Time Series Modeling of Solute Transport in an H_{II} Phase Lyotropic Liquid Crystal Membrane

Benjamin J. Coscia Michael R. Shirts

September 4, 2019

1 Introduction

Mathematical descriptions of transport in complex separations membranes are a powerful way to understand mechanisms and formulate membrane design principles.

- The complexity of a model generally parallels the complexity of the transport mechanism at hand, as well as the desired outputs.
- In dense homogeneous membranes, the solution-diffusion model can extract diffusion and partition coefficients and has successfully predicted solute transport rates.
- Analogously, pore-flow models yield predictions of diffusion coefficients and solute transport rates in nanoporous membranes. [**?**]
- Single particle tracking allows us to further characterize more complex diffusive behavior.
- Finally, Molecular simulation provides a framework for studying both single particle dynamics and bulk transport properties with atomic resolution.
- All of these approaches facilitate speculation on the molecular origins of separations and suggest experiments that should be run by attempting to give a more intuitive understanding of how solutes move as a function of their environment.

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

Amphiphilic molecules are capable of self-assembling into ordered nanostructures.

Lyotropic liquid crystals are a class of amphiphilic molecules that can be cross-linked into mechanically strong and highly selective membranes.

- H_{II} 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.
- Q_I phase LLCs consist of a tortuous network of 3D interconnected pores. They are easier to make.

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 our previous work, we studied the transport of 20 small polar molecules in an H_{II} phase LLC membrane.

- In general, we observed subdiffusive transport behavior characterized by intermittent hops separated by periods of entrapment.

- We identified three mechanisms responsible for the solute trapping behavior: entanglement among monomer tails, hydrogen bonding with monomer head groups, and association with sodium counter ions.

Unfortunately, the timescales that we can simulate with MD are insufficient to be able to make well-converged predictions of macroscopic transport properties traditionally used to characterize membranes in the lab.

- However, if we use descriptive stochastic models that sufficiently capture solute dynamics, then we could project long timescale behavior in addition to gaining a deeper understanding of solute behavior on short timescales.

Researchers have developed a rigorous theoretical foundation for describing the motion of particles exhibiting non-Brownian, or anomalous, transport behavior.

- The tools introduced by fractional calculus are instrumental to this theory.
- They allow us to generalize the normally linear diffusion equation to fractional derivative orders, providing descriptions of a much more diverse set of behavior. [15]

Three well-known classes of behavior leading to anomalous subdiffusion are continuous time random walks (CTRW), fractional Brownian motion (FBM) and random walks on fractals (RWF).[1]

- The theory is frequently used to describe single particle trajectories
- A CTRW is characterized by a distribution of hop lengths and dwell times, where each each hop is characterized by independent random draws from each distribution.[2, 3]
- FBM is common in crowded, viscoelastic environments where each jump comes from a Gaussian distribution but is anti-correlated to its previous step. [4, 5, 6]
- An RWF is imposed by a system's geometry. Systems with tortuous pathways and dead ends cause anti-correlated motion.[1, 7]
- The processes described above can happen alone or in combination.
- Meroz and Sokolov suggested a decision tree for determining which subdiffusive mechanisms apply to a system, adapted in Figure X.
- Of course, more complicated variations on these models can arise which we will incorporate as needed.

Our first modeling approach is based on the anomalous diffusion literature, applying the mathematical formalism describing subdiffusive transport to observed solute trajectories.

- We determine the most likely subdiffusive mechanisms at play using the flow chart in figure X.
- We then fit anomalous diffusion models in two ways: First, based on a single set of parameters. Second, we use a two-state approach incorporating two sets of parameters based on a solute's radial distance from the pore center.

While elegant, pure anomalous diffusion theory may in some cases give only a qualitative understanding of transport mechanisms.

- Complex temporal relationships in between hops may have a significant effect on the average MSD
- Inhomogeneity of a medium or multiple types of interactions may lead to a number of distinct diffusive regimes.
- In this case, a state based model may be useful.

