67 (m, 6 H, 5 x CH*H*-Ph, H-1D), 4.61 (d, ^2J = 11.8 Hz, 1 H, CH*H*-Ph), 4.58-4.48 (m, 10 H, 10 x CH*H*-Ph), 4.46-4.38 (m, 8 H, 8 x CH*H*-Ph), 4.37-4.28 (m, 6 H, 5 x CH*H*-Ph, H-6_aB), 4.21 (m, 3 H, 2 x CH*H*-Ph,

---

[7] **CD8** was isolated with 12% residual imine compound.

[8] **CD8** was isolated from a mixture of compounds [diphenyl compound (272 mg, 19%), impure AE-isomer (172 mg), and the starting **CD7** (264 mg, 20%).

H-6$_a$E), 4.19-4.12 (m, 4 H, H-3G, H-3C, H-6$_a$C, H-6$_a$G), 4.08-3.98 (m, 9 H, H-3D, H-3F, H-3A, H-3E, H-3B, H-4C, H-4G, H-5D, H-5G), 3.93-3.83 (m, 8 H, H-4F, H-4E, H-4B, H-5A, H-5E, H-5B, H-5C, H-6$_a$F), 3.82 (m, 1 H, H-5F), 3.72 (m, 1 H, H-6$_b$C), 3.67 (m, 1 H, H-6$_b$G), 3.65 (m, 1 H, H-6$_b$B), 3.62 (m, 1 H, H-6$_b$F), 3.60-3.55 (m, 4 H, H-2G, H-2C, H-4D, H-6$_b$E), 3.49 (dd, ^{3}J = 3.6 Hz, ^{3}J = 9.8 Hz, 1 H, H-2F), 3.46-3.42 (m, 2 H, H-2E, H-2B), 3.35 (dd, ^{3}J = 2.9 Hz, ^{3}J = 9.6 Hz, 1H, H-2D), 3.33-3.28 (m, 2 H, H-2A, H-4A), 2.88 (m, 1 H, H-6$_a$D), 2.62-2.61 (m, 2 H, H-6$_b$D, H-a$_a$), 2.58-2.56 (m, 2 H, H-6$_a$A, H-6$_b$A), 2.54-2.48 (m, 3 H, H-a$_b$, H-c$_a$, H-d$_a$), 2.44-2.39 (m, 2 H, H-g$_a$, H-g$_b$), 2.17-2.08 (m, 2 H, H-c$_b$, H-d$_b$), 1.76-1.66 (m, 2 H, 2 x H-b), 1.45 (m, 2 H, H-f$_a$, H-e$_a$), 1.33 (m, 1H, H-f$_b$), 1.21 (m, 1H, H-e$_b$).

NMR ^{13}C (150 MHz, CDCl$_3$, 300 K): δ = 142.45 (C$^{IV\text{-}Ph}$), 139.79-138.16 (19 x C$^{IV\text{-}Ar}$), 128.60-125.89 (100 x CH-Ar), 100.36 (C-1E), 99.66 (C-1A), 99.43 (C-1G), 99.36 (C-1B), 98.49 (C-1F), 98.45 (C-1D), 98.25 (C-1C), 81.87 (2 C, C-4A, C-4F), 81.82 (C-4B or C-4E), 81.76 (C-4E or C-4B), 81.63 (C-3G), 81.16 (C-4C), 80.98 (C-3C), 80.94 (2 C, C-3E, C-3B), 80.72 (C-3D), 80.51 (2 x C, C-2D, C-3A), 80.25 (C-3F), 79.92 (C-2A), 79.53 (C-2B), 79.37 (C-2F), 78.64 (2 x C, C-2G, C-4G), 78.44 (C-2E), 78.08 (C-2C), 77.46 (C-4D), 76.32 (*C*H$_2$-Ph), 76.26 (*C*H$_2$-Ph), 76.14 (*C*H$_2$-Ph), 76.04 (*C*H$_2$-Ph), 75.50 (*C*H$_2$-Ph), 74.12 (*C*H$_2$-Ph), 73.82 (*C*H$_2$-Ph), 73.63-73.36 (7 C, 7 x *C*H$_2$-Ph), 73.15 (2 C, 2 x *C*H$_2$-Ph), 73.11 (2 C, 2 x *C*H$_2$-Ph), 73.00 (2 C, 2 x *C*H$_2$-Ph), 72.67 (2 C, 2 x *C*H$_2$-Ph), 72.43 (2 C, 2 x *C*H$_2$-Ph), 72.21 (C-5B), 72.10 (C-5E), 71.85 (2 C, 2 x *C*H$_2$-Ph), 71.79 (2 C, C-5F, C-5C), 71.18 (C-5G), 70.32 (C-5A), 70.10 (C-5D), 69.57 (C-6B), 69.25 (C-6F), 69.05 (C-6C), 69.00 (C-6E), 68.90 (C-6G), 59.31 (C-6A), 54.30 (C-c), 53.92 (C-d), 51.97 (C-6D), 48.98 (C-g), 34.00 (C-a), 29.18 (C-b), 26.98 (C-f), 25.02 (C-e).

ESI-HRMS (m/z) calculated for $C_{16}H_{32}N_3$ $[M + H]^+$ 266.2591, found 266.2583 (err: 2.9 ppm)

Table 3.3.1-1 Total NMR analysis of compound **CD8**

Cycle	Position label	^{13}C / δ (ppm)	^{1}H / δ (ppm)	Correlation	Attribution
A	1	99.66	4.89	NOESY 3.85	^{A}H-1 ↔ ^{B}H-4
	2	79.92	3.32		
	3	80.53	4.05		
	4	81.87	3.30	HMBC 99.43	^{A}H-4 ↔ ^{G}C-1
	5	70.32	3.866		
	6	59.31	2.58 and 2.56		
B	1	99.37	4.95	NOESY 4.04	^{B}H-1 ↔ ^{C}H-4
	2	79.53	3.45		
	3	80.25 or 80.96 or 80.99	4.00		
	4	81.76 or 81.82	3.85	HMBC 99.66	^{B}H-4 ↔ ^{A}C-1
	5	72.21	3.91		
	6	69.57	3.65 and 4.35		
C	1	98.25	5.76	NOESY 3.60	^{C}H-1 ↔ ^{D}H-4
	2	78.08	3.59		
	3	80.98	4.14		
	4	81.16	4.04	HMBC 99.37	^{C}H-4 ↔ ^{B}C-1
	5	71.79	3.92		
	6	69.05	3.72 and 4.19		
D	1	98.45	4.73	NOESY 3.84	^{D}H-1 ↔ ^{E}H-4
	2	80.52	3.36		
	3	80.72	4.03		
	4	77.47	3.60	HMBC 98.25	^{D}H-4 ↔ ^{C}C-1
	5	70.10	4.05		
	6	51.97	2.88 and 2.61		
E	1	100.36	4.98	NOESY 3.85	^{E}H-1 ↔ ^{F}H-4
	2	78.44	3.43		
	3	80.25 or 80.96 or 80.99	4.00		
	4	81.76 or 81.82	3.84	HMBC 98.45	^{E}H-4 ↔ ^{D}C-1
	5	72.10	3.87		
	6	69.00	3.56 and 4.21		
F	1	98.49	5.27	NOESY 4.08	^{F}H-1 ↔ ^{G}H-4
	2	79.37	3.49		
	3	80.25 or 80.96 or 80.99	4.00		
	4	81.87	3.85	HMBC 100.36	^{F}H-4 ↔ ^{E}C-1
	5	71.79	3.82		
	6	69.25	3.62 and 3.88		
G	1	99.43	5.36	NOESY 3.30	^{G}H-1 ↔ ^{A}H-4
	2	78.65	3.56		
	3	81.63	4.16		
	4	78.65	4.08	HMBC 98.49	^{G}H-4 ↔ ^{F}C-1
	5	71.18	4.06		
	6	68.90	3.67 and 4.17		

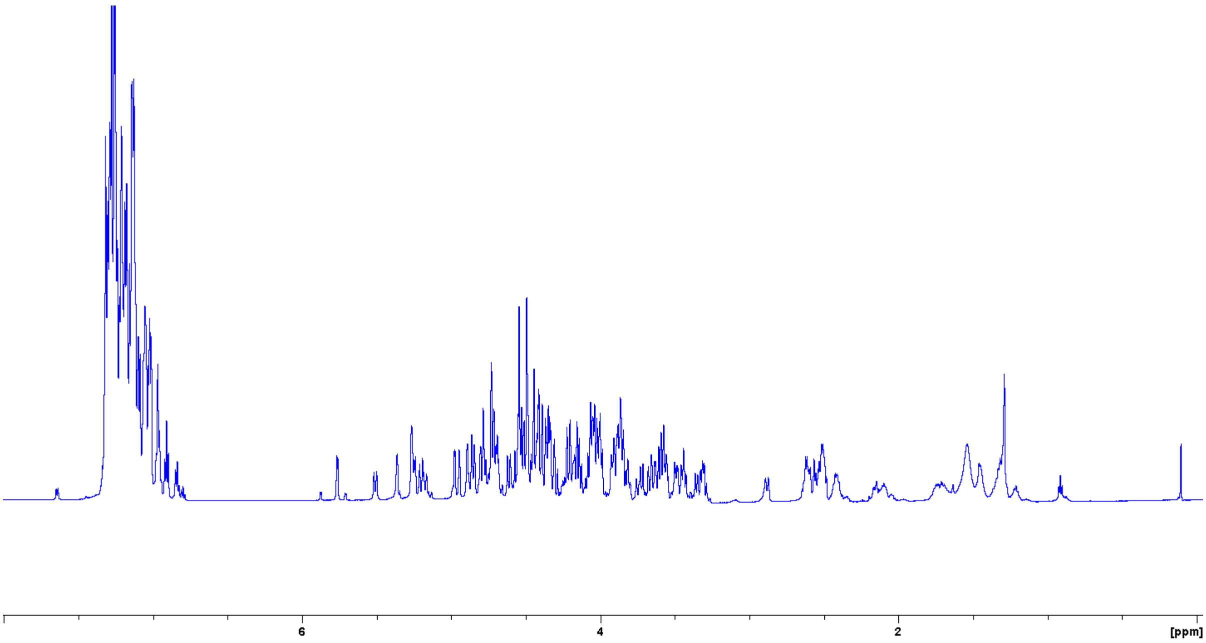


Figure S 25 NMR ^{1}H (600 MHz, $CDCl_3$, 300 K) of compound **CD8** (88% pure)

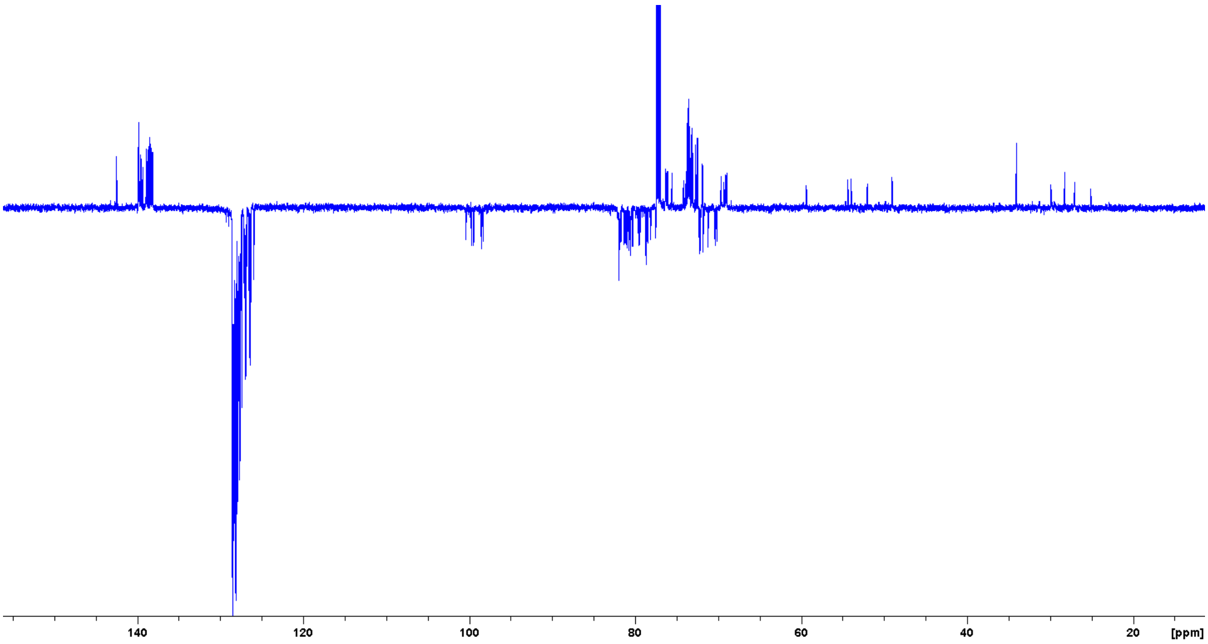
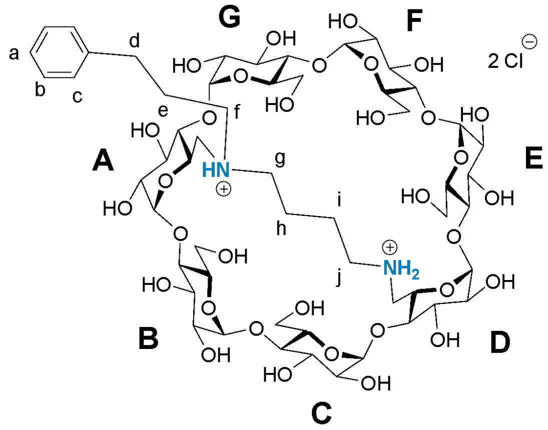


Figure S 26 NMR ^{13}C-Jmod (150 MHz, $CDCl_3$, 300 K) of compound **CD8**

S4.4.2 Synthesis of ***CD3***

6^A,6^D-Dideoxy-6^A,6^D-diamino-N,N-Bis(2-aminoethyl)-N-(3-phenyl)propyl)amine N'-ammonio-β-cyclodextrin chlorohydrate



The AD-isomer of **CD8** (130 mg, 0.043 mmol) was dissolved in a THF/H_2O mixture (9/3 mL) under argon in a 100 mL round bottom flask. TFA (32 µL, 0.43 mmol, 10 eq) and $Pd(OH)_2/C$ (130 mg, 0.19 mmol, 20% wt., 4.3 eq.) were added and the reaction mixture was stirred under a hydrogen atmosphere for 40 h. After 24 h, $Pd(OH)_2/C$ (70 mg, 0.10 mmol, 20% wt., 2 eq.) were added to speed up the reaction. Reaction was monitored carefully by ESI-TOF MS to check for undesired phenyl reduction. Once all the benzyl groups were removed, the resulting mixture was purged with argon, filtered through µpore filter, and rinsed with MeOH and water. Solvents were evaporated under reduced pressure and the resulting aqueous solution was lyophilized (m = 54 mg). The crude product was precipitated in acetone, filtered, and purified by flash chromatography on a reverse phase C-18 column with H_2O/CH_3CN (gradient in CH_3CN from 0% to 40%) to afford an amorphous powder.

The powder was redissolved in a minimum of water and eluted with water through an ion exchange column (Amberlite® IRA-410 Cl^-). After lyophilization the desired **CD3** was obtained as a white amorphous powder (21 mg, 37%).

NMR ^{1}H (600 MHz, D_2O, 300 K): δ =7.45 (m, 2 H, 2 x H-c), 7.36 (m, 3 H, H-a, 2 x H-b), 5.27 (d, ^3J= 3.6 Hz, 1 H, H-1C), 5.14 (d, ^3J = 3.9 Hz, 1H, H-1G), 5.13 (d, ^3J = 3.7 Hz, 1 H, H-1F), 5.11 (m, 3H, H-1A, H-1D, H-1B), 5.08 (d, ^3J = 3.4 Hz, 1 H, H-1E), 4.37 (m, 1 H, H-5A), 4.28 (m, 1 H, H-5D), 4.04-3.89 (m, 15 H, 7 x H-3, H-5F, H-5B, H-5E, H-6$_a$B, H-6$_a$C, H-6$_a$E, H-6$_a$F), 3.88-3.80 (m, 5 H, H-6$_a$G, H-6$_b$B, H-6$_b$C, H-6$_b$E, H-6$_b$F), 3.79- 3.61 (m, 13 H, 7 x H-2, H-4B, H-4E, H-4F, H-5C, H-5G, H-6$_b$G), 3.59-3.41 (m, 7 H, H-4A, H-4C, H-4D, H-4G, H-6$_a$A, H-6$_b$A, H-6$_a$D), 3.34 (m, 1 H, H-6$_b$D), 3.30-3.15 (m, 3 H, 2 x H-f, H$_a$-j), 3.10 (m, 3 H, 2 x H-g, H$_b$-j) 2.83 (m, 1 H, H$_a$-d), 2.71 (m, 1 H, H$_b$-d), 2.15 (m, 1 H, H$_a$-e), 2.02 (m, 1 H, H$_a$-e), 1.84 (m, 1 H, H$_a$-h), 1.74 (m, 1 H, H$_a$-i), 1.62 (m, 1 H, H$_b$-h), 1.47 (m, 1 H, H$_b$-i).

