

Sequence-complementarity dependent co-assembly of phosphodiester-linked aromatic donor-acceptor trimers

Nadeema Appukutti, Alex H. de Vries,^b Prashant G. Gudeangadi,^a Bini R. Claringbold,^a Michelle D. Garrett,^c Michael R. Reithofer^d and Christopher J. Serpell*^a

^a School of Physical Sciences, Ingram Building, University of Kent, Canterbury, Kent, CT2 7NH, UK

^b Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, the Netherlands

^c School of Biosciences, Stacey Building, University of Kent, Canterbury, Kent, CT2 7NJ, UK

^d Dept. of Inorganic Chemistry, University of Vienna, Währinger Straße. 42, 1090 Vienna, Austria

Supplementary Information

Contents

Materials and instrumentation	2
Synthesis and molecular characterisation	3
UV-visible spectroscopy	9
Atomic force microscopy	12
Transmission electron microscopy	18
Molecular dynamics	24
References	30

Materials and Instrumentation

All starting materials and solvents were commercially available and used without further purification. 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite and 5-(ethylthio)-1H-tetrazole reagents were purchased from LinkTech.

NMR spectra were obtained using Bruker AV2 400 MHz spectrometer. The data was processed using ACD labs software. NMR chemical shift values are reported in parts per million (ppm) and calibrated to the centre of the residual solvent peak set.

Electrospray mass spectra were recorded on a Bruker micrOTOF-Q II mass spectrometer. Samples were introduced into the mass spectrometer by on-line reverse-phase HPLC on a Phenomenex Nucleosil C18 column (3 μm , 120 \AA , 2.0 mm x 150 mm) running on an Agilent 1100 HPLC system at a flow rate of 0.2 mL/min using a short gradient from 10% B to 100% B (A: 15 mM TEA, 400 mM HFIP in water; B: 15 mM TEA, 400 mM HFIP in methanol). The eluent was monitored at 200-800 nm and then directed into the electrospray source, operating in negative ion mode, at 3.5 kV and mass spectra recorded from 400-3000 m/z. Data was analysed and deconvoluted to give uncharged masses with Bruker's Compass Data Analysis software.

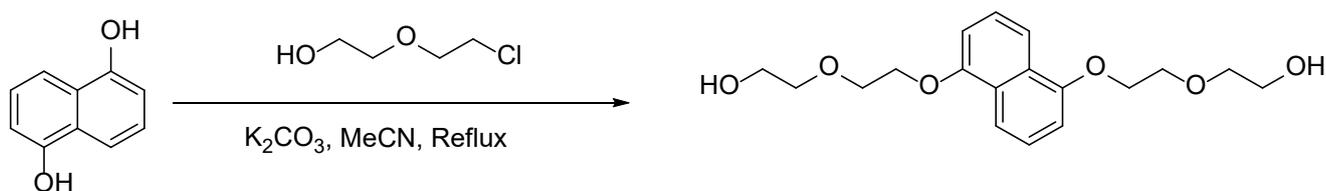
UV-visible absorption spectra of the samples were obtained using NanoDrop One spectrophotometer by Thermo Fisher Scientific or BMG Labtech Clariostar plate reader using MARS data analysis software.

Samples for atomic force microscopy (AFM) were prepared in water at 1 mM concentration. Samples were deposited on freshly cleaved mica sheets for one minute, followed by 5 water washes with 5 μL of water. Excess liquid was blotted off with the edge of a filter paper. Prepared mica plates were dried under vacuum for one hour. Measurements were conducted on a Bruker Multi-Mode microscope with a Quadrex Nanoscope III controller using Bruker ScanAsyst-Air silicon tip on nitride lever, with frequency of 70 Hz. Data were analyzed using Nanoscope Analysis 1.5 software (Bruker, CA, US).

Samples for transmission electron microscopy (TEM) were prepared in water and then dropped onto a carbon coated copper 200 mesh TEM grids and negatively stained using 2% uranyl acetate (5 μL , 30 seconds). The stain was wicked off. Samples were imaged on a Jeol 1230 TEM, operating at an accelerating voltage of 80 kV and the images were recorded with a Gatan Multiscan 790 digital camera.

Synthesis and Molecular Characterisation

DAN monomer 1a



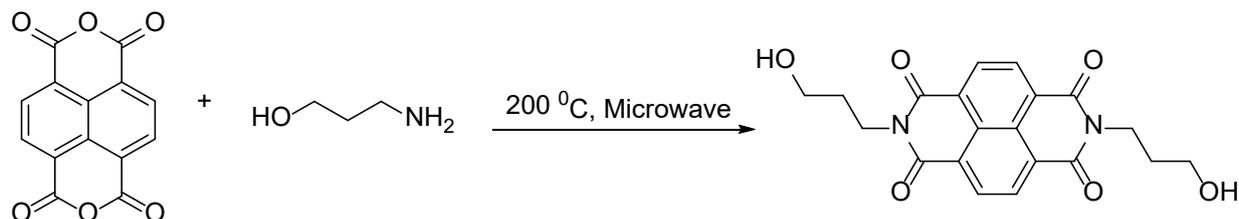
Potassium carbonate (187.3 mmol, 25.86 g) was suspended in acetonitrile and the solution degassed under nitrogen. 18.73 mmol (3.00 g) of 1,5-dihydroxynaphthalene was added to the solution. 41.2 mmol (4.35 mL) of chloroethoxyethanol was added next, and the solution was refluxed over three nights under a nitrogen atmosphere. The solvent was evaporated and the residue was dissolved in dichloromethane and washed with H₂O (x1), NaOH (x2) and brine water (x1) then dried over anhydrous MgSO₄. The filtered product was dried under vacuum and the light brown powder was obtained in 52 % yield.

¹H-NMR (400 MHz, CDCl₃, 25 °C): δ (ppm) 7.80 (d, 2H Ar), 7.33 (t, 2H Ar), 6.81 (d, 2H Ar), 4.16 (t, 2H), 2.01 (t, 2H)

¹³C-NMR (100 MHz, CDCl₃, 25 °C): δ (ppm) 153.96, 125.14, 122.15, 114.54, 195.03, 70.51, 66.97, 30.93, 24.79

HR ESI-MS: Found 337.1663 m/z. Calc'd for [M+H]⁺ 337.1646 m/z

NDI Monomer 1b



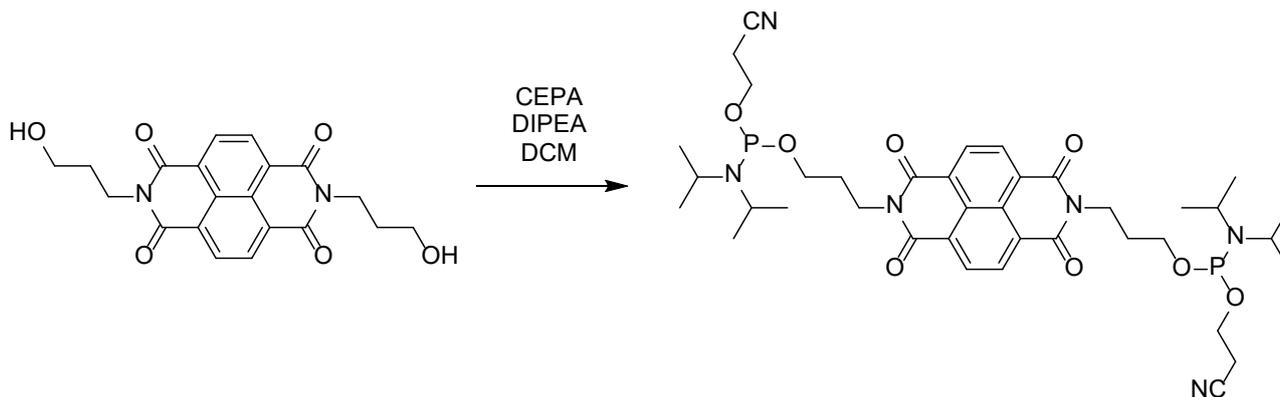
1,4,5,8-Naphthalenetetracarboxylic dianhydride (1.86 mmol, 0.5 g), water (20 mL) 3-aminopropanol (3.7 mmol, 282 μ L) were added to a microwave tube, heated at 200 °C for 30 minutes in the microwave. When the solution has reached the room temperature the precipitated solid was filtered and washed with water and diethyl ether, giving the pure product as white crystals in 97% yield.

¹H-NMR (400 MHz, CDCl₃, 25 °C): δ (ppm) 8.78 (s, 2H Ar), 4.37 (t, 2H), 3.63 (d, 2H), 2.62 (t, 2H), 2.01 (t, 2H)

¹³C-NMR (100 MHz, CDCl₃, 25 °C): δ (ppm) 131.41, 58.70, 37.85, 31.44, 2.02

HR ESI-MS: Found 383.1246 m/z. Calc'd for [M+H]⁺ 383.1238 m/z

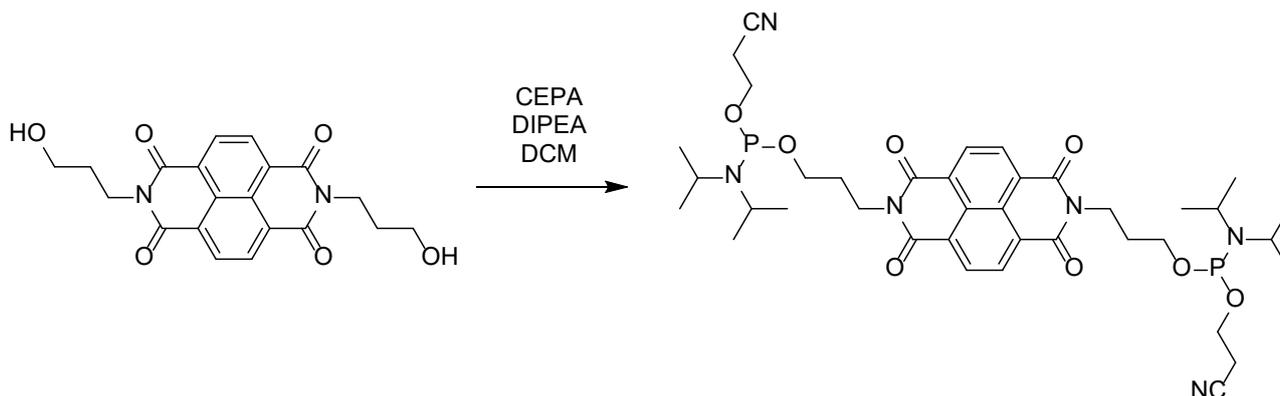
Double-phosphoramidite activated DAN 2a



DAN diol **1a** (2 mmol, 672 mg), and dimethylaminopyridine (75 mg) were dissolved in 100 mL of dry DCM, to which diisopropylethylamine (10 mmol, 1.432 mL) and 2-cyanoethyl diisopropylchlorophosphoramidite (6 mmol, 1.33 mL) were then added. The mixture was then stirred at room temperature under an atmosphere of nitrogen, for 2 hours. After the DCM was removed under vacuum, and a thick brown oil was obtained, which was used immediately after obtaining ^{31}P spectrum.

^{31}P NMR: 148.03 ppm

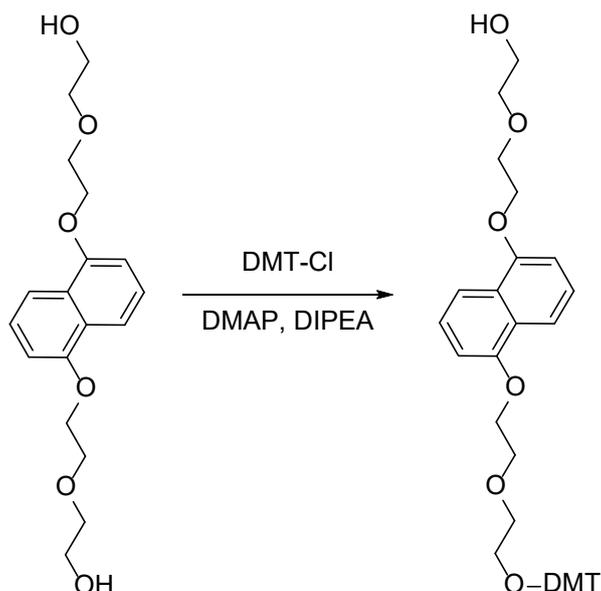
Double-phosphoramidite activated NDI 2b



NDI diol **1b** (2 mmol, 764 mg), and dimethylaminopyridine (75 mg) was dissolved in 100 mL of dry DCM, to which diisopropylethylamine (10 mmol, 1.432 mL) and 2-cyanoethyl diisopropylchlorophosphoramidite (6 mmol, 1.33 mL) were added. The mixture was then stirred at room temperature under an atmosphere of nitrogen, for 2 hours. After the DCM was removed under vacuum, and a thick brown oil was obtained, which was used immediately after obtaining ^{31}P spectrum.

^{31}P NMR: 147.14 ppm

Singly DMT-protected DAN monomer 3a



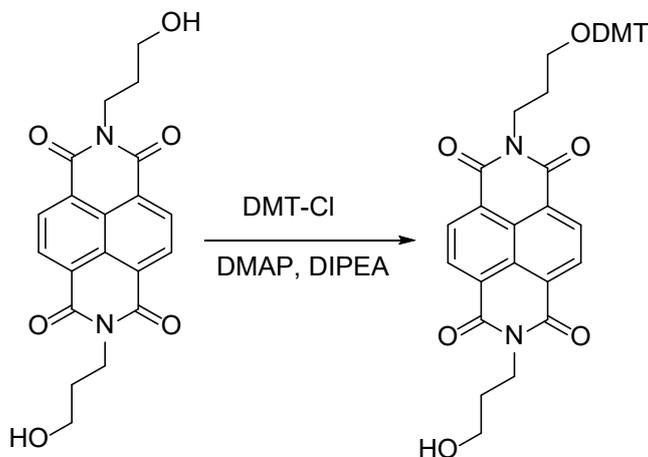
DAN diol **1a** (2.29 mmol, 770 mg) and dimethylaminopyridine (0.46 mmol, 56.1 mg) were dissolved in 200 mL of dry dichloromethane. Dimethoxytrityl chloride (2.29 mmol, 775 mg) was then added, followed by the dropwise addition of diisopropylamine (2.98 mmol, 0.519 mL) over a period of 30 minutes. The mixture was left stirring under an atmosphere of nitrogen for an additional 24 hours. After the dichloromethane was removed under vacuum, the crude product, a beige powder, was purified by column chromatography with a basic alumina stationary phase, eluted with dichloromethane:methanol (400:1). The dried product **3a** was obtained at 520 mg with 35% yield

¹H-NMR (400 MHz, DMSO-d₆, 25 °C): δ (ppm) 7.75 (t, 2H Ar), 7.40 (t, 3H Ar), 6.81 (d, 2H Ar), 4.16 (t, 2H), 2.01 (t, 2H)

¹³C-NMR (100 MHz, DMSO-d₆, 25 °C): δ (ppm) 158.42, 154.41, 145.48, 136.23, 130.12, 128.23, 128.18, 127.04, 126.48, 125.85, 113.55, 106.38, 85.74, 70.44, 69.61, 68.29, 63.42

HR ESI-MS: Found 661.2780 m/z. Calc'd for [M+NH₄]⁺ 661.2777 m/z

Singly DMT-protected NDI monomer 3b



NDI diol **1b** (2.29 mmol, 874 mg) and dimethylaminopyridine (0.46 mmol, 56.1 mg) were dissolved in 200 mL of dry dichloromethane. Dimethoxytrityl chloride (2.29 mmol, 775 mg) was then added, followed by the dropwise addition of diisopropylamine (2.98 mmol, 0.519 mL) over a period of 30 minutes. The mixture was left stirring under an atmosphere of nitrogen for an additional 24 hours. After the dichloromethane was

removed under vacuum, and the crude product, a brown powder, was purified by column chromatography using a basic alumina stationary phase, eluted with dichloromethane:methanol (400:1). The dried product **3b** was obtained at 485 mg with 30% yield.

