Supporting information

pH-tunable membrane-active polymers, NCMNP2a-x, and their potential membrane protein applications

Thi Kim Hoang Trinh,[a,b,1] Andres Jorge Cabezas,[c,d,1] Soumil Joshi[e,1], Claudio Catalano,[a,b,] Abu Bakkar Siddique,[a,b] Weihua Qiu,[a,b] Sanket Deshmukh,[e] Amedee des Georges,[c,d,f] Youzhong Guo[a,b,*]

[a]Department of Medicinal Chemistry, School of Pharmacy, Virginia Commonwealth University, Richmond, VA 23298, USA.

[b]Institute for Structural Biology, Drug Discovery and Development, School of Pharmacy, Virginia Commonwealth University, Richmond, VA 23219, USA.

[c]Structural Biology Initiative, CUNY Advanced Science Research Center, City University of New York, New York, New York, 10017, USA.

[d]Ph.D. Program in Biochemistry, The Graduate Center of the City University of New York, New York, New York, 10017, USA.

[e]Department of Chemical Engineering, Virginia Tech, Blacksburg, VA2 4060, USA.

[f]Department of Chemistry & Biochemistry, City College of New York, New York, New York, 10017, USA.

[1]Co-first Authors

*Email: yguo4@vcu.edu

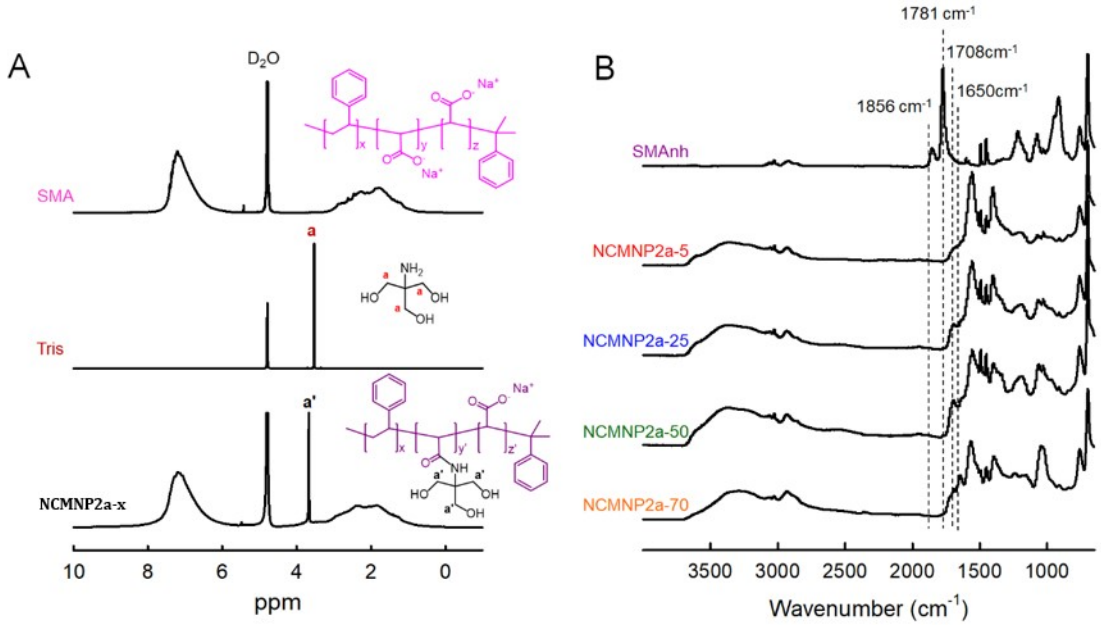
†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

Contents

I. Additional figures.....3

II. Additional tables.....14

I. Additional figures



C

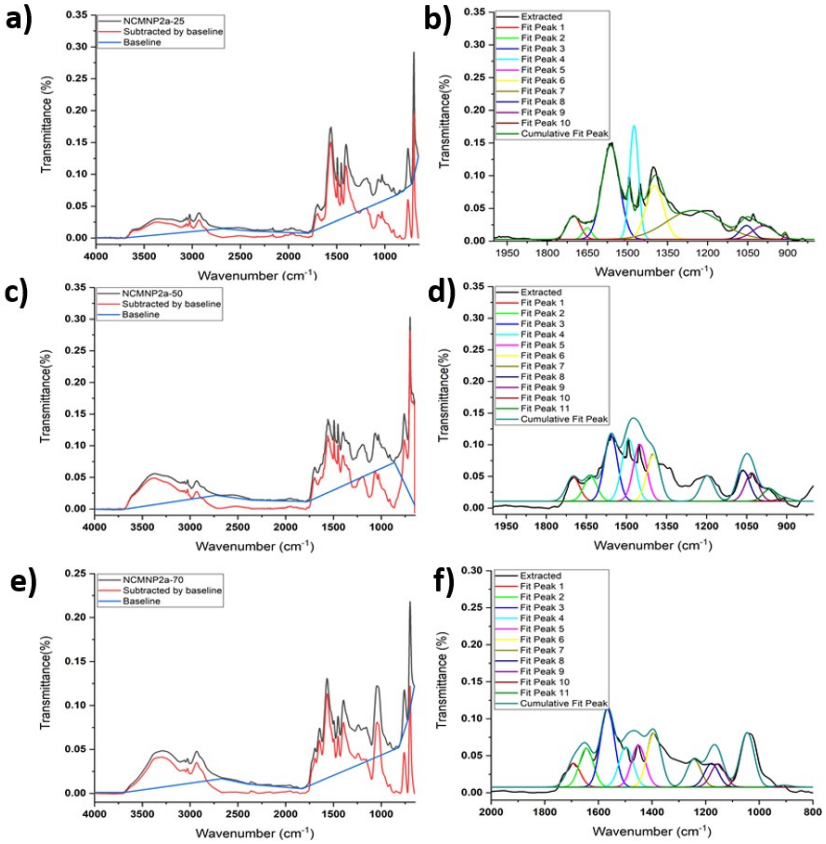


Fig. S1 Structural characterizations of NCMNP2a-x polymers. (A) ^{1}H-NMR spectra of non-modified SMA polymer, Tris and NCMNP2a-x representative in D_2O, 300 MHz. The SMA spectrum has no singlet signal of methylene protons from Tris (δ = 3.53 ppm). In contrast, it appears and shifts to 3.67 ppm in NCMNP2a-x, suggesting that the Tris grafts on individual MAnh units instead of free-form presence. (B) FT-IR spectra of SMAnh and its derivatives altered with different amounts of Tris, NCMNP2a-x. Observation of the broad peaks in 3000 – 3600 cm^{-1} region of NCMNP2a-x spectra indicates the target polymers obtained; this attributed to a combination of N-H stretching frequency from amide and O-H from the carboxylic acid and alcoholic group pendants. Moreover, the shift of peaks at 1856 and 1781 cm^{-1} (C=O anhydride stretching) to lower wavenumbers 1708 cm^{-1} (C=O acid stretching) and 1650 cm^{-1} (C=O amide stretching) illustrates for non-residue of unreacted MAnh and formation of amides on the polymer backbone. (C) Deconvolution of the FTIR data for NCMNP-2a-x (where x= 25, 50 and 70). The original curve, baseline and its substraction is represented by C(a) for NCMNP2a-25, C(c) for NCMNP2a-50 and C(e) for NCMNP2a-50 where the corresponding decolvolution were represented by C(b) for NCMNP2a-25, C(d) for NCMNP2a-50 and C(f) for NCMNP2a-50.