We can use the observations of transport from our previous work in order to create state-based models.

- We know that dynamics differ between solutes moving inside the pores and in the tails.
- We can use the anomalous diffusion approach to model our data as a function of radial distance from the closest pore center.
- Alternatively, we can abandon the anomalous diffusion theory and adapt the methods of Markov state modeling.

Markov state models (MSMs) are a popular class of models used to project long timescale behavior of molecules based on molecular simulation trajectories by identifying different dynamical modes and quantifying the rates of transitions between them.

- MSMs are frequently used to study systems with slow dynamics, such as protein folding. [8, 9]
- Researchers typically aim to come up with a low dimensional representation of the system based on features which preserve the process kinetics.
- This still often results in thousands of distinct states.

Our second modeling approach adapts the MSM framework to a relatively small set of known states based on the three observed solute trapping mechanisms.

- Standard MSMs assume first order kinetics and are typically applied towards determining equilibrium populations of various state configurations.
- We extend the framework to include dynamics within states and between jumps.
- That is we incorporate the magnitude and temporal dependence of the time series fluctuations between each transition to a different state.
- Temporal dependence has a significant influence on mean squared displacement.

In this work we attempt to model the long timescale behavior of the four fastest moving solutes studied in our previous work.

- Specifically, we study methanol, urea, ethylene glycol and acetic acid
- In addition to moving quickly, these solutes have a range of chemical functionality and experience each of the three trapping mechanisms to different degrees.
- It is our goal to reproduce the MSDs exhibited by each solute on MD timescales and then to project the long-term dynamics of each solute.
- We explore two modeling approaches

2 Methods

We ran all MD simulations and energy minimizations using GROMACS 2018. We performed all post-simulation trajectory using python scripts which are available online at
`https://github.com/shirtsgroup/LLC_Membranes`.

2.1 Molecular Dynamics Simulations

We studied transport of solutes in the H_{II} phase using an atomistic molecular model of four pores in a monoclinic unit cell with 10 % water by weight. (See Figure tbd)

- Approximately one third of the water molecules occupy the tail region with the rest near the pore center.
- We chose to study the 10 wt % water system because solutes move significantly faster than in the 5 wt % system studied previously.

- Appropriate stochastic modeling requires that solutes sample the accessible mechanisms with representative probability.

We chose to study a subset of 4 of the fastest moving solutes from our previous work: methanol, acetic acid, urea and ethylene glycol.

- In addition to exploring membrane structural space the most, these solutes have a relatively diverse set of chemical functionality.
- For each solute we created a separate system and to each system we added 6 solutes per pore for a total of 24 solutes.
- This number of solutes per pore provides a balance of a low degree of interaction between solutes and sufficient amount of data from which to generate statistics on the time scales which we simulate.
- Further details on the setup and equilibration of these systems can be found in our previous work.[10]

We extended the 1 μs simulations of our previous work to 5 μs in order to collect ample data.

- We simulated the system with a time step of 2 fs at a pressure of 1 bar and 300 K controlled by the Parinello-Rahman barostat and the v-rescale thermostat respectively.
- We recorded frames every 0.5 ns

2.2 The Anomalous Diffusion Model

Solutes in this system 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 follows a power law of the form:

$$\langle x^2(t)\rangle = K_\alpha t^\alpha \tag{1}$$

  where α is the anomalous exponent and K_α is the generalized diffusion coefficient.

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

We analyzed 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 [1] and can be written as

$$\langle x^2(t)\rangle = \langle x(t) - x(0)\rangle^2 \tag{2}$$

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

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

  where τ is the time lag and T is the length of the trajectory [1].

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.