NMR ^{13}C (150 MHz, D_2O, 300 K): δ = 140.47 (C^{IV}-Ar), 128.89 (C-c), 128.67 (C-b), 126.70 (C-a), 102.47 (C-1G), 101.95 (C-1F), 101.21 (2 x C-1C, C-1A or C-1B or C-1D), 101.07 (C-1E), 99.88 (C-1A or C-1B or C-1D), 99.54 (C-1A or C-1B or C-1D), 84.29 (C-4A), 81.70 (C-4G), 81.42 (C-4C), 81.22 (C-4D), 80.96 (C-4F), 79.05 (C-4E), 78.18 (C-4B), 73.12 (2 x C, C-3B, C-5B), 72.82 (C-3F), 72.72 (C-3G), 72.63 (3 x C, C-3A, C-3D, C-3E), 72.31 (2 x C, C-5C, C-2B or C-2E), 72.15 (2 x C, C-5G, C-5E), 72.08 (3 x C, C-3C, C-2A or C-2D or C-2F), 72.31 (C-2B or C-2E), 71.87 (C-5F), 71.50 (C-2A or C-2D or C-2F), 71.34 (C-2G), 70.69 (C-2C), 65.17 (C-5D), 64.28 (C-5A), 61.09 (C-6G), 60.70 (C-6F), 60.45 (C-6C), 60.32 (C-6E), 59.94 (C-6B), 53.85 (C-6A), 53.70 (C-f), 48.55 (C-j), 45.89 (C-6D), 43.53 (C-g), 31.59 (C-d), 24.32 (C-e), 22.07 (C-h), 18.69 (C-i).

ESI-HRMS (m/z) calculated for $C_{55}H_{89}N_2O_{33}$ $[M + H]^+$ 1305.5342, found 1305.5338 (err: 0.3 ppm)

$[\alpha]^{20}_D$ = + 145.3 (c = 0.821)

Table 3.3.2-1 Total NMR analysis of compound **CD3**

Cycle	Position label	^{13}C / δ (ppm)	^{1}H / δ (ppm)	Correlation	Attribution
A	1	101.21 or 99.87 or 99.55	5.11	NOESY 3.74	^{A}H-1 ↔ ^{B}H-4
	2	72.08 or 71.50 or 72.07	3.70 or 3.69 or 3.67		
	3	72.63	4.01		
	4	84.29	3.52		
	5	64.28	4.38		
	6	53.85	3.47 and 3.48		
B	1	101.21 or 99.87 or 99.55	5.11	NOESY 3.57	^{B}H-1 ↔ ^{C}H-4
	2	72.31 or 71.97	3.70 or 3.69 or 3.67		
	3	73.14	4.02		
	4	78.18	3.74		
	5	73.14	3.93		
	6	59.94	3.87 and 3.91		
C	1	101.21	5.27	NOESY 3.55	^{C}H-1 ↔ ^{D}H-4
	2	70.69	3.73		
	3	72.08	3.96		
	4	81.42	3.57		
	5	72.29	3.77		
	6	60.45	3.85 and 3.97		
D	1	101.21 or 99.87 or 99.55	5.11	NOESY 3.71	^{D}H-1 ↔ ^{E}H-4
	2	72.08 or 71.50 or 72.07	3.70 or 3.69 or 3.67		
	3	72.63	4.01		
	4	81.22	3.55		
	5	65.17	4.28		
	6	45.89	3.33 and 3.42		
E	1	101.07	5.08	NOESY 3.64	^{E}H-1 ↔ ^{F}H-4
	2	72.31 or 71.97	3.67		
	3	72.63	4.00		
	4	79.05	3.71		
	5	72.15	3.96		
	6	60.32	3.92 and 3.97		
F	1	101.95	5.13	NOESY 3.50	^{F}H-1 ↔ ^{G}H-4
	2	72.08 or 71.50 or 72.07	3.71		
	3	72.82	3.98		
	4	80.96	3.64		
	5	71.87	3.92		
	6	60.70	3.84 and 3.96		
G	1	102.47	5.14	NOESY 3.52	^{G}H-1 ↔ ^{A}H-4
	2	71.34	3.74		
	3	72.72	3.95		
	4	81.70	3.50		
	5	72.15	3.76		
	6	61.09	3.73 and 3.85		

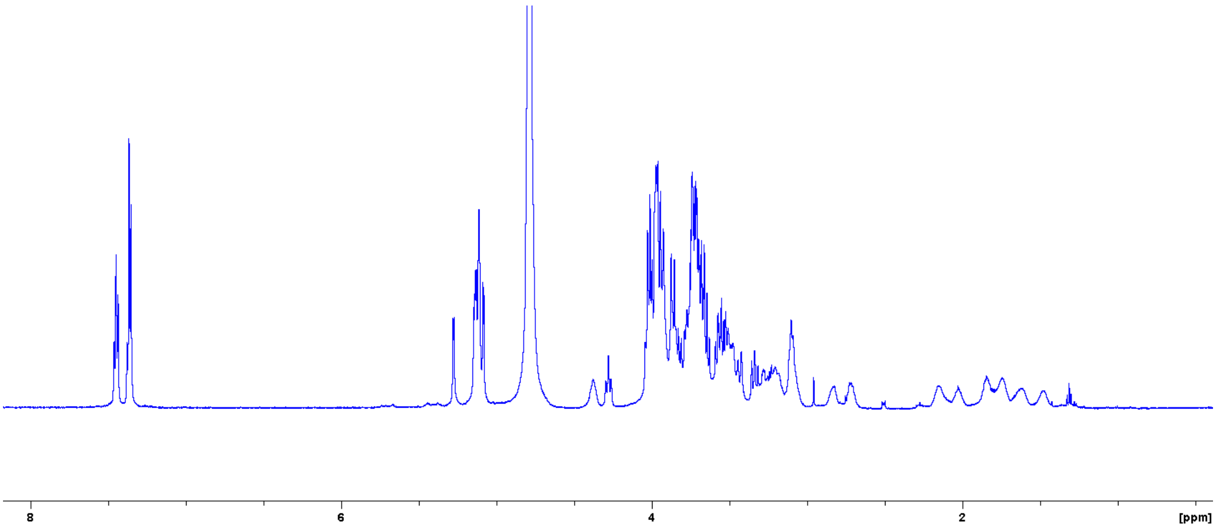


Figure S 27 NMR ^{1}H (600 MHz, D_2O, 300 K) of compound **CD3**

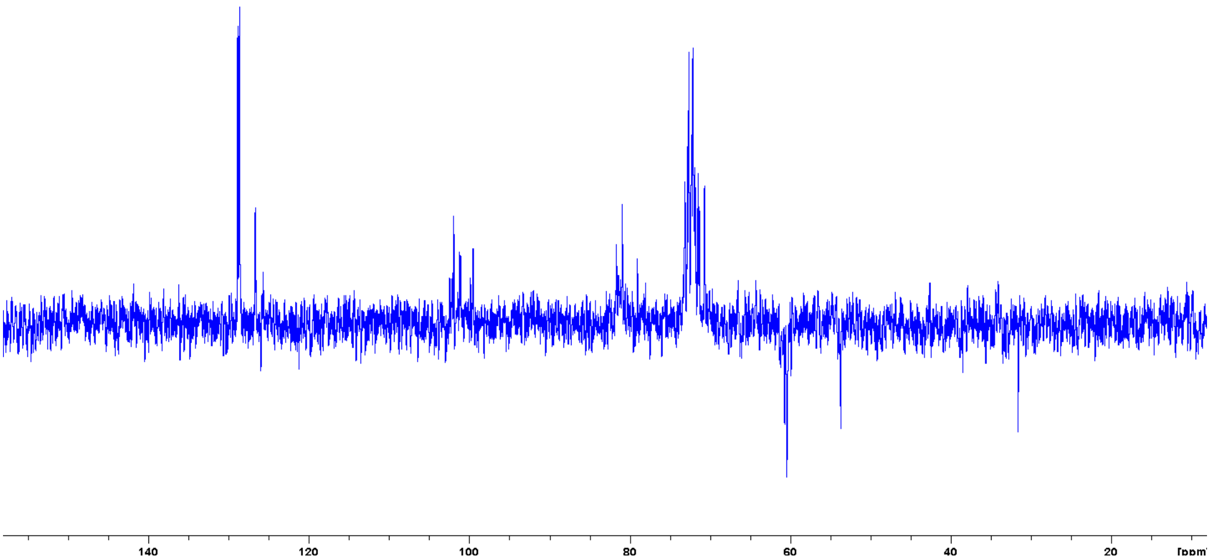
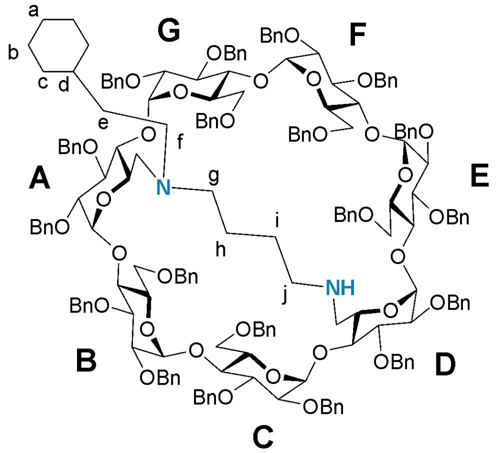


Figure S 28^{13}C-Jmod (150 MHz, D_2O, 300 K) of compound **CD3**

S4.5 Synthesis of the cyclohexyl CD (**CD4**)

S4.5.1 Synthesis of ***CD9***

6^{A},6^{D}-Dideoxy-6^{A},6^{D}-diamino-N,N'-butyl-N-(ethylcyclohexyl)-N'-amine-$2^{A\text{-}G}$,$3^{A\text{-}G}$,$6^{B\text{-}C;\ E\text{-}G}$-nonadeca-O-benzyl-β-cyclodextrin



To a solution of **CD7** (1.20 g, 0.41 mmol) in DCM (200 mL) under argon, sodium triacetoxyborohydride (614 mg, 2.90 mmol, 7 eq) was added. Then 2-cyclodexylacetaldehyde **6** (57 mg, 0.45 mmol, 1.1 eq.) was added at 0 °C dropwise over 1 h. The reaction was stirred overnight at RT. The reaction was monitored by ESI-TOF MS. Upon completion water (40 mL) was added, and layers were separated. The aqueous layer was extracted with DCM (2 x 30 mL). Organic layers were combined, washed with saturated $NaHCO_3$ aqueous solution (30 mL), brine (30 mL), dried with $MgSO_4$, filtered, and concentrated under reduced pressure. The resulting crude product was dissolved in toluene and purified on silica gel (80 g) by flash chromatography, eluted with cyclohexane/EtOAc (gradient from 95:5 to 1:1 with 1% Et_3N at 1:1) to afford the AD-isomer of **CD9** (326 mg, 26%)[9,10] as a white foam.

NMR ^{1}H (600 MHz, $CDCl_3$, 300 K): δ =7.41-6.81 (m, 95 H, 95 x H-Ar_{Bn}), 5.91 (d, ^{3}J = 4.0 Hz, 1 H, H-1G), 7.73 (d, ^{3}J = 4.1 Hz, 1 H, H-1C), 5.54 (app. d, ^{2}J = 10.3 Hz, 2 H, 2 x CH*H*-Ph), 5.38 (m, 2 H, 2 x CH*H*-Ph), 5.24 (d, ^{2}J = 10.6 Hz, 1 H, CH*H*-Ph), 4.95 (m, 2 H, CH*H*-Ph, H-1B), 4.89 (m, 3 H, CH*H*-Ph, H-1A, H-1E), 4.84-4.76 (m, 5 H, 5 x CH*H*-Ph), 4.74-4.63 (m, 6 H, 4 x CH*H*-Ph, H-1D, H-1F), 4.62-4.49 (m, 9 H, 9 x CH*H*-Ph), 4.48-4.38 (m, 7 H, 5 x CH*H*-Ph, H-6_aB, H-6_aE), 4.36-4.24 (m, 7 H, 5 x CH*H*-Ph, H-6_aC, H-6_aG), 4.22-3.94 (m, 18 H, 3 x CH*H*-Ph, 7 x H-3, H-4B, H-4C, H-4G, H-5A, H-5D, H-5E, H-5F, H-5G), 3.90-3.78 (m, 4 H, H-4E, H-5B, H-5C, H-6_aF), 3.74-3.58 (m, 6

---

[9] **CD9** adopts 2 conformations in NMR depending on the substitution conformation of the cyclohexyle unit : H-d equatorial or H-d axial. Only the H-d equatorial conformation was fully characterized and described below.

[10] **CD9** was isolated from a mixture of compounds, *i.e.* the dicyclohexyle compound (176 mg, 14%), an impure AE-isomer (118 mg) and the starting **CD7** (421 mg, 35%).

H, H-2C, H-2G, H-4F, H-6$_b$C, H-6$_b$G, H-6$_b$F), 3.52-3.42 (m, 6 H, H-2B, H-2E, H-4A, H-4D, H-6$_b$B, H-6$_b$E), 3.38 (dd, ^{2}J = 9.9 Hz, ^{3}J = 3.5 Hz, 1 H, H-2F), 3.35 (dd, ^{2}J = 9.9 Hz, ^{3}J = 3.4 Hz, 1 H, H-2A), 3.30 (dd, ^{2}J = 9.8 Hz, ^{3}J = 3.0 Hz, 1 H, H-2D), 2.77 (m, 2 H, H-6$_a$D, H-f$_a$), 2.65-2.34 (m, 7 H, H-6$_a$A, H-6$_b$A, H-6$_b$D, H-f$_b$, H-g$_a$, H-j$_a$, H-j$_b$), 2.30-2.17 (m, 3 H, 2 x H-c$_{eq}$, H-b$_{eq}$), 2.10 (m, 1 H, H-b$_{eq}$), 2.02 (m, 1 H, H-g$_b$), 1.91 (m, 1 H, H-a$_{eq}$), 1.74-1.21 (m, 11 H, **H-d$_{eq}$**, 2 x H-e, H-a$_{ax}$, 2 x H-b$_{ax}$, 2 x H-c$_{ax}$, H-i$_a$, 2 x H-h), 1.12 (m, 1 H, H-i$_b$).

NMR ^{13}C (150 MHz, CDCl$_3$, 300 K): δ = 139.97-138.09 (19 x C$^{IV\text{-}Ar}$), 128.56-126.11 (95 x CH-Ar), 102.27 (C-1E), 100.35 (C-1D), 99.90 (C-1F), 99.19 (C-1C), 98.49 (C-1B), 97.88 (C-1G), 96.55 (C-1A), 83.84 (C-4F), 88.44 (C-4E), 82.37 (C-3D), 82.23 (C-3A), 81.52 (C-3G or C-3C), 81.16 (3 C, C-3B or C-3F, C-4C or C-4B, C-3G or C-3C),), 80.95 (C-3B or C-3F), 80.88 (C-4G), 80.55 (C-3E), 80.39 (C-2D), 80.29 (C-2A), 79.77 (C-4C or C-4B), 79.48 (C-2B or C-2E), 79.34 (C-2F), 79.01 (C-4D), 78.54 (C-2C or C-2G), 78.43 (C-2B or C-2E), 77.90 (C-2C or C-2G), 76.75 (*C*H$_2$-Ph), 76.56 (*C*H$_2$-Ph), 76.52 (*C*H$_2$-Ph), 76.45 (*C*H$_2$-Ph), 76.25 (*C*H$_2$-Ph), 76.00 (C-4A), 74.43 (*C*H$_2$-Ph), 73.75 (*C*H$_2$-Ph), 73.62-73.32 (6 C, 6 x *C*H$_2$-Ph), 73.16 (*C*H$_2$-Ph), 72.96 (2 C, 2 x *C*H$_2$-Ph), 72.46 (*C*H$_2$-Ph), 72.27 (*C*H$_2$-Ph), 72.20 (*C*H$_2$-Ph), 72.10 (2 C, C-5B, C-5G), 71.80 (C-5C), 71.73 (C-5F), 71.57 (C-5E), 70.56 (C-5D), 70.39 (C-5A), 69.66 (C-6B), 69.52 (C-6F), 69.07 (C-6G), 68.98 (C-6C), 68.26 (C-6E), 57.95 (C-6A), 54.79 (C-g), 52.96 (C-6D), 52.83 (C-f), 49.06 (C-j), 37.23 (C-d), 36.81 (C-e), 34.41 (C-c), 34.22 (C-c), 27.31 (C-b), 27.20 (C-b), 27.06 (C-a), 25.74 (C-i), 24.72 (C-h).