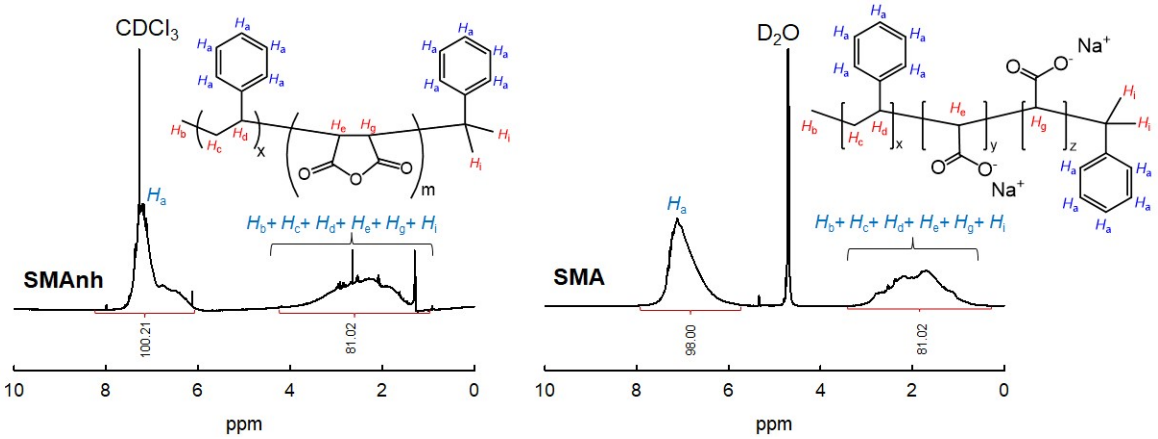


Fig. S2 [1]H-NMR spectra of commercial SMAnh and its hydrolysis SMA in $CDCl_3$ and D_2O, respectively.

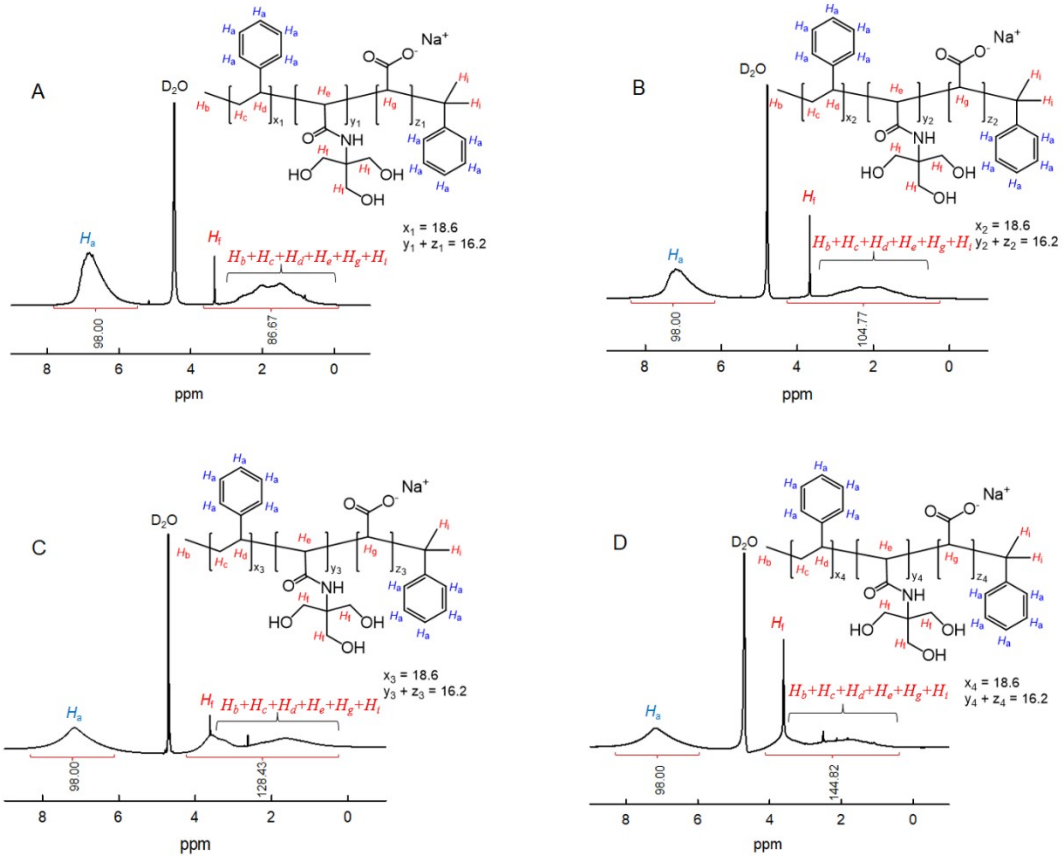


Fig. S3 [1]H-NMR spectra of NCMNP2a-x copolymers with full integration: (A) NCMNP2a-5, (B) NCMNP2a-25, (C) NCMNP2a-50 and (D) NCMNP2a-70.

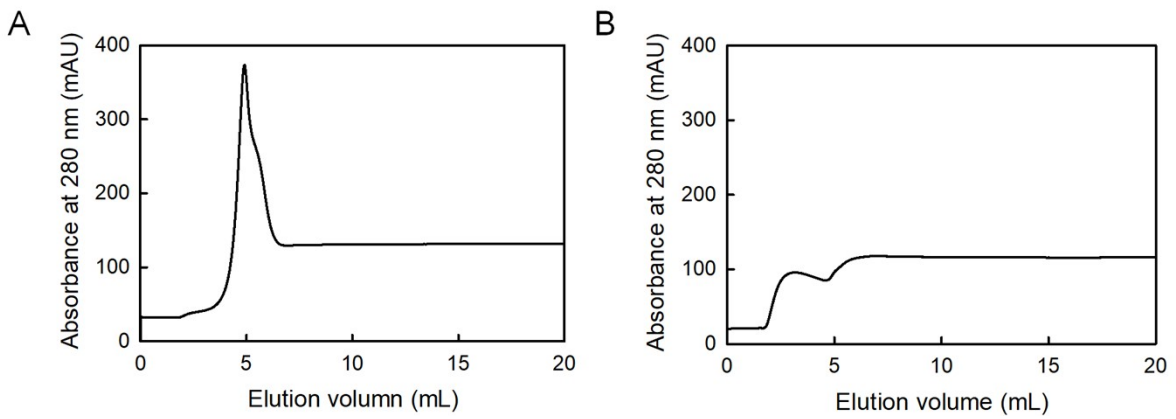


Fig. S4 Effect of Tris incorporation percentages on the protein purification efficiency from the native membranes. Purification profiles of AcrB extracted with (A) non-modified SMA polymer, SMA2000 (2.5 % w/v), and (B) NCMNP2a-70 (5 % w/v).