- Working down the flow chart...
- We observe non-stationary z-coordinate traces of each solute's center of mass (COM).
- Combined with the observation that the ensemble-averaged MSD differs from the time-averaged MSD, we have evidence of non-ergodicity in this system, a trait inherent to CTRWs but not FBM or RWFs. [18]

- The process is likely subordinated to another process because the hop lengths recorded after each dwell period are anti-correlated.
- 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

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

  where α is the anomalous exponent characteristic of the leading CTRW process and β 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 automatically identify changepoints in solute trajectories.[20] (See Supporting Information for more details on chosen parameters. i.e. type of cost function, cost function penalty tolerance, number of dimensions used)
- We used the corresponding hop lengths and dwell times between break points to construct empirical distributions.

For solutes in our system, the distribution of hop lengths appears to be well-represented by a Gaussian distribution. [21, 22, 23]

- We are most interested in the standard deviation, σ, of the hop length distribution.

The distribution of dwell times is expected to fit a power law distribution proportional to $t^{-1-\alpha}$. [1]

- 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 α parameter we calculated by maximizing the following likelihood function:

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

  where $\beta = 1+\alpha$, x_i are collected dwell time data points, n the total number of data points, and ζ is the Hurwitz zeta function where x_{min} is the smallest measured value of x_i. [24]
- We obtained distributions of the hop length standard deviations, σ, and α using statistical bootstrapping.[25]

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: [4]

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

  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.

In general, we observe different dynamical behavior when solutes move inside the pore versus in the tail region.

- This inspired two models of varying complexity.
- We created a simple, single mode model with a single α, σ and H parameter fit to each solute.
- Our second, two mode model assigns a set of parameters to each of 2 modes based on the solute's radial location.
- Solutes in mode 1 are in the pore region defined as less than 0.75 nm from any pore center and all else are in mode 2, the tail region.
- We determined this cut-off and described how we calculated radial distance from the pore center in our previous work [10]

For the two mode model, we needed to define a probability transition matrix describing transitions between the tail and pore region.
For each solute, we simulated 10000 5 μs sFBM trajectories.

- 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 equaled or exceeded 5 μs, then truncated the last data point so that the total time exactly equaled 5 μs.
- 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. [7, 22]
- We reported the MSD after 5 μs with corresponding 95 % intervals

2.3 Markov State Models

A Markov state model (MSM) decomposes a time series into a set of discrete states with transitions between states defined by a transition probability matrix, T, which describes the conditional probability of jumping to a specific state given the previously observed state. [11, 12]

- Software packages such as MSMbuilder [13] and pyEMMA [14] provide work flows capable of featurization and dimensional reduction.

In this work, we define a total of 8 discrete states based on the 3 trapping mechanisms observed in our previous work.

- Therefore, there is no need to subject our trajectories to any reduction in dimensionality.
- However, we will adopt the validation techniques used in the packages mentioned above.
- The states we've chosen include all combinations of trapping mechanisms in the pore and out of the pore:

1. In tails, not trapped	5. In pores, not trapped
2. In tails and hydrogen bonding	6. In pores and hydrogen bonding
3. In tails and associated with sodium	7. In pores and associated with sodium
4. In tails, hydrogen bonding and associated	8. In pores, hydrogen bonding and associated

Table 1

- These choices assume that there are no significant kinetic effects resulting from solute conformational changes or pore size fluctuations, an assumption that may be relaxed in future work.

We constructed a state transition probability matrix based on observed solute trajectories.

- Using methods described in our previous work, we determined which, if any, trapping mechanisms affected the solute at each time step, and assigned the observation to a specific state according to the definitions above. [10]
- Based on the current and previous state observation, we incremented the appropriate entry of an $n \times n$ count matrix by 1, where n is the number of states.
- For example, if we observed state 1 followed by a transition to state 3, we increment the $(1, 3)$ entry of the count matrix by 1.
- We generate the transition probability matrix from the count matrix by normalizing the entries in each row so that they summed to unity.

We recorded the z-direction displacement at each time step in order to construct individual emission distributions for each state and transition between states.

- This is where we have modified the standard MSM framework.
- While the dynamics of the states themselves are important, the majority of observations involve transitions between states, so properly modeling the transition dynamics is even more important.
- This results in 64 distinct emission distributions. Some are far more populated than others.

