

# Size, Shape and Functionality Dependence of Solute Transport Mechanisms in a Cross-linked H<sub>II</sub> Phase Lyotropic Liquid Crystal Membrane

Benjamin J. Coscia

Michael R. Shirts

June 26, 2019

## 1 Introduction

We need highly selective membranes in order to perform efficient separations.

H<sub>II</sub> phase lyotropic liquid crystals have densely packed, uniform sized pores and have the potential to disrupt conventional membrane separation techniques by being selective based not only on size and charge, but on chemical functionality as well.

We can only learn so much from experiment. MD can give us mechanistic insights with atomistic resolution so that we can intelligently design new membranes for solute-specific separations.

In previous work, we determined the most likely structure of the hexagonal phase formed by the monomer Na-GA3C11.

- We developed techniques for equilibrating the hexagonal phase made by neat monomer as well as with varying amounts of water in the pores.

In this work, we have determined the transport mechanisms and macroscopic transport properties exhibited by a number of polar solutes with varying size, chemical functionality and hydrophilic character.

- Many of the separations we are interested in involve polar organic compounds.

We have compared our calculated diffusion coefficients with experimental measurements made using DOSY NMR.

## 2 Methods

### 2.1 Molecular Dynamics Simulations

#### System Setup

There is a broad range of water concentrations which will form a stable H<sub>II</sub> phase with Na-GA3C11.

- In the literature this system is typically synthesized with close to 10 wt % water
- However, Resel et al. noted that the system is likely fully hydrated with less than 7 wt % water.
- We decided to test two different levels of water content: 5 and 10 wt %

We observed that water partitions into the tail region of our system and therefore built our initial configurations with water in both regions close to the expected equilibrium value.

- There is about 2:1 water in the pores versus in the tails for the 10 wt % system.
- The amount of water present in the tails may or may not be experimentally consistent but if we don't put it in, the results will not be thermodynamically consistent, which will give issues with measurements and calculations.

- See supporting info for water equilibration simulation data.
- We adjusted the pore radius in our systems so that the right amount of water fits in the pores without any vacuum using `gmx solvate`.
- We placed water molecules in the tail region one at a time in random locations with short energy minimizations between insertions.

We equilibrated an initial solvated configuration before adding solutes.

- We equilibrated the initial configuration using the ‘wet’ equilibration procedure described in our previous work [1].
- We cross-linked the equilibrated solvated configuration using the cross-linking procedure described in our previous work.

We added 6 solute molecules to each pore of the equilibrated cross-linked configuration.

- We equally spaced each solute in the pore
- 6 solutes per pore provided a balance of a useful amount of data for generating statistics and a low degree of interaction between solutes (reference to supporting information to show low degree of interaction)
- At each insertion point we placed a randomly oriented solute molecule then ran a short energy minimization.
- We allowed the solutes to equilibrate for 5 ns using berendsen pressure control
- We collected transport data using long simulations, on the order of 1 microsecond
- .mdp files are in the supporting information

## Modeling subdiffusion

Solutes in our H<sub>II</sub> LLC membrane exhibit subdiffusive behavior, a type of anomalous diffusion.

- During an anomalous diffusion process, the mean squared displacement (MSD) does not grow linearly with time, rather it is of the form:

$$\langle x^2(t) \rangle = K_\alpha t^\alpha \quad (1)$$

where  $\alpha$  is the anomalous exponent and  $K_\alpha$  is the generalized diffusion coefficient.

- A value of  $\alpha < 1$  indicates a subdiffusive process, while a value of  $\alpha = 1$  and  $\alpha > 0$  is characteristic of Brownian and superdiffusive motion respectively.

We calculated both the ensemble-averaged and time-averaged MSDs of the simulated trajectories.

- The ensemble-averaged MSD measures the displacement of a particle from its initial position [2] and can be written as

$$\langle x^2(t) \rangle = \langle x(t) - x(0) \rangle \quad (2)$$

- The time-averaged MSD measures the displacement between all possible time lags and can be written as

$$\overline{x^2(\tau)} = \frac{1}{T - \tau} \int_0^{T-\tau} (x(t + \tau) - x(t))^2 dt \quad (3)$$

where  $\tau$  is the time lag and  $T$  is the length of the trajectory [2].

Three common mathematical models for modeling anomalous subdiffusion processes include continuous time random walks (CTRW), fractional Brownian motion (FBM) and random walks on fractals (RWF).[2]

- FBM is common in crowded, viscoelastic environments where each step comes from a Gaussian distribution but is anti-correlated to its previous step. [3, 4, 5]

- A CTRW is characterized by a distribution of hop lengths and dwell times, where each step is characterized by independent random draws from each distribution.[6, 7]
- An RWF is imposed by a system's geometry. Systems with tortuous pathways and dead ends cause anti-correlated motion.[2, 8]
- The processes described above can happen alone or in combination.