¹H-NMR (400 MHz, DMSO-*d*₆, 25 °C): δ (ppm) 8.67 (s, 4H Ar), 7.24 (d, 2H Ar), 7.19 (t, 2H Ar), 7.13 (t, 1H Ar), 7.07 (d, 4H Ar), 6.69 (d, 4H Ar), 4.58 (t, 2H), 4.16 (p, 4H), 3.66 (s, 6H), 3.53 (q, 2H), 3.07 (t, 2H)

¹³C-NMR (100 MHz, DMSO-*d*₆, 25 °C): δ (ppm) 163.39, 159.03, 155.42, 148.30, 148.17, 140.22, 130.66, 129.55, 129.47, 128.17, 127.57, 126.96, 126.83, 112.84, 112.57, 106.85, 81.40, 59.95, 54.67, 53.84, 38.17, 31.04

HR ESI-MS: Found 685.2565 m/z. Calc'd for [M+H]⁺ 685.2544 m/z

Synthesis of Trimers

Synthesis was carried out as a one-pot reaction. Double phosphoramidite derivatives (either **2a** or **2b**) were dissolved in 50 mL of acetonitrile. 5 mL of 0.25M 5-ethylthio-1H-tetrazole activator was added to the mixture. Three molar equivalents of mono protected derivative (either **3a** or **3b**) was added, and reaction was stirred for 30 minutes. 5 mL of oxidation mix (0.015 M iodine in water/pyridine/THF 2/20/78) was added to the mixture and reaction was stirred for another 10 minutes. The products were vacuumed dried and column chromatography was carried out using a basic alumina stationary phase, eluted with hexane:ethyl acetate (1:1), and then again using basic alumina eluted with DCM: MeOH (3:1). The cyanoethyl protecting group was removed from the purified sample by adding 5 mL of concentrated ammonium hydroxide and heating it at 60 °C for 1 hour, with NH₄OH subsequently removed under vacuum. A 4:1 acetic acid:water mixture was then added to remove DMT protecting groups. After removal of solvent under vacuum, 20 mL of DCM was added to the solid sonicated for 2-5 minutes to remove soluble impurities (protecting groups). This process was repeated three times and the remaining residue was dried with heat (60 °C) and vacuum.

Each product, **DAN₃**, **NDI₃**, **DAN-NDI-DAN** and **NDI-DAN-NDI** gave 53%, 52 %, 37% and 42% yields, respectively.

NMR spectra of trimers

DAN₃

¹H NMR (400 MHz, DMSO-d₆, 25°C): 7.74 (d, 6H, Ar), 7.39 (t, 6H, Ar), 6.99 (d, 6H, Ar), 4.24 (t, 10H), 3.88 (t, 11H), 3.56 (m, 20H), 3.06 (m, 12H)

¹³C NMR (100 MHz, DMSO-d₆, 25°C): 154.32, 126.40, 125.92, 114.27, 106.37, 73.06, 69.45, 68.25, 60.75, 57.00, 52.55, 46.73, 46.13, 21.49, 19.21

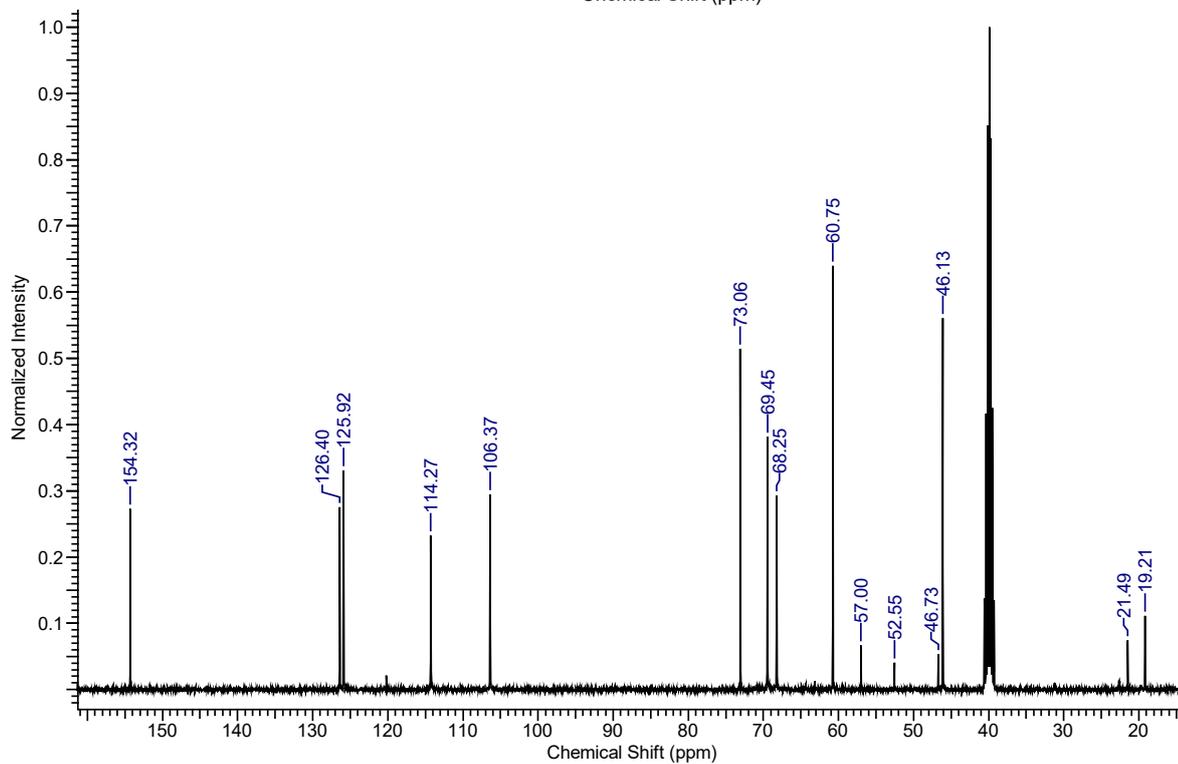
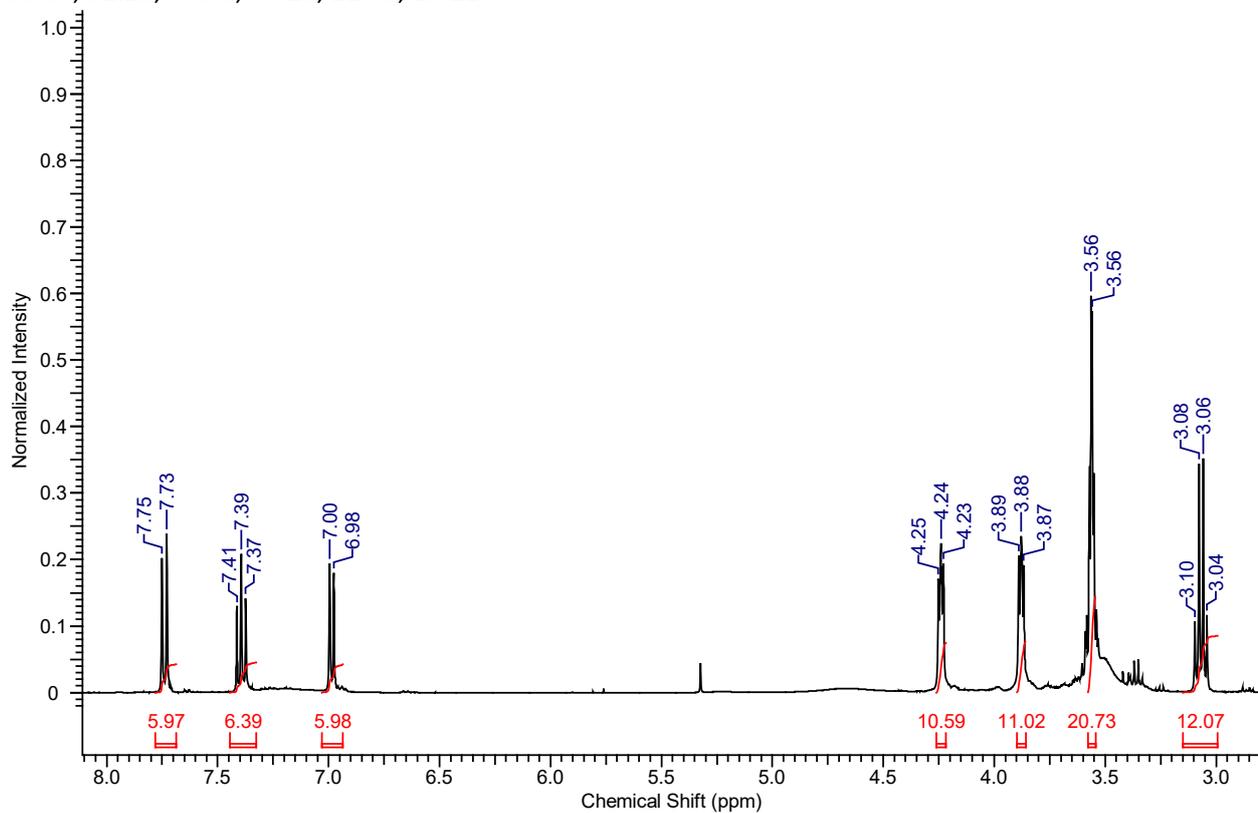


Figure S1. ¹H and ¹³C NMR spectra of DAN₃ (DMSO-d₆).

NDI₃

¹H NMR (400 MHz, D₂O, 25°C): 8.16 (s, 12H, Ar), 4.39 (t, 3H), 4.00 (m, 12H), 3.96 (m, 6H), 3.66 (m, 6H), 3.27 (s, 2H), 3.06 (t, 3H), 1.95 (m, 6H), 1.85 (m, 6H)

¹³C NMR (100 MHz, D₂O, 25°C): 163.01, 157.38, 156.53, 141.17, 138.01, 130.78, 125.25, 124.94, 117.86, 107.84, 106.63, 61.94, 59.36, 52.20, 48.80, 47.15, 39.45, 38.19, 29.77, 28.28, 19.21, 18.21

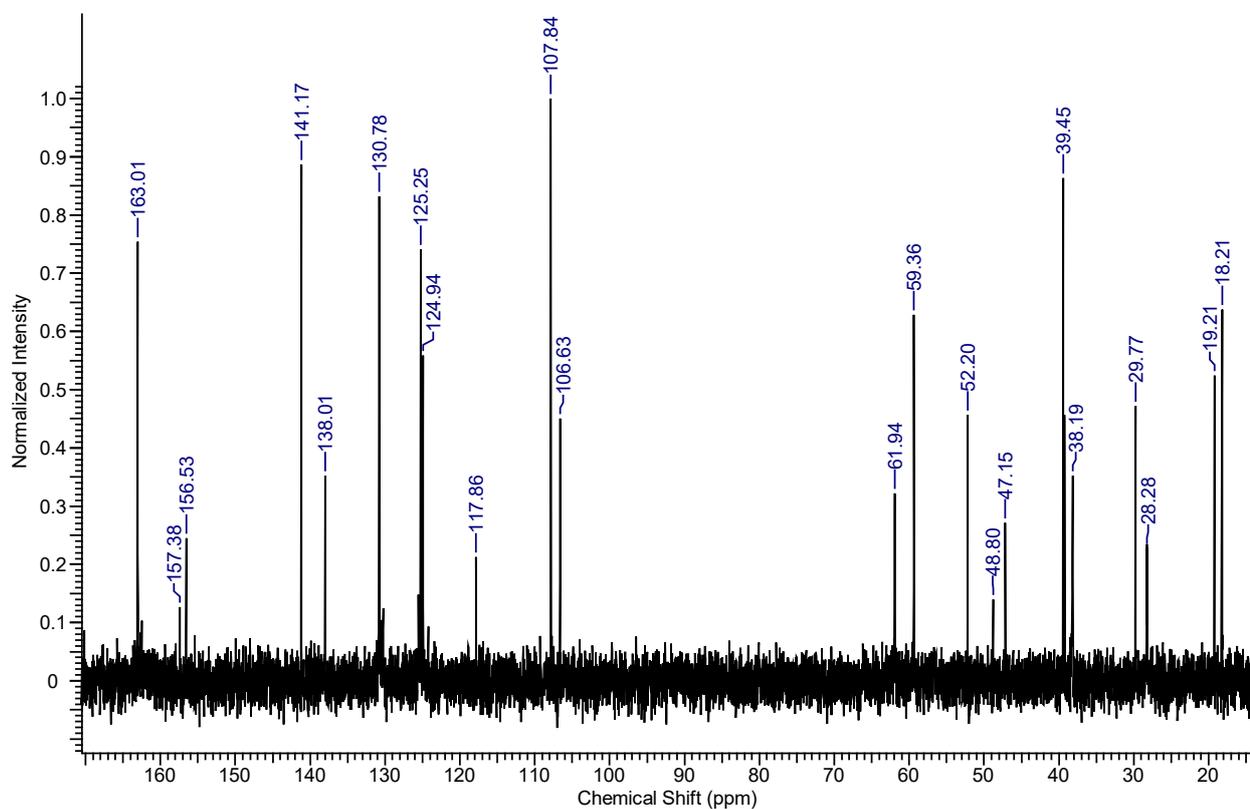
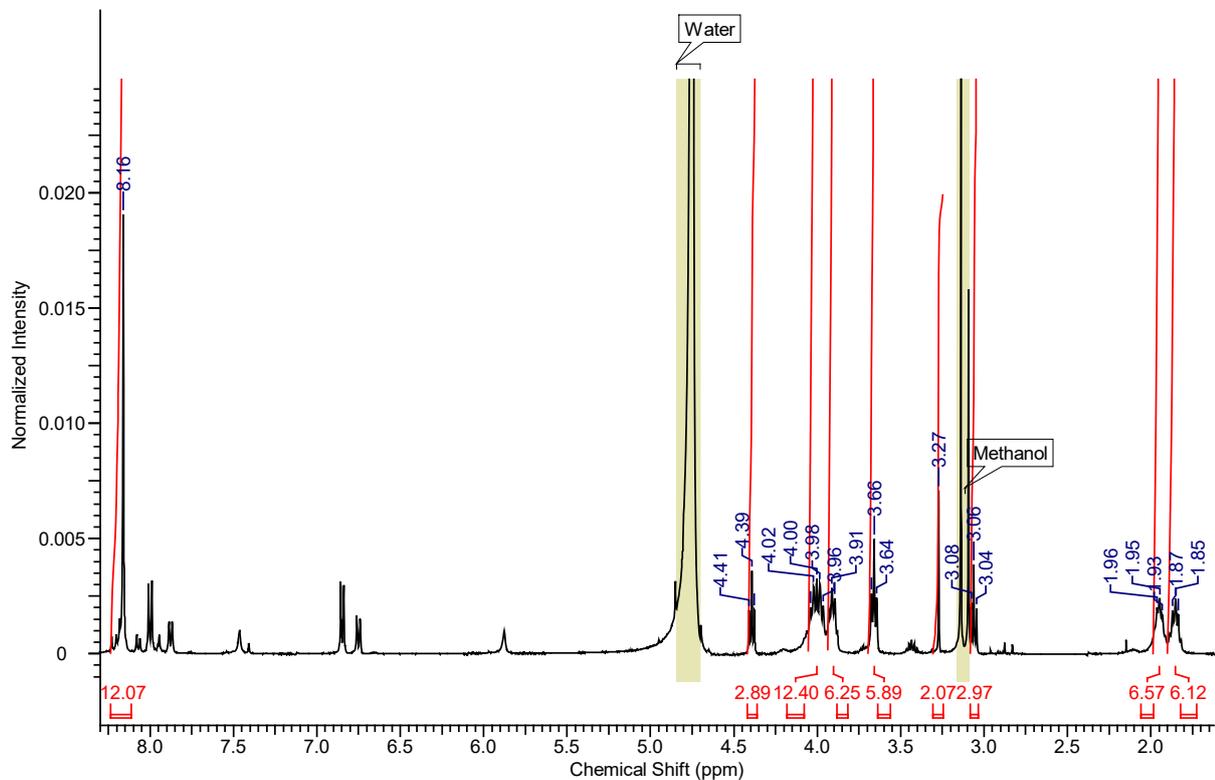


Figure S2. ¹H and ¹³C NMR spectra of NDI₃ (D₂O).