We modeled the emission distributions as Lévy stable distributions.

- For independent and identically distributed (iid) random variables, the generalized central limit theorem guarantees convergence of the associated probability distribution function (PDF) to a Lévy stable PDF. [15]
- The characteristic equation describing the Fourier transform of a Lévy stable PDF is given below:

$$p_{\alpha,\beta}(k;\mu,\sigma) = \exp\left[i\mu k - \sigma^{\alpha}|k|^{\alpha}\left(1 - i\beta\frac{k}{|k|}\omega(k,\alpha)\right)\right] \tag{7}$$

- Where ω is defined with something that is complicated to format and can be added later if I keep these formulas in.
- α is the index of stability or Lévy index, β is the skewness parameter, μ is the shift parameter and σ is a scale parameter.
- The most familiar case, and 1 of 3 that can be expressed in terms of elementary functions, is the Gaussian PDF ($\alpha = 2$).
- The more general family of Lévy stable distributions allows greater flexibility in defining the observed emission distribution PDFs.
- In the case that the empirical emission distributions are iid, sequential draws from a Lévy stable PDF defines a Lévy process.
- We can relax the iid assumption and maintain the flexibility of Lévy stable probability distributions if we instead consider the dynamics to be governed by a non-Gaussian Ornstein-Uhlenbeck (OU) process.[16]
- OU processes contain temporal dependence but the marginal observation distributions are consistent with the underlying stochastic driver. [17]
- We will discuss the implications of each.

Lévy stable distributions often have heavy tails and an undefined variance which can give rise to arbitrarily long particle displacements.

- We do not observe hops longer than x nm in our simulations which emphasizes that these distributions are only an approximation.
- We avoid enormous and unrealistic hops problem by truncating the tails of the distribution.
- Since solutes in our system cannot hop further than the length of our simulation unit cell, we truncate the tails of the distribution at this magnitude.

We simulated realizations of the stochastic process using the probability transition matrix and emission distributions.

- For each trajectory simulated, we chose an initial state randomly from a uniform distribution.
- We randomly drew subsequent state transitions and corresponding emissions from the rows of the probability transition matrix and the appropriate emission distribution respectively.

3 Results and Discussion

3.1 Molecular Dynamics Simulations

The MSDs for of the solutes calculated from 5 μs MD simulation trajectories are shown in Figure TBD.

- 2 options for a figure:
- 1) 2 panels: 1st panel: bar chart with final MSD value + error bars. 2nd panel: a selected MSD curve
- 2) 1 panel. All MSD curves. Overlapping errors might look bad

Methanol and acetic acid are not fully equilibrated, therefore we did not use them to test our models.

3.2 Subordinated Fractional Brownian Motion Modeling

Using the techniques from Section 2.2, we extracted values of σ, α, H and λ for each solute for the 1 and 2 mode models (see Table 2).

- In most cases, it is easy to relate the values of σ, α and H presented in Table 2 to the simulated MSD values.
- Higher values of σ indicate larger average hop lengths.
- Higher values of α mean that there will be less sampling of long dwell times.
- Higher values of λ truncates the power law distribution earlier preventing extremely long dwell times.
- Values of H near the Brownian limit of 0.5, indicate a lower degree of anti-correlation.
- All of which contribute to an overall increase in the MSD.

System	σ (nm)	α	H
Methanol	0.46	0.85	0.40
Urea	0.33	0.64	0.40
Ethylene Glycol	0.35	0.64	0.36
Acetic Acid	0.28	0.51	0.44

Table 2: We calculated values σ, α and H from MD simulation trajectories and then computed the average ensemble-averaged MSD of 10000 simulated trajectories.

We compared the MSD of simulated sFBM and sFLM trajectories to MD simulated MSDs.