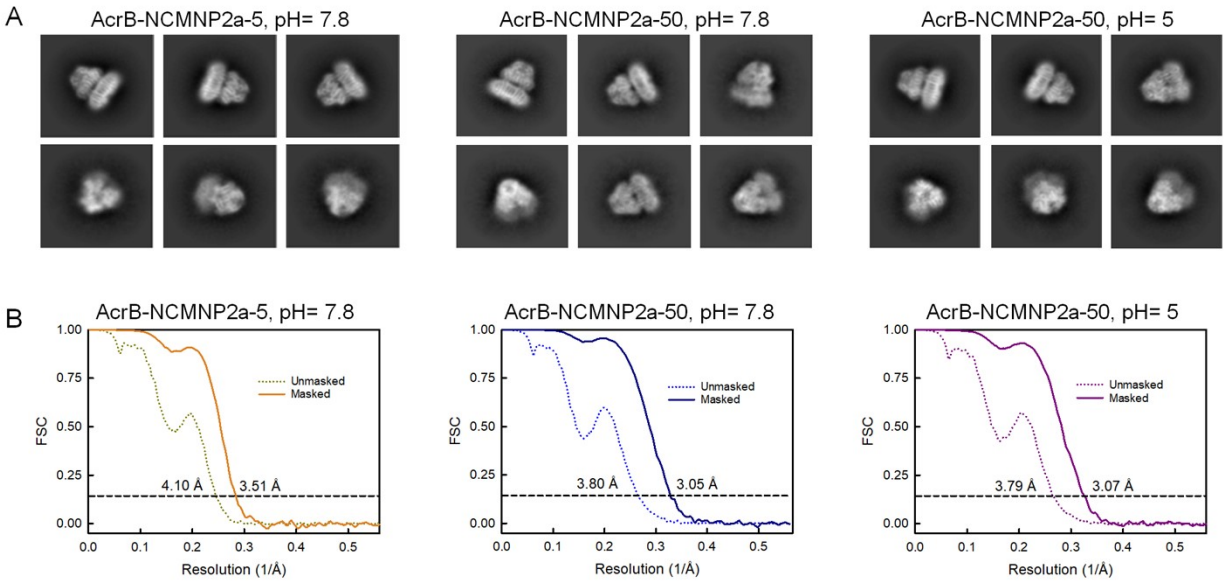


Fig. S5 Cryo-EM analysis of purified AcrB-NCMN particles. (A) Representative 2D classes of AcrB particles extracted from various polymers at different pH conditions. (B) The Fourier shell correlation (FSC) resolution curves for 3D reconstruction using the gold standard with FSC = 0.143.

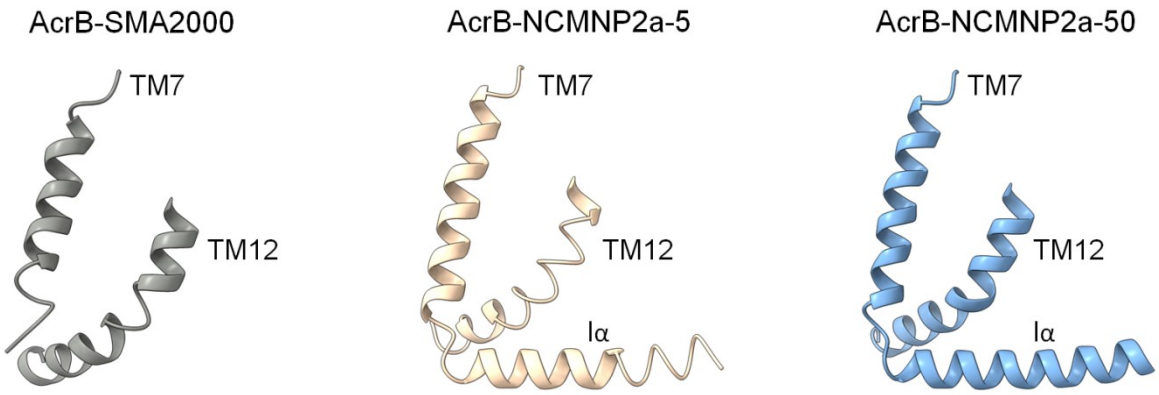


Fig. S6 Comparison of TM7, TM12, and Iα helix in protomer O of AcrB extracted by SMA2000 (6BAJ)[1] and our NCMN polymers (from current work) at same pH conditions, pH = 7.8.

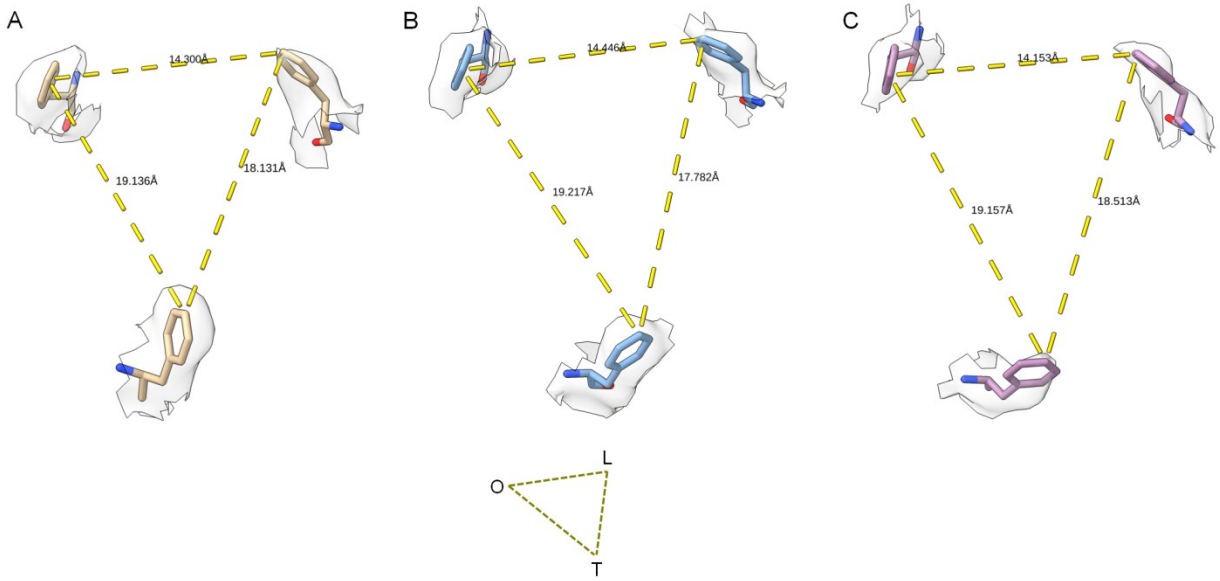


Fig. S7 Distances between F386 residues of wide-type AcrB in this work. Owing to the lipids preserved, the central cavity of AcrB-NCMN particles remains unchanged with distances of F386-

