

# Supporting Information : Understanding the nanoscale structure of hexagonal phase lyotropic liquid crystal membranes

Benjamin J. Coscia      Douglas L. Gin      Richard D. Noble      Joe Yelk  
Matthew Glaser      Xunda Feng      Michael R. Shirts

September 14, 2018

## 1 Further details regarding monomer parameterization

We parameterized monomers according to the following procedure:

1. *Create monomer structure file with connectivity* : We drew atomistic structures using MarvinSketch 17.13 [1] with all hydrogen atoms drawn out explicitly. We optimized the 3D geometry of the structure using the ‘Clean in 3D’ function of MarvinSketch. We saved the structure as a .mol file, then converted it to .pdb format using Open Babel 2.4.1 [2, 3].
2. *Assign GAFF atomtypes using antechamber* : Using the .pdb structure file as input, we ran `antechamber` [4] using the AM1-BCC charge model. The net charge on the monomer is input as -1 since the sodium ion is kept as a separate residue. We use `LEaP` [5] and the output of `antechamber` to create Amber topology files. A detailed tutorial can be accessed elsewhere [6].
3. *Create GROMACS topologies from Amber output* : The output of `LEaP` is a .inpcrd and a .prmtop file which are Amber topology files. Using `acpype.py` [7], we converted the `LEaP` output into GROMACS .gro and .top files.
4. *Perform a simulated annealing procedure on the monomer* : We created a cubic box around the monomer using the GROMACS command `gmx editconf`. The monomer was centered in the box with edges of the box spaced at least 3 nm from the monomer on all sides. We ran an energy minimization on the system with the steepest descent algorithm. Next we performed an NVT simulated annealing procedure. We linearly decreased the temperature of the system from 1000 K to 50 K over the course of 10 ns. We randomly chose a monomer configuration from the last 10 % of the trajectory.
5. *Reassign charges with molcharge* : With the monomer configuration taken from the annealed trajectory, we reassigned charges using `molcharge` with the am1bccsym method in order to ensure charges are symmetric. This condition is not guaranteed with `antechamber`. The charges in the GROMACS topology file (.top) were replaced with the new charges calculated by `molcharge`.
6. *Anneal again to get final structure* : We repeated the same simulated annealing procedure using the monomer topology with `molcharge` charges. A random monomer configuration was pulled from the last 10 % of the trajectory and was used to build all assemblies reported (Figure 1).

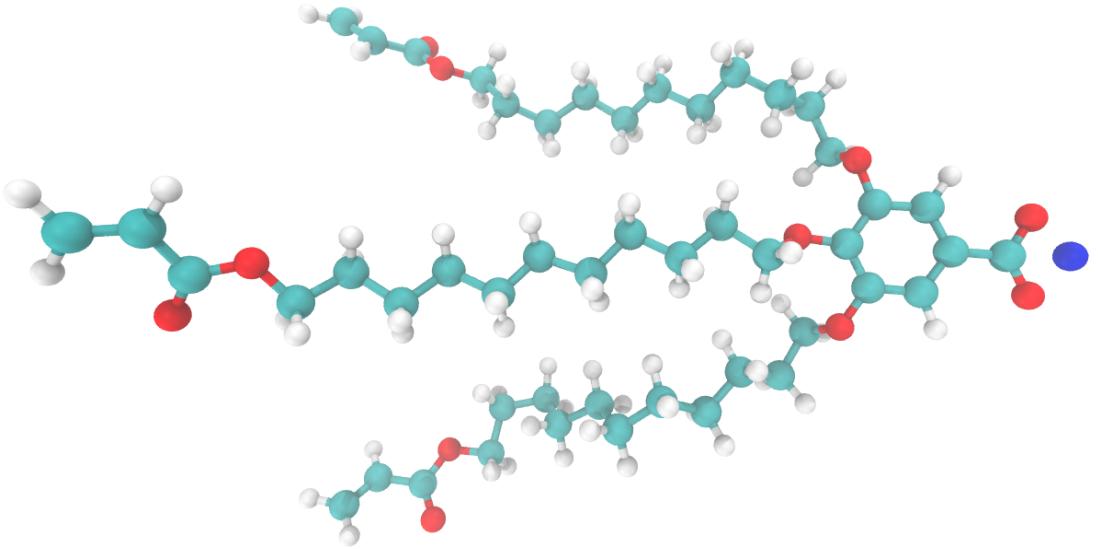


Figure 1: Atomistic representation of the monomer Na-GA3C11. White atoms represent hydrogen, cyan atoms represent carbon, red atoms represent oxygen and the blue atom is sodium.

## 2 Attempted Self-assembly

We attempted self assembly of Na-GA3C11 monomers by using an isotropic configuration generated using Packmol [8] (Figure 2a). The input file given to Packmol is shown below.

```

tolerance 2.0
output packed.pdb
filetype pdb
structure monomer.pdb
    number 400
    inside box 0. 0. 0. 84.3 84.3 84.3
end structure

```

We used the same number of monomers (400) that we used in all other simulations. Since Packmol does not have the capability to create monoclinic boxes, we used a cubic box with a volume of  $660\text{ nm}^3$ , 25% larger than those used to create ordered unit cells we studied. We tested systems with semi-isotropic and anisotropic pressure coupling since the shape of the unit cell will likely need to change in order to accommodate hexagonally packed pores.

We use the nematic order parameter as defined in Section 12 to observe any progress towards system ordering. In both cases, we see that the nematic order parameter stays close to zero for the duration of the simulation time, indicating an isotropic arrangement of head groups (Figure 2b). The values are slightly negative since there are only 400 monomers. Equation 3 is bounded between -1 and 1, however, with a proper choice of nematic director, measured values will range between 0 and 1. With a large number of samples, an isotropic system would have an order parameter of 0.

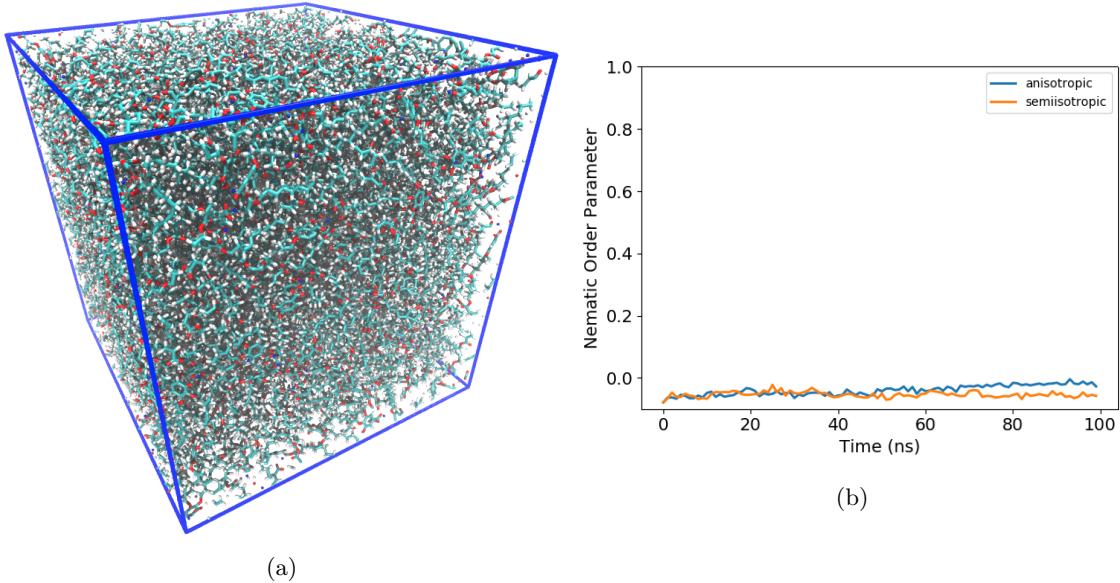


Figure 2: (a) We created a box of isotropically packed monomers and allowed it to simulate for 100 ns using isotropic and anisotropic pressure coupling. (b) The nematic order parameter hovers close to zero for the duration of the simulation meaning the system maintains its isotropic alignment.

### 3 Number of monomers per column

We chose to build all systems using 20 monomers per column. We made this choice by considering a number of factors that directly affect the quality of our results. Clearly, choosing a small number of layers will cause the system to suffer from finite size effects. Even in our 20 monomer-per-column system, axial correlations between stacked monomers persist throughout the entire correlation function (see Figure 10 of the main text). The correlation function of the system which we built with 40 monomers per column is very similar to that created with 20 monomers per columns (continues to oscillate for its full length (Figure 3b)). If we continue adding monomers, the correlation function will eventually fully decay but it is not worth the computational expense. Regardless, the correlation lengths calculated from 20 monomer per column simulations are in sufficient agreement with experiment.

We may be able to simulate systems using less than 20 monomers per column, but chose to use 20 since it gives fairly good resolution when simulating X-ray diffraction patterns. The size of the Fourier space bin in each dimension is determined by  $\frac{2\pi}{L}$  where  $L$  is the length of a given box vector. The  $z$ -box vector for 20 monomer-per-column systems is  $\approx 85 \text{ \AA}$  which means we will see a reciprocal space resolution of  $0.074 \text{ \AA}^{-1}$  in the  $z$  direction. Further decreasing the resolution would increase the noise in the simulated patterns and make it more difficult to measure the intensity of reflections of interest.

### 4 Initial Configuration Dependence

We addressed any major dependence on initial configuration in the main text. There we showed that systems are stable when made with 4, 5, 6, 7, and 8 columns per pore. We also showed that systems are stable when monomer head groups are oriented in the parallel displaced and sandwiched configurations (see Figure 8). Our model is most consistent with experiment when built with 5 columns per pore, with monomers stacked  $3.7 \text{ \AA}$  apart in the parallel displaced configuration.