Table 3.3.2-2 NMR analyses of **CD9**

Cycle	Position label	^{13}C / δ (ppm)	^{1}H / δ (ppm)	Correlation	Attribution
A	1	96.55	4.90	NOESY 4.11	^{A}H-1 ↔ ^{B}H-4
	2	80.29	3.36		
	3	82.23	4.04		
	4	76.00	3.50		
	5	70.39	4.05		
	6	57.95	2.37 and 2.56		
B	1	98.49	4.95	NOESY 4.11	^{B}H-1 ↔ ^{C}H-4
	2	79.48 or 78.43	3.48		
	3	81.16 or 80.95	3.99		
	4	81.17 or 79.78	4.11		
	5	72.10	3.86		
	6	69.66	3.45 and 4.44		
C	1	99.19	5.73	NOESY 3.47	^{C}H-1 ↔ ^{D}H-4
	2	77.90 or 78.55	3.61		
	3	81.16 or 81.59	4.02		
	4	81.17 or 79.78	4.11		
	5	71.80	3.89		
	6	68.98	3.71 and 4.31		
D	1	100.37	4.73	NOESY 3.85	^{D}H-1 ↔ ^{E}H-4
	2	80.39	3.30		
	3	82.37	3.99		
	4	79.01	3.47		
	5	70.56	3.97		
	6	52.96	2.62 and 2.81		
E	1	102.27	4.88	NOESY 3.66	^{E}H-1 ↔ ^{F}H-4
	2	79.48 or 78.43	3.48		
	3	80.55	4.05		
	4	82.44	3.85		
	5	71.57	3.96		
	6	68.26	3.48 and 4.42		
F	1	99.90	4.73	NOESY 4.01	^{F}H-1 ↔ ^{G}H-4
	2	79.34	3.39		
	3	81.16 or 80.95	3.99		
	4	83.84	3.66		
	5	71.73	3.99		
	6	69.52	3.73 and 3.79		
G	1	97.88	5.91	NOESY 3.51	^{G}H-1 ↔ ^{A}H-4
	2	78.55 or 77.90	3.61		
	3	81.16 or 81.59	4.02		
	4	80.88	4.01		
	5	72.10	4.20		
	6	69.07	3.61 and 4.27		

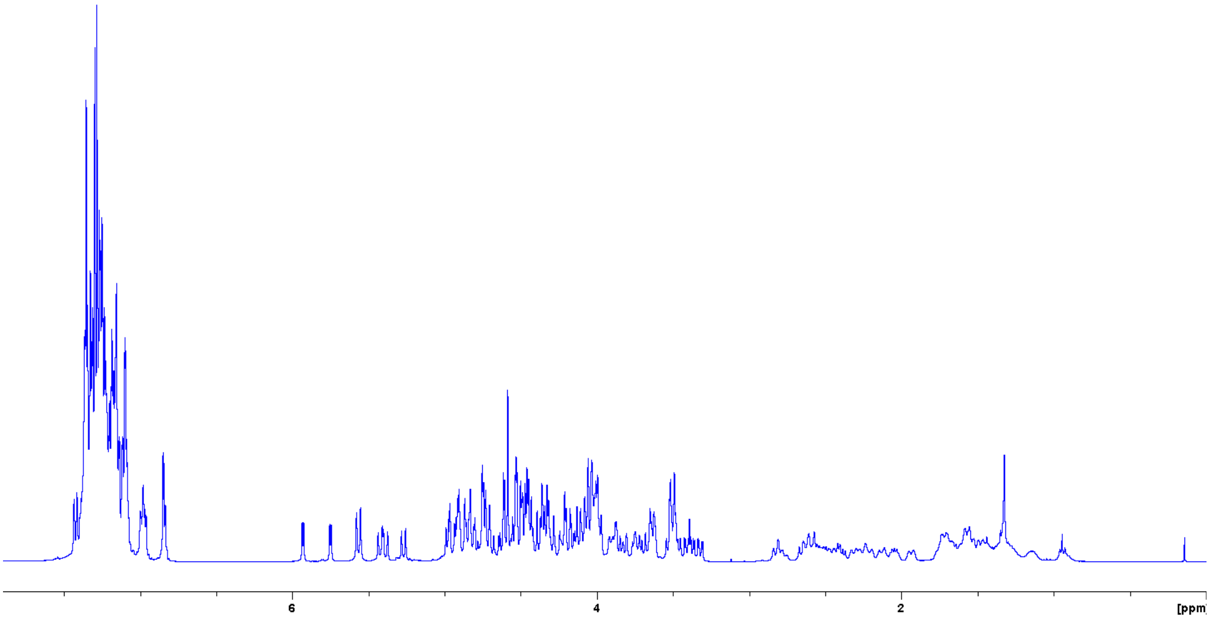


Figure S 29 NMR ^{1}H (600 MHz, CDCl$_3$, 300 K) of compound **CD9**.

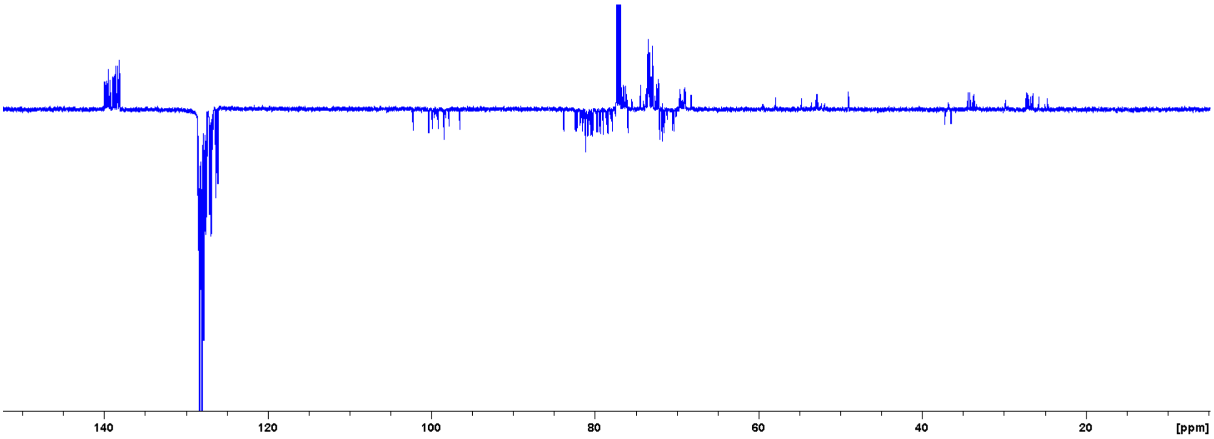
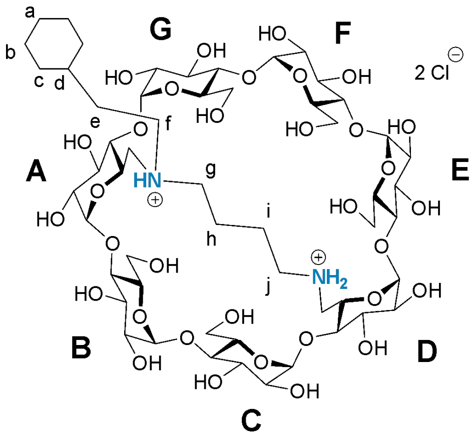


Figure S 30 ^{13}C-J$_{mod}$ (150 MHz, CDCl$_3$, 300 K) of compound **CD9.**

*S4.5.2 Synthesis of **CD4***

6^A,6^D-Dideoxy-6^A,6^D-diamino-N,N-Bis(2-aminoethyl)-N-(2-cyclohexyl)ethyl) amine N'-ammonio-β-cyclodextrin chlorohydrate



The perbenzylated bridged-mono-cyclohexyl cyclodextrin **CD9** (200 mg, 0.066 mmol) was dissolved in a THF/H_2O mixture (18/6 mL) under argon in a 100 mL round bottom flask. TFA (51 µL, 0.66 mmol, 10 eq) and $Pd(OH)_2/C$ (200 mg, 0.285 mmol, 20% wt., 4.3 eq.) were added and the reaction mixture was stirred under a hydrogen atmosphere for 40 h. Reaction was monitored by ESI-TOF MS. Once all the benzyl groups were removed, the resulting mixture was purged with argon, filtered through Celite, and rinsed with MeOH and water. Solvents were evaporated under reduced pressure and the resulting aqueous solution was lyophilized (m = 150 mg). The crude product was purified by flash chromatography with a reverse phase C-18 column with H_2O/CH_3CN (gradient in CH_3CN from 0% to 40%) to afford an amorphous powder.

The powder was redissolved in a minimum of water and eluted with water through an ion exchange column (Amberlite® IRA-410 Cl^-). After lyophilization the desired bridged-mono-cyclohexyl-β-cyclodextrin **CD4** was obtained as a white amorphous powder (52 mg, 60%).

NMR ^{1}H (600 MHz, D_2O, 300 K): δ = 5.28 (d, ^3J= 3.7 Hz, 1 H, H-1C), 5.18 (d, ^3J = 3.8 Hz, 1H, H-1G), 5.14 (d, ^3J = 3.7 Hz, 1 H, H-1F), 5.12 (m, 3H, H-1A, H-1D, H-1B), 5.08 (d, ^3J = 3.5 Hz, 1 H, H-1E), 4.27 (m, 1 H, H-5A), 4.20 (m, 1 H, H-5D), 4.04-3.89 (m, 17 H, 7 x H-3, H-5F, H-5B, H-5E, H-6$_a$B, H-6$_b$B, H-6$_a$C, H-6$_a$E, H-6$_a$F, H-6$_a$G, H-6$_b$G), 3.88-3.80 (m, 2 H, H-6$_b$C, H-6$_b$E), 3.79- 3.60 (m, 13 H, 7 x H-2, H-4B, H-4E, H-4F, H-5C, H-5G, H-6$_b$F), 3.59-3.53 (m, 5 H, H-4A, H-4C, H-4D, H-6$_a$A, H-6$_b$A), 3.50 (m, 1 H, H-4G), 3.44 (m, 1 H, H-6$_a$D), 3.39 (m, 1 H, H-6$_b$D), 3.36-3.25 (br., 4 H, 2 x H-f, H$_a$-g, H$_a$-j), 3.15-3.06 (br., 2 H, H$_b$-g, H$_b$-j), 1.87-1.72 (m, 9 H, H$_a$-h, H$_b$-h, H$_a$-i, H$_b$-i, 2 x H$_a$-b, 2 x H$_a$-c, H$_a$-e), 1.67 (m, 2 H, H$_b$-a, H$_b$-e), 1.46 (m, 1 H, H-d), 1.36-1.22 (m, 3 H, H$_b$-a, 2 x H$_b$-b), 1.10 (m, 2 H, 2 x H$_b$-c).

NMR ^{13}C (150 MHz, D_2O, 300 K): δ = 102.50 (C-1G), 101.99 (C-1F), 101.40 (2 x C-1C, C-1A or C-1B or C-1D), 101.21 (C-1E), 99.82 (C-1A or C-1B or C-1D), 99.58 (C-1A or C-1B or C-1D), 84.30 (C-4A), 81.84 (C-4G), 81.53 (2 x C, C-4C, C-4D), 81.14 (C-4F), 78.83 (2 x C, C-4E, C-4B), 73.39 (C-3G), 73.26 (C-3F), 73.00 (C-3C), 72.87 (C-5E), 72.82 (C-5B), 72.71 (C-3A or C-3B or C-3D), 72.70 (C-3E), 72.68 (C-3A or C-3B or C-3D), 72.57 (C-3A or C-3B or C-3D), 72.50 (C-5G), 72.29 (C-2E), 72.18 (C-2A or C-2D or C-2B), 72.11 (C-5F), 72.02 (C-5C), 71.95 (C-2F), 71.93 (C-2A or C-2D or C-2B), 71.39 (C-2A or C-2D or C-2B), 71.28 (C-2G), 70.62 (C-2C), 65.19 (C-5D), 64.48 (C-5A), 61.31 (C-6G), 60.80 (C-6C), 60.50 (C-6B), 60.23 (C-6E), 59.98 (C-6G), 53.37 (C-f), 53.05 (C-6A),49.47 (C-j or C-g), 45.70 (C-6D), 43.36 (C-g or C-j), 35.28 (C-d), 32.45 (C-c), 30.70 (C-e), 26.40 (C-a), 25.91 (C-b), 22.53 (C-i or C-h), 19.22 (C-i or C-h).

ESI-HRMS (m/z) calculated for $C_{54}H_{93}N_2O_{33}$ $[M + H]^+$ 1297.5655, found 1297.5682 (err: -2.0 ppm)

$[\alpha]^{20}_D$ = +160.0 (c = 0.77)

Table 3.3.3-1 : Total NMR analysis of **CD4**

Cycle	Position label	^{13}C / δ (ppm)	^{1}H / δ (ppm)	Correlation	Attribution
A	1	101.40 or 99.82 or 99.58	5.12	NOESY 3.74	^{A}H-1 ↔ ^{B}H-4
	2	72.18 or 71.93 or 71.39	3.74 or 3.72 or 3.68		
	3	72.71 or 72.68 or 72.57 or 72.70	3.99		
	4	84.30	3.59		
	5	64.48	4.27		
	6	53.05	3.57 and 3.59		
B	1	101.40 or 99.82 or 99.58	5.12	NOESY 3.57	^{B}H-1 ↔ ^{C}H-4
	2	72.18 or 71.93 or 71.39	3.74 or 3.72 or 3.68		
	3	72.71 or 72.68 or 72.57 or 72.70	3.99		
	4	78.85	3.74		
	5	72.93	3.97		
	6	60.50	3.94 and 4.00		
C	1	101.40	5.28	NOESY 3.58	^{C}H-1 ↔ ^{D}H-4
	2	70.62	3.76		
	3	72.99	3.93		
	4	81.53	3.57		
	5	72.03	3.72		
	6	60.80	3.82 and 3.99		
D	1	101.40 or 99.82 or 99.58	5.12	NOESY 3.74	^{D}H-1 ↔ ^{E}H-4
	2	72.18 or 71.93 or 71.39	3.74 or 3.72 or 3.68		
	3	72.71 or 72.68 or 72.57 or 72.70	3.99		
	4	81.55	3.58		
	5	65.19	4.20		
	6	45.70	3.39 and 3.44		
E	1	101.21	5.08	NOESY 3.63	^{E}H-1 ↔ ^{F}H-4
	2	72.29	3.68		
	3	72.71 or 72.68 or 72.57 or 72.70	3.98		
	4	78.85	3.74		
	5	72.87	3.92		
	6	60.23	3.86 and 3.94		
F	1	101.99	5.14	NOESY 3.50	^{F}H-1 ↔ ^{G}H-4
	2	71.95	3.73		
	3	73.26	3.97		
	4	81.14	3.63		
	5	72.11	3.91		
	6	61.31	3.77 and 3.96		
G	1	102.50	5.18	NOESY 3.57	^{G}H-1 ↔ ^{A}H-4
	2	71.28	3.77		
	3	73.39	3.92		
	4	81.84	3.50		
	5	72.51	3.78		
	6	59.98	3.89 and 3.96		

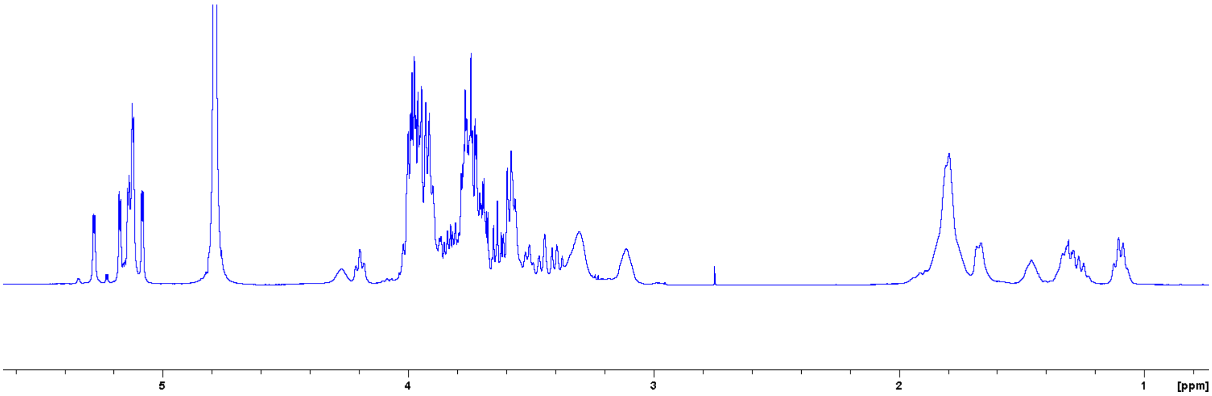


Figure S 31Figure 3.4-5 NMR ^1H (600 MHz, D_2O, 300 K) of **CD4**

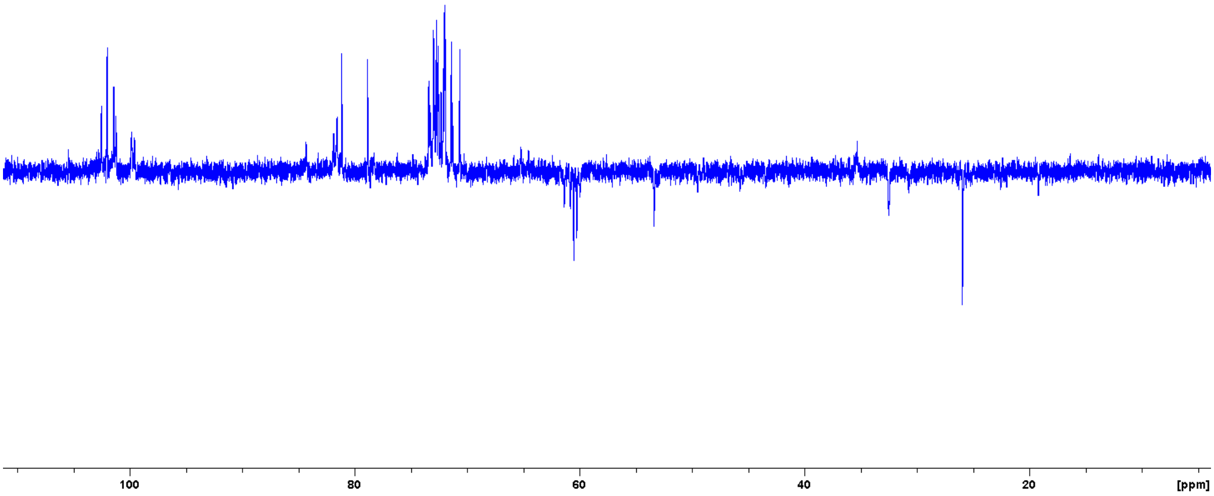


Figure S 32 Figure 3.4-6 ^{13}C-DEPT135 (150 MHz, D_2O, 300 K) of **CD4**

S5 Surface Plasmon Resonance (SPR)

S5.1 Supplementary Text

To thermodynamically investigate the interaction between **CD1** and dsDNA, we conducted Surface Plasmon Resonance (SPR) experiments using SPR sensor chips grafted with dsDNA of 10, 20 or 60 base pairs and by injecting solutions of **CD1** of 10 nM to 100 µM at 50 µL min^{-1}. Analysis of the sensorgrams allows us to assess the Bmax (maximum number of bound **CD1** per supported dsDNA), the K_d (apparent dissociation constant of **CD1** protofibrils with the supported dsDNA) and the Hill coefficient (apparent degree of cooperativity). The Bmax values for the supported $dsDNA_{10}$, $dsDNA_{20}$, $dsDNA_{60}$ correspond to approximately 21, 42 and 166 CDs per DNA, respectively. These values, when related to the length of the dsDNA, are consistent with between 6 and 8 **CD1** protofibrils surrounding the dsDNA.