DAN-NDI-DAN

^1H NMR (400 MHz, DMSO- d_6 , 25°C): 8.56 (s, 4H, Ar), 7.69 (d, 4H, Ar), 7.36 (t, 4H, Ar), 6.96 (d, 4H, Ar), 4.67 (t, 4H), 4.56 (t, 2H), 4.24 (t, 8H), 4.11 (t, 4H), 3.89 (t, 8H), 3.57 (m, 18H), 3.35 (s, 2H), 1.83 (quin., 4H)

^{13}C NMR (100 MHz, DMSO- d_6 , 25°C): 162.12, 153.40, 129.85, 125.69, 125.45, 124.97, 113.36, 105.42, 72.33, 68.59, 67.31, 59.90, 58.59, 37.68, 30.40

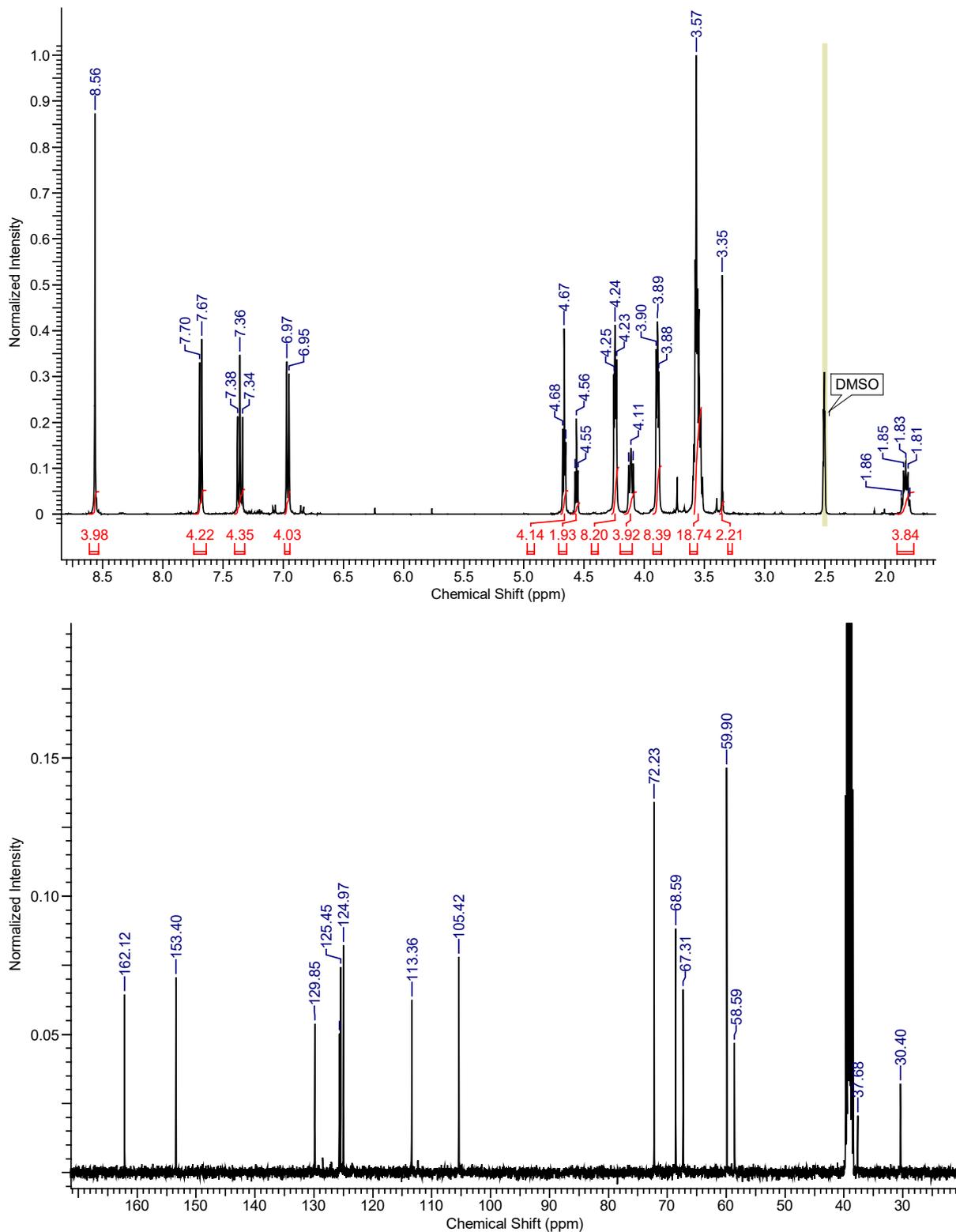


Figure S3. ^1H and ^{13}C NMR spectra of DAN-NDI-DAN (DMSO- d_6).

NDI-DAN-NDI

^1H NMR (400 MHz, DMSO- d_6 , 25°C): 8.95 (s, 8H, Ar), 8.20 (d, 2H, Ar), 7.40 (m, 2H, Ar), 6.98 (m, 2H, Ar), 5.90 (s, 2H), 4.24 (t, 1H), 3.88 (t, 1H), 3.57 (m, 12H), 3.30 (m, 9H), 3.17 (s, 6H), 3.10 (m, 6H), 2.59 (t, 6H)

^{13}C NMR (100 MHz, DMSO- d_6 , 25°C): 157.32, 139.57, 125.91, 120.24, 107.42, 73.06, 69.46, 60.72, 56.97, 53.62, 46.39, 41.96, 21.43, 19.01, 12.61

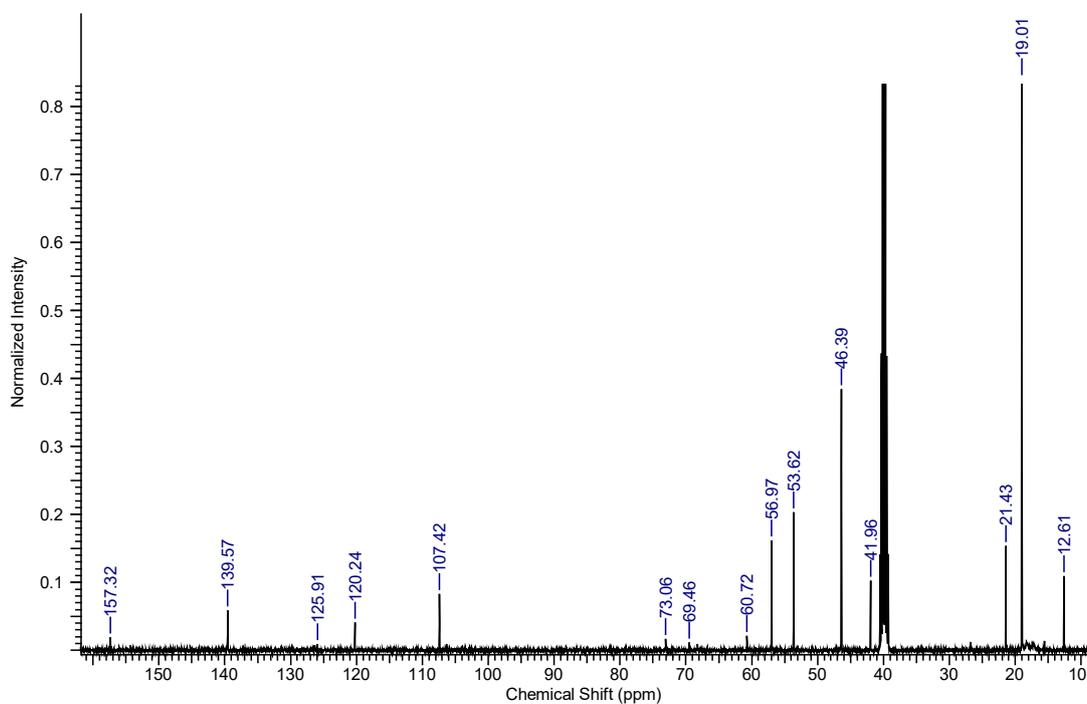
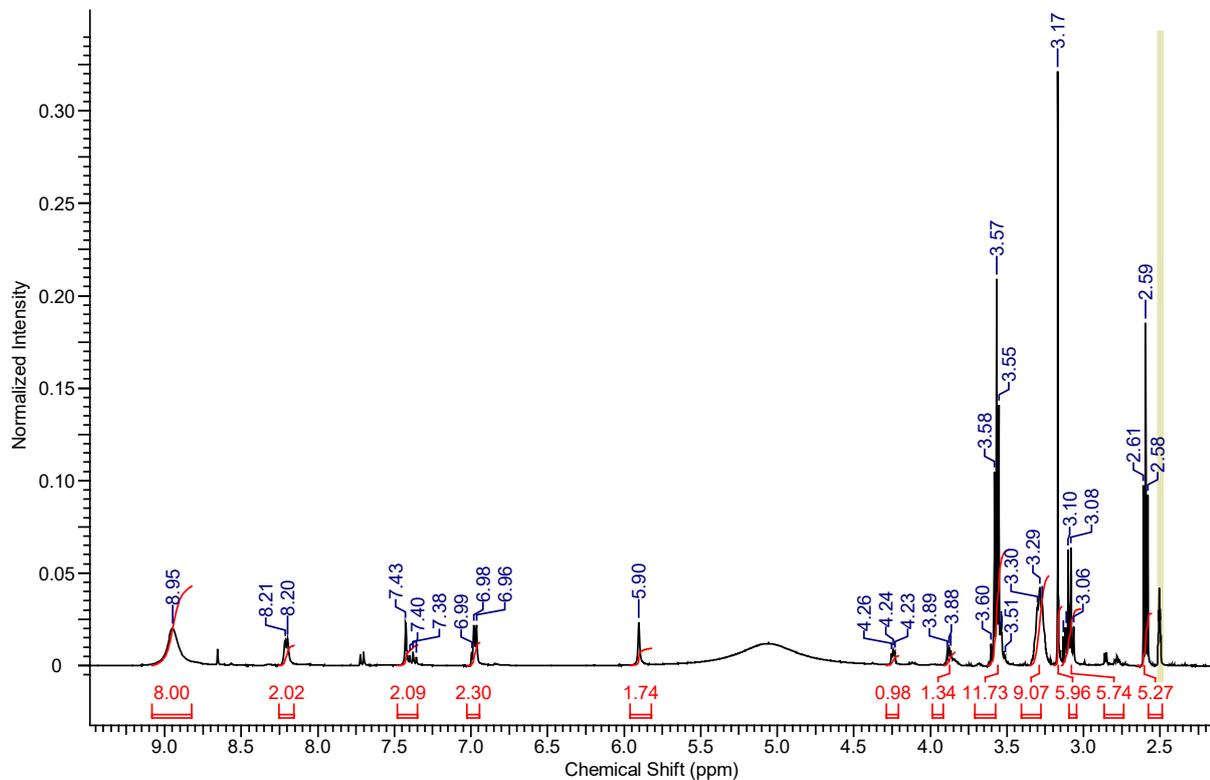


Figure S4. ^1H and ^{13}C NMR spectra of NDI-DAN-NDI (DMSO- d_6).

Mass spectra of trimers

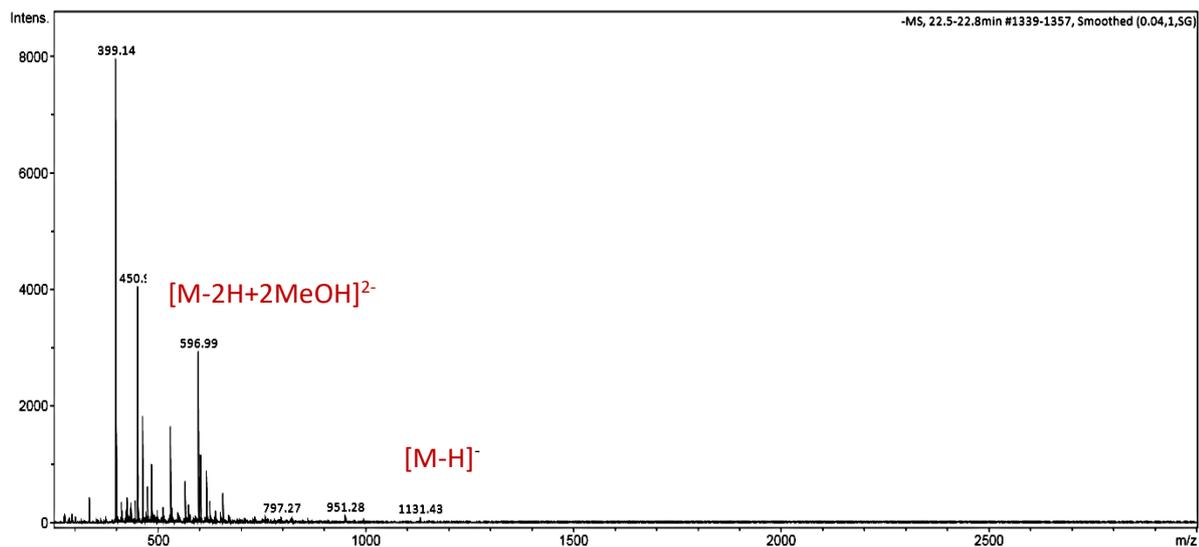


Figure S5. Mass spectrum of DAN₃. [M-H]⁻ calc. m/z = 1131.38. [M-2H+2MeOH]²⁻ calc. m/z = 597.21 (bottom).

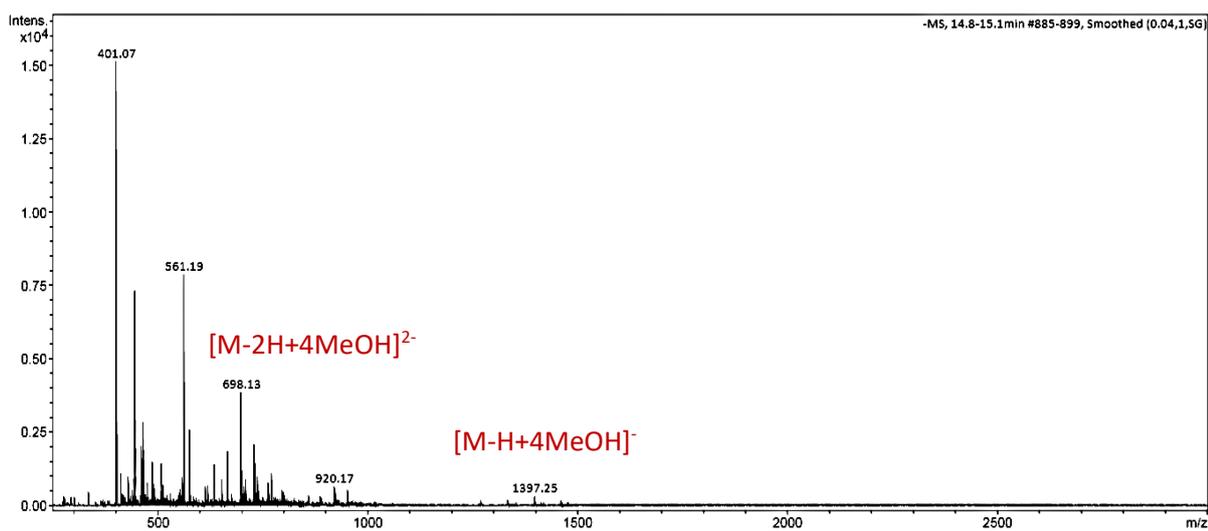


Figure S6. ¹H and ¹³C NMR spectra of NDI₃ in D₂O (top). Mass spectrum of NDI₃. [M-H+4MeOH]⁻ calc. m/z = 1397.36. [M-2H+4MeOH]²⁻ calc. m/z = 698.18

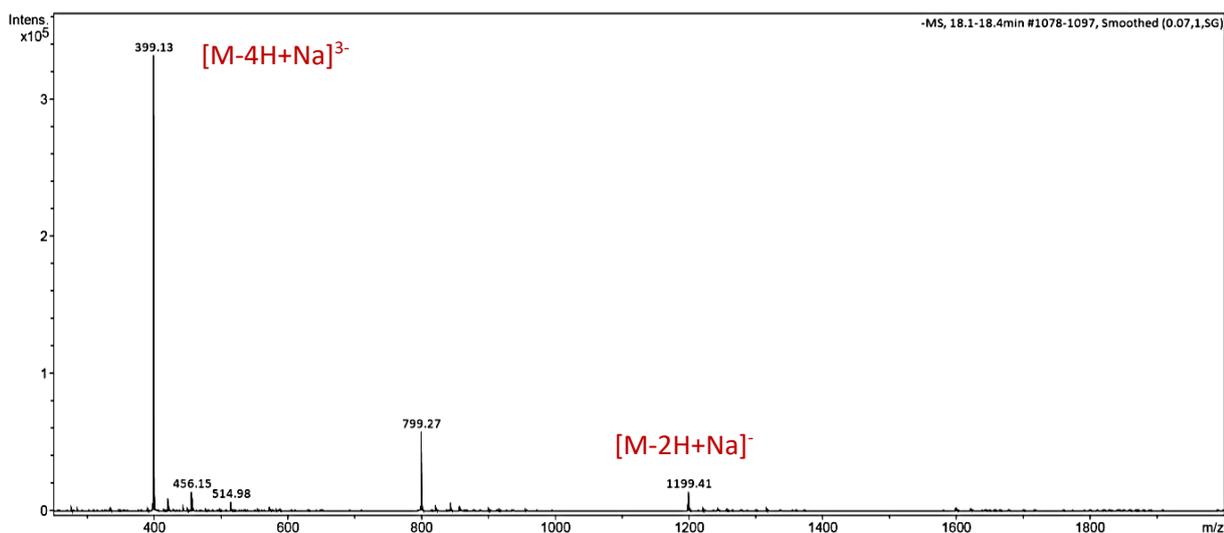


Figure S7. Mass spectrum of **DAN-NDI-DAN**. $[M-2H+Na]^{-}$ calc. $m/z = 1199.32$. $[M-4H+Na]^{-}$ calc. $m/z = 399.10$.