F386 residues similar to AcrB structure in SMA2000 (6BAJ). Whereas, in a detergent system where all the supporting lipids were removed, such distances were reduced to around 6.7 Å.[3]

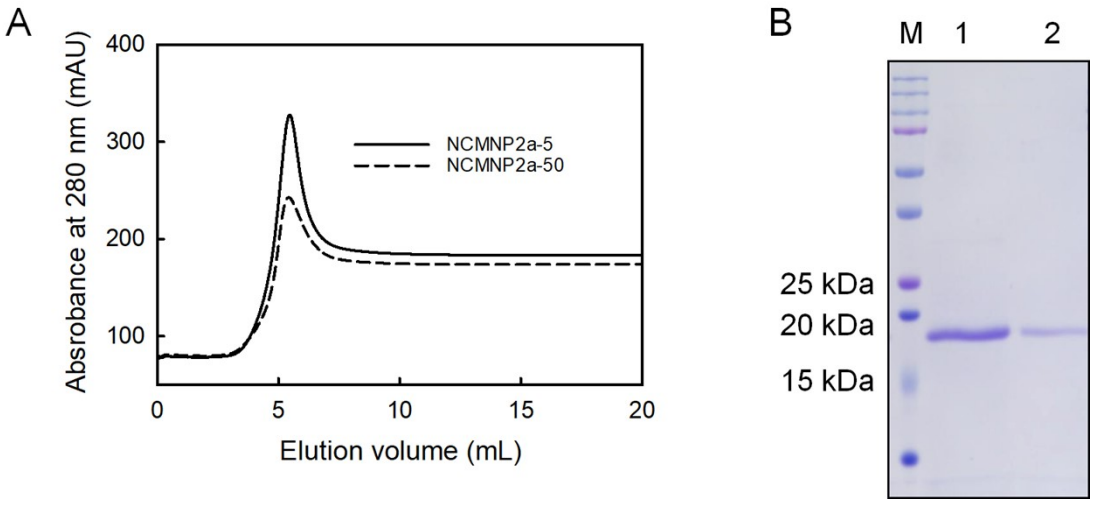


Fig. S8 Effect of Tris incorporation percentages on the protein purification efficiency from the native membranes. (A) Comparative purification profiles of *Bc*TSPO solubilized by NCMNP2a-5 and NCMNP2a-50. (B) Visualization of purified *Bc*TSPO particles on Coomassie-stained SDS-PAGE compared to protein ladder (M). Lanes 1 and 2 correspond to *Bc*TSPO particles extracted by NCMNP2a-5 and NCMNP2a-50, respectively.

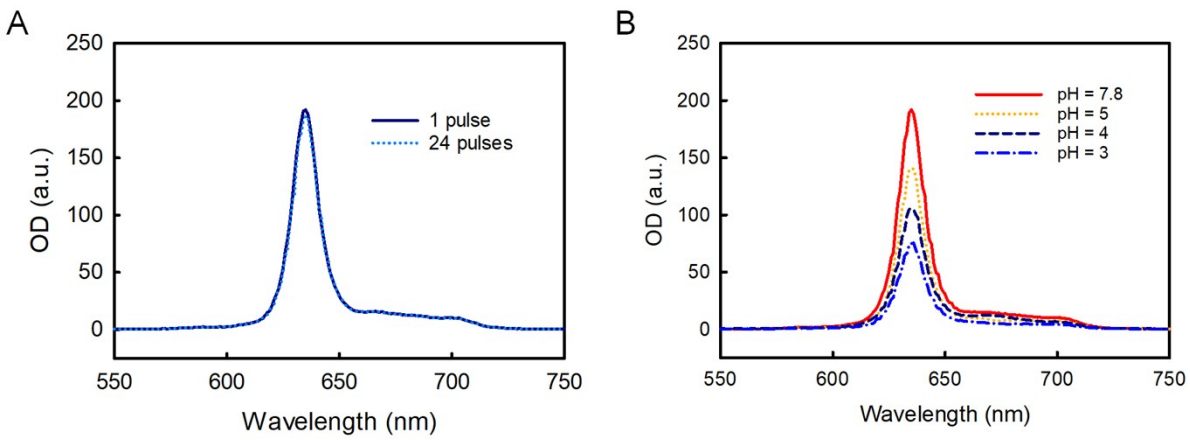


Fig. S9 Fluorescence emission spectra of PpIX upon excitation at 410 nm. (A) Spectra of PpIX only at different laser pulses. Quenching of fluorescence was undetectable, suggesting that the PpIX was stable under such experimental conditions. (B) Spectra of PpIX/ *Bc*TSPO-NCMNP2a-50 solutions at different pH upon excitation at 410 nm. At pH 3 – pH 5, the fluorescence intensity of PpIX solution displays a less intense signal compared to pH = 7.8, indicating higher-order aggregates of PpIX.[2]

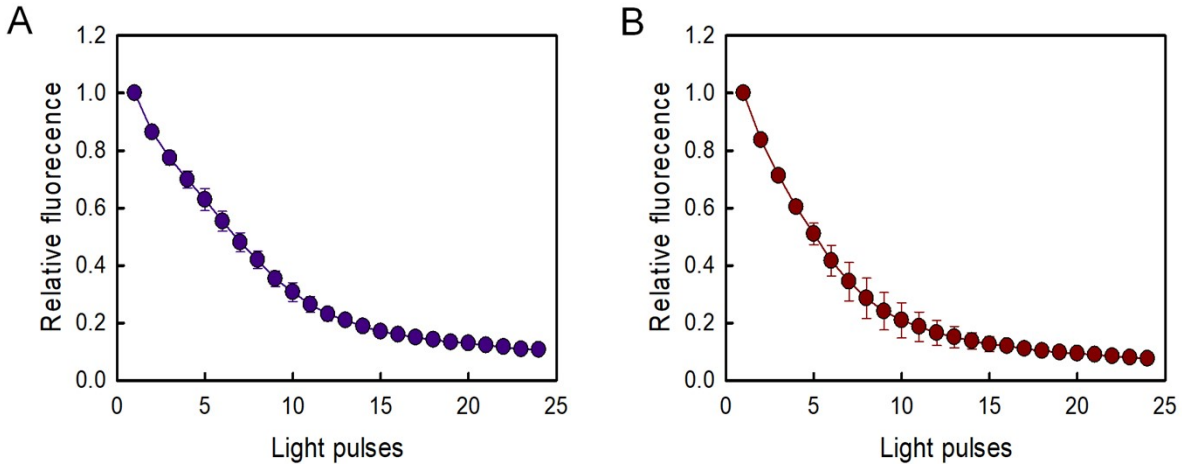


Fig. S10 Photodegradation of PpIX in the presence of *Bc*TSPO upon excitation at 410 nm. Decays of PpIX at 632 nm as a function of the light pulse catalyzed by *Bc*TSPO extracted by NCMNP2a-5 (A) and NCMNP2a-50 (B) at pH = 7.8.