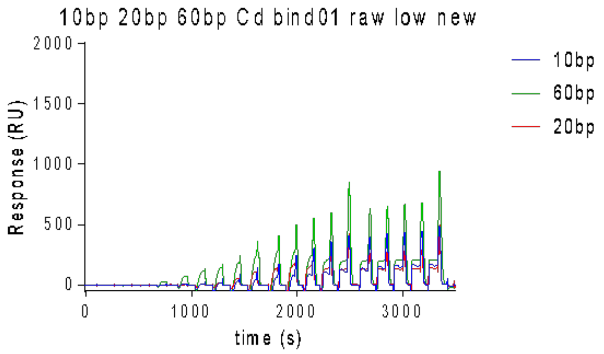
S5.2 Material and methods

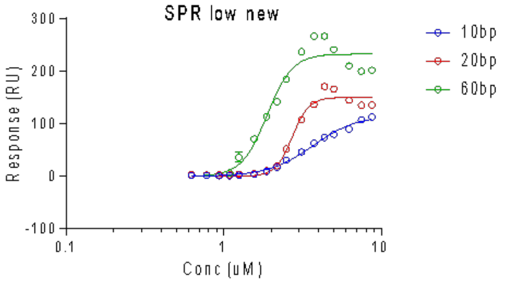
All SPR experiments were carried out at the National Cancer Institute (Amsterdam, Netherlands) using Biacore t200 (GE Healthcare) at 25°C. Biotinylated dsDNAs (10bp, 20bp, 60bp) were immobilized on the streptavidin SA sensor chip in HEPES buffer (20 mM HEPES, 0.05% Tween 20, pH 7.4) at different densities. The densities were calculated using RU difference between work cell and the reference cell. **CD1** solutions were injected with a flow rate 50 µL/min and the concentration range was 10 nm-100 µM.

Data analyses was performed using GraphPad Prism 8.0. The curves were fitted using Hill-Langmuir model. dsDNA was immobilized as following:

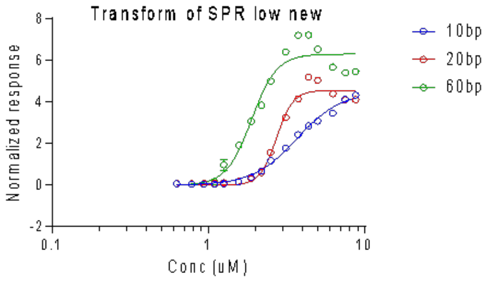
- *Low density* ($dsDNA_{10}$ - 26 RU, $dsDNA_{20}$ - 33 RU, $dsDNA_{60}$ - 37 RU)
- *Medium density* ($dsDNA_{10}$ - 156 RU, $dsDNA_{20}$ – 167 RU, $dsDNA_{60}$ - 177 RU)
- *High density* ($dsDNA_{10}$ - 292 RU, $dsDNA_{20}$ - 414 RU, $dsDNA_{60}$ - 407 RU).

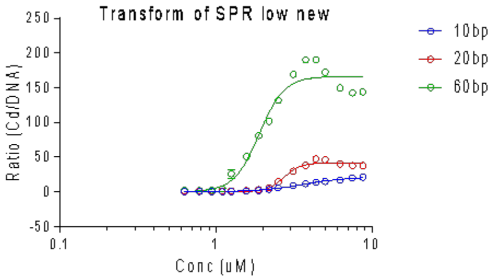
S5.3 Low density dsDNA





	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	116.216	2.636	3.077	0.141	3.685	0.084
20bp	149.953	2.703	8.137	0.839	2.731	0.045
60bp	232.641	5.225	5.379	0.661	1.888	0.049



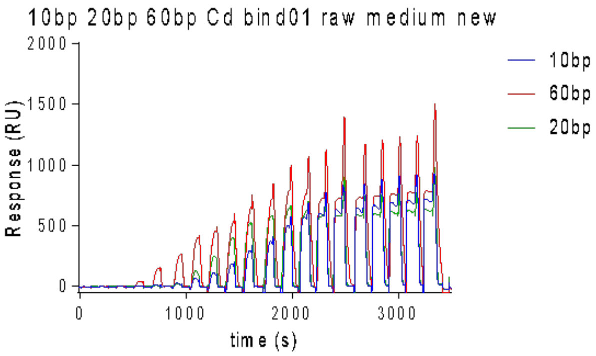


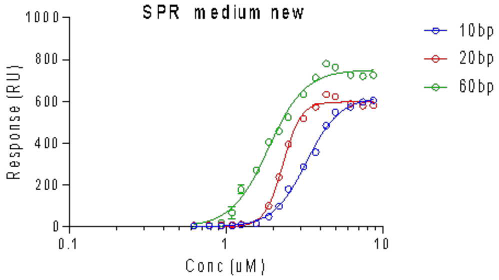
	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	4.470	0.101	3.077	0.141	3.685	0.084
20bp	4.544	0.082	8.137	0.839	2.731	0.045
60bp	6.288	0.141	5.379	0.661	1.888	0.049

	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	21.522	0.488	3.077	0.141	3.685	0.084
20bp	41.654	0.751	8.137	0.839	2.731	0.045
60bp	166.172	3.732	5.379	0.661	1.888	0.049

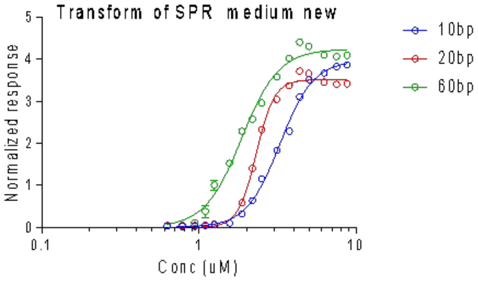
Figure S 33

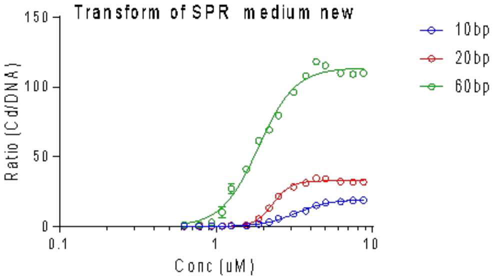
S5.4 Medium density dsDNA





	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	619.347	6.724	4.069	0.130	3.271	0.034
20bp	597.660	4.384	7.503	0.361	2.313	0.016
60bp	751.069	9.197	3.686	0.185	1.870	0.030



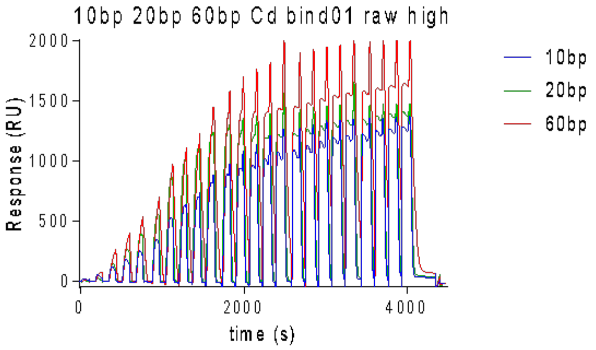


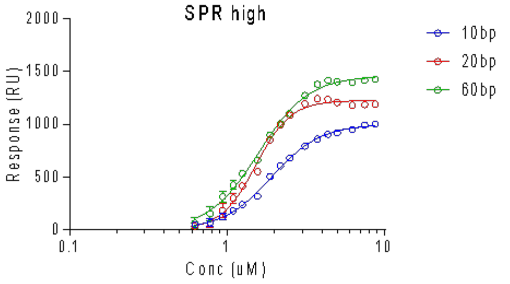
	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	3.970	0.043	4.069	0.130	3.271	0.034
20bp	3.516	0.026	7.503	0.361	2.313	0.016
60bp	4.243	0.052	3.686	0.185	1.870	0.030

	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	19.294	0.209	4.069	0.130	3.271	0.034
20bp	32.659	0.240	7.503	0.361	2.313	0.016
60bp	113.798	1.394	3.686	0.185	1.870	0.030

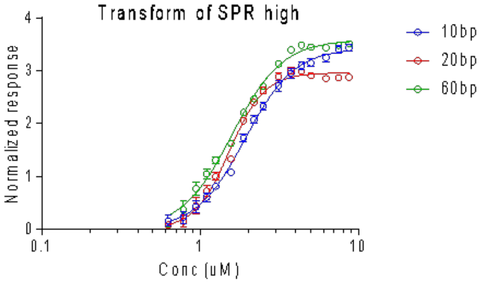
Figure S 34

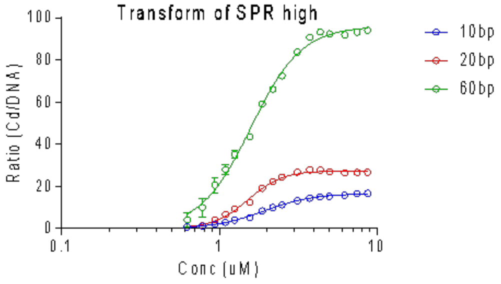
S5.5 High density dsDNA





	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	1000.355	9.771	2.797	0.087	1.923	0.026
20bp	1225.950	11.529	3.919	0.170	1.527	0.020
60bp	1459.242	13.844	2.732	0.092	1.589	0.023





	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	3.426	0.033	2.797	0.087	1.923	0.026
20bp	2.961	0.028	3.919	0.170	1.527	0.020
60bp	3.585	0.034	2.732	0.092	1.589	0.023

	Bmax		h		Kd	
	Value	SEM	Value	SEM	Value	SEM
10bp	16.645	0.163	2.797	0.087	1.923	0.026
20bp	27.488	0.259	3.919	0.170	1.527	0.020
60bp	96.003	0.911	2.732	0.092	1.589	0.023

Figure S 35

S6 cryo-EM for 3D reconstruction

S6.1 Sample preparation for data collection

For cryo-EM data collection, three microliters of the assembly formed by **CD1** (0.7 mM) and dsDNA$_{18}$ (6 µM) in TBE were placed on glow discharged (25 s) Lacey Carbon supported Cu 300 mesh grids. The grids were blotted for 3.0 s using a blotting force of -5, and flash-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fischer Scientific) operated at 4°C and 100% humidity. A total of 20,091 micrographs were collected on a Titan Krios G3 microscope at NanoImaging Core Facility at Pasteur Institute (Paris) operating at 300 kV equipped with a Gatan K3 Direct Detector and Bioquantum/K3 energy Filter. The automation of the data collection was done with the software EPU. Movies were recorded at 105,000x nominal magnification in electron-counting mode using exposures of 2s dose-fractionated into 36 frames with a dose rate of 1.7139 electrons/Å^2, resulting in a total dose of 61.7 electrons/Å. A defocus range of – 0.8 to – 2.6 µm was used. The pixel size was 0.86 Å per pixel.

S6.2 Cryo-EM data processing

After a visual inspection to remove poor-quality micrographs (containing thick ice, low contrast, abnormal background, and overlapping fibres), a total of 2,785 micrographs were selected for further cryo-EM processing. The movies were motion-corrected and dose-weighted using MotionCor2[11] with the removal of the two first frames. The contrast function parameters were estimated using CTFFIND4.[12] Given the severe problem of overlapping fibres (Figure S 37A), a manual picking was performed, so most of the fibres were separated from each other's, using Xmipp 3.0[13] within the Scipion 3 software framework,[14] resulting in 346,248 selected particles. The subsequent image processing steps were performed using RELION[15] within Scipion3. Several box sizes, ranging from 516 Å to 55 Å, were tested to extract particles from the images. The best results were obtained from two different box sizes of 130 Å and 198 Å (150 and 230 pixels, respectively) which were then subjected to one round of 2D classification. With the box size of 130 Å, a total of 215,869 particles were

---

[11] S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

[12] A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol*. **192**, 216–221 (2015).

[13] J. M. de la Rosa-Trevin, J. Otón, R. Marabini, A. Zaldívar, J. Vargas, J. M. Carazo, C. O. Sorzano, Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J. Struct. Biol*. **184**, 321-328 (2013).

[14] J. M. de la Rosa-Trevin, A. Quintana, L. del Cano, A. Zaldívar, I. Foche, J. Gutiérrez, J. Gómez-Blanco, J. Burguet-Castell, J. Cuenca-Alba, V. Abrishami, J. Vargas, J. Otón, G. Sharov, J. L. Vilas, J. Navas, P. Conesa, M. Kazemi, R. Marabini, C. O. Sorzano, J. M. Carazo, Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol*. **195**, 93-99 (2016).

[15] J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl, S. H. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

extracted from nine 2D class averages and used for 3D initial model, with no symmetry applied (C_1). We noticed that the power spectra of the 2D class averages were of poor quality and were not compatible with the presence of helical symmetry of the fibre. When applying high-resolution 3D refinement using 3D auto-refine in Relion, we implemented C_1 (no symmetry), C_6 and C_7 symmetry.

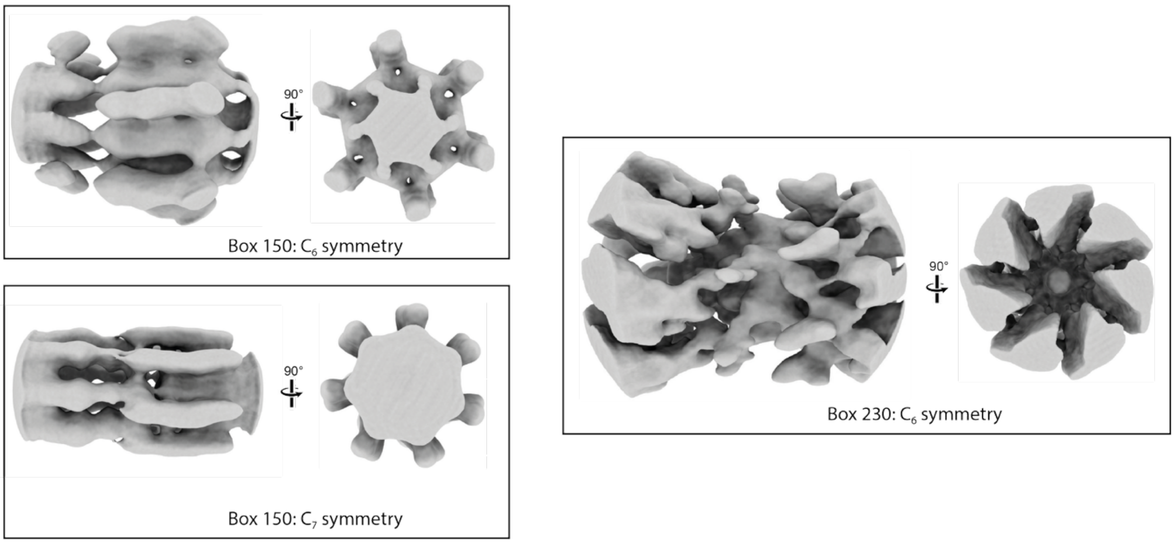


Figure S 36. Cryo-EM maps obtained applying C_6 and C_7 symmetries.
(left) Particles extracted using a box size of 150 pixels (130 Å).
(right) Particles extracted using a box size of 230 pixels (198 Å).