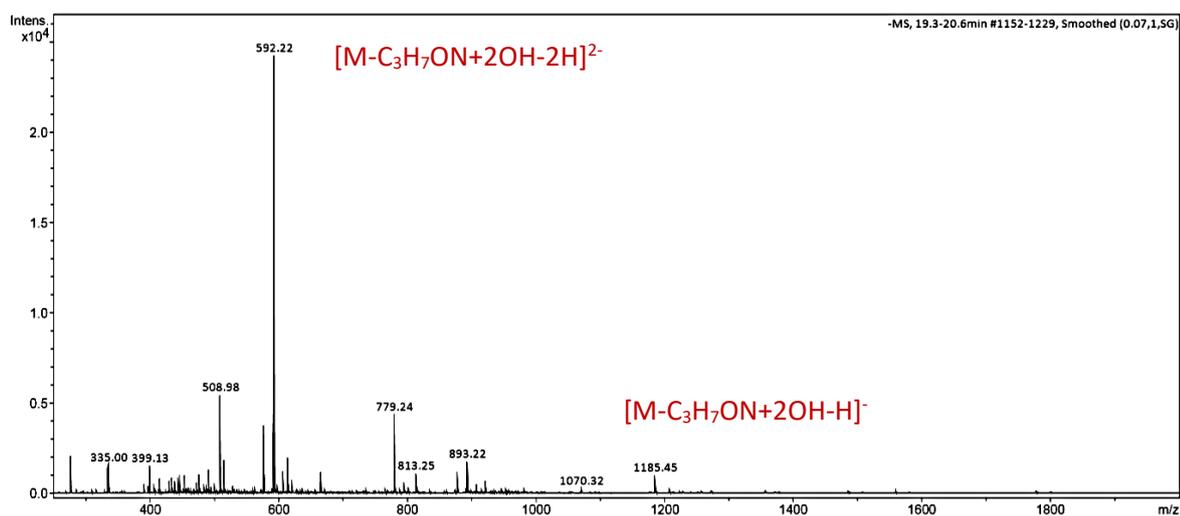


Figure S8. Mass spectrum of **NDI-DAN-NDI**. $[M-C_3H_7ON+2OH-H]^{-}$ calc. $m/z = 1184.25$. $[M-C_3H_7ON+2OH-2H]^{2-}$ calc. $m/z = 591.62$. This corresponds to hydrolysis of a single imide within the mass spectrometer.

UV-visible spectroscopy

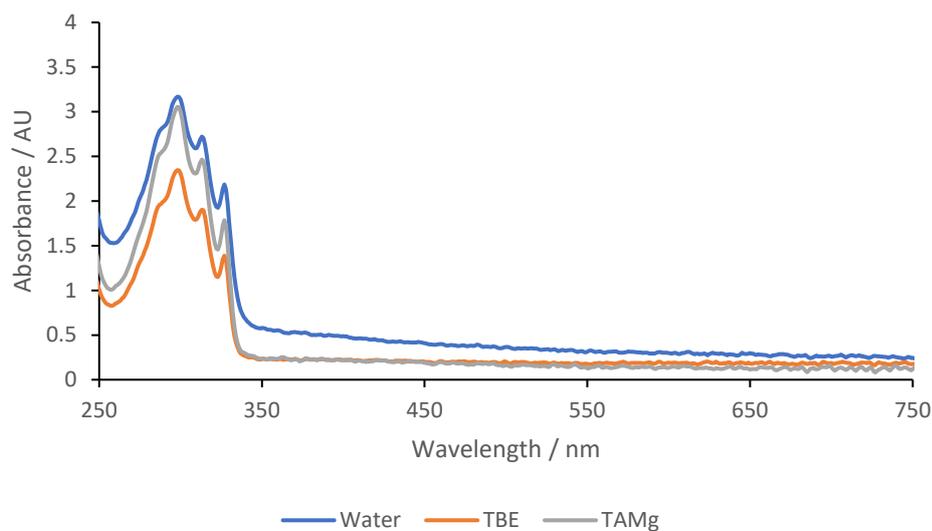


Figure S9. UV-visible absorption spectrum of **DAN₃** at 250 μM in water, and TBE and TAMg buffer.

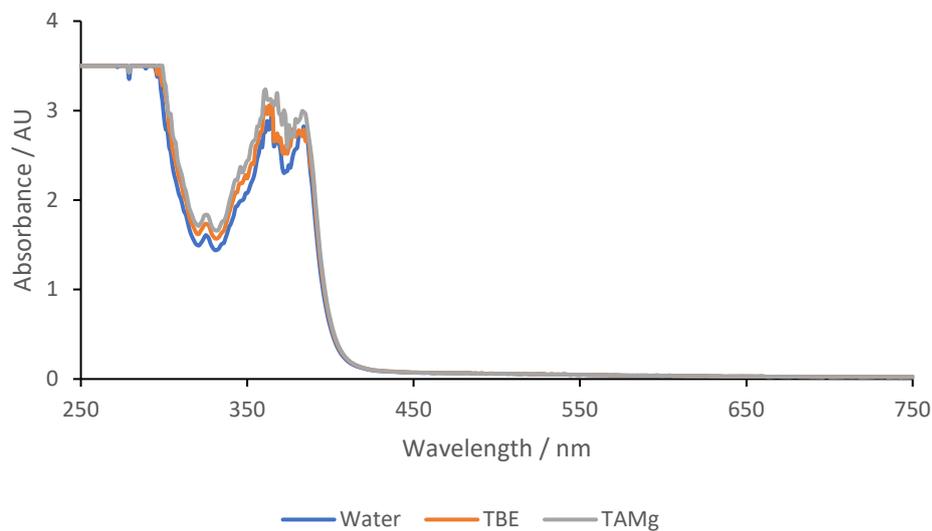


Figure S10. UV-visible absorption spectrum of **NDI₃** at 250 μM in water, and TBE and TAMg buffer.

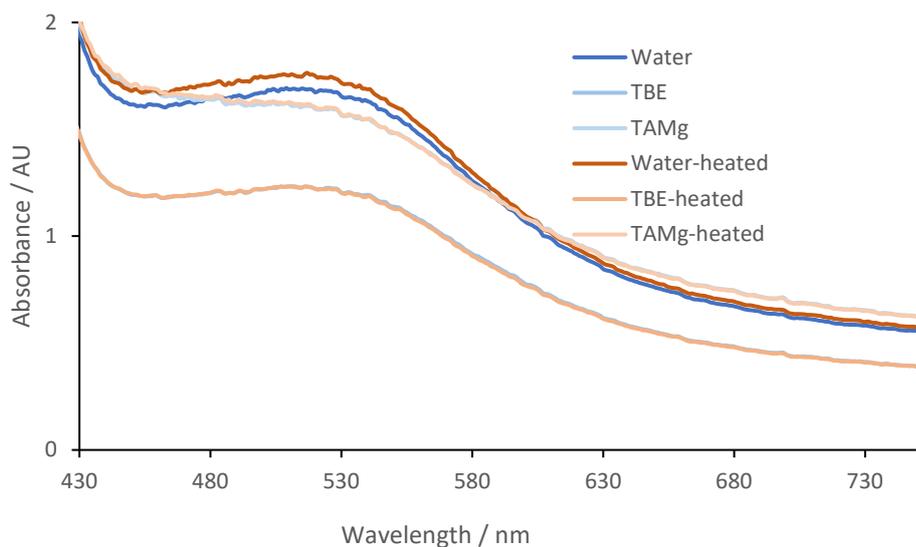


Figure S11. UV-visible absorption spectra of 1:1 $\text{DAN}_3\text{:NDI}_3$ (10 μM each) at 25 °C and 45 °C in different buffer conditions at 10 mM, focusing on the charge transfer band.

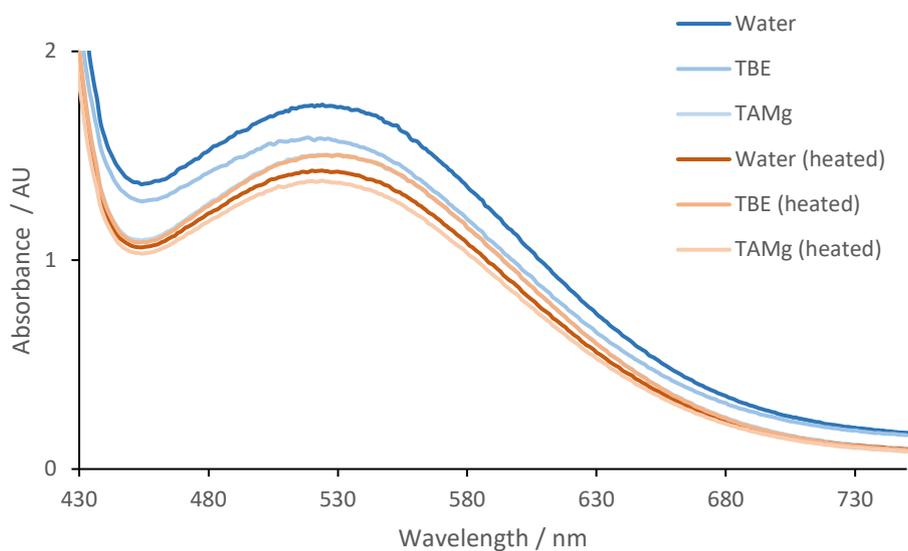


Figure S12. UV-visible absorption spectra of DAN-NDI-DAN at 25 °C and 45 °C in different buffer conditions at 10 mM, focusing on the charge transfer band.

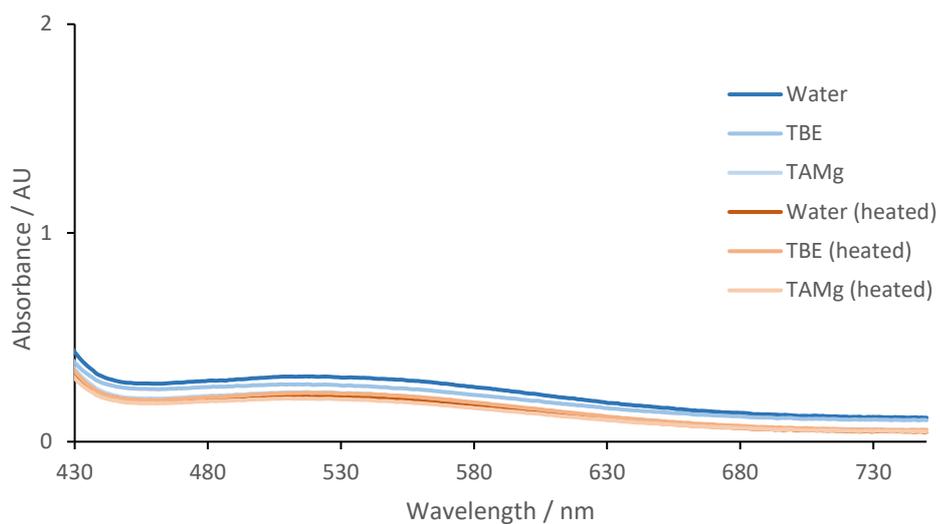


Figure S13. UV-visible absorption spectra of **NDI-DAN-NDI** at 25 °C and 45 °C in different buffer conditions at 10 mM, focusing on the charge transfer band.



Figure S14. Visible colour change observed in 1:1 **DAN-NDI-DAN:NDI-DAN-NDI** at room temperature (left) and above 45 °C (right)

Atomic force microscopy

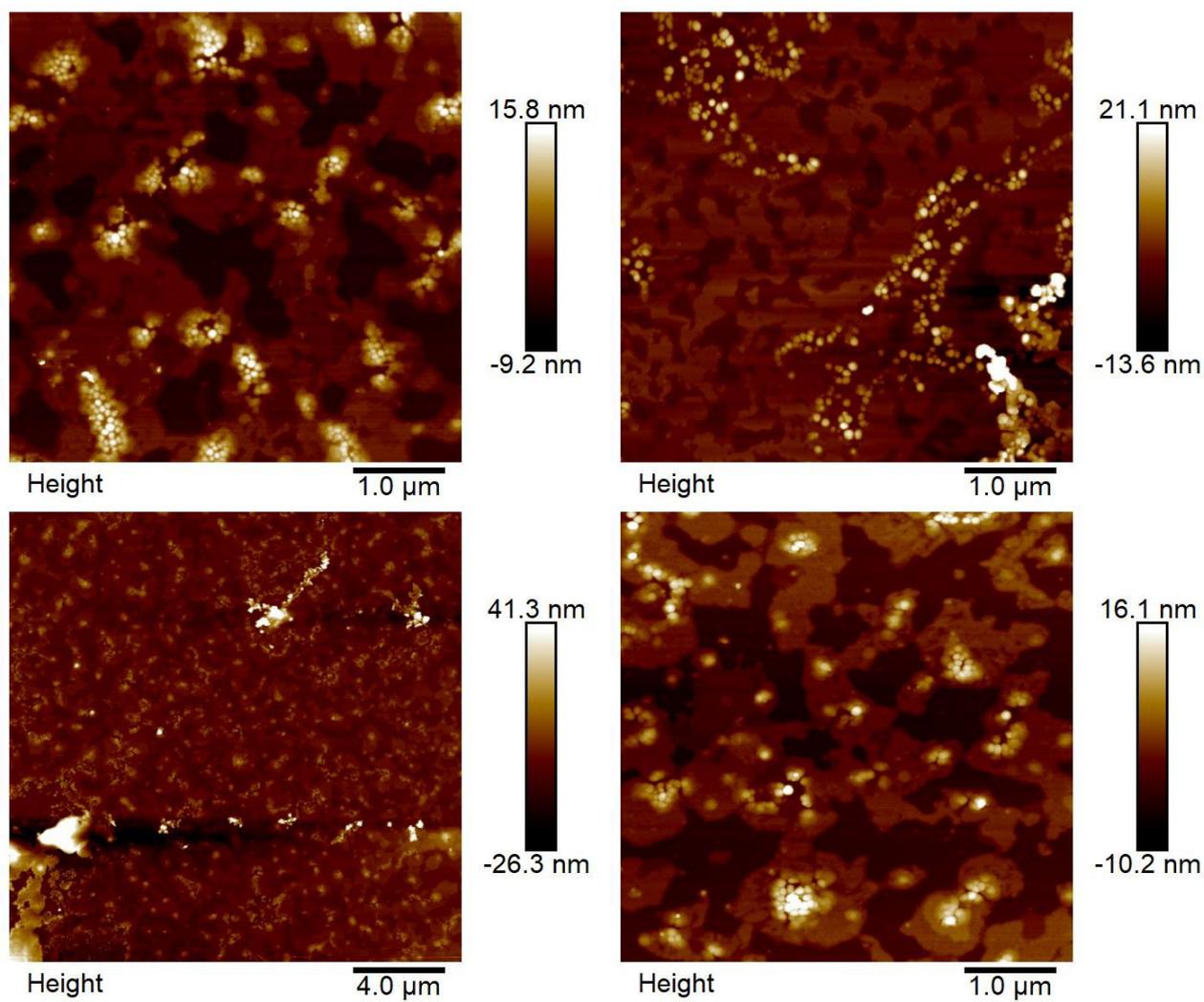


Figure S15. Supplementary AFM images of DAN₃.