There are three other parameters whose choice may influence the equilibrium structure: initial pore spacing, initial pore radius and initial stacking distance between monomers. Here we show the results of a sensitivity analysis performed on the three parameters. To reduce the size of the sensitivity analysis, we only

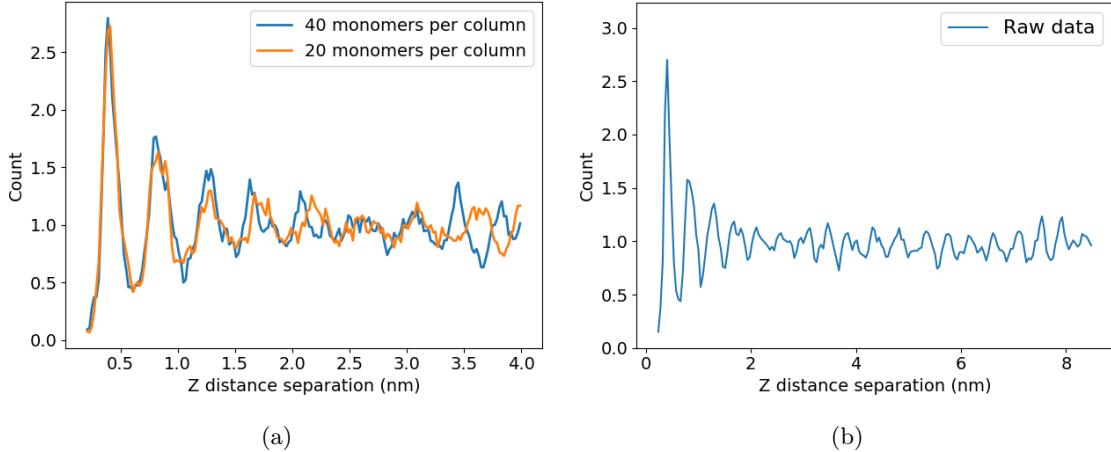


Figure 3: (a) The correlation functions generated from ordered basin sandwiched configurations are nearly the same when the  $z$ -dimension of the system is doubled. (b) Oscillations persist throughout the length of the correlation function for the sandwiched system in the ordered basin built with 40 monomers per column. The full system is ca. 17 nm tall so the full correlation function covers half this length since periodicity forces the remaining length to be its mirror image.

tested systems built with 5 columns per pore with monomers stacked in the parallel displaced configuration. We equilibrated all systems according to the dry equilibration procedure.

#### 4.1 Initial pore spacing

We tested five different initial pore spacings, defined as the distance between the central axis of each pore with all others (Figure 10). To reduce the number of variables, we held the pore radius constant, at 5 Å and the distance between layers at 3.7 Å since those were the values used in our optimal system in the main text. We prioritized ensuring that resulting configurations maintain the expected hexagonal symmetry. If initial pore spacing is too small, we observe repulsion between columns which disrupts equilibration. If the pore spacing is too large the pores squeeze together, but in a distorted hexagonal array because the  $xy$  initial translation of the pores is somewhat erratic.

1. 39 Å : We tested a pore spacing of 39 Å in order to have a test system with an initial spacing below the experimental value. As soon as the restraining potential switches to  $56 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , the columns are able to repel resulting in a large jump in pore spacing (Figure 4b).
2. 41 Å : We chose to test a pore spacing of 41 Å since it closely matches the experimental pore spacing. Again, we observe abrupt repulsion of columns once the restraining potential is switched to  $56 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  (Figure 4c).
3. 45 Å : A pore spacing of 45 Å is about 10 % larger than the experimental value. This system exhibits the smallest response to our equilibration procedure (Figure 4d) and we chose to use this value for all our simulations in the main text.
4. 50 Å : We tested a pore spacing of 50 Å, about 20 % larger than the experimental value. Once the restraining potential reaches  $56 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , the pore spacing begins to decrease rapidly, following a linear trend (Figure 4e).
5. 55 Å : We tested a pore spacing of 55 Å which is at a distance where monomers in each pore no longer intersect those of adjacent pores. Once the restraining potential reaches  $56 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , the pore spacing changes erratically until it begins to settle when the force constants are below  $3 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  (Figure 4f). We recommend avoiding a system such as this where vacuum gaps between pore columns are introduced unnecessarily.

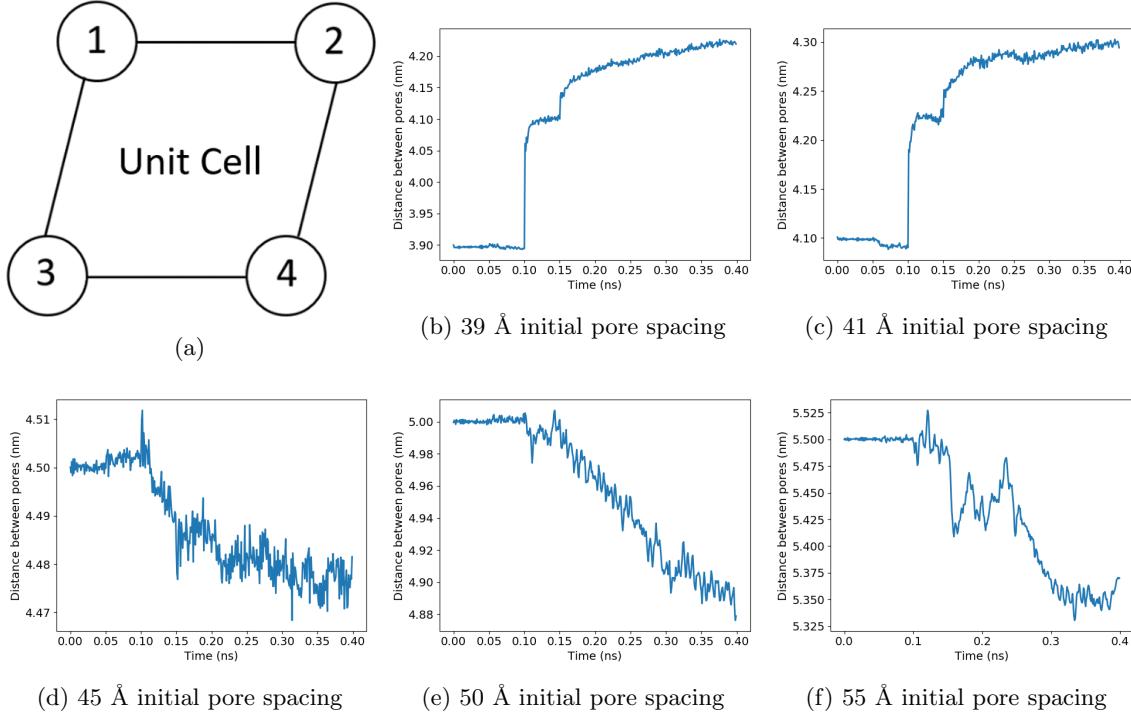


Figure 4: (a) In a perfect hexagonal, there are 5 distances that should be exactly equal. As pictured in this diagram the distance from 1-2, 1-3, 1-4, 2-4, and 3-4 are the same. All plots in (b)–(f) are the average of these distances and represent the pore spacing during the restrained portion of the dry equilibration procedure. Every 50 ps (0.05 ns) we reduce the force constant on the position restraints according to the sequence: 1000000, 3162, 56, 8, 3, 2, 1, 0  $\text{kJ mol}^{-1} \text{ nm}^{-2}$ . When we chose an initial pore spacing below experimental value (b) or at the experimental value (c), there is an abrupt change in pore spacing when the position restraints are reduced to 56  $\text{kJ mol}^{-1} \text{ nm}^{-2}$ . When we initially position pores 45 Å apart (d) the pore spacing remains relatively stable. When we initially place pores 50 Å apart (e), the pore spacing decreases nearly linearly once the restraints are reduced to 56  $\text{kJ mol}^{-1} \text{ nm}^{-2}$ . When we initially place pores 55 Å apart (f), so that monomers do not intersect with adjacent pores, and position restraints are reduced to 56  $\text{kJ mol}^{-1} \text{ nm}^{-2}$ , the pore spacing changes erratically before stabilizing when force constants are reduced below 3  $\text{kJ mol}^{-1} \text{ nm}^{-2}$ .

## 4.2 Initial pore radius

We tested 3 different pore radii, defined as shown in Figure 5. For each system, we held the initial pore spacing constant at 45 Å and the initial stacking distance between monomers constant at 3.7 Å. Equilibrated values of pore radii are presented in Table 1. The pore radius in each simulation frame is calculated as the average distance of all carbonyl carbons (see Figure 5) from their associated pore center. We generated statistics for the pore radii reported from the time series representing the average pore radius at each frame. We detected equilibration as described in Section 2.5.1 of the main text. We calculated the average and standard deviation of pore radii using only data points collected after equilibration was detected.

1. 2.5 Å : The smallest pore radius that we can achieve before energy minimization becomes problematic is 2.5 Å. After being run through the full dry equilibration procedure, the average pore radius is  $0.40 \pm 0.01$  Å.
2. 5 Å : We tested a pore radius of 5 Å because it is slightly larger than the equilibrated pore radius of the system simulated with an initial pore radius of 2.5 Å. After being run through the full dry equilibration procedure, the average pore radius is  $0.42 \pm 0.01$  Å, which agrees with the 2.5 Å configuration within uncertainty.

3. 8 Å : The largest pore radius that can achieved before energy minimization becomes problematic is 8 Å. One should use caution with such a structure because of the relatively large vacuum space that is created in the pore region of the initial configuration. After being run through the full dry equilibration procedure, we see a combination of cylindrical and slit-like pores (Figure 6). Measuring the pore radius of this system does not have a concrete meaning since slits do not have a single radius, but its calculated value is still reported in Table 1. We are wary of such a non-symmetric structure and choose not to use a pore radius of 8 Å for our starting configurations.

Initial Pore Radius	Equilibrated Pore Radius
2.5 Å	$4.0 \pm 0.1$ Å
5 Å	$4.2 \pm 0.1$ Å
8 Å	$6.9 \pm 0.1$ Å

Table 1: The average pore radii of systems built with an initial pore radius of 2.5 Å and 5 Å equilibrate to values that agree within error. If the pore radius is too large, slit pores may form.