The C_1 (no symmetry) provided the most satisfactory results. Similarly, using a box size of 198 Å, a total of 272,287 particles were extracted from three 2D class averages and subjected to 3D initial model and 3D refinement following the same procedure as for the smaller box. Figure S 37B shows the cryo-EM processing workflow. The Fourier shell correlation (FSC) curves indicate the global resolution for the final cryo-EM maps from particles extracted from box sizes of 130 Å and 198 Å of 3.3 Å (Figure S 38A left) and 5.0 Å (Figure S 38B left), respectively (gold-standard FSC = 0.143). The final maps were subjected to B-factor map sharpening according to the resolution values obtained using the Auto-sharpen tool available in the PHENIX software suite (Liebschner et al., 2019) applying B-factor values of 194.09 Å^2 and 95.93 Å^2 for the box sizes of 130 Å and 198 Å, respectively. Local resolution maps were obtained using PHENIX software and are showed in Figure S 38 right panels.

S6.3 Model building and refinement

The crystal structure of **CD1** (CCDC 2465040) was manually placed into the final cryo-EM maps in Coot[16] followed by Dock in Map tool available in PHENIX.[17] Then, the dsDNA$_{18}$ molecule (sequence $^{5'}$AATCCTGTACTGTCTAC$^{3'}$ and its complementary strand) was built into the final electron density map using Coot. This fitting was followed by several iterative cycles of real space refinement in PHENIX and manual adjustments in Coot. Note that the **CD1** is an abiotic molecule and is not in the library of Coot. Therefore, for the step of the real space refinement, we created a CIF file with the geometry restraints deduced from the **CD1** crystal structure (CCDC 2465040). This allowed us to calculate the atomic structure of the complex. The cross-correlation (CC) values calculated using Chimera between the atomic structures obtained after real refinement and the associated cryo-EM maps of **CD1** and dsDNA$_{18}$ fibres from box size of 130 Å and 198 Å are 0.75 and 0.72, respectively. All figures were generated using Chimera, ChimeraX and PyMOL software. Cryo-EM density maps and coordinates of the self-assembly between **CD1** and dsDNA$_{18}$ from particles extracted using box sizes of 130 Å and 198 Å have been deposited in the Electron Microscopy Data Bank (EMDB) with the ID codes EMD-71427 and EMD-71428, respectively.

---

[16] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* **66**, 486-501 (2010).

[17] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using x-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol*. **75**, 861–877 (2019).

Results

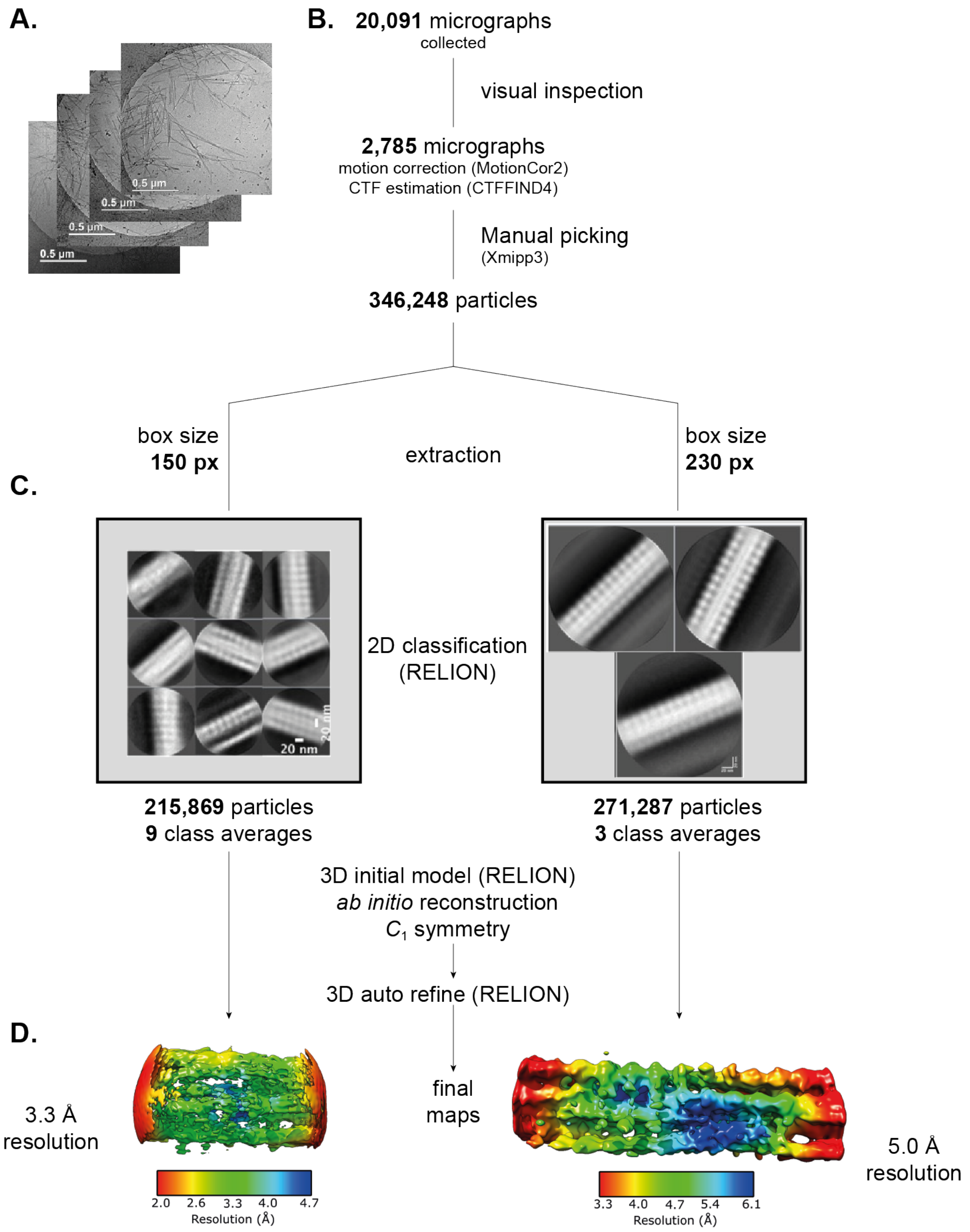


Figure S 37 Data processing workflow for the cryo-EM data processing of the CD1/dsDNA$_{18}$ assembly. **A**) Representative micrographs of the **CD1**/dsDNA$_{18}$ assembly obtained with Titan Krios G3 microscope operating at 300 kV equipped with a Gatan K3 Direct Detector and Bioquantum Imaging Filter. **B**) All steps performed to calculate the final structure of the **CD1**/dsDNA$_{18}$ assembly from cryo-EM maps obtained from particles extracted using a box size of 150 and 230 pixels (130 Å and 198 Å respectively). **C**) 2D-class averages used for particle extraction and (**D**) the corresponding sharpened maps colored accordingly to the local resolution.

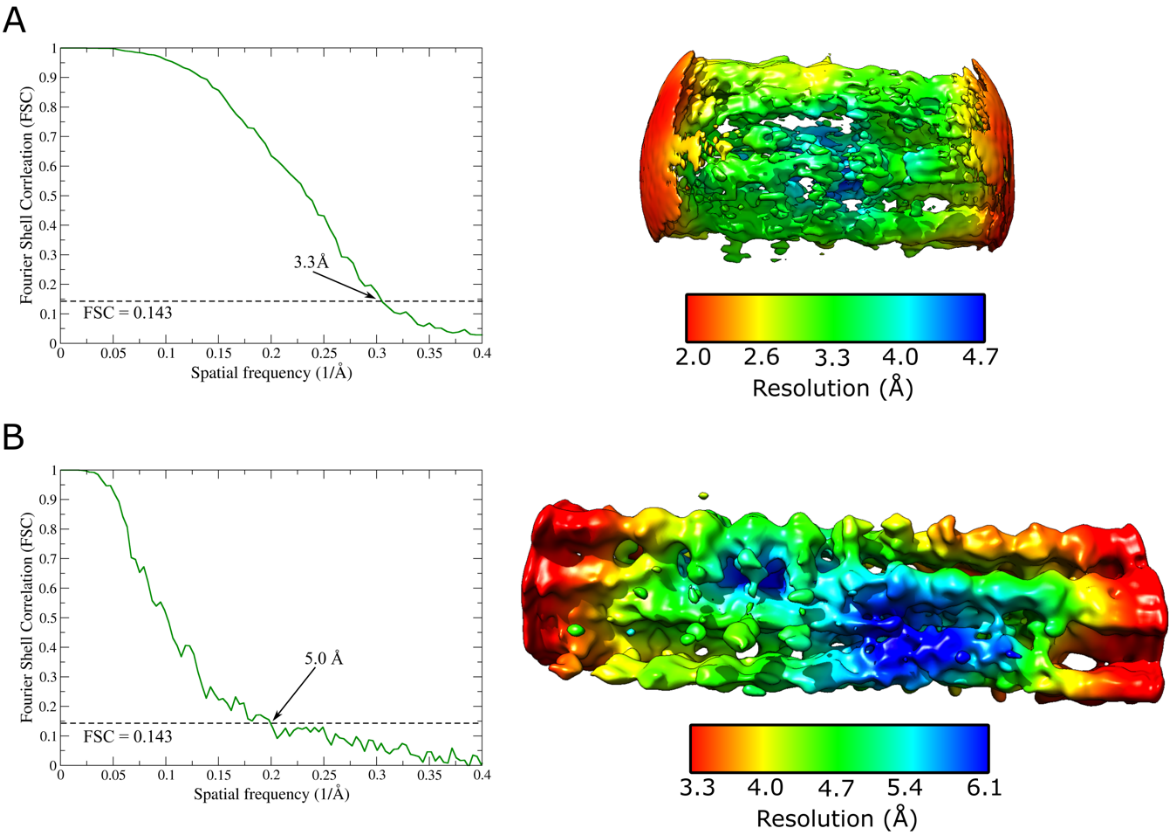


Figure S 38. Fourier shell correlation (FSC) curves obtained for **CD1**/dsDNA$_{18}$ assembly. The resolution of the final cryo-EM maps were estimated by the half-maps FSC curve using the "gold-standard" cutoff value (0.143). (A) Cryo- EM map from the particles extracted using a box size of 130 Å (150 pixels). (B) Cryo-EM map from the particles extracted using a box size of 198 Å (230 pixels). The sharpened maps are colored accordingly to the local resolution calculated using PHENIX.

S7 Crystal Structure of CD1

S7.1 Methods

A crystallization screen (Jena JCSG++ HTS) was performed with a mixture of 1 mM *Ralstonia solanacearum* lectin and 5 mM **CD1**. Crystals were obtained in condition B1, 0.8 M ammonium sulfate, 0.1 M sodium citrate at pH 4. X-ray diffraction data were collected at SOLEIL synchrotron, beamline PROXIMA 2A. The diffraction frames were integrated using XDS[18] and the output reflection file was converted using XPREP.[19] *Ab initio* structure determination was performed using SHELXD,[20] the initial model was manually improved using COOT[16] and refined with SHELXL.

S7.2 Crystal structure and packing of **CD1**

The crystals were space group $P2_12_12_1$ (a = 14.473 Å, b = 23.430 Å, c = 24.150 Å) with an asymmetric unit comprising **CD1**, one sulfate counterion and 17 water molecules (CCDC 2465040). Figure S28 shows the ORTEP diagram, as well as sticks representations of **CD1** and the crystal packing. The cyclodextrin forms supramolecular polymers in which the adamantyl group of one **CD1** inserts into the cavity of the adjacent **CD1**.

CD1 units are separated 11.7 Å apart within the solid state supramolecular polymer (measured between two consecutive adamantyl groups).

---

[18] Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr.* **66**,125-132 (2010).

[19] XPREP 2014; X-ray data Preparation and Reciprocal space Exploration Program. Bruker AXS Inc., Madison, Wisconsin, USA.

[20] Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr.* **58**, 1772-1779 (2002).

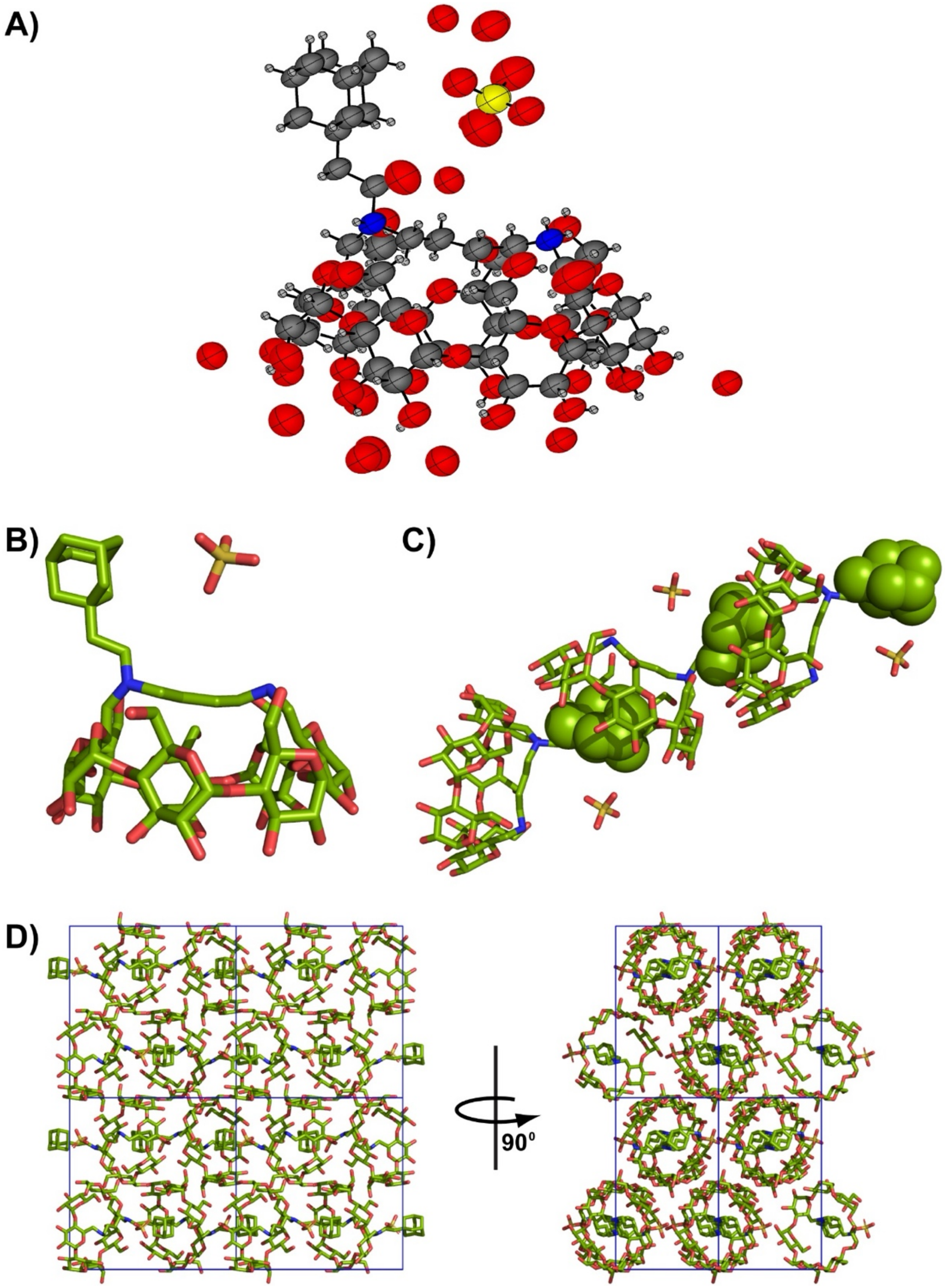


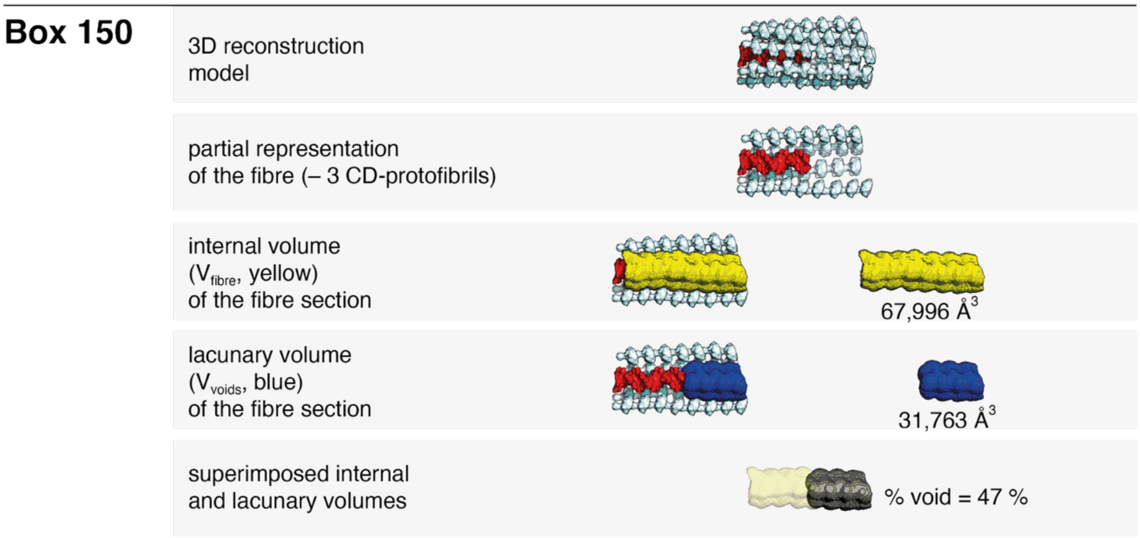
Figure S 39 **CD1** crystal structure showing **A)** ORTEP diagram of the asymmetric unit with thermal ellipsoids drawn at 50 % probability, **B)** sticks representation of **CD1** and the sulfate counterion, **C)** a trimeric assembly illustrating encapsulated adamantyl groups (spacefill), and **D)** crystal packing in 2 x 2 x 2 unit cells (blue axes).