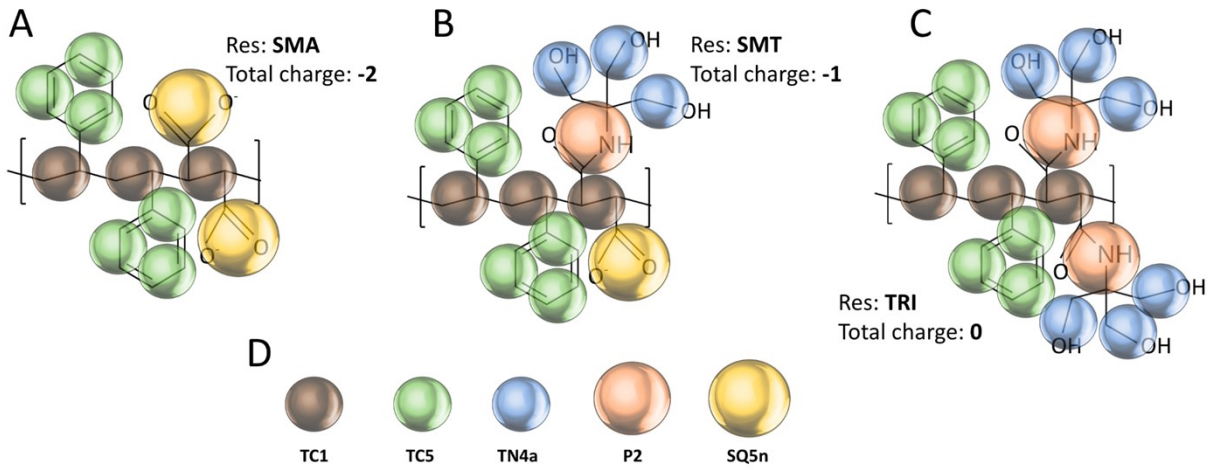


Fig. 11 Mapping schemes used to model the three residues - A) **SMA**, B) **SMT**, and C) **TRI**, combinations of which were used to generate different Tris-substituted polymer models. D) MARTINI 3.0 beads used in the simulations to model polymers.

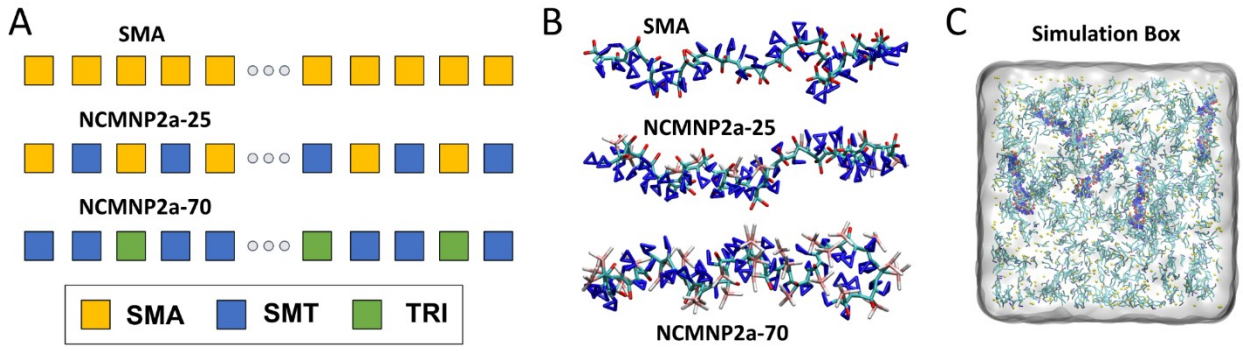


Fig. 12 A) Schematic showing the generation of polymer chains using the three modeled residues, B) Snapshots of relaxed polymer chains, C) Representative snapshot of simulation box for self-assembly simulations.

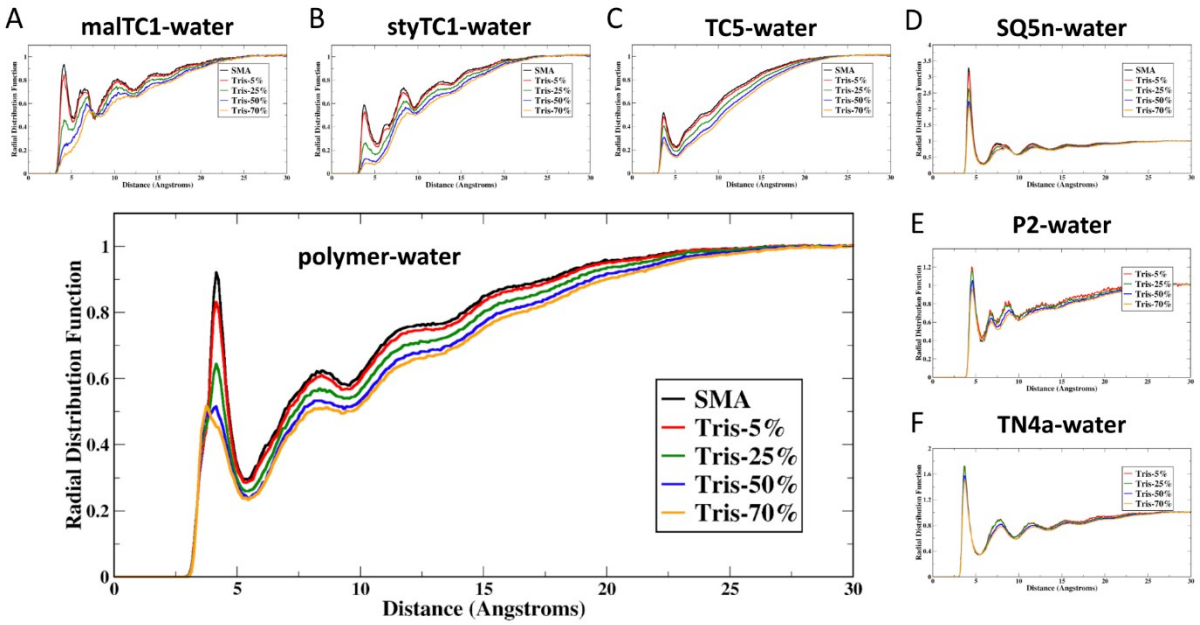


Fig. S13 Averaged RDF between water and polymer components in charged simulations including A) TC1 (backbone) beads attached to maleic acid, B) TC1 beads attached to styrene, C) TC5 (styrene) beads, D) SQ5n (charged COO^-) beads, E) P2 (amide group from Tris) beads, and F) TN4a (hydroxyl group from Tris) beads. Also shown is the RDF between water and the entire polymer chains to identify the contribution of these individual groups on the overall RDF. The legends indicate different polymers based on their percentage of Tris substitution for easy understanding: Tris-5 % corresponds to NCMNP2a-5, Tris-25 % corresponds to NCMNP2a-25, and so on.