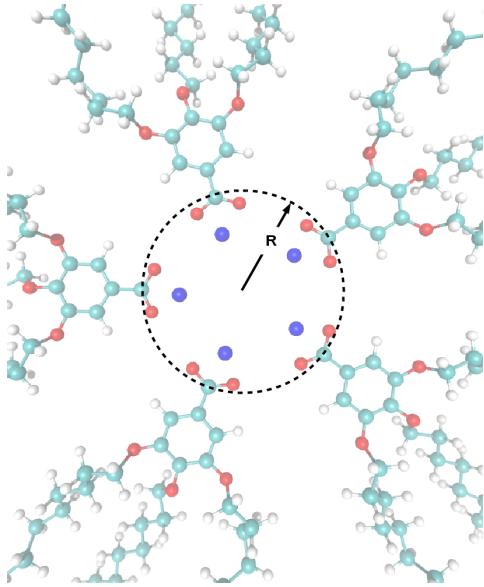


Figure 5: When creating an initial configuration, we define the pore radius,  $\mathbf{R}$ , based on the distance of the carbonyl carbon from the pore's central axis.

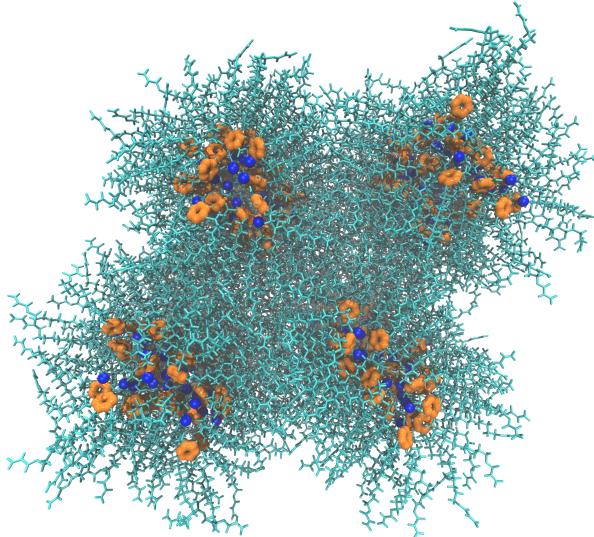


Figure 6: A system that was built with an initial pore radius of 8 Å equilibrates to a structure that exhibits both cylindrical and slit-like pores. As pictured here, sodium ions are colored blue, carbon atoms in the aromatic ring of the head group are colored orange and all else is colored cyan.

### 4.3 Initial stacking distance between monomers

We tested 3 different initial monomer spacings, defined as the  $z$ -distance between the planes of aromatic rings. Systems built with layers stacked 3.7 Å and 5 Å apart are discussed extensively in the main text. We

also tested a system built with layers stacked 10 Å apart.

Figure 7 shows the structure of an assembly built with an initial layer spacing of 10 Å immediately after the restrained portion of the equilibration procedure. Since we used position restraints, the simulations were run in the NVT ensemble. When layer spacing is large, such as this situation, there is a significant amount of vacuum space which the monomer attempts to fill. Even if turning pressure control on allows the system to recover the geometry of the hexagonal phase, we would likely need much longer equilibration times, and it will almost certainly get trapped in a metastable configuration that bears no resemblance to the experimental profile.

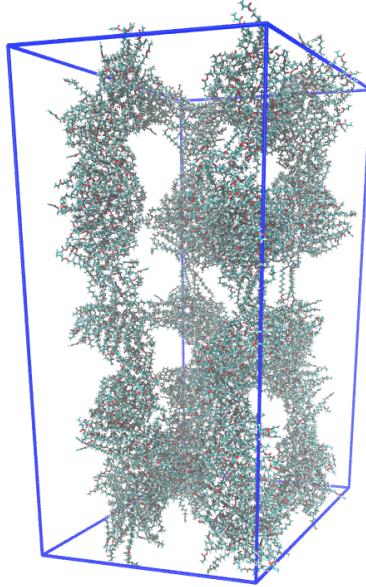


Figure 7: When layers are initially stacked 10 Å apart and the system is equilibrated using the dry equilibration procedure, large vacuum gaps form as the monomers attempt to fill space.

## 5 Placement of monomer head groups

The monomer head groups consist of a substituted benzene ring which may participate in  $\pi - \pi$  stacking interactions. There are three  $\pi - \pi$  stacking modes known to occur in nature: sandwiched, parallel displaced and T-shaped (Figures 8a-8c) [9]. We ruled out the T-shaped configuration because its vertical stacking distance is inconsistent with experimental spectroscopy. Additionally, steric interactions between the bulky alkane tails make this motif unfavorable. We chose to explore systems in both the sandwiched and parallel displaced configurations since their stacking distances are both consistent with experimental diffraction patterns.

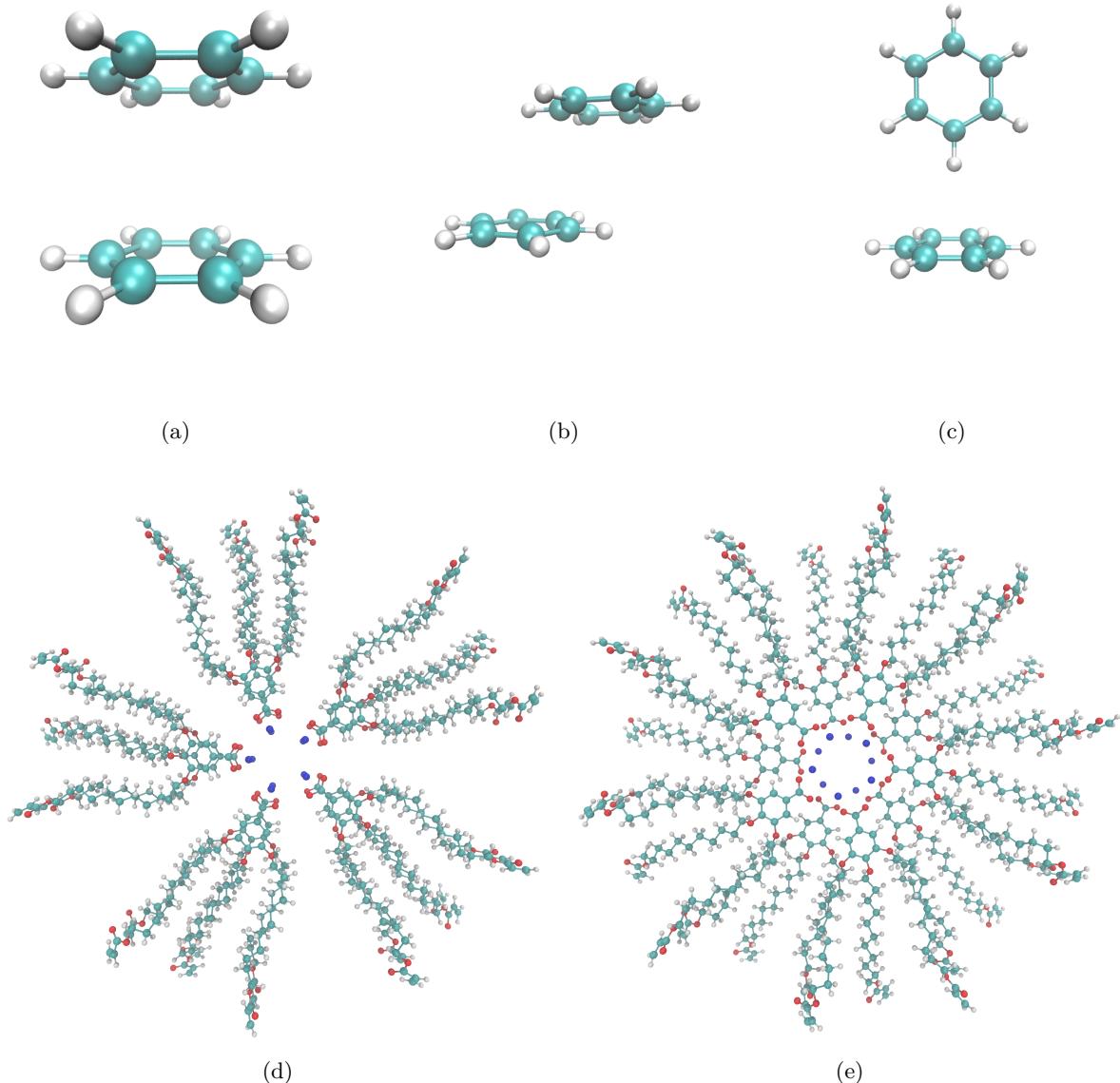


Figure 8: (a) Sandwiched benzene dimers stack 3.8 Å apart. (b) Parallel-Displaced benzene dimers stack 3.4 Å vertically and 1.6 Å horizontally apart. (c) T-shaped benzene dimers stack 5.0 Å apart. (d) Monomers stacked in the sandwiched configuration (e) Monomers stacked in the parallel-displaced configuration.

Based on evidence from section 3.2.4 of the main text, we believe it is most likely that the head groups

stack somewhere between the sandwiched and parallel displaced modes which we studied (Figure 9).