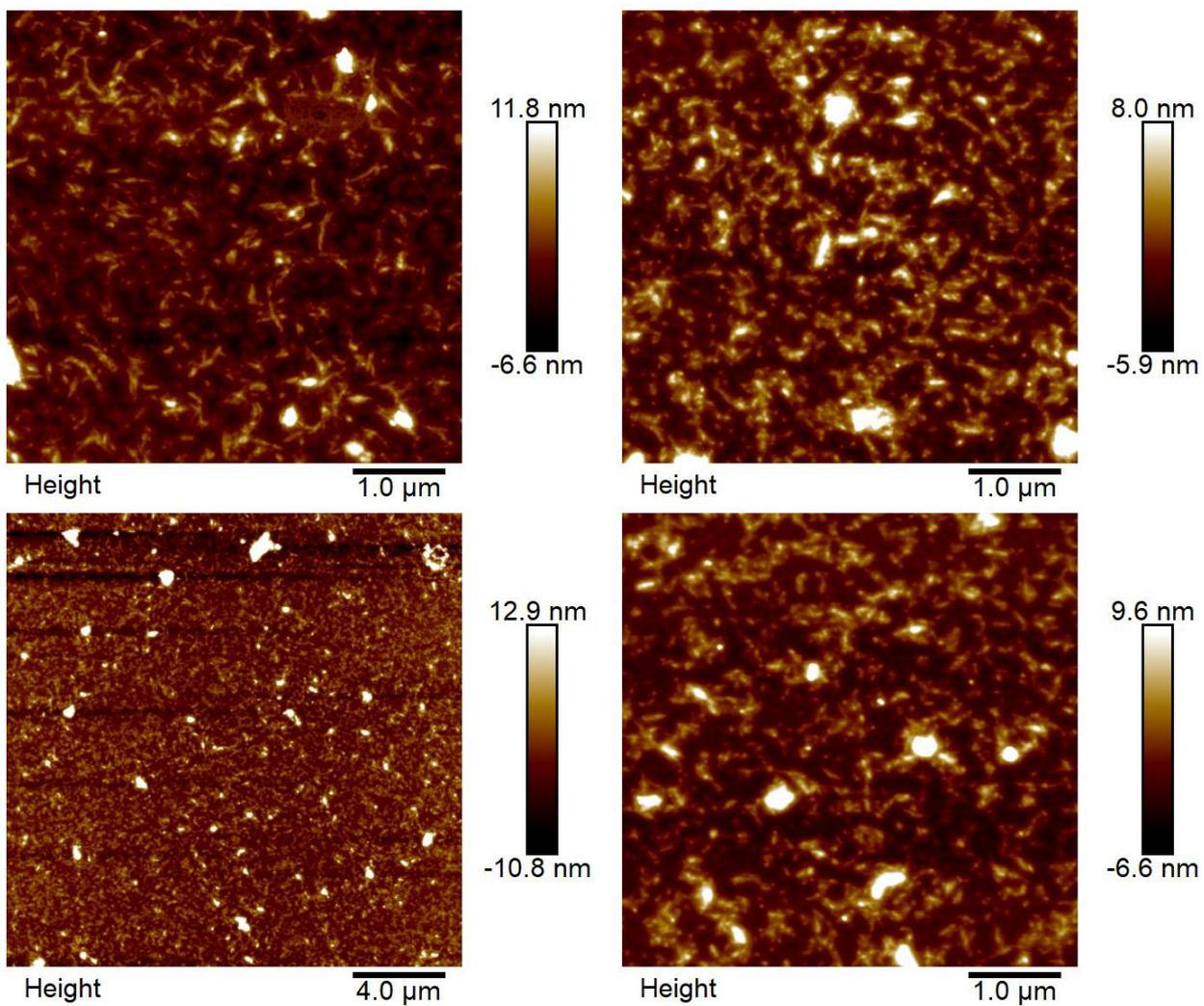


Figure S16. Supplementary AFM images of NDI₃.

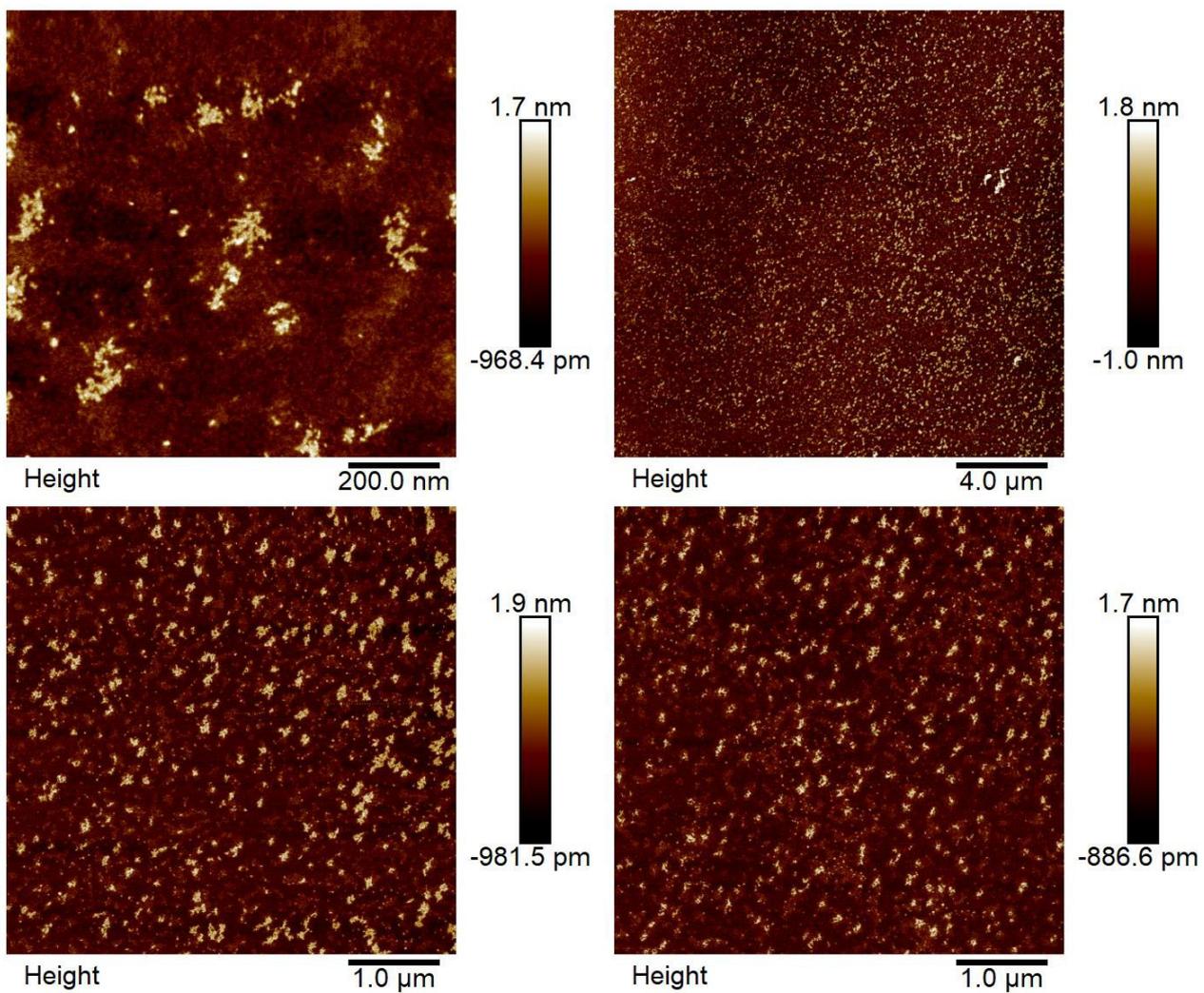


Figure S17. Supplementary AFM images of **DAN-NDI-DAN**.

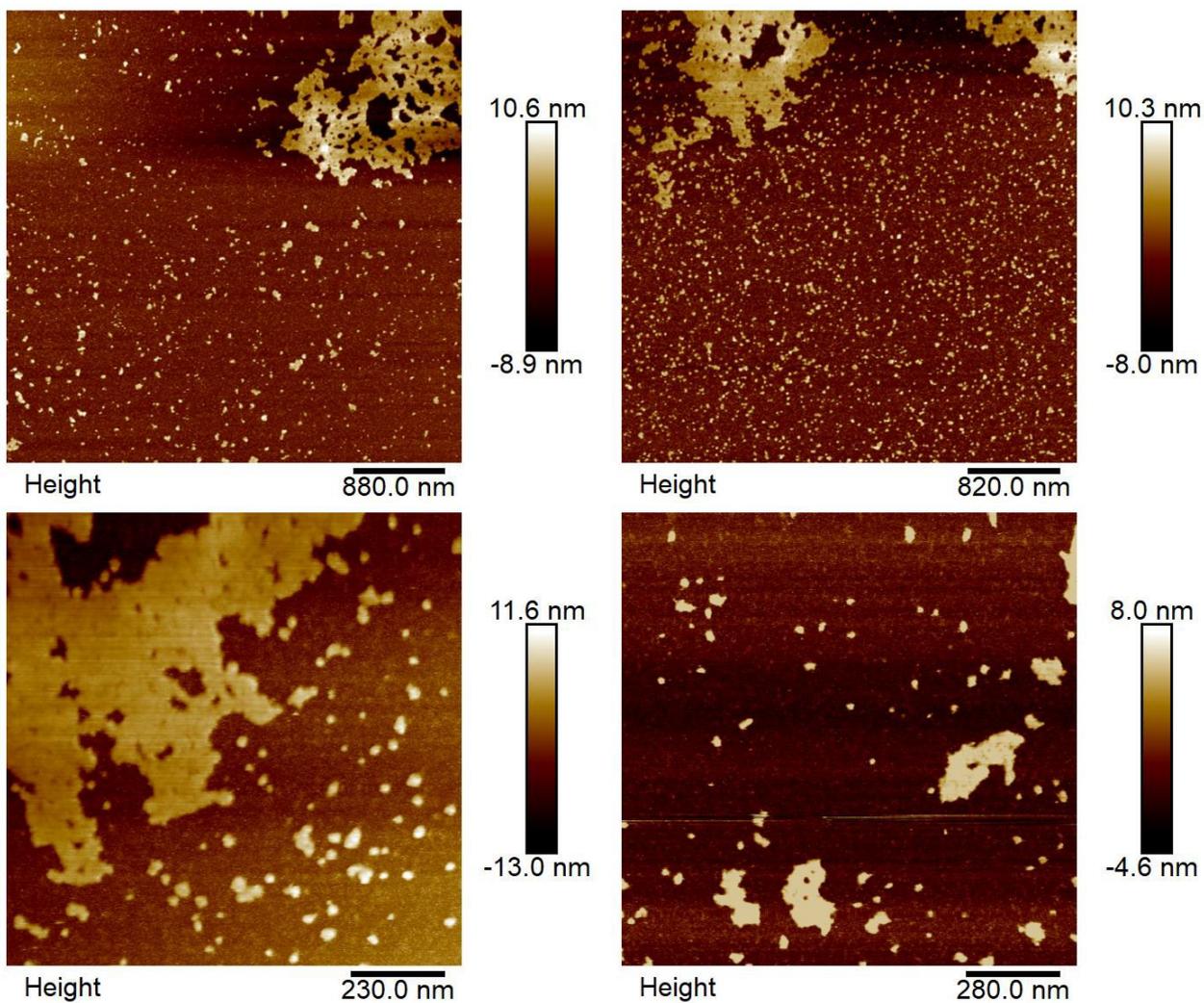


Figure S18. Supplementary AFM images of **NDI-DAN-NDI**.

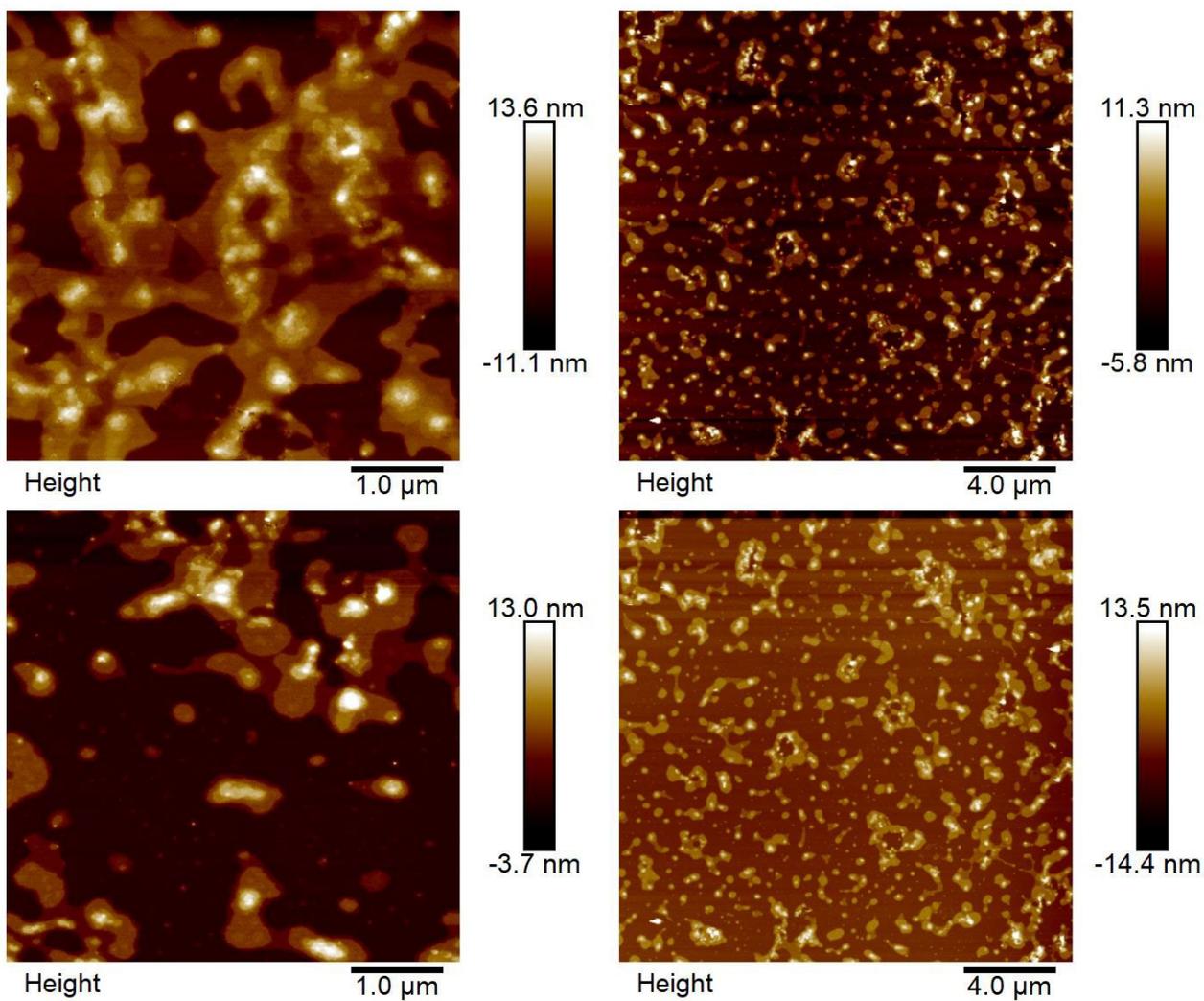


Figure S19. Supplementary AFM images of 1:1 $\text{DAN}_3:\text{NDI}_3$

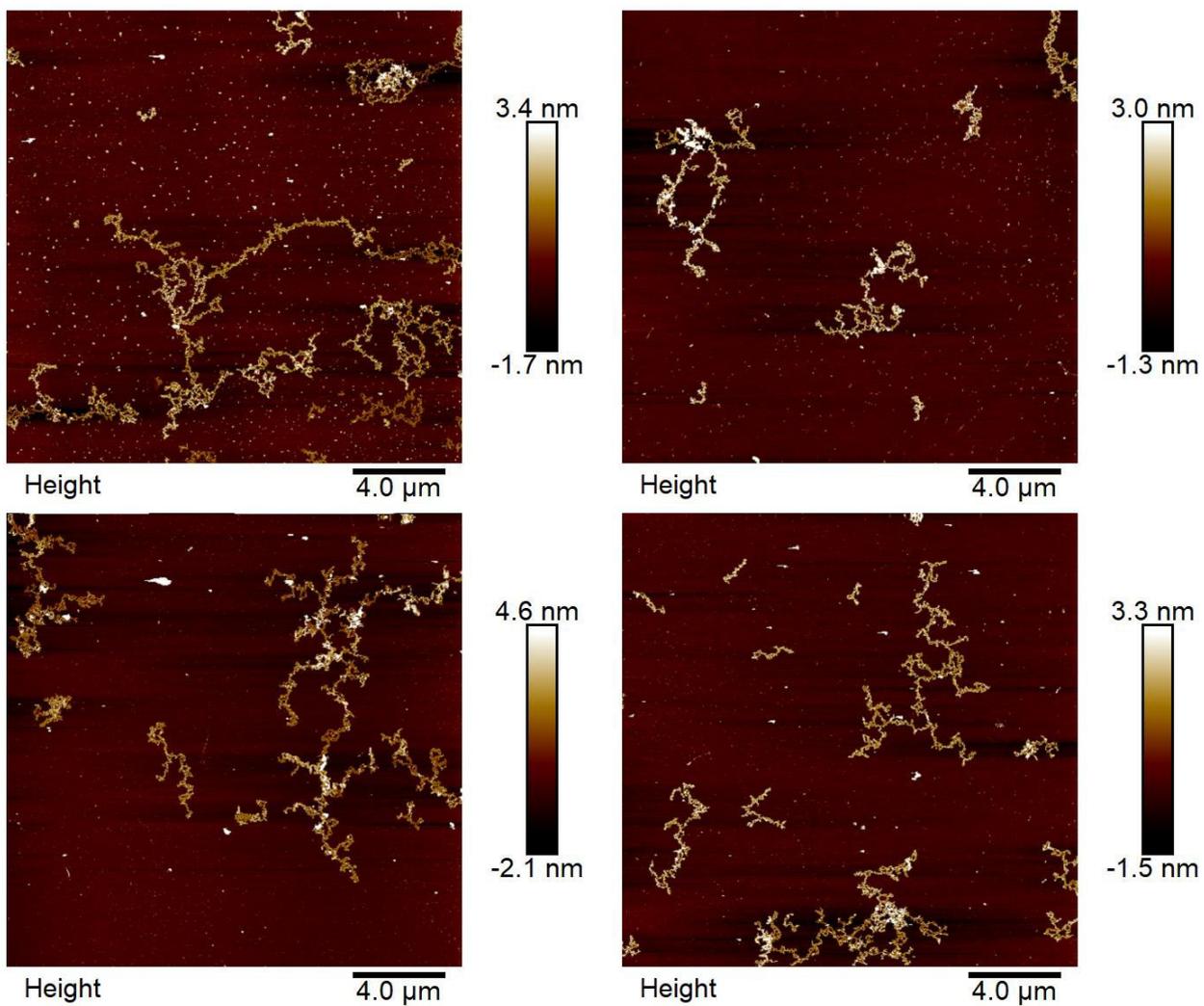


Figure S20. Supplementary AFM images of 1:1 DAN-NDI-DAN:NDI-DAN-NDI.