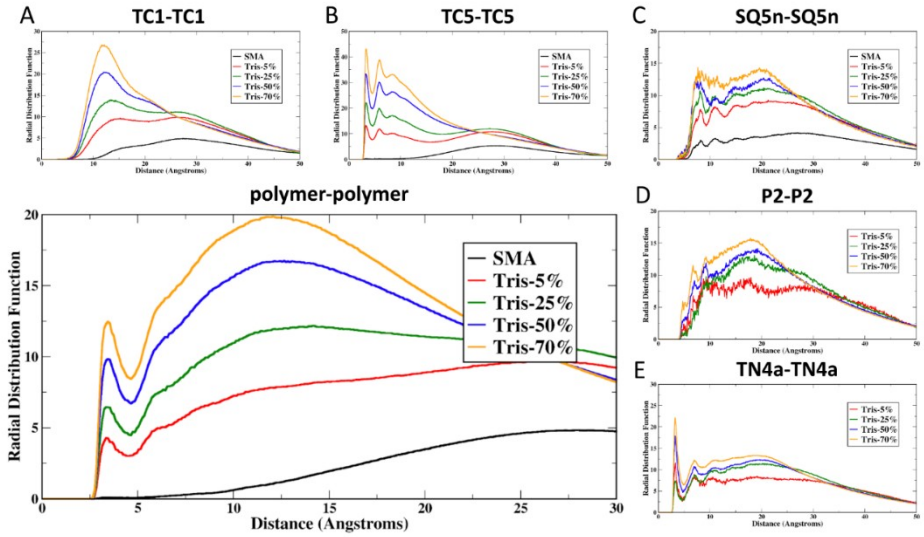


Fig. S14 Averaged RDF between polymer components from multiple chains including A) TC1 (backbone) beads, B) TC5 (styrene) beads, C) SQ5n (charged COO^-) beads, D) P2 (amide group from Tris) beads, and E) TN4a (hydroxyl group from Tris) beads. Also shown is the RDF between the whole polymer chains to identify the contribution of these individual groups on the overall RDF. The legends indicate different polymers based on their percentage of Tris substitution for easy understanding: Tris-5 % corresponds to NCMNP2a-5, Tris-25 % corresponds to NCMNP2a-25, and so on.

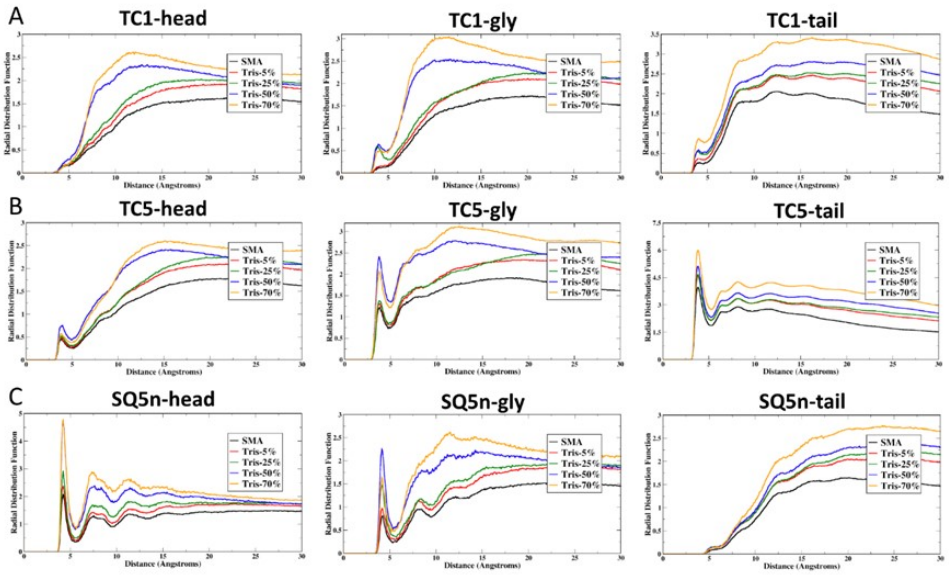


Fig. S15 Averaged RDF between polymer components including A) TC1 (backbone) beads, B) TC5 (styrene) beads, and C) SQ5n (charged COO^-) beads, with lipid head (column 1), glycerol (column 2), and tail (column 3) groups. The legends indicate different polymers based on their percentage of Tris substitution for easy understanding: Tris-5 % corresponds to NCMNP2a-5, Tris-25 % corresponds to NCMNP2a-25, and so on.

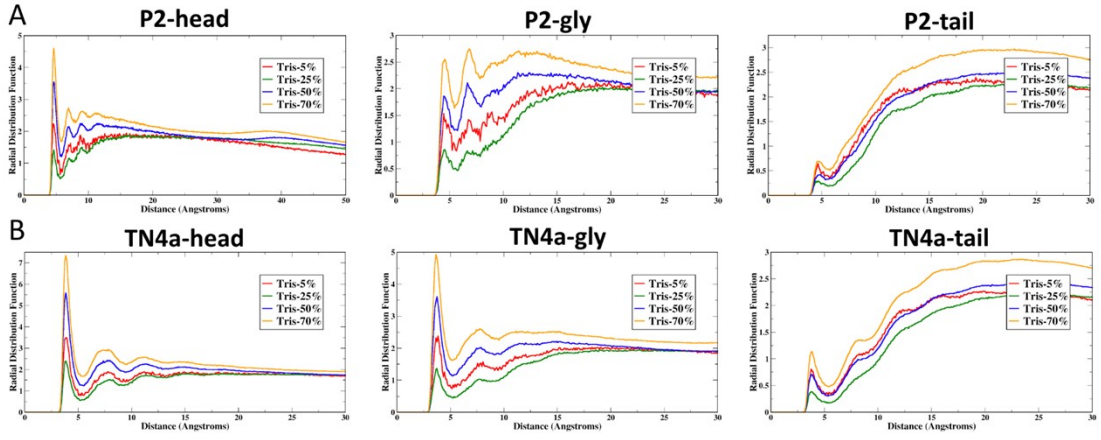


Fig. S16 Averaged RDF between polymer components including A) P2 (amide group from Tris) beads, and B) TN4a (hydroxyl group from Tris) beads, with lipid head (column 1), glycerol

(column 2) and tail (column 3) groups. The legends indicate different polymers based on their percentage of Tris substitution for easy understanding: Tris-5 % corresponds to NCMNP2a-5, Tris-25 % corresponds to NCMNP2a-25, and so on.