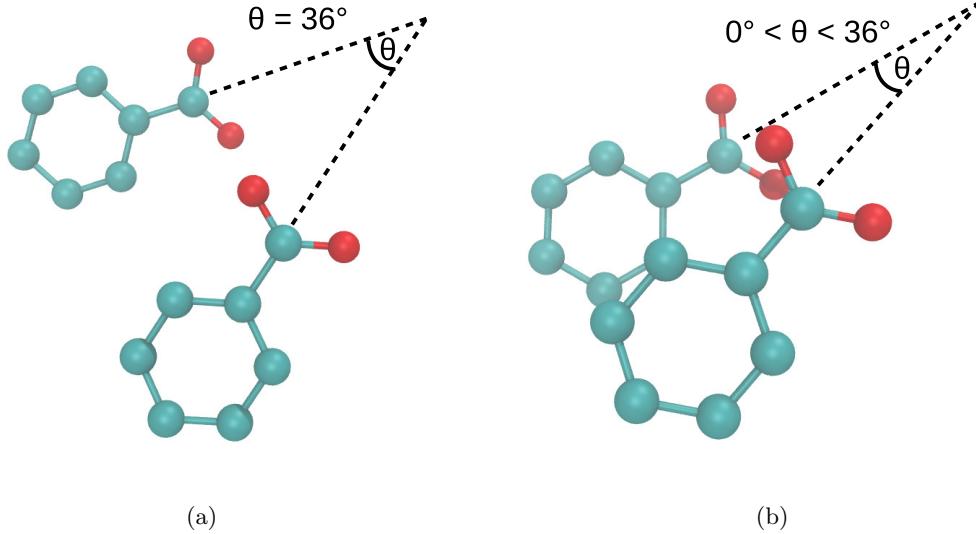


Figure 9: Head groups are likely stacked in a configuration between sandwiched and parallel displaced. (a) In our parallel displaced configurations with 5 columns per pore, we placed head groups so the angle,  $\theta$ , between vectors extending from vertically stacked head groups to the pore center equals  $36^\circ$  ( $(360^\circ / n_{col}) / 2$ , where  $n_{col}$  is the number of columns per pore). (b) In the true configuration,  $\theta$  is likely distributed between  $0^\circ$  and  $36^\circ$ . When  $\theta=0^\circ$ , the system architecture is the same as a sandwiched configuration.

## 6 Equilibration Details

We developed equilibration schemes to create dry and wet configurations. Both schemes start with an initial configuration generated according to the previous guidelines. To create a dry configuration, we fix monomer head groups in the sandwiched or parallel-displaced configuration using position restraints with a force constant of  $10^6$  kJ mol $^{-1}$  nm $^{-2}$ . We run a 50 ps simulation in the NVT ensemble which allows the monomer tails to settle without disrupting the ordering of the head groups. Doing so also mitigates system dependence on initial monomer configuration. Every 50 ps, we reduce the force constants by the square root of its previous value. Once the force constant is below 10 kJ mol $^{-1}$  nm $^{-2}$ , we reduce the restraints in a sequence with values of 8, 3, 2, 1, and 0 kJ mol $^{-1}$  nm $^{-2}$  respectively. We allow the resulting unrestrained structure to equilibrate for 5 ns in the NPT ensemble with pressure controlled by the Berendsen barostat. Next, we run long NPT equilibration simulations for at least 400 ns using the Parrinello-Rahman barostat with a time constant of 10 ps.

In order to create a “wet” system, we solvated an initial configuration with water using `gmx solvate`. We remove all water molecules placed outside the pore region. Then we randomly remove water molecules inside the pore region until the pores reach the desired concentration of water. The remainder of the equilibration follows the same procedure as the dry system.

## 7 Calculation of pore-to-pore spacing statistics

We are interested in 5 pore-to-pore distances which should all be equal in a perfect hexagonal array (Figure 10). Only 4 of the 5 distances are independent. We can calculate a trajectory of spacing versus time for each of the 5 distances. We calculated the average pore-to-pore spacing and its uncertainty according to the following procedure:

1. We calculated the time when each of the pore-to-pore distances were equilibrated using `pymbar.timeseries.detectEquilibration()` [10, 11]. We began calculations after the largest of the five values.
2. We calculated how long it takes for the data in each of the 5 trajectories to become uncorrelated using `pymbar.timeseries.integratedAutocorrelationTime()` [10, 11].
3. We broke the full equilibrated trajectory into blocks of length  $\tau$ , where  $\tau$  is the maximum of the five autocorrelation times calculated. Each block contains five sub-trajectories of pore-to-pore spacings.
4. We generate statistics using the bootstrapping technique. For each bootstrap trial, we reconstruct an equilibrium trajectory by randomly sampling from the trajectory blocks.
5. The average pore-to-pore distance is the mean of all pore spacings among all bootstrap trials.
6. To calculate the uncertainty in pore-to-pore distance, we calculate the average pore spacing for each pore over all bootstrap trials. Using the 5 average pore-to-pore distances, we calculate the spread with:

$$s = \sqrt{\frac{1}{4} \sum_{i=1}^5 (x_i - \bar{x})^2} \quad (1)$$

where  $\bar{x}$  is the average pore-to-pore distance. We divided the sum in equation 1 by 4 since there are only 4 independent pore spacings.

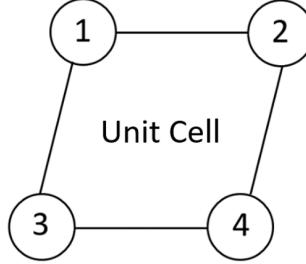


Figure 10: There are five pore-to-pore distances that are equal in a perfect hexagon. If each number in the diagram represents the center of a pore in a hexagonal unit cell, then the distance from 1 to 2, 2 to 4, 4 to 3, 3 to 1 and 1 to 4 should be equal. Only 4 of the distances are independent. For example, the distance from 1 to 4 is defined by the location of all other pore centers.

Alternatively, one can calculate the pore spacing as half of the  $x$  and  $y$  box vectors, however this technique does not fully capture the uncertainty in the pore-to-pore distances.

System	Pore Spacing (nm)	Pore Spacing (nm) $\frac{x \text{ box vector}}{2}$ (nm)
	Pore centers	
Sandwiched, Ordered	$4.18 \pm 0.13$	$4.16 \pm 0.01$
Parallel Displaced, Ordered	$4.17 \pm 0.12$	$4.18 \pm 0.01$
Sandwiched, Disordered	$4.05 \pm 0.14$	$4.01 \pm 0.01$
Parallel Displaced, Disordered	$4.03 \pm 0.07$	$4.01 \pm 0.01$

Table 2: The pore spacing measured in two ways agree within uncertainty. In the main text, we report values calculated by measuring the average distance between pore centers. Alternatively, one can estimate the pore spacing as half of the  $x$  and  $y$  box vectors. Both methods agree within uncertainty, but measuring the distance between pore centers does a better job of showing the spread of pore-to-pore distances.

## 8 Component radial density functions

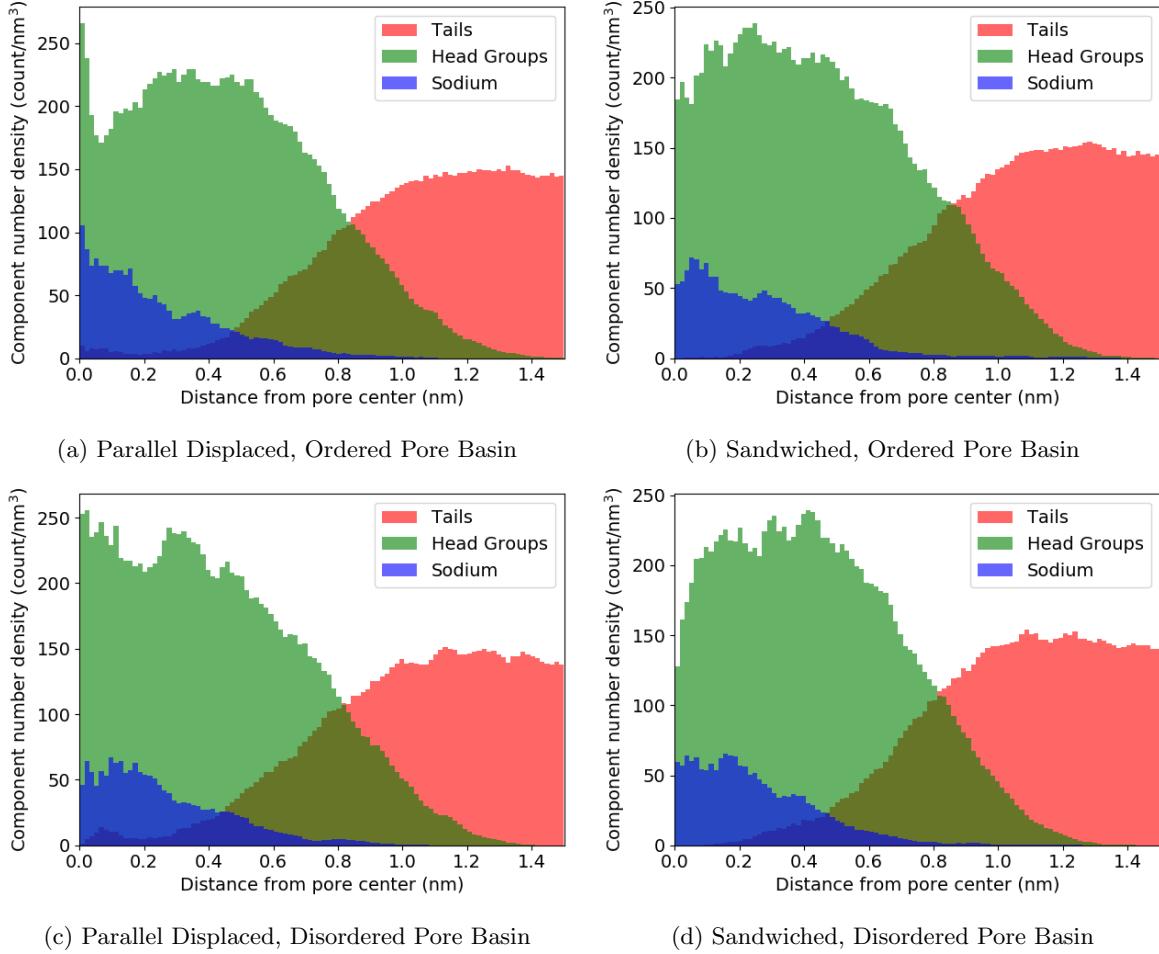


Figure 11: In all cases, the component radial distribution functions exhibit a composition gradient transitioning from the hydrophilic to the hydrophobic regions. Systems with layers initially spaced 3.7 Å apart in the parallel displaced (a), and sandwiched (b) configurations both show the highest concentrations of ions and head groups away from the pore center. Systems with layers initially spaced 5 Å apart (c) and (d), both show the highest concentrations of ions and head groups near the pore center, implying a more uniform, disordered pore.