Transmission electron microscopy

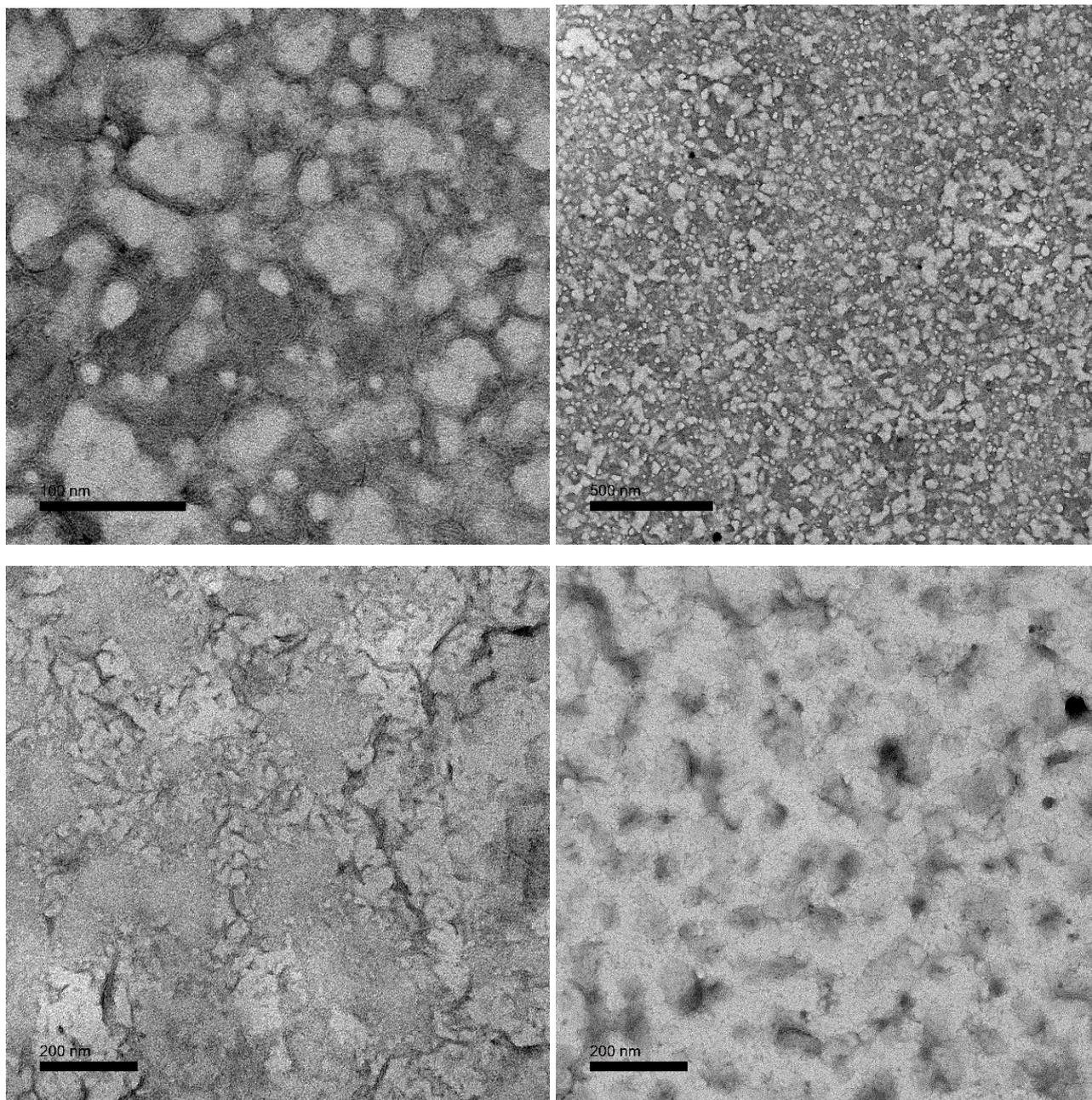


Figure S21. Supplementary TEM images of DAN₃.

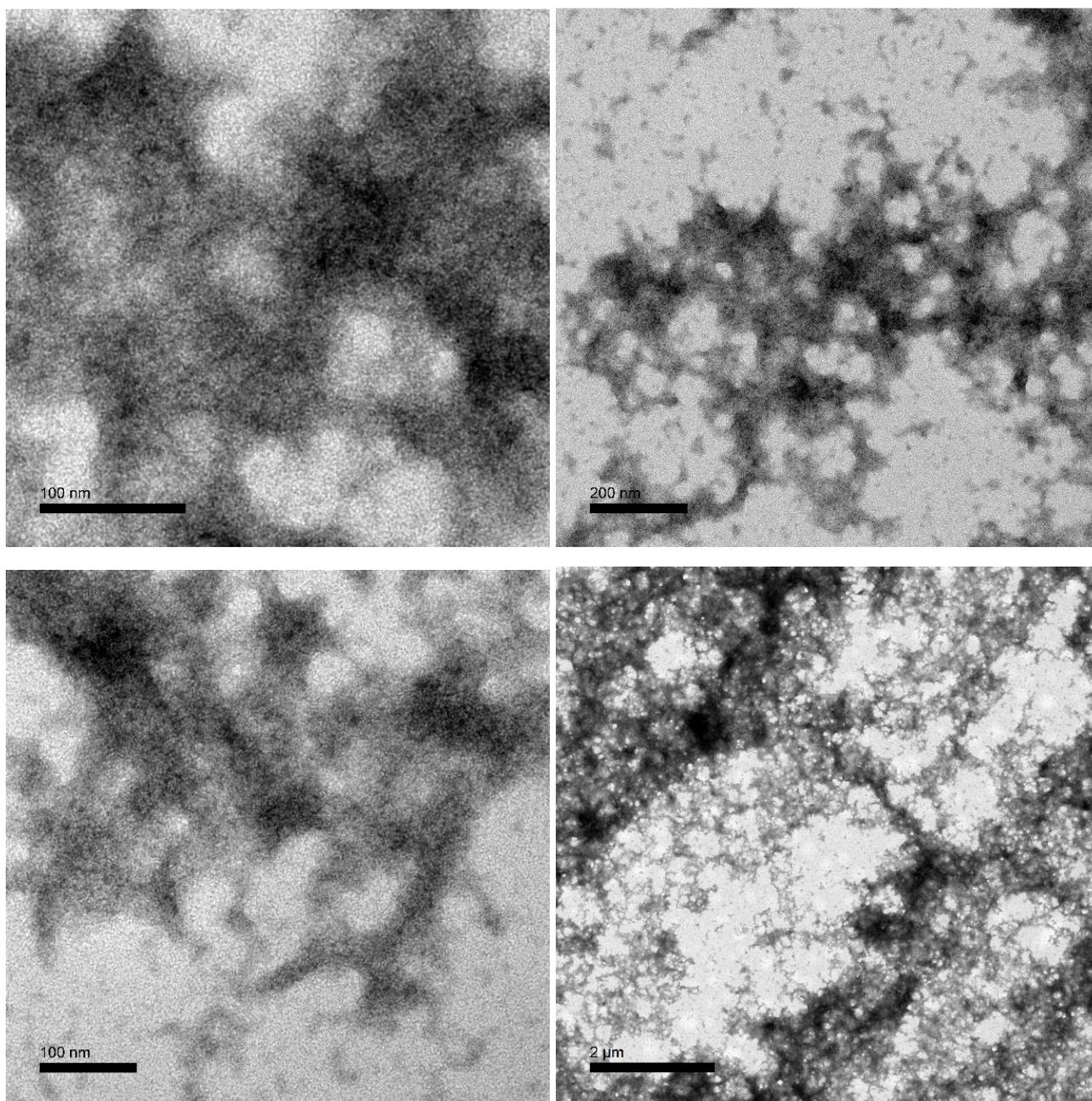


Figure S22. Supplementary TEM images of NDI₃.

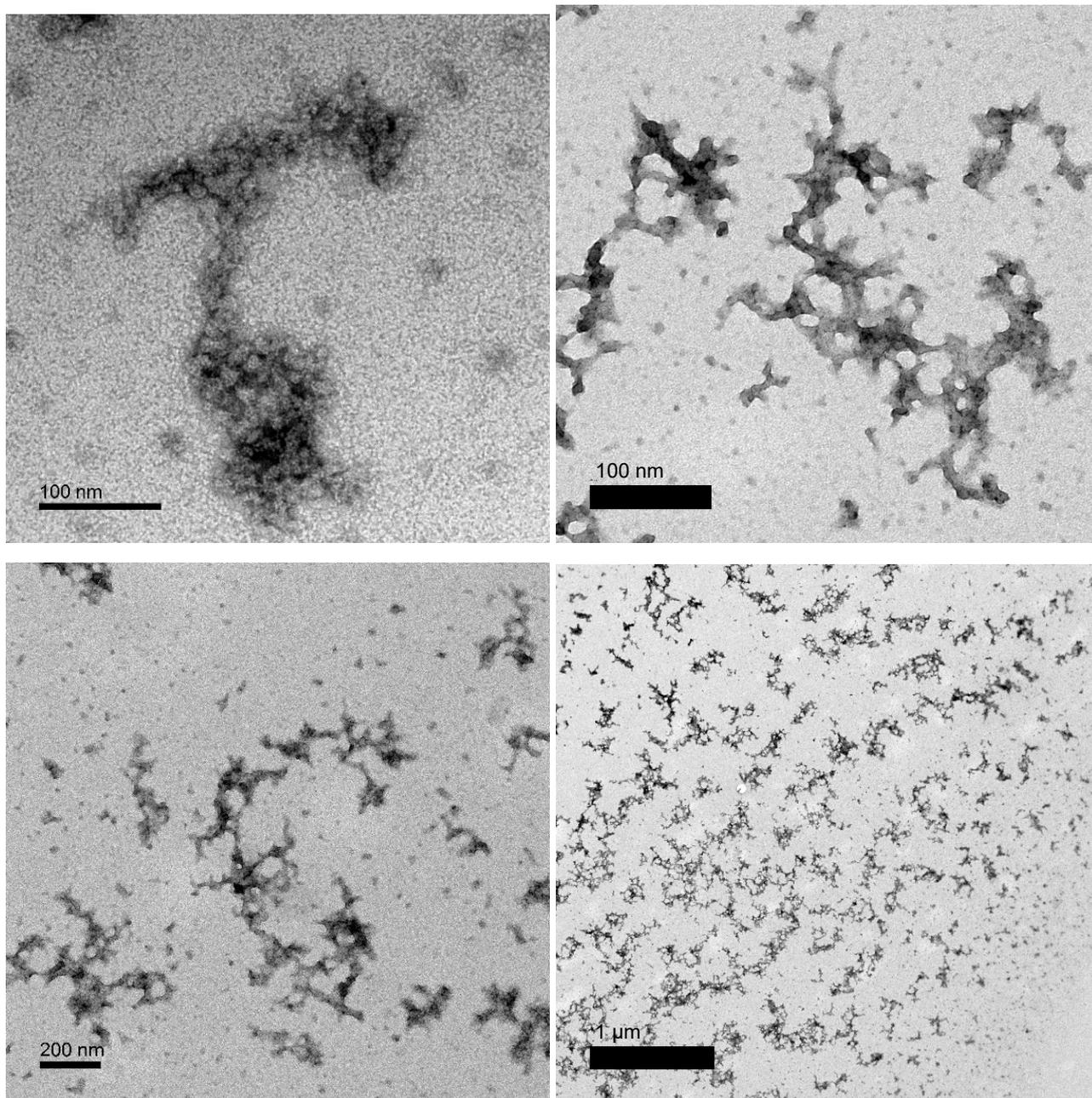


Figure S23. Supplementary TEM images of DAN-NDI-DAN.

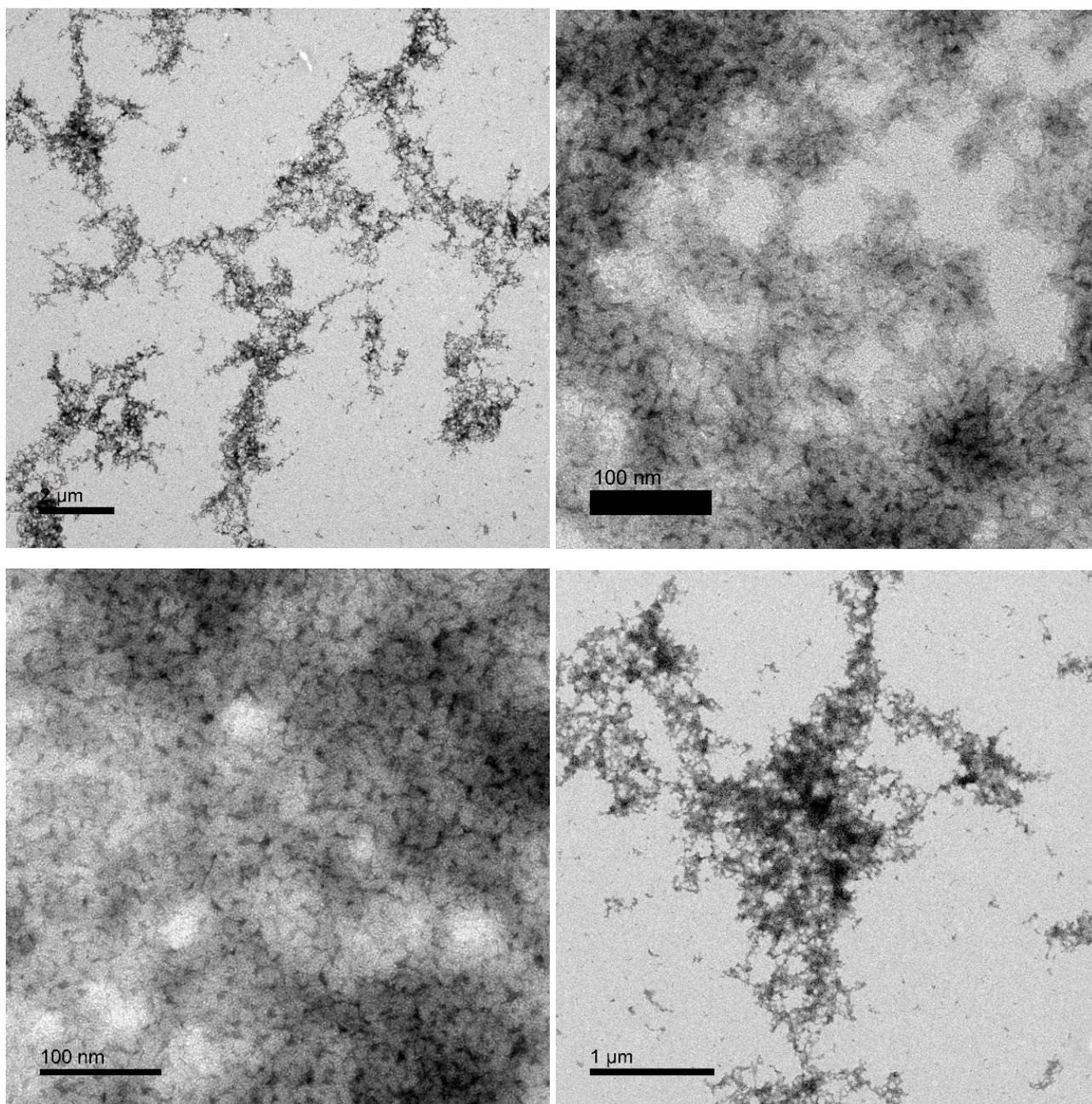


Figure S24. Supplementary TEM images of NDI-DAN-NDI.