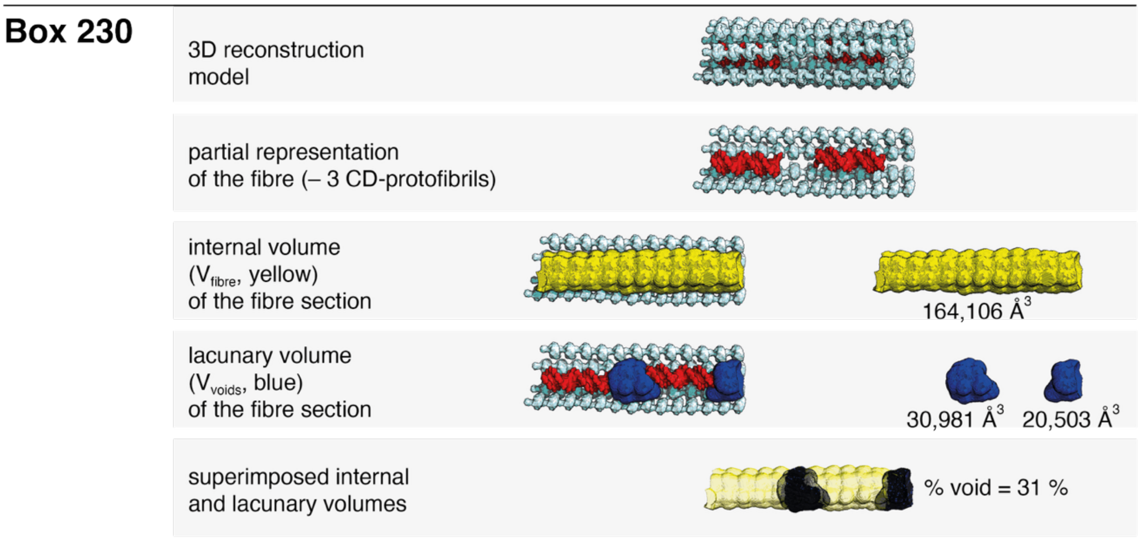
S8 Estimation of volumes and packing

Internal (V_{fibre}) and lacunary volumes (V_{voids}) of the fibres were estimated using MoloVol[21] software with a two probes mode using the two 3D reconstruction models and the following parameters:

Elements radii:	
O :	1.5 Å
H :	1.2 Å
N :	1.66 Å
C :	1.77 Å

Small probe radius: 7 Å
Large probe radius: 20 Å
Grid resolution: 0.7 Å
Optimization depth: 4





---

[21] J. B. Maglic, R. Lavendomme, MoloVol: An Easy-to-Use Program for Analyzing Cavities, Volumes and Surface Areas of Chemical Structures. *J. Appl. Crystallogr.* **55**, 1033–1044 (2022). https://doi.org/10.1107/S1600576722004988.

The linear packing section was calculated using the following equation :

$$\%\ linear\ packing = \frac{V_{fiber} - V_{voids}}{V_{fiber}}$$

The estimation vary between 53 % and 69 % for box 150 and 230 respectively. This corresponds to statistical mechanical calculations providing a 50% linear packing (see text and S8).

S9 Ionic strength for standard cryoEM conditions

The ionic strength for standard cryoEM conditions (*i.e.* [**CD1**] = 0.7 μM, [dsDNA$_{18}$] = 6 μM in 1X TBE buffer) was calculated as follow :

- Contributions from the **CD1**•HCl

Both ions have a charge magnitude of |z| = 1. Given the concentration of **CD1**•HCl is 0.7 mM (0.0007 M), the ionic strength contribution is:

$$I_{\mathbf{CD1}\bullet\mathrm{HCl}} = \frac{1}{2}\left\{\left([\mathbf{CD1}\bullet\mathrm{H}^{+}]\times|z|^{2}\right)+\left([\mathrm{Cl}^{-}]\times|z|^{2}\right)\right\}$$

$$I_{\mathbf{CD1}\bullet\mathrm{HCl}} = \frac{1}{2}\{(0.0007\times1^{2})+(0.0007\times1^{2})\}$$

$$I_{\mathbf{CD1}\bullet\mathrm{HCl}} = 0.0007\ \mathrm{M}$$

- Contributions from dsDNA$_{18}$•36Na

Each base pair of DNA contributes approximately 2 negative charges from phosphate groups. For an 18 bp DNA molecule, the total charge is:

$$z_{\mathrm{dsDNA}^{18}} = -2\times18$$

$$z_{\mathrm{dsDNA}^{18}} = -36$$

The concentration of DNA is 6 μM (6×10^{-6} M), so its ionic strength contribution is:

$$I_{\mathrm{dsDNA}^{18}\bullet36\mathrm{Na}} = \frac{1}{2}\left\{\left([\mathrm{dsDNA}^{18}]\times|z|^{2}\right)+\left([\mathrm{Na}^{+}]\times|z|^{2}\right)\right\}$$

$$I_{\mathrm{dsDNA}^{18}\bullet36\mathrm{Na}} = \frac{1}{2}\{(6\times10^{-6}\times36^{2})+(6\times10^{-6}\times36\times1^{2})\}$$

$$I_{\mathrm{dsDNA}^{18}\bullet36\mathrm{Na}} = 0.004\ \mathrm{M}$$

- Contributions from 1X TBE Buffer

Tris base (0.089 M), Borate (0.089 M) and EDTA (0.002 M) dissociate into : Tris^{+}, B(OH)$_4^{-}$ and EDTA^{2-}.

$$I_{\mathrm{TBE}} = \frac{1}{2}\left\{\left([\mathrm{TRIS}]\times|z|^{2}\right)+\left([\mathrm{B(OH)}_4^{-}]\times|z|^{2}\right)+\left([\mathrm{EDTA}^{2-}]\times|z|^{2}\right)\right\}$$

$$I_{\mathrm{TBE}} = \frac{1}{2}\{([0.089]\times1^{2})+([0.089]\times1^{2})+([0.002]\times2^{2})\}$$

$$I_{\mathrm{TBE}} = 0.093\ \mathrm{M}$$

- Total Ionic Strength Calculation

$$I = I_{\mathbf{CD1}\bullet\mathrm{HCl}} + I_{\mathrm{dsDNA}^{18}\bullet36\mathrm{Na}} + I_{\mathrm{TBE}}$$

$$I = 0.0007 + 0.004 + 0.093$$

$$I = 0.0977\ \mathrm{M}$$

The ionic strength of the solution is *ca*. **0.098 M**

S10 Molecular dynamics simulations

S10.1 Methods of simulation

The initial structure of the **CD1** was built using the crystal structure of **CD1** and a model protofibril of eight **CD1** in a supramolecular polymer was built by translation using the Avogadro program. The dsDNA$_{18}$ with was generated with the Nucleic Acid Builder (NAB) from the AMBER package.[22] Seven **CD1** protofibrils were then distributed around the dsDNA$_{18}$ at equal distances between protofibrils, starting in a C7-symmetry for the model co-assembly. All simulations were performed using the AMBER suite.[22] The ParmBSC1 force field was employed to describe the nucleic acid,[23] and the **CD1** protofibrils were treated with the q4MD-CD force field.[24] GAFF parameters with AM1-BCC charges were applied to the additional fragment (adamantane moiety) using ANTECHAMBER module from AMBER. The system was energy-minimized using first the steepest descent algorithm and then using the conjugate gradient. The production MD simulations were each carried out for 250 ns at 300.0 K using Langevin thermostat with a collision frequency of 1.0 ps^{-1}.[25] All bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm,[25] allowing for an integration time step of 2.0 fs. Given the size of the co-assembly, explicit water box solvation was found computationally costly and instead an implicit description of water was considered using the Generalized Born solvent model, adding chloride ions at the periphery of the co-assembly to reach electroneutrality of the system.[26] The electrostatic interactions were treated with a virtually infinite cut-off (9999 Å). Five MD simulations tests, each of 100 ns, were carried out with or without restraints on **CD1** protofibrils and CD-DNA distances, to find the settings to avoid the collapse of the co-assembly during the MD simulation. Given the results, a set of restraints (25 kcal/mol) was applied between each **CD1** protofibril and between **CD1** protofibril-dsDNA$_{18}$, to avoid a too important diffusion of CDs from DNA. These restraints were set in such a way that the potential energy of the system is not modified when a **CD1** protofibril is close to dsDNA$_{18}$, as the restraint function only consists in a parabola at longer distance ('DISANG' option in AMBER input files). Spring forces were only applied if the distance between **CD1** protofibril and DNA is larger than 12 Å. The MD production phases were carried out for 250 ns, with a timestep of 2 fs and snapshots extracted every 0.5 ns of the simulation.

---

[22] Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruz, V. W. D., Forouzesh, N., Giese, T. J., Götz, A. W., Gohlke, H., Izadi, S., Kasavajhala, K., Kaymak, M. C., King, E., Kurtzman, T., Lee, T.-S., Li, P., Liu, J., Luchko, T., Luo, R., Manathunga, M., Machado, M. R., Nguyen, H. M., O'Hearn, K. A., Onufriev, A. V., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Risheh, A., Schott-Verdugo, S., Shajan, A., Swails, J., Wang, J., Wei, H., Wu, X., Wu, Y., Zhang, S., Zhao, S., Zhu, Q., Cheatham, T. E. III, Roe, D. R., Roitberg, A., Simmerling, C., York, D. M., Nagan, M. C., Merz, K. M. Jr. AmberTools. *J. Chem. Inf. Model.* **63**, 6183-6191 (2023).

[23] Ivani, I., Dans, P., Noy, A., et al. Parmbsc1: A refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016). https://doi.org/10.1038/nmeth.3658

[24] Cézard, C., Trivelli, X., Aubry, F., Djedaïni-Pilard, F., Dupradeau, F.-Y. DOI: 10.1039/C1CP20854C. *Phys. Chem. Chem. Phys*. **13**, 15103-15121 (2011).

[25] Ryckaert, J. P., Ciccotti, G., Berendsen, H. J. C. *J. Comp. Phys*. **23**, 327-341 (1977).

[26] Onufriev, A. V., Case, D. A. *Annu. Rev. Biophys*. **48**, 275-296 (2019).

Analysis of simulations

The structural parameters such as distances and dihedral angles were monitored throughout the MD simulations using CPPTRAJ AMBER module. Detection of pores between DNA and CDs were performed with ChExVis server with a sphere probe of 1.4 Å, *i.e.* the average radius of a water molecule.[27]

S10.2 Bending of protofibrils

In Figure S 40 **CD1** protofibrils tubes show substantial bending, as quantified by a linearity score $\left(LS = \left|\cos\frac{\pi\times\alpha}{180}\right|\right)$ estimated from the dihedral angles (α) between four consecutive CD monomers for each CD protofibril (averaged on all snapshots of the MD simulation), see Figure S 40. The profiles show that the linearity score varies along a given CD protofibril and in between the different CD protofibrils. The time-profiles of CD linearity scores constantly evolve, with a standard deviation up to of 0.3 (on a scale from 0 to 1) for the three main dihedral angles of a particular CD tube. The bending dynamics is therefore an important factor of the co-assembly in aqueous solution.

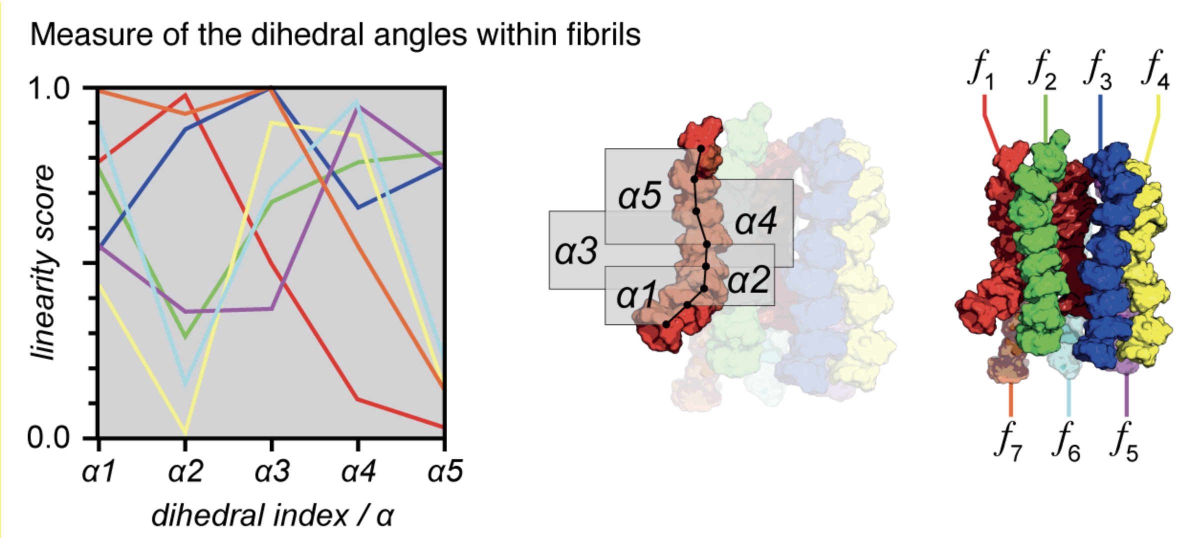


Figure S 40 Linearity score (LS) for each individual **CD1** protofibril (color code on the right).

S10.3 Distances between neighboring **CD1** protofibrils

At the static scale probed by cryo-EM, the distances between CD protofibrils are not uniform, evolving from 13 Å to 18 Å. In the MD simulations, the proximity of a CD protofibril to an adjacent CD protofibril is higher near the extremity of the DNA fragment, while larger distances between CD protofibrils are observed in the middle sections of the DNA double-helix. This mirrors previous studies on DNA-nanoparticle interactions,[28] where the end regions of DNA are often more flexible and can be prone

---

[27] T. Bin Masood, S. Sandhya, N. Chandra, V. Natarajan, *BMC Bioinformatics*, **16**, 119 (2015).

[28] P. Yin, H. Yan DNA-based nanostructures for molecular recognition. *Nat. Commun*. **7**, 13499 (2016). https://doi.org/10.1038/ncomms13499

to tighter packing with surrounding molecules. The increased proximity at a DNA end is also related to the local structural distortions or bending at the termini, which favor tighter interactions with the CD protofibrils. Figure S 41 shows a set of distances between pairs of adjacent CDs belonging to two consecutive protofibrils. In the simulations, such distances between secondary nitrogen atoms vary, in average, between 13 Å to 28 Å.