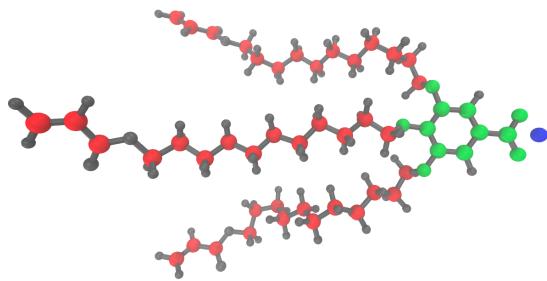


Figure 12: The groups used for radial distribution calculations. Red atoms are in the tails group. Green atoms are in the head group region. The blue atom is sodium.

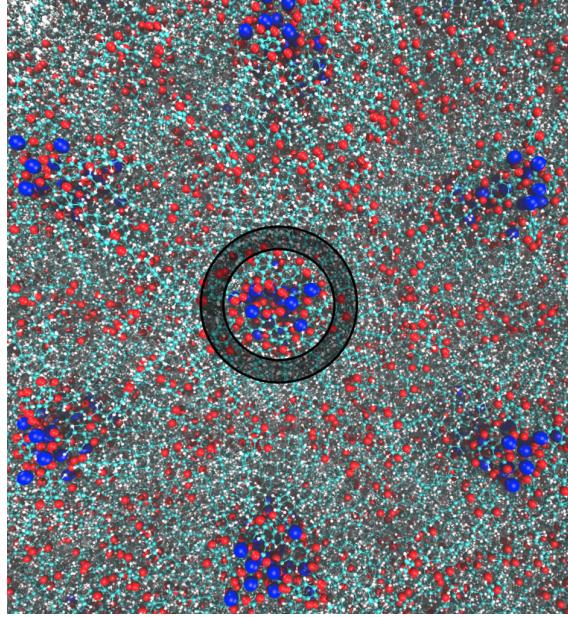


Figure 13: Looking down onto the plane of membrane, we binned the radial distance of all atoms in chosen groups from pore centers. The bins are defined by the annulus bounded by concentric circles centered at the pore centers, as shown. We normalize by dividing the count of atoms within each bin annulus by its volume.

## 9 2D Small Angle X-ray Scattering Data

## 10 Ionic Conductivity

In the main text, we calculated ionic conductivity using the Nernst-Einstein relationship. We also measured ionic conductivity using a second method, termed the Collective Diffusion model, for robustness. The collective diffusion model measures the movement of the collective variable,  $Q$ , which is defined as the amount of charge transfer through the system and can be thought to represent the center of charge of the system. The conductance,  $\gamma$ , of the system can be calculated as:

$$\gamma = \frac{D_Q}{k_b T} \quad (2)$$

$D_Q$  is the diffusion coefficient of the collective variable  $Q$  and is calculated using the Einstein relation. We convert the resulting value to ionic conductivity by multiplying by channel length and dividing by the membrane cross sectional area. One can access a detailed derivation of the model elsewhere[12].

## 11 Crosslinking algorithm details

Crosslinking of the monomer Na-GA3C11 occurs via a UV initiated free radical polymerization (FRP) (Figure 16). Head-to-tail addition takes place between terminal vinyl groups on each of the monomer tails. We only considered head-to-tail addition since it is the dominant propagation mode in the real system.

We based our cross-linking algorithm on the known reaction mechanism. FRPs require an initiator which bonds to the system, meaning new atoms are introduced into the system. For simplicity, we simulated the initiator as hydrogen and made it present in the simulation by including them as dummy atoms in all possible

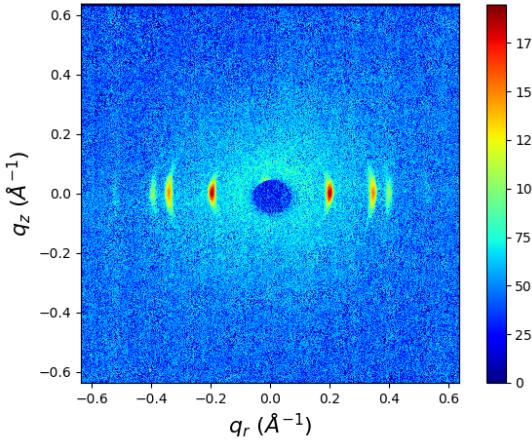


Figure 14: Two dimensional small angle X-ray scattering experiments show reflections along the  $q_r$  axis at  $q_z=0$ . The spacing between the reflections is characteristic of a hexagonal phase and is directly related to the distance between pores. The reflections are arced because the pores are not perfectly aligned. A perfectly aligned system would show dots at each peak location. An isotropically aligned system would show concentric rings about the origin which pass through each peak location. The  $d_{100}$  and  $d_{110}$  peaks (the first two pairs of reflections from origin located at  $|\mathbf{q}| \approx 0.18 \text{ \AA}^{-1}$  and  $|\mathbf{q}| \approx 0.31 \text{ \AA}^{-1}$ ) are not visible in the 2D WAXS pattern. The edges of the  $d_{200}$  reflection (the third pair of reflections located at  $|\mathbf{q}| \approx 0.36 \text{ \AA}^{-1}$ ) are slightly visible in the 2D WAXS pattern, so we used them to approximate the experimental intensity of R-pores.

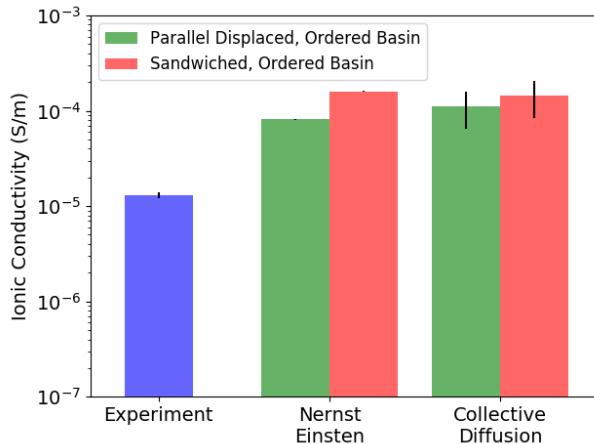


Figure 15: The Collective Diffusion model and the Nernst-Einstein relation yield agreeing values of ionic conductivity for both types of ordered basin systems simulated. Both methods estimate the value of the ionic conductivity to be an order of magnitude higher than the experimental value. There is more noise in the collective diffusion model because there is inherently less data that can be used for its calculation. For that reason, we use only the Nernst-Einstein relationship in the main text.

locations where an addition could occur. We carry out the cross-linking procedure iteratively. During each iteration, the algorithm selects eligible bonding carbon atoms based on a distance cut-off. The algorithm then updates the topology with new bonds and changes dummy hydrogen atoms to appropriate hydrogen atom types. We only considered head-to-tail addition due to its dominance in the real system [13]. We did not consider direction of attack because the resultant mixture is racemic.

Our implementation requires long simulation times to achieve high cross-link densities. A typical cross-

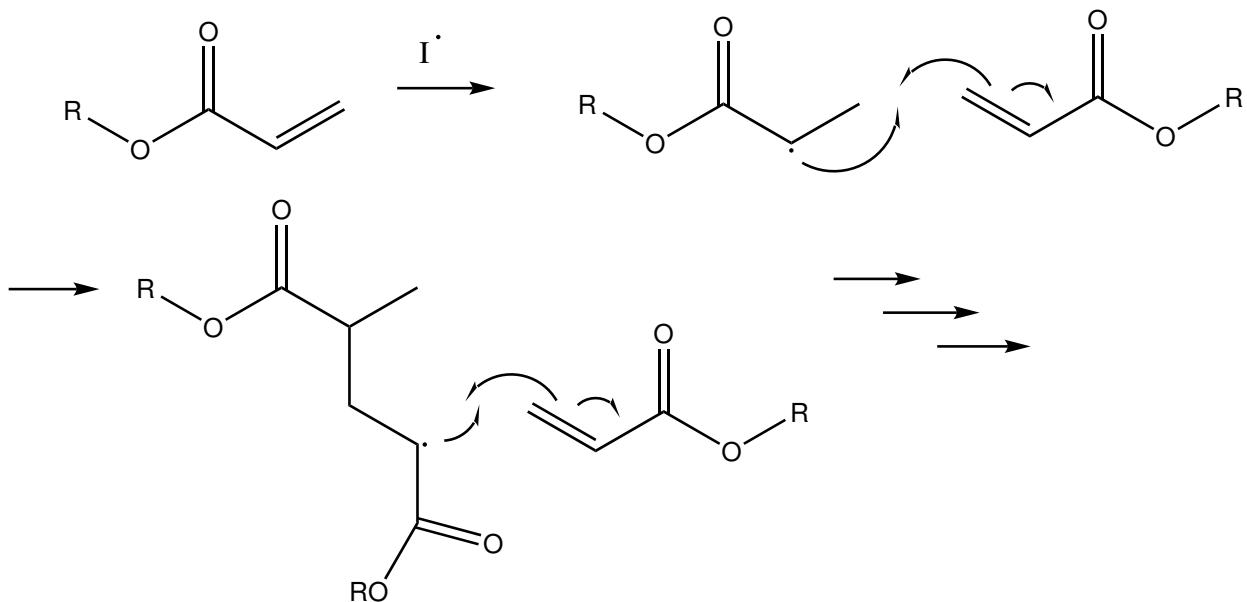


Figure 16: Terminal vinyl groups present on each monomer tail react with free radical initiators to create monomers with terminal vinyl radicals. Vinyl radicals react with the vinyl groups of other monomers in order to propagate crosslinking.

linking procedure can take up to 24 hours. In order to collect equilibrated data, further NPT simulation is necessary. We typically run a cross-linked system for an additional 100 ns to allow the system to readjust. For those reasons we did not cross-link all systems tested.