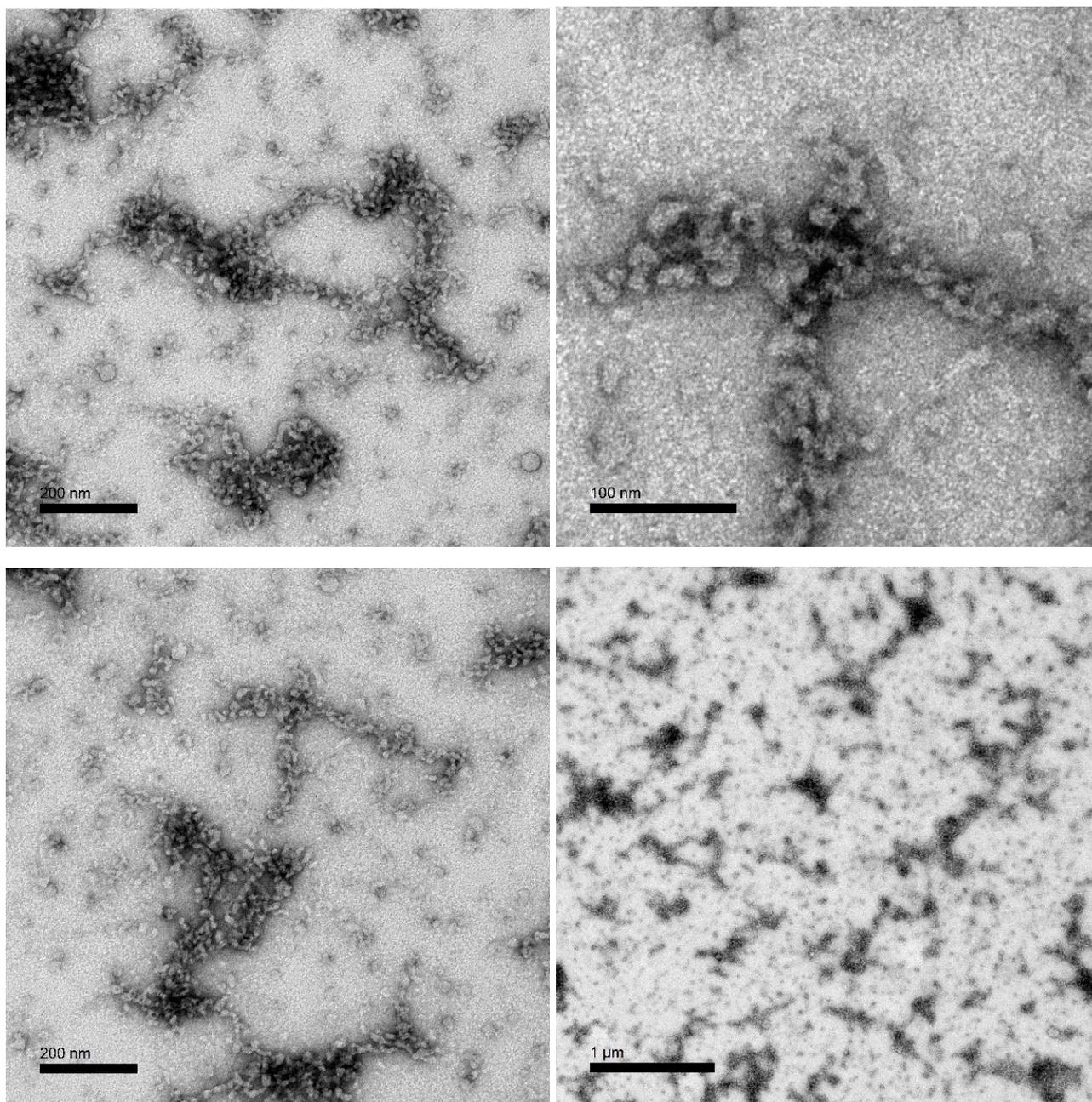


Figure S25. Supplementary TEM images of 1:1 $\text{DAN}_3\text{:NDI}_3$.

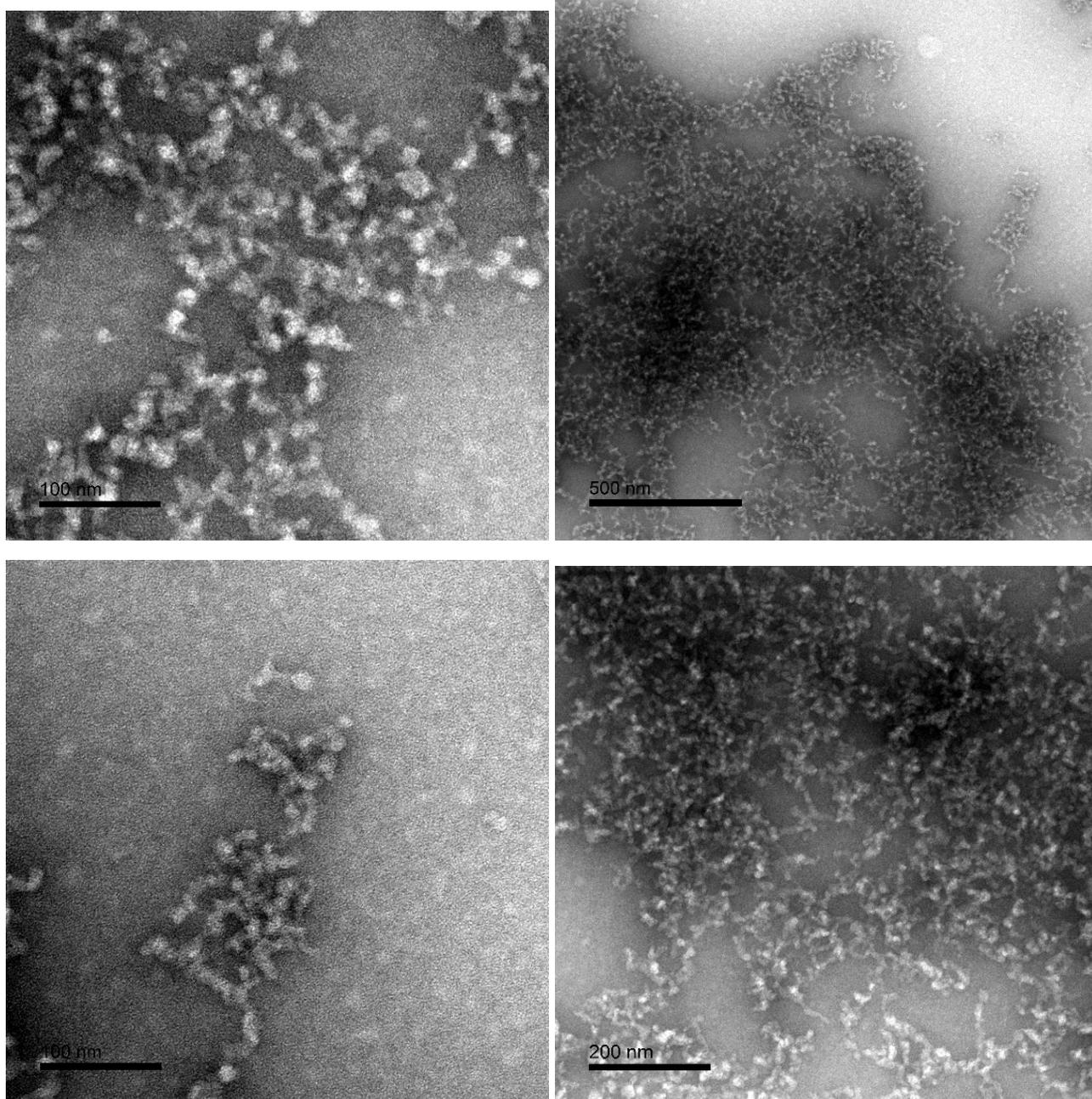


Figure S26. Supplementary TEM images of 1:1 DAN-NDI-DAN:NDI-DAN-NDI.

Molecular Dynamics model, simulations, system set-up, and analysis

Model and simulation parameters

The models for the trimer species used in the Molecular Dynamics (MD) simulations are united-atom models, and were based on the GROMOS force fields.^[1] Gromos models are modular, i.e. molecular fragments can easily be combined to create larger molecules using a list of standard bonded parameters (bond lengths, angle values, etc.) and atom types (e.g. ether O, CH₂ group, aromatic C, etc.). The 2016H66 version^[2] was used because of its superior hydration free energy of ethylene glycol chains. Because partial charges on the DAN and NDI cores are not defined in this force field, they were calculated using the quantum chemistry GAMESS-UK package^[3] by performing the Dipole Preserving Charge (DPC) analysis^[4] on the charge density calculated in vacuo using Density Functional Theory (DFT). The level of DFT employed here used the PBE correlation functional and the PBE0 exchange functional, with 25% Hartree-Fock exchange,^[5] employing the 6-31g* basis set. This approach is the same as used recently in the development of aedamer models within the Martini 3 coarse-grained model,^[6] in which the thermodynamics of the stacking interaction between DAN and NDI moieties in water was determined by potential of mean force (PMF) calculations.

MD simulations were run with the GROMACS package (version 2016.5).^[7] The equations of motion were integrated with the leap-frog algorithm, employing a time-step of 2.5 fs. Full periodic boundary conditions were used, with simulation boxes being either cubic or dodecahedral. Temperature was maintained using the Canonical Sampling Velocity-Rescaling (CSVR) thermostat algorithm^[8] (invoked by the GROMACS keyword `v-rescale`), with a coupling constant of 0.3 ps. Pressure (isotropic coupling to 1 bar) was maintained using the Berendsen coupling scheme,^[9] with coupling time-constant of 0.5 ps; the compressibility was set to $4.6 \cdot 10^{-5} \text{ bar}^{-1}$. Translation of the center of mass was removed every 10 steps. The Simple Point Charge (SPC) model^[10] was used for water, and counter ions (Na⁺) were added to neutralize the system. The SETTLE algorithm^[11] was used to treat SPC water^[10] as rigid. All bond lengths were constrained using the LINCS algorithm,^[12] with settings `lincs-order=4` and `lincs-iter=1`. The Verlet update scheme^[13] was used with a buffer tolerance of $0.005 \text{ kJ.mol}^{-1}.\text{ps}^{-1}$, updating the neighbor list every 20 steps. The cut-off is 1.4 nm for both Lennard-Jones (LJ) and coulomb interactions. The coulomb interactions were treated with the reaction field (RF) modification due to Tironi et al.^[14] This method alleviates cut-off artifacts by ensuring that both potential and force decay smoothly to zero at the cut-off of 1.4 nm. The parameter for the bulk dielectric constant (`epsilon_rf` in GROMACS) was set to 62.

Systems simulated and their set-ups

All four trimers (**DAN₃**, **NDI₃**, **DAN-NDI-DAN (DND)**, and **NDI-DAN-NDI (NDN)**) were simulated in water, either individually as a monomer (examples of structures are shown in the main text, Figure 1), or as aggregates of multiple molecules (examples are shown in the main text, Figures 3 and 4). The phosphate in the phosphodiester linkage was modeled with its formal negative charge. In all cases, the system was made neutral by adding an appropriate number of Na⁺ ions to compensate for the phosphate negative charge. The molecule or molecules were randomly placed in a simulation box (cubic or dodecahedron shaped), water was added, and counter ions placed using standard GROMACS tools (`gmx insert-molecules`, `gmx solvate`, `gmx genion`). For the larger assemblies, this was done in several steps, adding dispersed molecules in the solution containing an already formed assembly. Energy minimization (up to 1000 steepest descent steps), and short runs (100 ps) at constant volume (NVT) and constant pressure (NPT) were performed to equilibrate the systems before production. Production runs lasted up to several hundreds of nanoseconds, during which the molecules explored their conformation space, aspects of which were subsequently analyzed.

Aggregation was studied at different numbers of molecules present in the system, with a maximum of 48 molecules, which means 48 identical molecules for pure compounds or 24 of each type in mixtures (**DAN₃**-**NDI₃** and **DND-NDN**).

Analysis

Molecular and system structures as shown in main text Figures 1, 3, and 4 were visualized using Visual Molecular Dynamics (VMD) software, version 1.9.4.^[15]

The propensity of individual monomers to fold in water at 298 K (i.e., at infinite dilution) was investigated by conformational clustering. This was performed by using the GROMACS tool `gmx cluster`, using the GROMOS algorithm,^[16] based on the RMS deviation matrix of the positions of selected atoms of the aromatic cores after best fitting of the structures on each other. Briefly, the number of neighbors of all structures is determined, a neighbor being defined as a structure that has an RMSD lower than a chosen cut-off value. The structure with the largest number of neighbors is chosen as the center of a cluster and all structures within the cut-off are added to this cluster, and subsequently eliminated from the pool. The process is then repeated until all structures are depleted. Suitable cut-off values were chosen to limit the number of clusters to a maximum of around 100. For each monomer, 10,000 structures taken from a 200 ns trajectory were analyzed. The molecules are quite dynamic multiple transitions between clusters are observed. One representative structure from each cluster (the middle structure in the cluster, i.e. the one with the smallest RMSD values to all other structures in the same cluster) was saved for characterization.

Figure 1 in the main text shows the representative structure of the largest cluster found for each monomer, and indicates the percentage of structures that belong to this cluster. It should be noted that similar stacking patterns can be achieved in different ways which may however be classified as different clusters. Figure S27 shows a few more examples of top-ranked structures of monomers in water and illustrates this point.

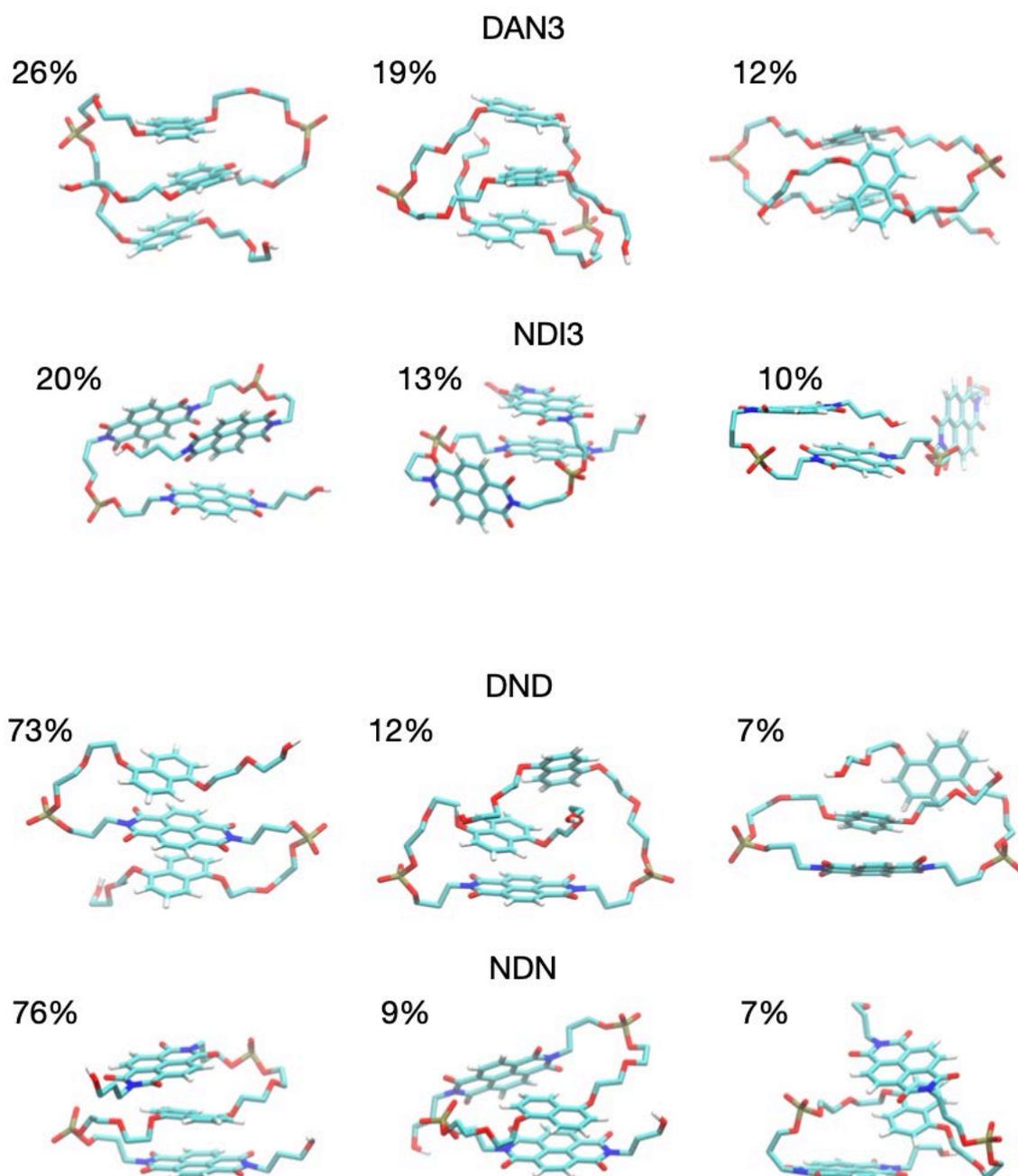


Figure S27. Licorice representation of representative structures of the top three clusters for each monomer in water. Water and ions are not shown. Colors: H: white; C: cyan; N: blue; O: red, P: gold. The top structure is shown in Figure 1 in the main text, highlighting the aromatic cores entirely in red (DAN) or blue (NDI), and with thicker bonds.