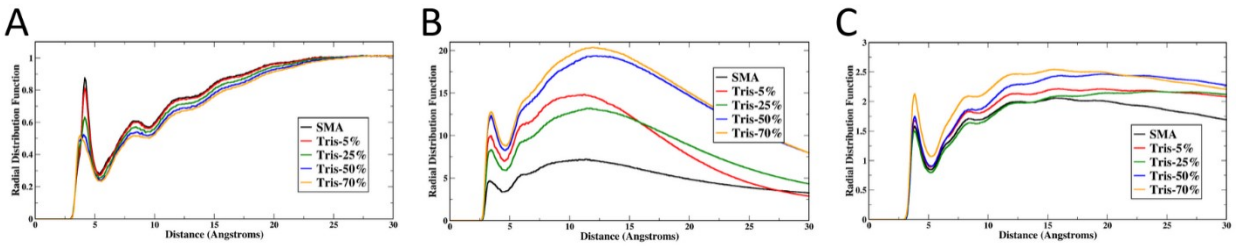


Fig. S17 Averaged RDF calculated using uncharged simulation trajectories between A) polymer-water, B) polymer-polymer, and C) polymer-lipids. The legends indicate different polymers based on their percentage of Tris substitution for easy understanding: Tris-5 % corresponds to NCMNP2a-5, Tris-25 % corresponds to NCMNP2a-25, and so on.

II. Additional tables

Table S1. Calculation of sulfonating degree of NCMNP2a-x and summary for QSAR analysis of NCMNP2a-x polymers toward their stability and performance in AcrB solubilization.

Polymer	Amine/ MAnh feeding ratio[a]	$\sum Aiphatic\ protons$	y_n	z_n	Tris-grafting degree (%)[b]	Optimum pH[c]	Tolerance to [Ca^{2+}] (mM)[c]	NCMN particle morphology[d]
NCMNP2a-5	0.05	86.67	0.95	15.25	5.9	> 4	< 1.5	Single particle
NCMNP2a-25	0.25	104.77	3.96	12.24	24.4	> 3	< 1.5	Single particle
NCMNP2a-50	0.5	128.43	7.90	8.30	48.8	> 2	< 2	Single particle
NCMNP2a-70	0.7	144.82	10.64	5.56	65.7	> 1	< 5	Large patch

[a] Molar ratio. [b] Experimental grafting percentage of amine calculated based on ^{1}H-NMR characterization. [c] Stability of NCMNP2a-x polymers in various buffer environments determined based on turbidity experiments. [d] Morphology of AcrB-NCMN particles visualized using negative-TEM analysis.

Table S2. Cryo-EM data collection and processing

	AcrB-NCMNP2a-5, pH= 7.8	**AcrB-NCMNP2a-50, pH= 7.8**	**AcrB-NCMNP2a-50, pH= 5**
Microscope	Titan (FEI)	Titan (FEI)	Titan (FEI)
Voltage (kV)	300	300	300
Detector	Gatan K3 (6k × 4k)	Gatan K3 (6k × 4k)	Gatan K3 (6k × 4k)
Nominal magnification	37,000 ×	37,000 ×	37,000 ×
Electron exposure (e^{-} Å^{-2})	0.84	0.94	0.95
Defocus range (μm)	0.7 – 2.0	0.7 – 2.0	0.7 – 2.0
Pixel size (Å^2 per pixel)	0.85	0.85	0.85
Dose rate (e^{-}/s/pixel)	12.65	8.02	14.28
Exposure time (s)	3.995	3.995	3.995
Movies stacks (no.)	3375	3031	3700
Boxsize (pixels)	360	360	360
Final particle images (no.)	72,707	100,418	73,352
Symmetry imposed	C1	C1	C1
Map resolution	3.51 Å	3.05 Å	3.07 Å
FSC threshold	0.143	0.143	0.143

Table S3. Structure refinement and validation statistics

	AcrB-NCMNP2a-5, pH= 7.8	AcrB-NCMNP2a-50, pH= 7.8	AcrB-NCMNP2a-50, pH= 5
PDB code	7RR7	7RR8	7RR6
EMDB Entry ID	EMD-24654	EMD-24655	EMD-24653
Model composition			
Non-hydrogen atoms	23,740	23,922	23,818
Protein residues	3,061	3,075	3,063
Ligands	27	26	35
R.m.s. deviations			
Bonds (Å)	0.005	0.004	0.006
Angles (°)	0.541	0.502	0.581
Validation			
MolProbity score	1.79	1.60	1.90
Clashscore	7.13	7.00	7.65
Poor rotamers (%)	1.78	1.20	2.14
Ramachandran plot			
Most favoured (%)	96.69	97.13	96.46
Allowed (%)	3.31	2.87	3.54

Table S4. Backbone RMSD comparison of AcrB structures

Structure	Reference structure	RMSD (Å)
NCMNP2a-5, pH = 7.8	SMA2000, pH=7.8	0.778
NCMNP2a-50, pH = 7.8		0.788
NCMNP2a-50, pH = 5		0.897
NCMNP2a-5, pH = 7.8	NCMNP2a-50, pH= 7.8	0.425
NCMNP2a-50, pH = 5		0.575

Table. S5. List of all CG MD simulations performed

System bundle	Systems studied, replicas	Total run time
Single polymer chain in water box	**5** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **3**	**30 µs** (2 µs per simulation)
Multiple polymer chains in water box	**5** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **2**	**20 µs** (2 µs per simulation)
Self-assembly of nanoparticles with 6 polymer chains	**5** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **2**	**30 µs** (3 µs per simulation)
Self-assembly of nanoparticles with 12 polymer chains	**5** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **2**	**30 µs** (3 µs per simulation)
Self-assembly of nanoparticles with 18 polymer chains	**5** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **2**	**30 µs** (3 µs per simulation)
Uncharged simulations - single polymer, multiple polymers, and self-assembly	**15** (SMA, NCMNP2a-5, NCMNP2a-25, NCMNP2a-50, NCMNP2a-70) × **3** × **1**	**30 µs** (2 µs per simulation)

Table S6. Number of sodium ions within 5 Å of the polymer chains when simulated individually or with other polymer chains in the absence of lipids. Data obtained by averaging the values from all replica simulations for the last 1 μs of the simulation trajectory.