## 12 The nematic order parameter

We calculated the nematic order parameter for our system in order to understand the degree of ordering among monomer head groups. Typically, the nematic order parameter is calculated for nematic liquid crystal systems which are characterized by unidirectional ordering of liquid crystal monomers. The preferred direction of monomers is defined by the unit director vector,  $\mathbf{n}$ . Assuming a single preferred direction of alignment, the nematic order parameter,  $S$ , is defined as: [14]

$$S = \frac{1}{2} \langle (3 \cos^2 \theta - 1) \rangle \quad (3)$$

where  $\theta$  is the angle between the molecular long axis and  $\mathbf{n}$ . In a perfectly ordered nematic liquid crystal system, the molecular axis of each monomer should be aligned with  $\mathbf{n}$  and give an order parameter of  $S = 1$ . We are interested in quantifying the degree of monomer head group alignment between systems. We use Equation 3 to accomplish this by defining  $\mathbf{n}$  as the  $z$ -axis (or pore axis), and then measuring the angles,  $\theta$ , between  $\mathbf{n}$  and the vectors perpendicular to the plane of the aromatic head groups ( $\mathbf{v}$  in Figure 17).

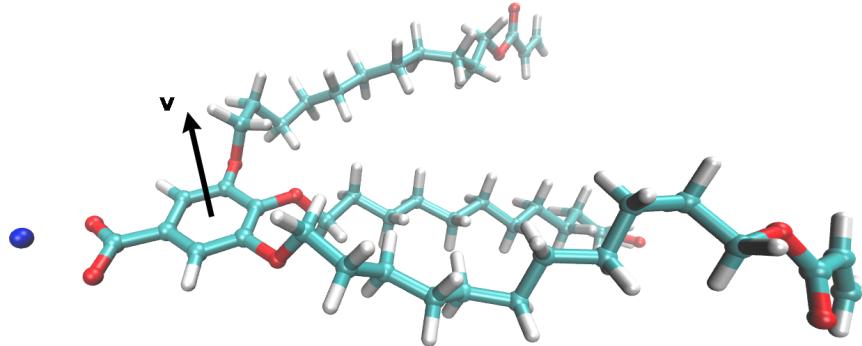


Figure 17: We calculate the nematic order parameter by measuring the angle between  $\mathbf{v}$  and the  $z$ -axis.

The nematic order parameter is highest in sandwiched systems. Figure 18 shows the distribution of angles between the nematic director and the vector perpendicular to the monomer head groups. Systems in the ordered basin have higher nematic order parameters than their disordered basin counterparts. Both sandwiched system have higher nematic order parameters than both parallel displaced systems. This may occur because monomers in the sandwiched configuration have their motion restricted by vertically adjacent monomers, while those in the parallel displaced configuration have more room to rotate.

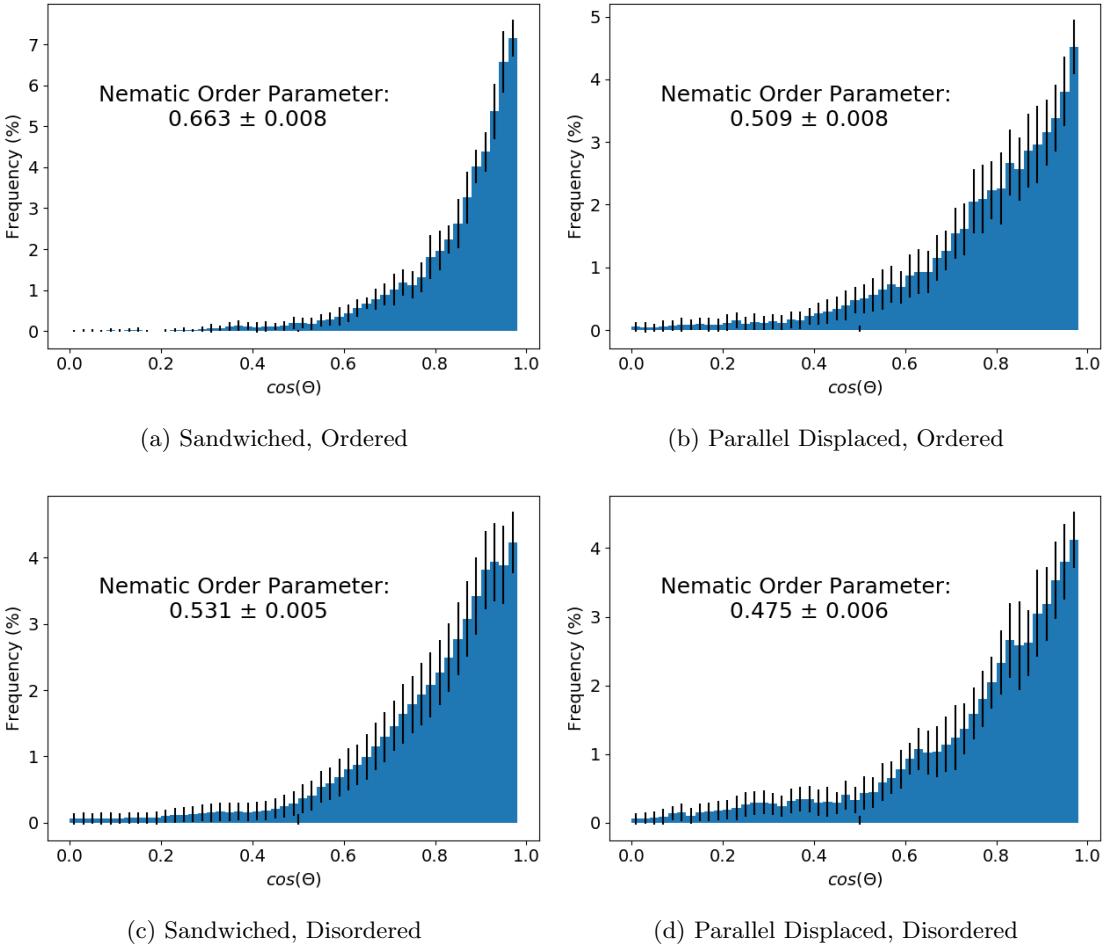


Figure 18: The distribution of angles between nematic director vector (see Figure 17) and the  $z$ -axis averaged over the equilibrated portion of each trajectory. The ordered sandwiched configuration has a higher nematic order parameter than the ordered parallel displaced configuration. The disordered sandwiched configuration has a higher nematic order parameter than the disordered parallel displaced configuration.

### 13 Noise in simulated diffraction patterns

The magnitude of noise in the simulated diffraction patterns decreases as the number of independent configurations increases (Figure 19). We simulated the structure factor of 100 randomly placed particles and varied the number of independent configurations ( $N$ ). It is expected that the structure factor should resemble a two-dimensional Gaussian. We cannot sample enough independent configurations of our system to completely eliminate noise.

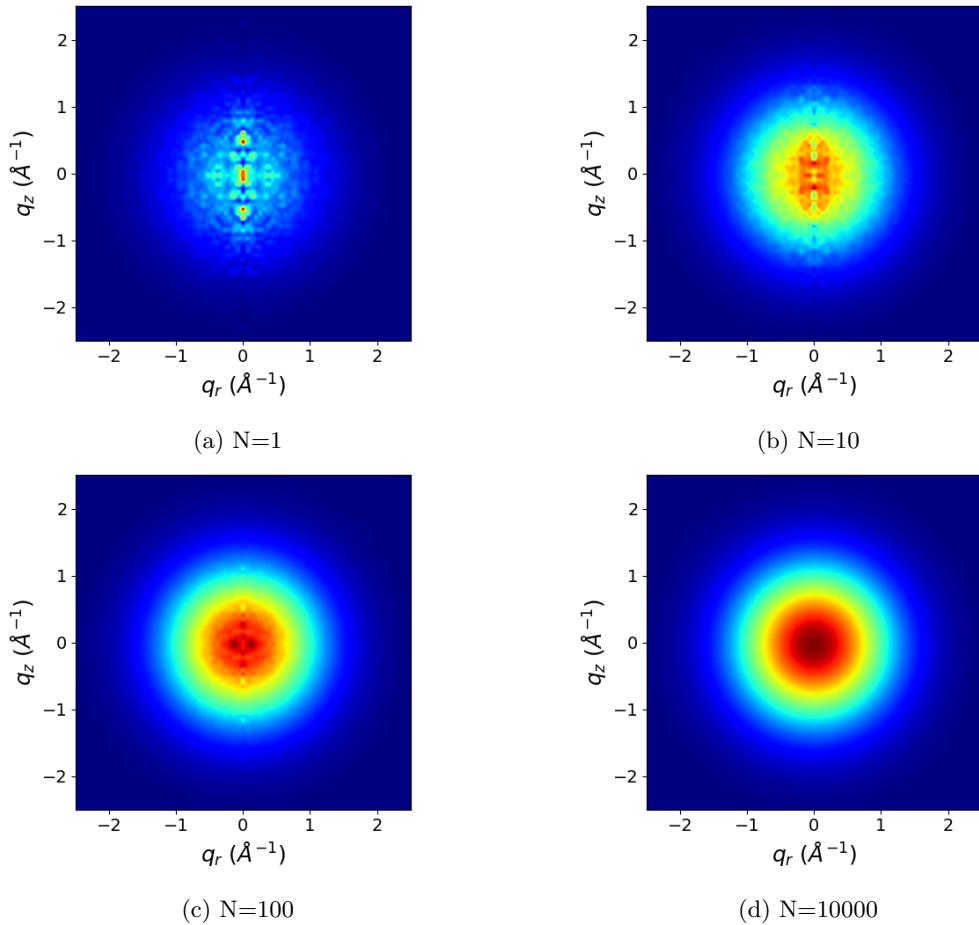


Figure 19: The magnitude of noise, especially along  $q_r=0$ , decreases as the number of independent configurations ( $N$ ) increases.