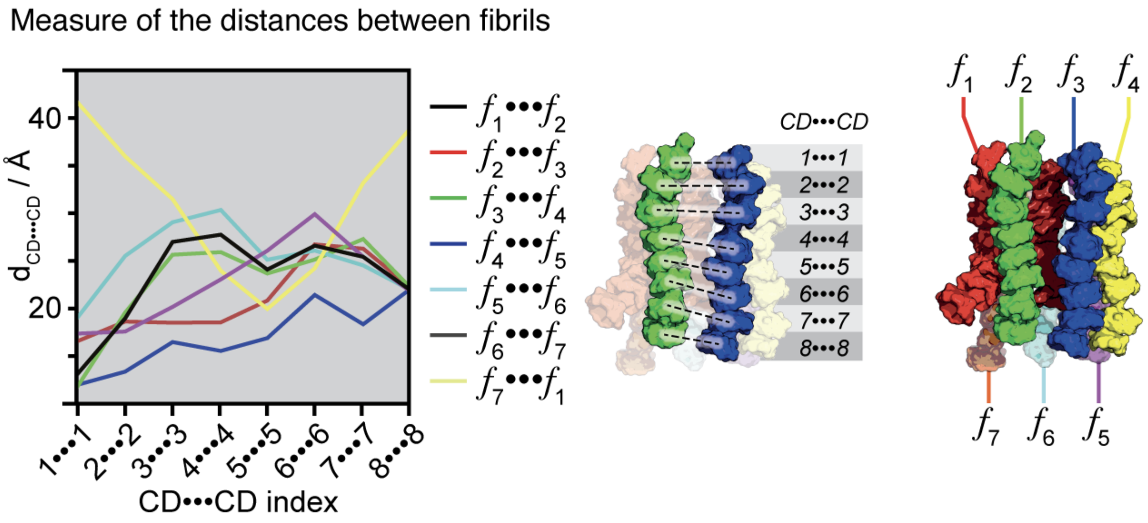


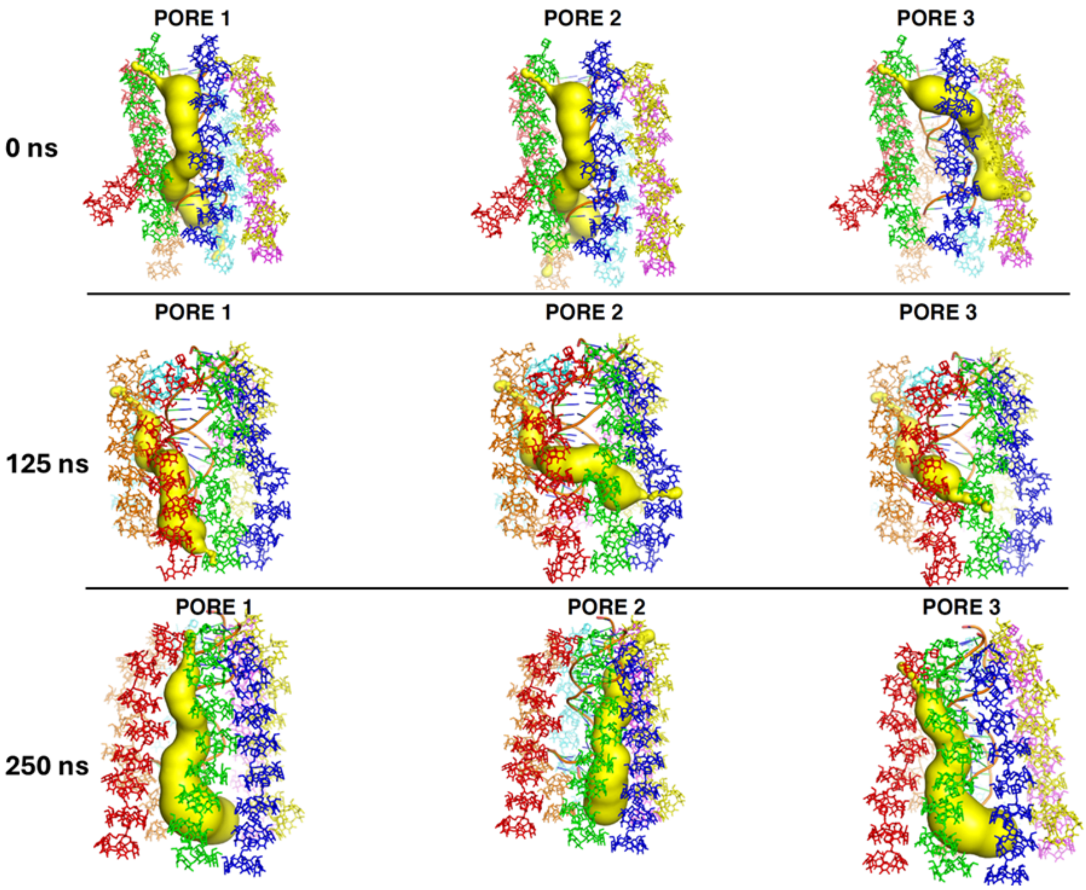
Figure S 41 Distances between pairs of adjacent **CD1** belonging to two consecutive protofibrils, noted as *f1* to *f7* (color code on the right).

S10.4 Voids between dsDNA$_{18}$ and **CD1** protofibrils

Figure S 42

Conformation	Pore # (rank)	Length (Å)	Straightness	Score
0 ns	1	129	0.78	0.92
	2	130	0.74	0.90
	3	112	0.76	0.87
125 ns	1	101	0.86	0.79
	2	110	0.69	0.75
	3	83	0.79	0.73
250 ns	1	83	0.86	0.86
	2	85	0.83	0.85
	3	84	0.83	0.80

Figure S 42 shows selected frames with voids (pores) in yellow, using a sphere probe of 1.4 Å. The table below reports the length and the straightness of the corresponding pores in sizes, which vary during the simulation. It is to stress that throughout the MD simulation, the overall connectivity of the pores is preserved, *i.e.* these pores remain extended along the dsDNA$_{18}$ double-helix. The length of the pores varies between 80 Å to 130 Å, with a straightness in the range of 0.70 to 0.85 (on a scale from 0 to 1).



Conformation	Pore # (rank)	Length (Å)	Straightness	Score
0 ns	1	129	0.78	0.92
	2	130	0.74	0.90
	3	112	0.76	0.87
125 ns	1	101	0.86	0.79
	2	110	0.69	0.75
	3	83	0.79	0.73
250 ns	1	83	0.86	0.86
	2	85	0.83	0.85
	3	84	0.83	0.80

Figure S 42 Examples of voids (or pores, in yellow) in the co-assembly for selected MD snapshots (at 0 ns, 125 ns, 250 ns). For each snapshot, the three highly ranked pores are shown. Each **CD1** protfibril is color coded as in Fig S 41. The main descriptors of the pores are reported in the table below (with the corresponding scores from ChExVis).

S11 Statistical mechanical calculations

S11.1 Summary of findings

We model the encapsulated DNA strands as tubes of CDs containing one or more strands. These tubes can be shorter than a single DNA, or larger than the one or more DNAs they encapsulate. The calculations focus on mass action on the one hand, and on the role of electrostatic interactions between the DNA and the protofibrils, and bending fluctuations of the DNA, on the other. We ignore the bending fluctuations of the protofibrils and the precise structure of the tubes, and incorporate these in effective values of interaction free energies. (We comment on the impact of the bending fluctuations of the CD protofibrils and their contribution to the optimal tube width below.) The mass action model hinges on two free energy parameters, one associated with the interaction between the DNA and the CDs making up the tube, and one with that between the CDs forming the protofibrils. The latter is presumed to be not large enough to force the CDs to form strands in free solution at the concentrations probed in the experiments, so without the presence of DNA. Electrostatic binding of the CDs by the DNAs locally enhances their concentration, allowing mass action to drive the assembly of the protofibrils. The tubes of CD protofibrils also form due to the interaction with the DNA - no attraction between them is needed for this. According to the model calculations, the tube diameter and the mean number of protofibrils in the cross section is set by the attractive screened Coulomb interaction between DNA and CDs and repulsive entropic interaction between them, caused by a suppression of bending fluctuations once encapsulated in a tube on account of steric and possibly also electrostatic repulsive interactions. It turns out to be a complicated function of the linear charge densities of the DNAs and CD protofibrils, the persistence length of the DNA and the Debye screening length that depends on the ionic strength of the solution. According to the calculations, the gap between a surface of a DNA segment and that of the protofibrils must be of the order of a Debye length, as one would perhaps expect. The co-assembly of the DNA and CDs we find to be co-operative, because both partially covered DNAs and fully filled tubes are strongly suppressed, in essence by the loss of co-translational entropy. This also drives the tubes to be only partially filled by only 50% according to the model. For assembly to take place, a large excess of CDs in the solution is required. In that case almost all DNA is encapsulated, and growth of the tubes only depends on the concentration of DNA and the binding free energy between the CDs in the protofibrils — the larger this binding free energy the longer the tubes become. The growth law, describing the power-law relation between the concentration and mean length of the tubes, has a growth exponent of 2/3 rather than the value of 1/2 one usually finds deep into the polymerised regime for self-assembling fibres, also indicating the strongly co-operative character of the co-assembly process.The existence of critical solution concentrations also point in this direction.

S11.2 Mass action model

To theoretically investigate the co-assembly of the short dsDNA fragments and CDs, we rely on a highly idealized and abstract model in which the CDs assemble into tubes of varying length that contain one or more DNA molecules. The tube elements are units of height h and the DNAs rods of length $l \gg h$. For our model, we do not need to specify how the tubes and tube elements are precisely built up from the individual CDs. We do need to separate the case where one or more DNAs are encapsulated in a tube that is at least as long as the combined length of the DNAs, and that where a single DNA is partially or wholly covered by a short tube. See Figure S43.

Let us first focus on the former, and let $\rho(N,n)$ denote the mole fraction of assemblies consisting of $N \geq 1$ DNA molecules and $n \geq N\,l\,/h$ tube elements. We ignore any empty tubes as the CDs on themselves do not self-assemble into tubes (or fibres) under the conditions of the experiments. The mole fractions of free DNA is $\rho(1,0)$ and that of free tube units (representing the CDs) is $\rho(0,1)$. Presuming the solution to be dilute, we write $\rho(1,0) = \lambda_{DNA}$ and $\rho(0,1) = \lambda_{CD}$ in terms of their fugacities λ_D and λ_{CD} implying a grand canonical ensemble that we later on convert to a canonical ensemble by imposing mass conservation. From statistical mechanics, we know that $\rho(N,n) = \lambda_{DNA}^N \lambda_{CD}^n Z(N,n)$ with $Z(N,n)$ the partition function (or configurational integral) of an assembly.[29,30]

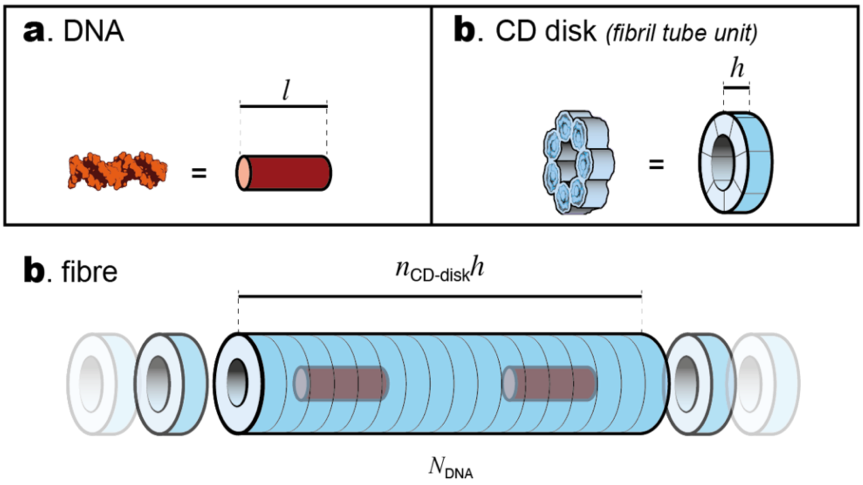


Figure S43. Schematic representation of the model. In blue the strands of DNA, the transparent cylinders the tubes consisting of CDs. The length of a strand of DNA is equal to , and that of the CD tube consisting of elementary units where . Top: only part of the DNA has been encapsulated. Bottom: two DNAs are encapsulated in a tube that is much longer as the combined length of the DNAs. See also the main text.

For the partition function, we make use of the well-known partition function $(nh - Nl)^N / \Lambda^N N!$ of a Tonks ``gas'' of N hard (non-overlapping) lines of length l confined to a line piece of length nh. Here, Λ denotes a microscopic length scale that is equal to the thermal wavelength if considering a gas. However, since our particles are dissolved in water we take for it the largest length scale ignored in our coarse-grained description, i.e., the size of a water molecule. Noting that the Tonks gas model demands that $nh \geq Nl$, a tube must be at least as long as the total length of the encapsulated DNAs. It is the Tonks gas contribution to the partition function, accounting for the impact of the drive for increasing the free "volume" (or in this case length) of the DNAs in the tubes, that (i) gives rise to the cooperativity of the co-assembly of DNAs and CDs, and (ii) works against head-to-tail packing of the DNAs in the tubes.

If a tube consist of a sequence of n tube elements, then the free energy of binding must be equal to $(n-1)\varepsilon$ where $\varepsilon < 0$ is the free energy of a single bond (in units of thermal energy). We expect that $\varepsilon \approx O(-1)$ for interactions between pairs of CD. If that tube holds N DNA molecules, we must have a total binding free energy equal to Ng with $g < 0$ the free energy of interaction of a single DNA molecule with the tube elements it is in contact with (also in units of thermal energy). It includes any loss of entropy associated with suppressed bending fluctuations of the DNA and attenuated conformational fluctuations of the CDs that make up the tubes.[32] Since a 18 nt dsDNA makes contact with about 40 CDs, we estimate that $g \approx O(-40)$. Note that we expect g to scale linearly with the length of the DNA. Consequently, we conclude that we must have $\rho(N,n) = \lambda_{DNA}^{N}\lambda_{CD}^{n}(nh - Nl)^{N}\ exp\ (-Ng)\ exp\ (-(n-1)\varepsilon)\ /\Lambda^{N}N!$.

We now describe the DNAs of length l covered by a tube of length $nh \leq l$. The mole fraction of such assemblies obeys $\rho(1,n) = \lambda_{DNA}\lambda_{CD}^{n}Z(1,n)$ with $Z(1,n)$ again a partition function. If $n = 0$ the DNA is naked (see above). We can again make use of a one-particle Tonks ``gas’’ partition function $(l - nh)/\Lambda$ for the tube on the DNA, noting that we do not allow the tube not to completely be in contact with the DNA. Including the binding free energies, we now obtain $\rho(1,n) = \lambda_{DNA}\lambda_{CD}^{n}\ exp\ exp\ (-ghn/l)\ exp\ exp\ (-(n-1)\varepsilon)\ (l - nh)/\Lambda$, where we account for the fact that the binding free energy between DNA and tube is reduced in proportion to the fraction of the DNA not covered by the tube. This indicates already that the probability that a DNA is partially covered must be very much smaller than that of the case where this single DNA is part of a tube that is longer. Indeed, if we compare the probability of a tube consisting of n elementary tube units covering a single DNA molecule, $P(t) = P(nh/l)$, we find that this ratio is very much larger for $nh > l$ than for $nh < l$. See Figure S44.There is a distinct minimum for $t = 1$ or $nh = l$, caused by the loss of translational entropy of one component relative to the other. Because of the suppression of partially covered DNAs, we ignore them completely in our further analysis and shall be presuming that for the assemblies $n \geq N\ l\ /h$. Incidentally, the statistical suppression of intermediates for $1 \leq n \leq N\ l\ /h$ explains why the co-assembly in our model becomes nucleated and is so co-operative.[29]

---

[29] P. van der Schoot, Nucleation and Co-Operativity in Supramolecular Polymers, *Adv. Chem. Eng*. **35**, 45-77 (2009).

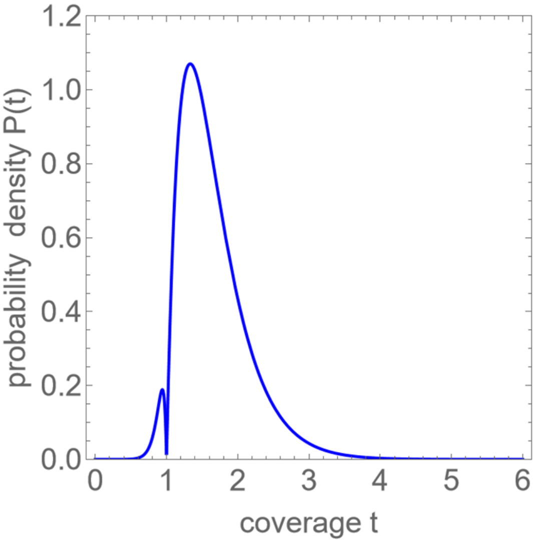


Figure S44. Probability density $P(t)$ that a single DNA is encapsulated in CD tube units equivalent to a coverage of $t \equiv nh/l$. For illustrative purposes we set the parameter combination $(l/h)\, ln\ (\lambda_{CD}\, exp\ (-\varepsilon))$ equal to -3 and presumed that $g = -20$. Notice that for $t = 1$ the entire DNA is covered by CD units, whilst for $t > 1$ the self-assembled tube is longer than the DNA. For the given parameter values, we have $\langle t \rangle = 1.6$. The probability of finding partial coverage can be ignored.

If the tube elements are not to self-assemble into empty tubes not containing any DNA, we must insist that $\phi_{CD}\, exp\ (-\varepsilon)$ must smaller than, say, about unity, where ϕ_{CD} is the overall mole fraction of tube elements.[30] To calculate the expectation value of the number of DNA molecules in tubes we need to evaluate the ratio

$$\langle N \rangle = \frac{\sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad N\, \rho(N,n)}{\sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad \rho(N,n)}, \tag{t1}$$

while the expectation value of the number of tube elements in tubes follows from

$$\langle n \rangle = \frac{\sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad n\, \rho(N,n)}{\sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad \rho(N,n)}. \tag{t2}$$

The sums are most easily evaluated using the well-known Stirling approximation $N! \sim \sqrt{2\pi N}\, N^N\, exp\ (-N)$, which formally is asymptotically exact for $N \gg 1$ but turns out to be excellent for $N \geq 1$, giving

$$\langle N \rangle = \frac{Li_{-1/2}(Y)}{Li_{+1/2}(Y)}, \tag{t3}$$

where $Li_{\alpha}(Y)$ denotes the polylogarithm of order α that is a function of the variable $0 \leq Y \equiv \lambda_{DNA}\, exp\ (-g)\ \times (\lambda_{CD}\, exp\ (-\varepsilon)\,)^{\frac{2l}{h}} \times \left(e\frac{l}{\Lambda}\right) < 1$ with e Euler's number. The variable Y combines two mass action variables, namely $\lambda_{DNA}\, exp\ (-g)$ and $\lambda_{CD}\, exp\ (-\varepsilon)$ into a composite one. In the limit $Y \to 0$, $\langle N \rangle \to 1 + \frac{Y}{\sqrt{2}} + \cdots$ and the average number of DNAs in tubes tends to unity, whilst for $Y \to 1$, that number diverges as $\langle N \rangle \to \frac{1}{2}(1 - Y)^{-1} + \cdots$. Interestingly, we exactly have

---

[30] P. van der Schoot Theory of supramolecular polymerization. In A. Ciferri (Ed.), *Supramolecular polymers*, 2nd Ed., Boca Raton: CRC Press, 77-106, (2005).

$$\frac{\langle n\rangle}{\langle N\rangle} = 2\left(\frac{l}{h}\right) \tag{t4}$$

for all values of Y. This implies that for our model 50 percent of the length of the tubes are not occupied by DNA. The reason for this is the translational entropy of the DNAs inside the tubes that very strongly suppresses complete filling.