The aggregation behavior of pure compounds as well as the **DAN₃-NDI₃** and **DND-NDN** mixtures was studied for different numbers of molecules in a simulation cell and visualized using VMD to obtain insight in the nature of the aggregates. For larger numbers of molecules, in all cases, aggregation propensity is high; once bound, molecules are almost never observed to dissociate from the assembly. In systems with larger numbers of molecules, smaller assemblies form first and then form larger assemblies when they meet after diffusion. Some of the assemblies tend to be elongated structures, which led to an assembly being connected to itself through the periodic boundary condition. Such connection may stabilize the elongated structure artificially. In such cases, the simulation box was enlarged and solvent added. All reported assemblies are assuredly not interacting with their periodic images. Because of avoiding interactions across the periodic boundaries was prevented, the systems get quite large and simulation expensive. The larger assemblies of 48 molecules were arrived at by running several consecutive simulations in which monomers dispersed in the solution were added to an existing smaller assembly. The total simulation time differed per system, but was approximately 500-800 ns. Although dynamic, with rearrangement of molecules with respect to each other within the assembly, on the timescales simulated are likely too short to see fully converged structural ensembles. In the final structures shown in main text Figures 3 and 4, it can be seen that aromatic cores are stacking, but this stacking is nowhere persistent over long distances; the maximum stack size is somewhere around 6, depending on how strictly stacking is defined. In view of the still limited sampling, we did not pursue extensive analysis of the stacking statistics, but preliminary statistics of the organization of **DAN-NDI** stacked pairs in terms of their mutual orientation does not reveal clear differences that may readily explain the observed differences in the UV-VIS charge transfer band intensities. This may reflect an imperfect force field, insufficient simulation time, too small assemblies, too crude analysis (quantum mechanical calculation of the spectra based on the structure should be able to pick up the sensitivity of oscillator strength to stacking patterns). Figures S28 and S29 show side-view and top-view of snapshots of the final assemblies, containing a total of 48 molecules. The side-views are shown in Figures 3 and 4 of the main text, together with AFM and TEM images.

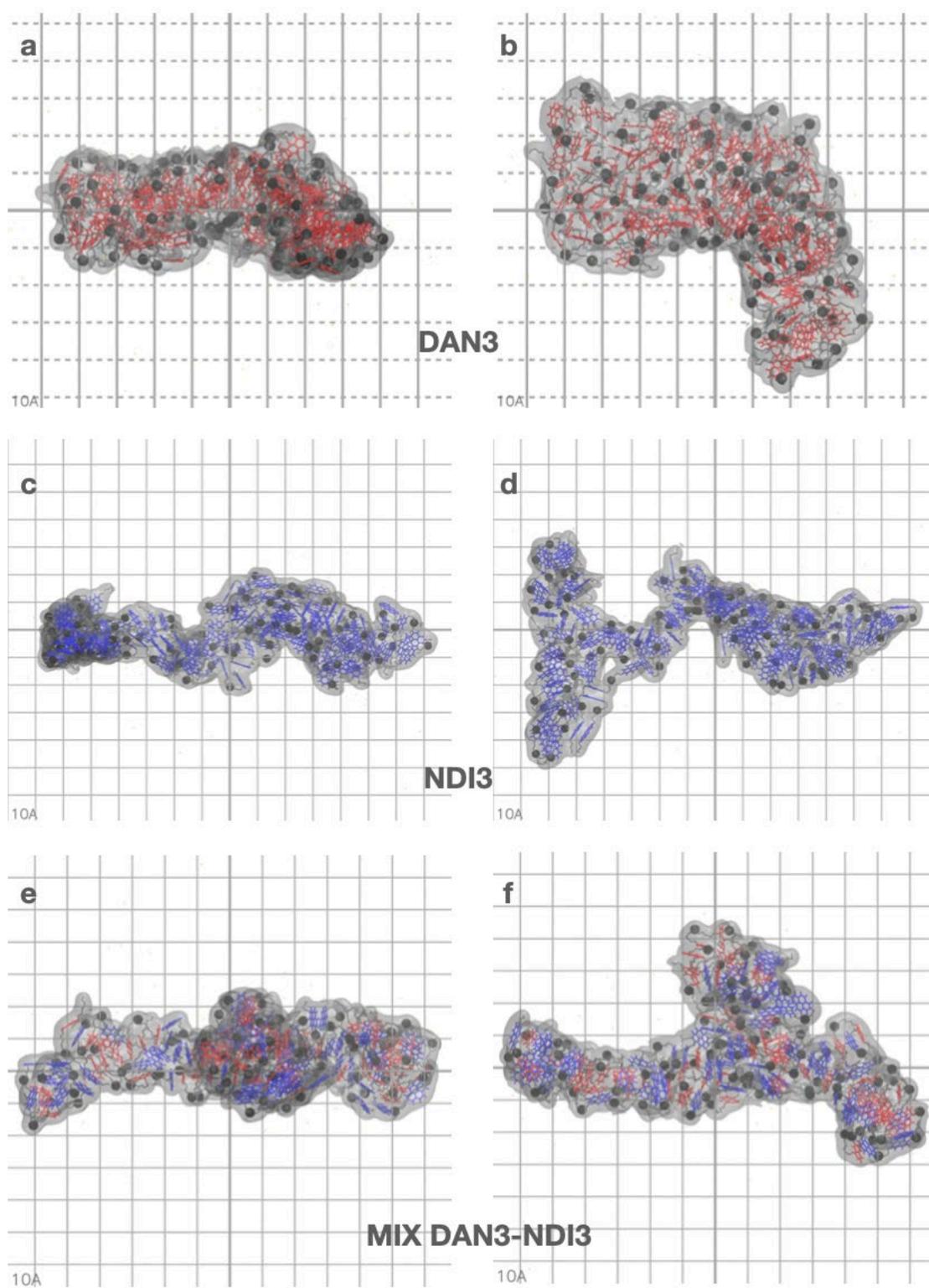


Figure S28. Side view (a, c, e) and top view (b, d, f) of the final structures of self-assembled aggregates of diester-linked aedamer molecules. (a, b) 48 molecules **DAN₃**; (c, d) 48 molecules **NDI₃**; (e, f) 24 **DAN₃** and 24 **NDI₃** molecules. The molecules are shown in a bond representation, coloring the linker atoms gray and the DAN and NDI cores red and blue, respectively. The phosphor atoms are shown as large dark spheres. The vanderWaals surface of the assembly is shown in transparent grey. Water and Na⁺ ions are not shown.

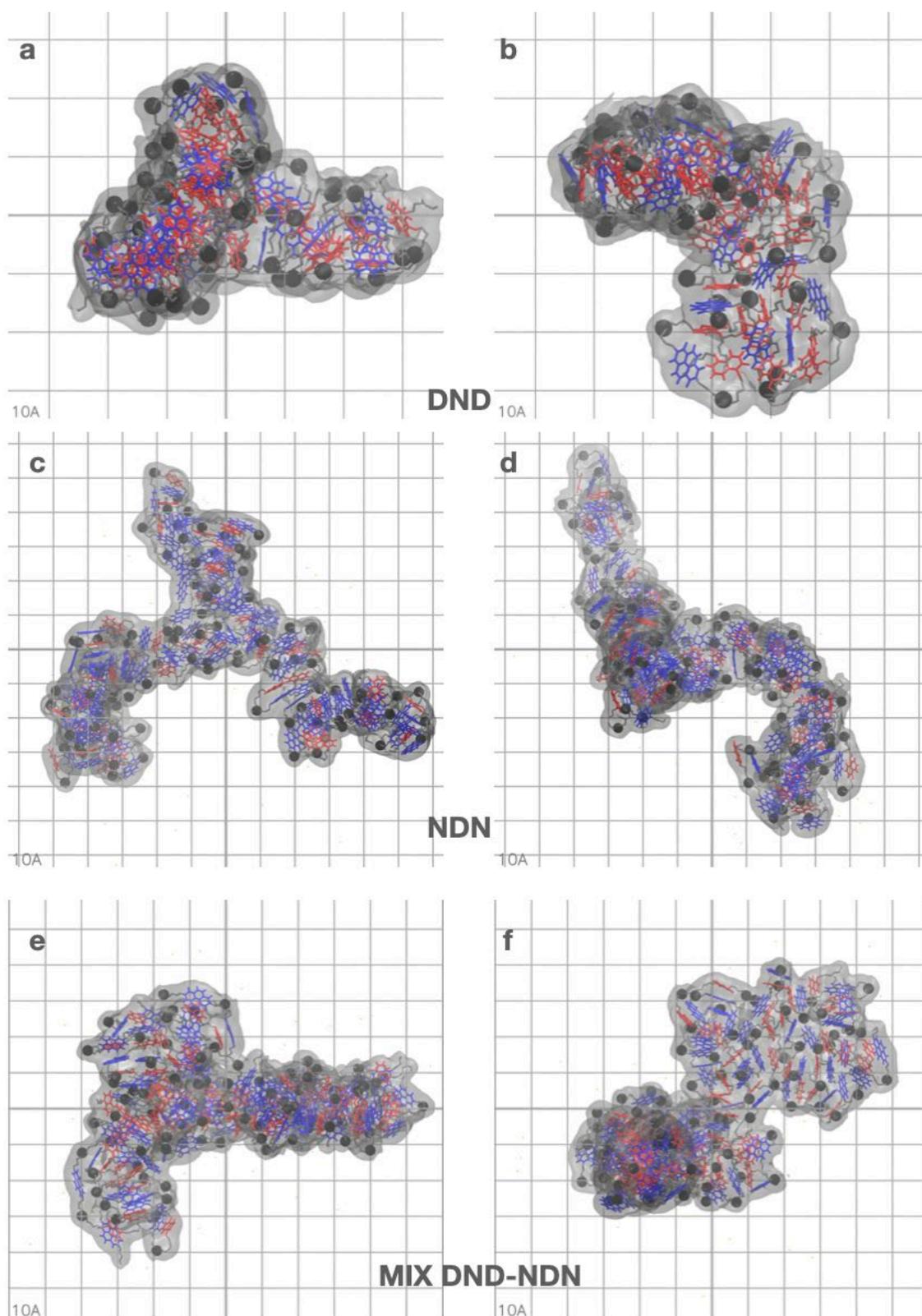


Figure S29. Side view (a, c, e) and top view (b, d, f) of the final structures of self-assembled aggregates of phosphodiester-linked aromatic donor-acceptor molecules. (a, b) 48 molecules **DND**; (c, d) 48 molecules **NDN**; (e, f) 24 **DND** and 24 **NDN** molecules. The molecules are shown in a bond representation, coloring the linker atoms gray and the DAN and NDI cores red and blue, respectively. The phosphor atoms are shown as large dark spheres. The van der Waals surface of the assembly is shown in transparent grey. Water and Na⁺ ions are not shown.

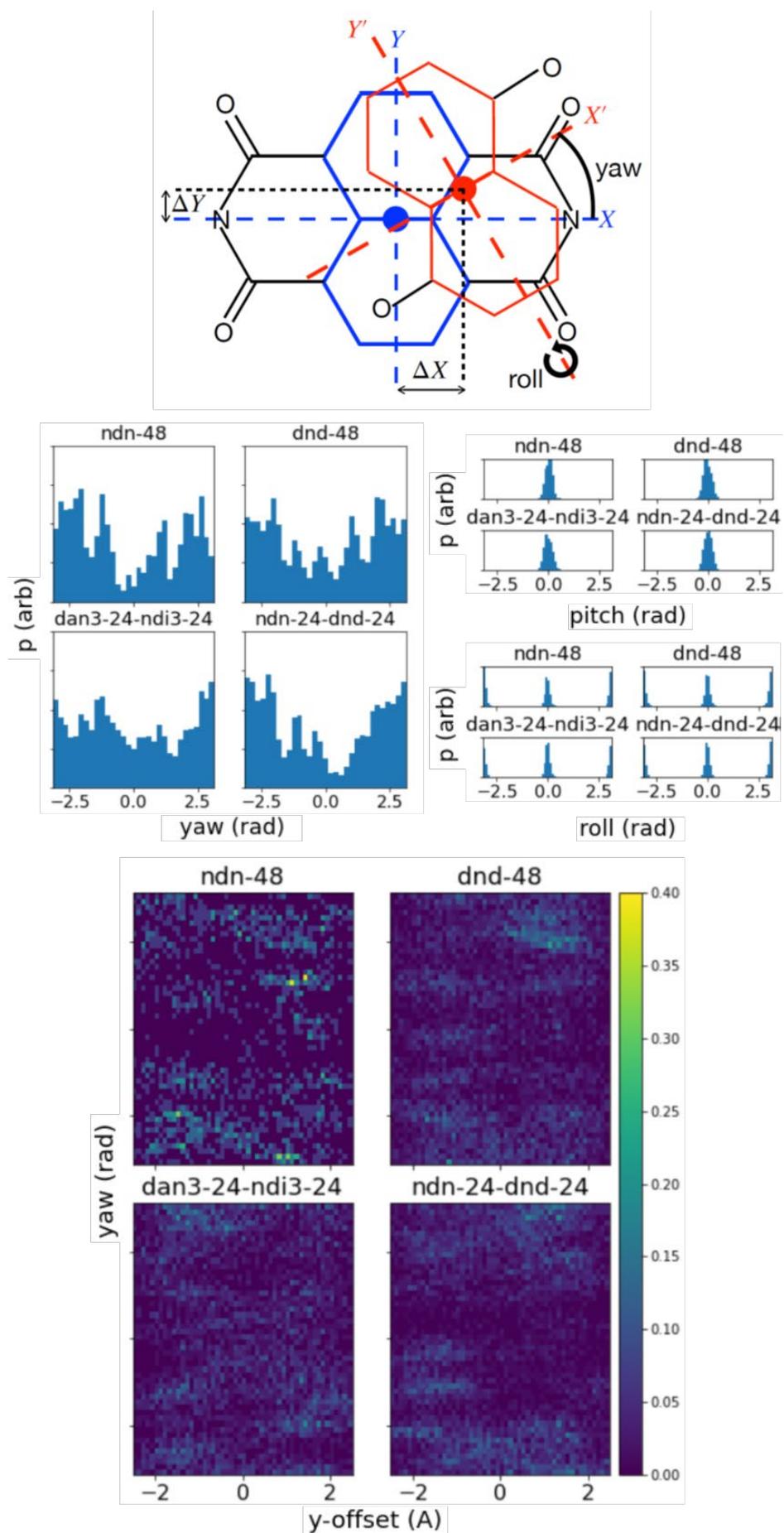


Figure S30. Analysis of geometry of aromatic interactions in self-assembled aggregates.

References

- [1] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- [2] B. A. C. Horta, P. T. Merz, P. F. J. Fuchs, J. Dolenc, S. Riniker, P. H. Hünenberger, *J. Chem. Theory Comput.* **2016**, *12*, 3825–3850.
- [3] M. F. Guest, I. J. Bush, H. J. J. Van Dam, P. Sherwood, J. M. H. Thomas, J. H. Van Lenthe, R. W. A. Havenith, J. Kendrick, *Mol. Phys.* **2005**, *103*, 719–747.
- [4] B. T. Thole, P. Th. van Duijnen, *Theor. Chim. Acta* **1983**, *63*, 209–221.
- [5] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [6] P. C. T. Souza, R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B. M. H. Bruininks, T. A. Wassenaar, P. C. Kroon, J. Melcr, V. Nieto, V. Corradi, H. M. Khan, J. Domański, M. Javanainen, H. Martinez-Seara, N. Reuter, R. B. Best, I. Vattulainen, L. Monticelli, X. Periole, D. P. Tieleman, A. H. de Vries, S. J. Marrink, *Nat. Methods* **2021**, *18*, 382–388.
- [7] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, *SoftwareX* **2015**, *1–2*, 19–25.
- [8] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
- [9] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [10] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, in *Intermolecular Forces B Pullman Ed*, Reidel, Dordrecht, **1981**, p. 331.
- [11] S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 952–962.
- [12] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- [13] S. Páll, B. Hess, *Comput. Phys. Commun.* **2013**, *184*, 2641–2650.
- [14] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren, *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- [15] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.
- [16] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240.