## 14 The effect of imposed tail tilt

Monomer tails prefer to lie flat rather than tilt with respect to the membrane plane. We created a monomer whose tails tilted ca.  $30^\circ$ . We added position restraints to the ends of the monomer tails with the same force constant as the head groups (see section 6). As we slowly reduce the force constants, the tails begin to tend towards a negligible tilt angle (Figure 20).

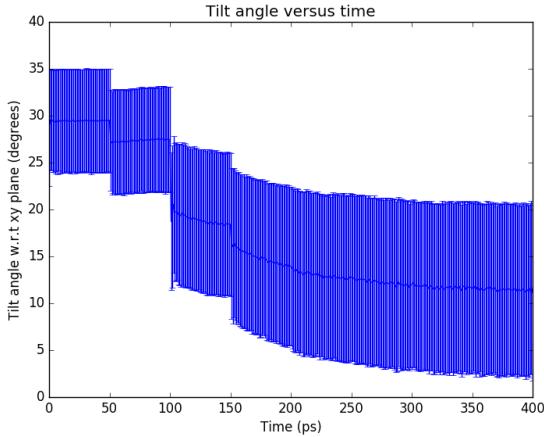


Figure 20: As we reduce the force constants applied to the end of each tail of an initially tilted monomer configuration, the average tilt angle with respect to the membrane plane decreases until it is nearly negligible.

## 15 Tail organization

In order to understand the origin of R-spots, we looked for ordering among the tails. We divided the tail into three sections: tail-fronts, tail-middles and tail-ends. The sections are illustrated in Figure 21.

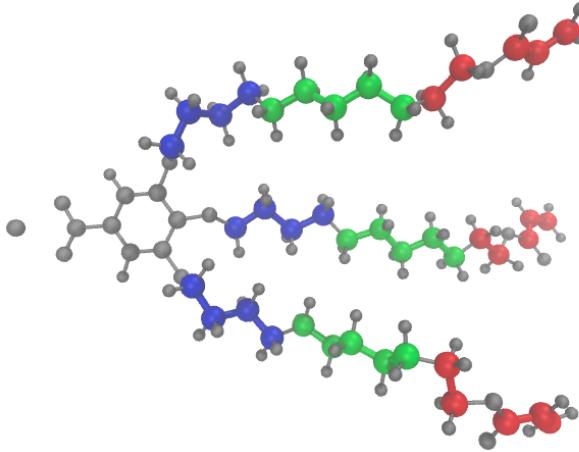


Figure 21: We divided the tail into three regions: tail-fronts (blue), tail-middles (green) and tail-ends (red).

We plotted the center of masses of each tail section and measured the angle between each center of mass and its nearest neighbor center of masses of the same tail section. For example, Figure 22a shows a 3D plot of the center of masses of each tail front. We wanted to know the average location of nearest neighbors to each center of mass. The red bead in Figure 22a is a randomly selected center of mass. The green beads are the nearest neighbor center of masses to the red bead. We measured the angle between the red bead and

each green bead with respect to the  $xy$  plane. We repeated the same calculation for each center of mass at each simulation frame and were able to plot the average distribution shown in Figure 22b. We repeated the calculation for the tail middles (Figure 22c) and the tail ends (Figure 22d). The amount of order decreases as we explore packing of tails farther from the head group.

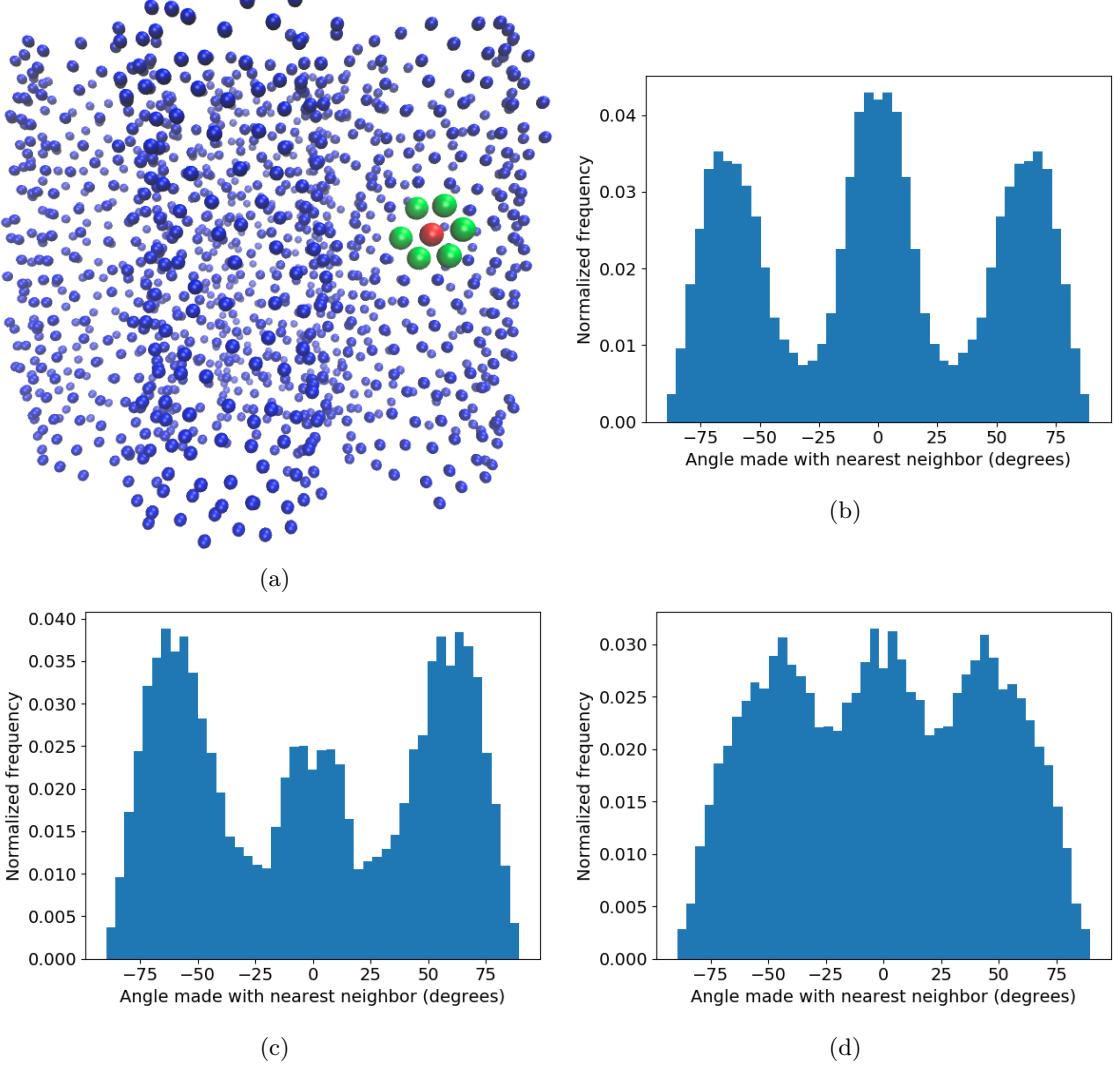


Figure 22: Monomer tails pack together hexagonally. (a) The center of mass of each tail-front (See Figure 21) is visualized as a blue sphere. The red sphere highlights an example of a tail-front center of mass with its nearest neighbors (green spheres) surrounding it in a hexagonal pattern. (b) We calculated the angle between each tail-front center of mass and its nearest neighbor tail-front center of masses and plotted the distribution. There are distinct peaks ca.  $-60^\circ$ ,  $0^\circ$  and  $60^\circ$  which is indicative of hexagonal packing. (c) We repeated the same procedure outlined in (a) and (b) with the tail-middles. There is still a high degree of order. (d) We repeated the procedure again with the tail-ends. There is far less order at the ends of the tails where there is more space to fill.

## 16 Carboxylate Dihedrals

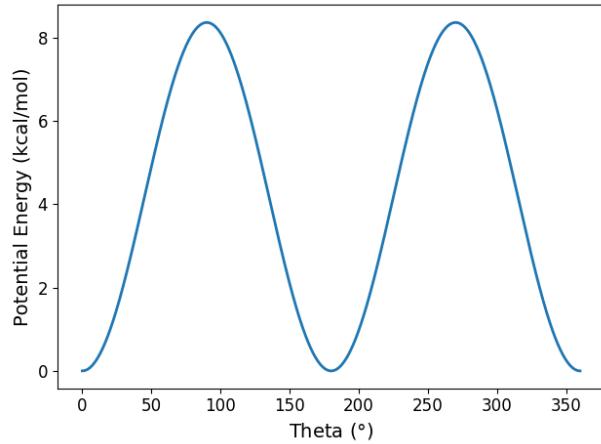


Figure 23: A potential energy diagram for the dihedral angle about the carbon-carbon bond of a carboxylate group attached to an aromatic ring, using parameters from the generalized amber forcefield. It is energetically favorable for the carboxylate group to stay in plane with the aromatic ring. There is a large energy penalty as the dihedral becomes antiparallel to the plane of the aromatic ring.