To obtain very large assemblies, we need to insist that $Y \rightarrow 1$. To understand what this means, we need to translate the fugacities λ_{DNA} and λ_{CD} to mole fractions. Let ϕ_{DNA} be the overall mole fraction of DNA molecules and recall that ϕ_{CD} is the overall mole fraction of tube elements. Mass conservation of DNA now dictates that

$$\phi_{DNA} = \lambda_{DNA} + \sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad N\,\rho(N,n), \tag{t5}$$

while mass conservation of tube elements demands that

$$\phi_{CD} = \lambda_{CD} + \sum_{N=1}^{\infty} \quad \sum_{n=Nh/l}^{\infty} \quad n\,\rho(N,n). \tag{t6}$$

From these expressions, we conclude that $\lambda_{DNA} < \phi_{DNA}$ and $\lambda_{CD} < \phi_{CD}$ and, as a consequence, we must also conclude that $Y < (\phi_{DNA}\, exp\ (-g)\,) \times (\phi_{CD}\ exp\ (-\varepsilon)\,)^{\frac{2l}{h}} \times \left(e\frac{l}{\Lambda}\right)$. For the tube elements to not form empty tubes (or fibres) in the absence of DNA, we had to insist that $\phi_{CD}\ exp\ (-\varepsilon)\ < 1$. Because $l \gg h$, this implies that $\phi_{DNA}\, exp\ (-g)$ has to be very large for the assemblies to form. It also explains why the formation of fibres is so sensitive to overall dilution even by a relatively modest amount and why so much excess CD is required, as in fact we have seen in the experiments where $N \gg P$.

This sensitivity suggests a highly co-operative assembly process, as can in fact also be deduced from a growth law that we may obtain in the limit $Y \rightarrow 1$, deeply in the polymerised regime where $\langle N\rangle \gg 1$. Under conditions of a vast excess of CDs, we may presume that $\phi_{CD} \simeq \lambda_{CD}$ in which case

$$\langle N\rangle \sim ((\phi_{DNA} - \phi_*)\ exp\ (-\varepsilon)\,)^{2/3} \tag{t7}$$

where

$$\phi_* = (\Lambda/el)\ exp\ (+g)(\phi_{CD}\ exp\ (-\varepsilon)\,)^{-2l/h} \tag{t8}$$

acts as a critical concentration (mole fraction), noting that we expect $\Lambda/el \ll 1$ as in our case Λ is of the order of the size of a solvent molecule, which is much smaller than the length l of an encapsulated DNA strand. From the theory outlined above, we deduce that for ϕ_{DNA} below $\approx \phi_*, \langle N\rangle = 1$ and assemblies do not form in appreciable numbers. See Figure S45for a phase diagram. For $\phi_{DNA} \gg \phi_*$, the relevant mass action variable is $\phi_{DNA}\, exp\ (-\varepsilon)$ involving the concentration of DNA and, interestingly, the binding free energy between the tube units (the model CDs). What happens here, is that all DNAs have been encapsulated and increasing the binding energy of a DNA to the tube, g, no longer impacts upon the mean aggregate size. The aggregates can now grow further by increasing the magnitude of $\varepsilon < 0$. This is also what we see in the experiments. Notice also the two-thirds growth exponent, which is larger than the usual exponent of one-half, typical of linear aggregation.[30] In mixed linear aggregates, both larger and smaller exponents have been reported, depending on the concentration regime of both components.[31]

---

[31] Theory of supramolecular co-polymerization in a two-component system, S. Jabbari-Farouji and P. van der Schoot, *J. Chem. Phys*. **137**, 064906 (2012).

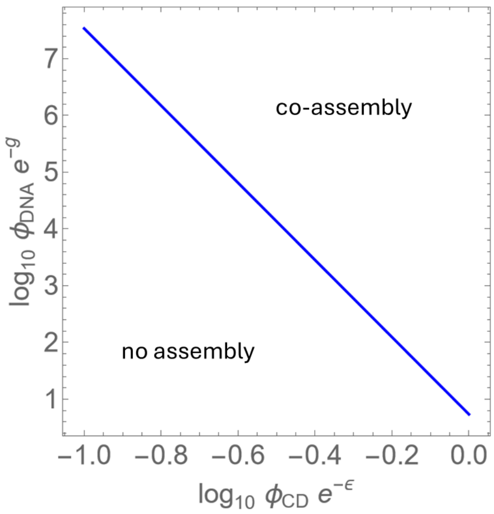


Figure S45 Model phase diagram for CD and DNA mixtures in terms of the mass action variables $\phi_{CD}\ exp\ (-\varepsilon)$ and $\phi_{DNA}\ exp\ (-g)$. We set $\Lambda = 0.28$ nm (the size of a water molecule) and $l = 4.1$ nm (the length of DNA18) .

From the expression for the critical concentration ϕ_*, we can read off a few things that are relevant to our experiments. First, for assemblies to form requires ϕ_{DNA} to be larger than ϕ_*. The critical concentration is very sensitive to the concentration of CDs, ϕ_{CD}, because the ratio $l/h \gg 1$ enters as an exponent and $\phi_{CD}\ exp\ (-\varepsilon)$ is to be not very large for otherwise protofibrils would form in the absence of DNA. In fact, if ϕ_{CD} $exp\ (-\varepsilon)$ is smaller than unity, would require a very strong attraction between the DNAs and the CDs, $g \ll -1$, for assembly to be possible in the first place. All of this explains why diluting the CDs has such a dramatic effect: it drives the critical concentration below the concentration of DNA in the solution very strongly due to the large exponent. Further, the impact of the length of the DNA enters the critical concentration in two ways. We recall that the binding free energy $g \ll -1$ scales linearly with the length of the DNA (see the section 8.3 below). Using this fact, we conclude from our expression for the critical concentration (mole fraction) that it may either increase, decrease or even exhibit non-monotonic behaviour as a function of the length of the DNA, all this depending on the solution conditions and the concentration of CDs. Finally, we expect the cooperativity of the coassembly transition to go up with increasing length of the DNA, as the length of the suppressed partially covered DNAs increases concomitantly.

S11.3 Bending fluctuations, Coulomb interactions and optimal tube radius.

The structure of the tubes is complex, and consist of on average 7 chains (protofibrils) of stacked CDs. The chains are arranged more or less parallel along the main axis of the tube and are not necessarily of the same length as the tube, so the tubes contain ends and beginnings of protofibrils. This in fact is what we expect as they are a form of equilibrium polymers. In the absence of DNA, the protofibrils do not form under the solution conditions investigated. This means the binding to DNA is crucial for the formation of the protofibrils — in essence the binding to DNA increases the local concentration of CD that drives assembly by mass action. The fibres that make up the tubes meander in a way along the surface of the tubes as can be inferred from Figure 2. This is arguably caused by thermally induced bending fluctuations of the protofibrils. Unfortunately, the persistence length (a measure for the bending stiffness) of the protofibrils is not known. The binding to the DNA and the interaction with neighbouring protofibrils suppresses bending fluctuations, which reduces their entropy.[32]

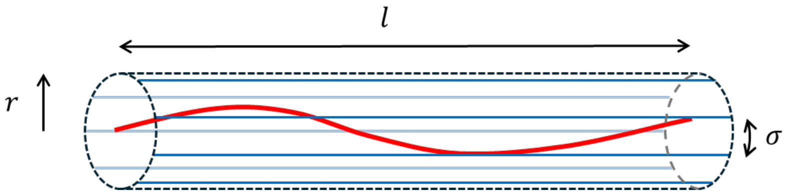


Figure S46. Schematic representation of a semi-flexible DNA segment (red) encapsulated by CD protofibrils (blue) that make up a tube (dashed black). The length of the tube is l, the radius r and the distance between the CD protofibrils σ. See also the main text.

The DNA fragments that are encapsulated by the tubes also undergo bending fluctuations and these fluctuations are also attenuated by their encapsulation into the tubes of protofibrils. Short DNA fragments are thought to exhibit a shorter persistence length than the value of 50 nm usually quoted, and may be as small as 32 nm for a 18 nucleotide sequence.[33] The persistence length of DNA does not only depend on its length but also the ionic strength and type of ions present in the solution.[34] In fact, the electrostatic interactions between the protofibrils and between the protofibrils and the DNA may also be affected by the bending fluctuations.[35]

It is also because of the role of entropy that the binding strengths g and ε of our mass action model are free energies rather than energies, and that the width of the tubes is not only set by the diameter of the DNA but also by the strength of the attraction between DNA and protofibrils and the repulsion between the protofibrils. Both are in part due to electrostatic interactions. A (very) crude estimate for the interaction between a DNA and the tube of protofibrils as well as the optimal tube diameter may be gotten from a combination of the Debye-Hückel theory of electrostatics and Odijk's semi-flexible polymer confinement theory.

---

[32] T. Odijk, On the Statistics and Dynamics of Confined or Entangled Stiff Polymers, *Macromolecules* **16**, 1340-1344 (1983).

[33] Yuan-Yan Wu, Lei Bao, Xi Zhang, and Zhi-Jie Tan, Flexibility of short DNA helices with finite-length effect: From base pairs to tens of base pairs, *J. Chem. Phys*. **142**, 125103 (2015)

[34] Sébastien Guilbaud, Laurence Salomé, Nicolas Destainville, Manoel Manghi, and Catherine Tardin, Dependence of DNA Persistence Length on Ionic Strength and Ion Type, *Phys. Rev. Lett*. **122**, 028102 (2019).

[35] T. Odijk, Undulation-enhanced electrostatic forces in hexagonal polyelectrolyte gels, *Biophys. Chem.* **46**, 69-75 (1993).

In our highly idealised model, we consider a semi-flexible chain in a cylindrical tube formed by perfectly parallel chains that act as a model for the protofibrils. We ignore the bending fluctuations of the protofibrils and focus on those of the DNA and pretend the DNA is a semi-flexible line piece. See Figure S46. Let the linear charge density of the DNA be equal to ν_- and that of the protofibrils be ν_+ in units of numbers of charges per unit length. We estimate that $\nu_- \simeq 5.9$ nm^{-1} and $\nu_+ \simeq 0.8$ nm^{-1}. The dimensionless Debye-Hückel potential acting between two point charges at distance r is $\Psi_\pm = \pm\frac{\lambda_B}{r}\, exp\ (-r/\lambda_D)$ where the $-$ sign holds if the charges have opposite sign and the $+$ sign if the charges have equal sign. Here, $\lambda_B \simeq 0.7$ nm denotes the so-called Bjerrum length and λ_D the Debye screening length that depends on the ionic strength of the solution. For 1-1 salts, $\lambda_D \simeq 0.3/\sqrt{I}$ nm with I the ionic strength in moles per litre.

In our model, the protofibrils are arranged perfectly parallel on a cylinder of radius r and length l, positioned at equal distances of σ. We will not specify this length scale here but discuss it below. If we ignore end effects, so presume that $l \gg \lambda_{D,}$ and also treat the DNA as a straight line piece, we find by summing over all pair interactions the following free energy associated with the interaction between the DNA and the tube of protofibrils

$$\frac{F_{Coulomb}}{k_B T} \simeq -4\pi\, l\, r\, \nu_-\, \nu_+ \left(\frac{\lambda_B}{\sigma}\right) K_0\left(\frac{r}{\lambda_D}\right) \tag{t9}$$

in units of thermal energy $k_B T$, where $K_n(x)$ is a modified Bessel function of the second kind with $n = 0$. This free energy is most negative for $r \simeq 0.6\, \lambda_D$, so we would naively expect this to be the equilibrium radius of the tube. In reality, this radius will be larger, even for a line charge, due to thermal bending fluctuations and because of repulsive interactions between the protofibrils that make up the tube. To account for the entropy loss associated with the confinement-induced suppression of the bending fluctuations of the DNA segment. A useful estimate for this is due to Odijk,[32] and given by

$$\frac{F_{bend}}{k_B T} \simeq \frac{l}{2^{4/3} P^{1/3} r^{2/3}} \tag{t10}$$

where P is the persistence length of the encapsulated DNA. Notice that both contributions to the free energy are extensive, that is, proportional to the length of the DNA.

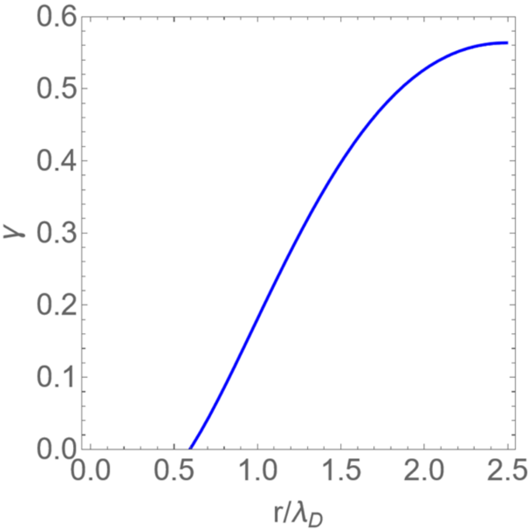


Figure S47. Scaled entropy loss parameter γ as a function of the optimal fraction scaled radius r/λ_D of the tubes with λ_D the Debye screening length. See also the main text.

The optimal radius r of the tube minimises the total free energy $F = F_{Coulomb} + F_{bend}$. Putting $\partial F/\partial r = 0$ gives the following equation

$$\gamma = -\left(\frac{r}{\lambda_D}\right)^{5/3} K_0\left(\frac{r}{\lambda_D}\right) + \left(\frac{r}{\lambda_D}\right)^{8/3} K_1\left(\frac{r}{\lambda_D}\right) \tag{t11}$$

with

$$\gamma = \left(\frac{1}{3\pi\, 2^{7/3}}\right) \frac{\sigma}{P^{1/3}{\lambda_D}^{5/3}\nu_-\nu_+\lambda_B} \tag{t12}$$

a measure for how strong the interaction is able to overcome the configurational entropy loss upon confinement of the DNA by the CD protofibrils. The smaller the value of $\gamma \geq 0$, the larger the interaction energy compared to the entropy loss. The solution to this equation that minimises the free energy is plotted in Figure S47; solutions for r/λ_D larger than 2.5 are free energy maxima, not minima. For $\gamma \to 0$, $r \simeq 0.6\,\lambda_D$, whilst for $\gamma \to 0.56$, $r \simeq 2.5\,\lambda_D$. For $\gamma > 0.56$, no minimum solution can be found, implying that binding does not take place. In other words, the interaction has to be sufficiently strong for binding to take place. In conclusion, the distance between the protofibrils and the DNA segments is on the order of the Debye length, which seems sensible as it is the only relevant length scale in the problem.

The optimal number of protofibrils that make up a tube is equal to $2\pi r/\sigma$. The length σ is arguably set by a combination of the bending fluctuations between the protofibrils and the Coulomb interactions between them. In principle, we could calculate it by adding a bending term for the protofibrils and a Coulomb term to the overall free energy, and simultaneously minimise this free energy with respect to this length and the tube radius. We deem it not sensible to actually pursue this calculation given the crudity of the model in which both the DNA and protofibrils are represented by line segments without a volume. In our opinion a more advanced theory is needed to deal in full with the problem, e.g., along the lines of the work of Odijk for gels of charged semi-flexible.[35] We leave this for future work.

On a final note, we point out that the model presumes the mobile ions to be point particles and ignores the fact that they take up a finite volume. Actually, hydration water increases the radius of an ion by 0.15-0.3 nm.[36] If some or all of the counterions from the DNAs and CDs are released because of the little room remains in the space between them and mobile ions from the bulk solution cannot enter easily for the same reason, then the electrostatic interactions between DNA and CDs are not fully screened. In that case the background ionic strength has less impact on the binding free energy than our simplified model predicts, noting also that the Debye-Hückel approximation that we make use of in our calculation should prove even less accurate than it already is in the present calculation. In principle, we would need to also account for counterion condensation. Again, we leave this for future work.

---

[36] J. Israelachvili, Intermolecular and Surface Forces (Academic Press, London, 1998)