	SMA	**NCMNP2a-5**	**NCMNP2a-25**	**NCMNP2a-50**	**NCMNP2a-70**
Simulation type: Single polymer chain in water (charged)					
Sodium per polymer chain	13.2 ± 6.13	13.44 ± 4.21	10.99 ± 3.56	4.64 ± 2.01	2.28 ± 1.48
Simulation type: Multiple polymer chains in water (charged)					
Sodium per polymer chain	18.13 ± 2.12	18.23 ± 1.54	12.42 ± 1.24	6.79 ± 0.78	3.14 ± 0.61
Simulation type: Self-assembly of nanodisks (charged)					
Sodium per polymer chain	22.29 ± 1.17	16.85 ± 1.45	15.82 ± 0.91	7.46 ± 0.83	4.09 ± 0.65

Table S7. Radii of Gyration (in Angstroms) for polymer chains (and polymer backbones where specified) in various simulated systems. Data obtained by averaging the values from all replica simulations for the last 1 μs of the simulation trajectory.

	SMA	**NCMNP2a-5**	**NCMNP2a-25**	**NCMNP2a-50**	**NCMNP2a-70**
Simulation type: Single polymer chain in water					
$\mathbf{R_g}$ (Å) charged - full	10.68 ± 0.38	10.66 ± 0.29	10.77 ± 0.19	11.14 ± 0.19	11.40 ± 0.21
$\mathbf{R_g}$ (Å) uncharged - full	10.26 ± 0.21	10.38 ± 0.19	10.66 ± 0.20	11.08 ± 0.19	11.33 ± 0.18
$\mathbf{R_g}$ (Å) charged - bb	10.67 ± 0.34	10.49 ± 0.26	10.30 ± 0.23	10.31 ± 0.22	10.28 ± 0.20
$\mathbf{R_g}$ (Å) uncharged - bb	10.24 ± 0.23	10.23 ± 0.21	10.15 ± 0.23	10.29 ± 0.22	10.24 ± 0.22
Simulation type: Multiple polymer chains in water					
$\mathbf{R_g}$ (Å) charged - full	10.6 ± 0.36	12.59 ± 2.72	13.23 ± 3.05	14.59 ± 2.49	16.67 ± 2.25
$\mathbf{R_g}$ (Å) uncharged - full	11.24 ± 1.12	13.44 ± 2.03	12.53 ± 1.37	16.25 ± 2.97	16.57 ± 3.03
Simulation type: Self-assembly of nanodisks					
$\mathbf{R_g}$ (Å) charged - full	17.10 ± 0.59	17.65 ± 1.24	17.53 ± 1.93	19.50 ± 1.43	19.65 ± 1.95
$\mathbf{R_g}$ (Å) uncharged - full	16.42 ± 1.02	18.08 ± 0.94	19.44 ± 1.54	18.21 ± 1.76	19.51 ± 1.68

Table S8. Solvent accessible surface areas (SASA) for hydrophobic beads in different polymers when simulated individually in the absence of lipids. Data obtained by averaging the values from all replica simulations for the last 1 μs of the simulation trajectory.

	SMA	**NCMNP2a-5**	**NCMNP2a-25**	**NCMNP2a-50**	**NCMNP2a-70**
Simulation type: Single polymer chain in water					
$SASA_{TC1}$ ($Å^2$) charged	1331.76 ± 39.41	1305.40 ± 42.99	1224.64 ± 45.71	1209.25 ± 43.01	1203.77 ± 43.34
$SASA_{TC1}$ ($Å^2$) uncharged	1257.99 ± 42.84	1243.08 ± 47.48	1199.07 ± 47.32	1200.67 ± 42.64	1200.47 ± 42.99
$SASA_{TC5}$ ($Å^2$) charged	1685.94 ± 78.21	1684.98 ± 82.64	1693.40 ± 80.22	1701.01 ± 82.11	1745.03 ± 80.39
$SASA_{TC5}$ ($Å^2$) uncharged	1671.58 ± 79.15	1680.69 ± 76.42	1689.58 ± 80.14	1706.12 ± 82.63	1745.86 ± 81.84

Table S9. Solvent-accessible surface areas (SASA) for benzenes comprising N1-N2-N3 and N4-N5-N6 beads in different polymers when simulated individually and in the presence of lipids. Data obtained by averaging the values from all replica simulations for the last 1 μs of the simulation trajectory.

	SMA	**NCMNP2a-5**	**NCMNP2a-25**	**NCMNP2a-50**	**NCMNP2a-70**
Simulation type: Single polymer chain in water					
$SASA_{N1\text{-}N2\text{-}N3}$ ($Å^2$)	878.57 ±	883.59 ±	870.23 ±	866.78 ±	884.33 ±

charged	62.11	65.29	68.59	66.06	64.64
SASA$_{N1-N2-N3}$ (Å^2) uncharged	869.57 ± 57.29	873.19 ± 62.27	872.89 ± 63.35	867.39 ± 67.63	890.61 ± 62.20
SASA$_{N4-N5-N6}$ (Å^2) charged	807.37 ± 63.18	801.39 ± 64.66	823.19 ± 66.73	834.24 ± 66.35	860.7 ± 69.06
SASA$_{N4-N5-N6}$ (Å^2) uncharged	802.34 ± 65.73	807.49 ± 62.37	816.69 ± 66.96	838.73 ± 67.97	855.25 ± 67.37
Simulation type: Self-assembly of nanodisks					
SASA$_{N1-N2-N3}$ (Å^2) charged	1032.81 ± 15.55	1022.52 ± 15.55	1026.83 ± 16.01	1023.13 ± 14.49	1055.47 ± 20.48
SASA$_{N1-N2-N3}$ (Å^2) uncharged	1029.64 ± 28.15	1011.11 ± 21.07	1029.84 ± 20.46	999.61 ± 18.36	1073.33 ± 20.64
SASA$_{N4-N5-N6}$ (Å^2) charged	970.7 ± 14.76	961.38 ± 14.75	1001.23 ± 15.31	1010.38 ± 15.29	1037.2 ± 19.15
SASA$_{N4-N5-N6}$ (Å^2) uncharged	982.13 ± 26.41	966.25 ± 20.78	1007.51 ± 20.11	989.93 ± 18.56	1056.77 ± 21.68

Table S10. Quantification of the Tris grafting on the polymer using FTIR data.

Polymer	Number of points	Degrees of freedom	Adj. R-square	Area (Acid)	Area (Amide)	Grafting (%)
NCMNP2a-5						
NCMNP2a-25	2490	2459	0.98261	2.24634	0.7273	24.46
NCMNP2a-50	2491	2457	0.52879	2.40271	2.43477	50.33
NCMNP2a-70	624	590	0.73583	2.09926	3.42499	61.999