## 17 All solvated systems

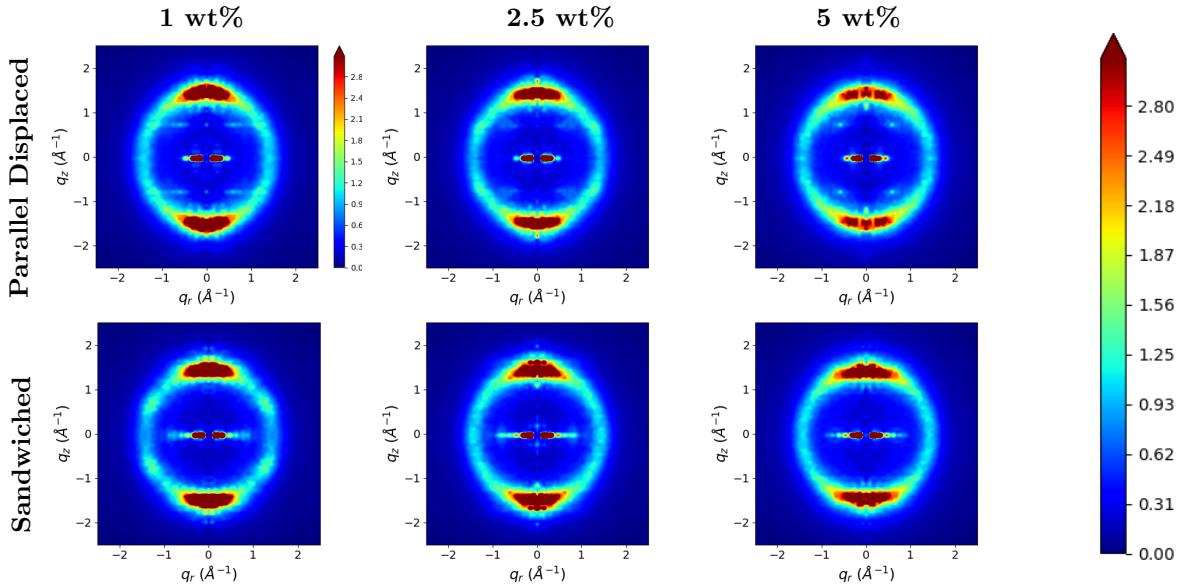


Figure 24: Adding 1 wt% water to the ordered basin dry systems causes an increase in the intensity R-spots. R-double becomes visible in the parallel displaced configuration. Additional amounts of water increase discrepancies with experiment

## 18 The role of water in the appearance of R-double

The appearance of R-double is likely a consequence of structure induced when vertically stacked monomer pairs share hydrogen bonds with a common water molecule (See Section 3.2.5 of the main text). We quantified the occurrence of these pairing interactions by binning the  $z$ -positions of the center of mass of the monomer head groups into 20 bins, the same number as the monomers per column. This works relatively well since there is a high degree of correlation between columns in our system. Peaks in the distributions represent the scenario where a monomer head group shares a water molecule with a head group vertically above it (See Figure 25a). We calculated the discrete Fourier transform of the distributions in Figure 25a which we used as a rough indicator of whether the head group arrangement will lead to R-double (Figure 25b). We see the strongest indicator of asymmetric monomer stacking in Pore 1 where the peaks at  $0.75 \text{ \AA}^{-1}$  are clearly distinguishable, meaning there is periodicity at twice the  $\pi$ -stacking distance.

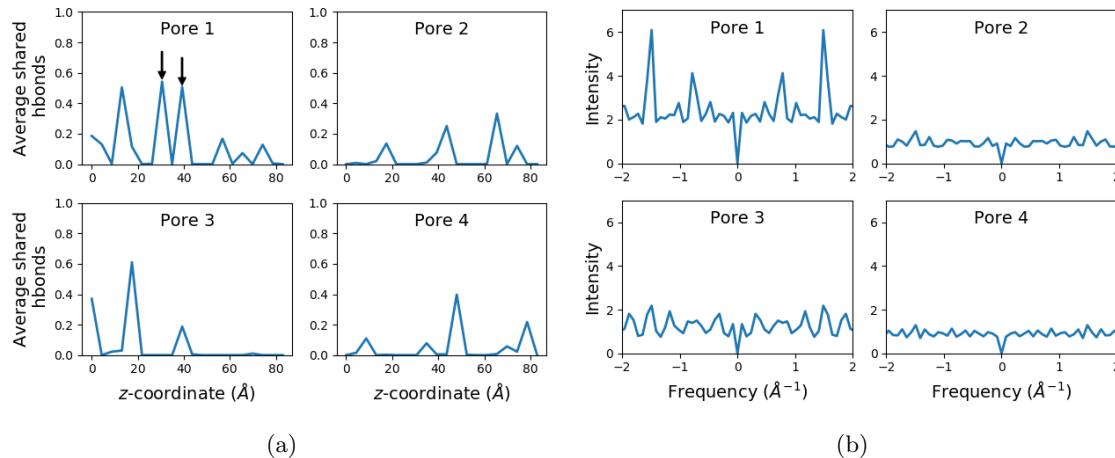


Figure 25: (a) We group the  $z$ -coordinate of the center of mass of monomer head groups into 20 bins. Peaks in the distributions indicate a shared hydrogen bonded water molecule between a monomer head group and one in the bin vertically above it. The distance between the peaks roughly corresponds to the distance between pairs of monomers. The distance between peaks in pore 1 (indicated by arrows) is equal to two times the  $\pi$ -stacking distance. In this scenario, monomers are no longer evenly spaced in the  $z$ -direction since they are pulled closer together by hydrogen bonds. (b) We performed discrete Fourier transforms on each distribution in (a). In Pore 1, there is periodicity every  $8.3 \text{ \AA}$ , twice the  $\pi$ -stacking distance, as indicated by the intensity at  $0.76 \text{ \AA}^{-1}$ . In this case, the center of mass of two adjacent pairs is separated by twice the average monomer stacking distance, while the four monomers involved are unequally spaced. These conditions seem likely to give rise to R-double.

## 19 Diffusion of monomers

We calculated the diffusion coefficient of the monomer head groups by measuring the slope of the linear region of their mean squared displacement curves accumulated over the length of the equilibration simulations (Table 3).

## 20 Autocorrelation of monomer dihedrals

In section 3.4 of the main text, we calculated the autocorrelation function of the outer tail either dihedrals pictured in Figure 26.

System	Diffusion Constant ( $m^2/s$ )
Sandwiched, Ordered	$1.4 \times 10^{-14}$
Parallel Displaced, Ordered	$2.0 \times 10^{-14}$
Sandwiched, Disordered	$2.0 \times 10^{-14}$
Parallel Displaced, Disordered	$3.5 \times 10^{-14}$

Table 3: The diffusion constants for all systems are on the order of  $10^{-14}m^2/s$ .

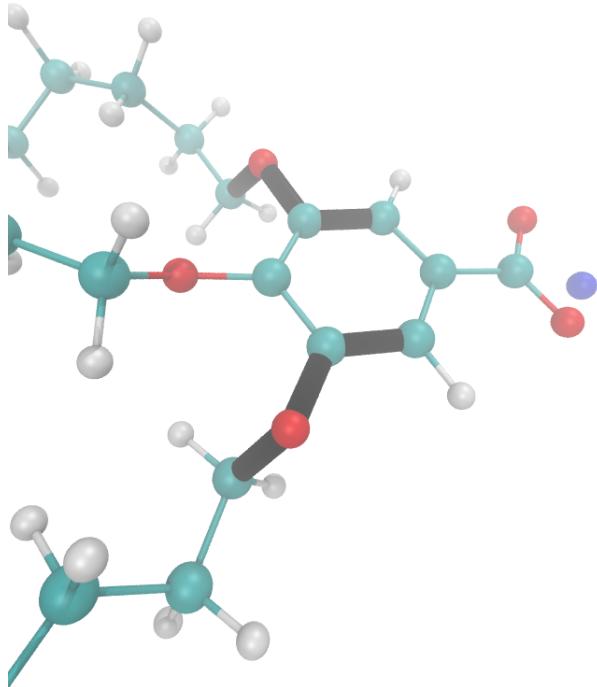


Figure 26: We measured the autocorrelation functions of the dihedrals highlighted in black. We only included dihedrals associated with the outer tails in our calculation since the position of the center tail leads to fundamentally different interactions and dynamics.

## References

- [1] ChemAxon, “MarvinSketch 17.13 2017 ChemAxon (<http://chemaxon.com>),” 2017.
- [2] N. M. O’Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, “Open Babel: An open chemical toolbox,” *Journal of Cheminformatics*, vol. 3, p. 33, Oct. 2011.
- [3] “The Open Babel Package, version 2.4.1 <http://openbabel.org>.”
- [4] J. Wang, W. Wang, P. A. Kollman, and D. A. Case, “Automatic atom type and bond type perception in molecular mechanical calculations,” *Journal of Molecular Graphics and Modelling*, vol. 25, pp. 247–260, Oct. 2006.
- [5] D. Case, R. Betz, W. Botello-Smith, D. Cerutti, T. Cheatham, III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, D. York, and P. Kollman, “AmberTools16,” Apr. 2016.
- [6] R. Walker and S. Tang, “Antechamber Tutorial ([ambermd.org/tutorials/basic/tutorial4b/](http://ambermd.org/tutorials/basic/tutorial4b/)),”
- [7] A. W. Sousa da Silva and W. F. Vranken, “ACPYPE - AnteChamber PYthon Parser interfacE,” *BMC Research Notes*, vol. 5, p. 367, July 2012.
- [8] L. Martnez, R. Andrade, E. G. Birgin, and J. M. Martnez, “PACKMOL: A package for building initial configurations for molecular dynamics simulations,” *J. Comput. Chem.*, vol. 30, pp. 2157–2164, Oct. 2009.
- [9] M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, “Estimates of the Ab Initio Limit for Interactions: The Benzene Dimer,” *J. Am. Chem. Soc.*, vol. 124, pp. 10887–10893, Sept. 2002.
- [10] J. D. Chodera, “A Simple Method for Automated Equilibration Detection in Molecular Simulations,” *Journal of Chemical Theory and Computation*, vol. 12, pp. 1799–1805, Jan. 2016.
- [11] M. R. Shirts and J. D. Chodera, “Statistically optimal analysis of samples from multiple equilibrium states,” *The Journal of Chemical Physics*, vol. 129, p. 124105, Sept. 2008.
- [12] Y. Liu and F. Zhu, “Collective diffusion model for ion conduction through microscopic channels,” *Biophys. J.*, vol. 104, pp. 368–376, Jan. 2013.
- [13] R. Young and P. Lovell, *Introduction to Polymers*. CRC Press, 3 ed., 2011.
- [14] P. Chaikin and T. Lubensky, *Principles of condensed matter physics*. Cambridge University Press, 1 ed., 1995.