We believe that solutes in the system studied here exhibit subordinated fractional Brownian motion (sFBM) where the parent process is FBM and the leading process is a CTRW.

- The ensemble-averaged MSD differs from the time-averaged MSD, which is indicative of non-ergodicity, a trait inherent to CTRWs but not FBM or RWFs. [9]
- We also observe non-stationary  $z$ -coordinate traces of each solute's center of mass (COM).
- For a pure CTRW, the time-averaged MSD should be linear. [8, 10]
- However, a typical time-averaged solute MSD is sublinear (see supporting information), which suggests that there is another underlying subdiffusive mechanism.
- The hop lengths recorded after each dwell period are anti-correlated (See supporting information)
- Given the viscoelastic nature of the monomers in our system, we believe the hop lengths can be modeled with FBM.
- For subordinated FBM, it can be shown that

$$\langle x^2(t) \rangle \simeq t^{\alpha\beta} \quad (4)$$

where  $\alpha$  is the anomalous exponent characteristic of the leading CTRW process and  $\beta$  is the anomalous exponent characteristic of the parent FBM process.

We can characterize a CTRW process using the parameters which describe its dwell time and hop length distribution.

- We used the `ruptures` python package in order to identify breakpoints in solute trajectories.[11] (See Supporting Information for more details on chosen parameters. i.e. type of cost function, cost function penalty tolerance, number of dimensions used)
- The corresponding hop lengths and dwell times between break points were used to construct empirical distributions.
- For solutes in our system, the distribution of hop lengths is well-represented by a Gaussian distribution. [12, 13, 14]
- We are most interested in the standard deviation,  $\sigma$ , of the hop length distribution.
- The distribution of dwell times is expected to fit a power law (or heavy-tailed) distribution proportional to  $t^{-1-\alpha}$ . [2]
- Because we are limited to taking measurements at discrete values dictated by the output frequency of our simulation trajectories, we fit the empirical dwell times to a discrete power law distribution whose maximum likelihood  $\alpha$  parameter we calculated by maximizing the following likelihood function:

$$\mathcal{L}(\beta) = -n \ln \zeta(\beta, x_{min}) - \beta \sum_{i=1}^n \ln x_i \quad (5)$$

where  $\beta = 1 + \alpha$ ,  $x_i$  are collected dwell time data points,  $n$  the total number of data points, and  $\zeta$  is the Hurwitz zeta function where  $x_{min}$  is the smallest measured value of  $x_i$ . [15]

- We obtained distributions of the hop length standard deviations,  $\sigma$ , and  $\alpha$  using statistical bootstrapping.[16]

FBM processes can be described using the Hurst parameter,  $H$ , where  $H = \beta/2$ .

- Brownian motion is recovered for  $H = 0.5$
- The autocovariance function of hop lengths has the analytical form: [3]

$$\gamma(k) = \frac{1}{2} \left[ |k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H} \right] \quad (6)$$

where  $k$  is the number of increments between hops.

- We obtained  $H$  by performing a least squares fit of Equation 6 to the empirically measured autocovariance function.
- We used statistical bootstrapping to generate a distribution of  $H$  values.

For each solute, we simulated 10000 sFBM trajectories for 1  $\mu$ s each.

- We constructed trajectories by simulating sequences of dwell times and correlated hop lengths generated based on parameters randomly chosen from our bootstrapped parameter distributions.
- We propagated each trajectory until the total time reached 1  $\mu$ s, and truncated the last data point so that the total time exactly equaled 1  $\mu$ s.
- Valid comparisons are only possible between fixed length sFBM simulations. The power law dwell time behavior gives rise to the aging phenomenon, embodied by a decrease in MSD with measurement time. [8, 13]
- We reported the MSD after 1  $\mu$ s with corresponding 95 % intervals

### Radial Distribution Functions

We measured the average radial distance of each solute of interest from the pore centers.

- We binned the radial distances and then normalized by the volume of the annulus defined by the bin edges.
- Although the pores are often described as straight, they have a small degree of tortuosity which disrupts the RDF calculation
- We obtain the best RDF by constructing splines that run through the pore centers.
- We construct the splines by dividing the membrane into 20 slices in the  $z$ -direction. Within each slice, we calculate the location of the pore centers based on the average location of the aromatic rings that make up the monomer head groups.
- When calculating the RDF, the radial distance from the pore center is based on the distance between the solute center-of-mass and the  $(x, y)$  coordinates of appropriate point on the spline.

### Coordination number

We quantified the coordination of solutes with surrounding molecules.

- For each frame, we counted the identities and number of coordinated molecules to a given solute based on a distance cut-off.
- We found that this approach is more useful than calculating the 3D spherical radial distribution function because it gives detailed frame-by-frame information rather than an average.