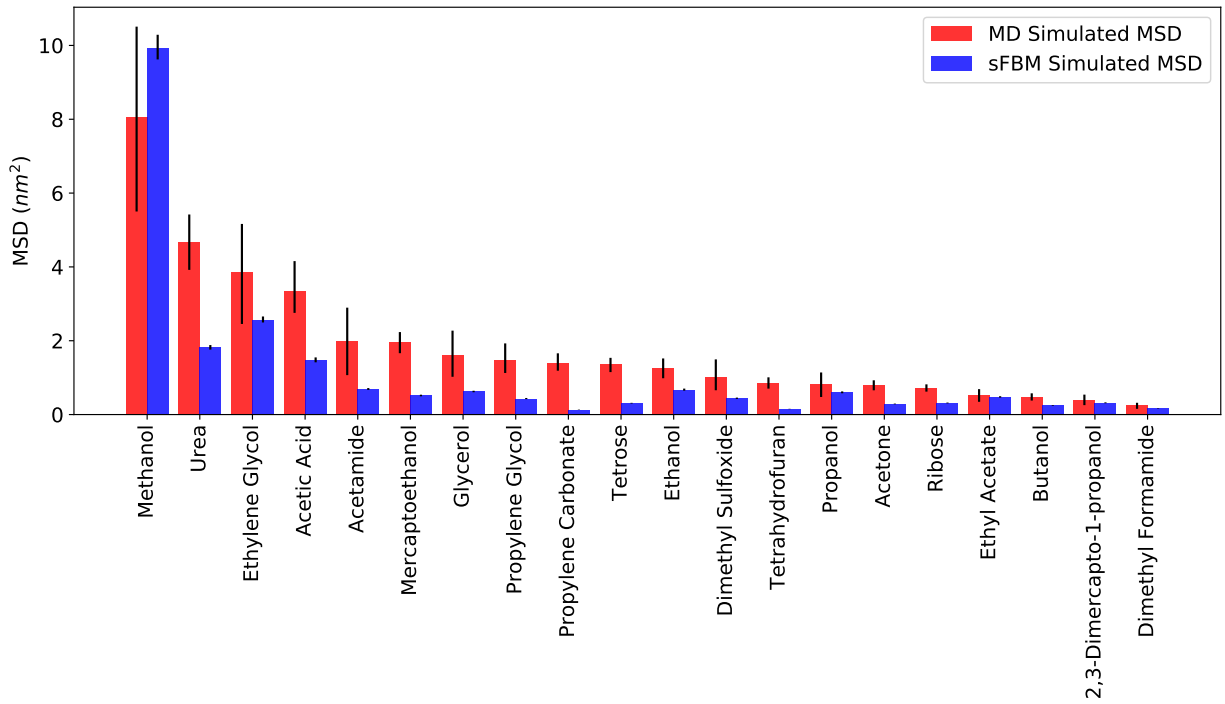

Figure 1

- We simulated 10000 sFBM trajectories of the same length as our MD simulations, as described in Section 2.2 of the Methods.
- The final MSDs of the sFBM trajectories are compared to those calculated directly from MD simulations in Figure 1.
- We would like to emphasize that we rely on the MD MSD values in order to define trends in the total MSD, while the sFBM trajectories and parameter values allow us to speculate as to the reasons for the observed trends.

The power law with exponential cut-off has the largest effect on the on the match between MD MSD and sFLM trajectories.

- Using the power law with cutoff produces predictions in relatively good agreement with MD.

The choice between Levy and Gaussian distributions has only a small effect, but the Levy appears to fit emission distributions better.
We can use the model to project our simulations out to very long timescales.

- We can not definitively say which combination of distributions are best to use.
- Although the longest observed dwell times and hop lengths are less than the length of the simulated trajectories and box lengths respectively, it is still possible that that extremely long dwell times and hops are observed experimentally in rare but significant cases.
- We present long timescale projections using combinations of each distribution in the table below

3.3 Markov State Model Fits

We constructed emission distributions which correspond to each entry in the transition matrix.

- Some states are sampled far more frequently than other leading to better converged distributions.

The emission distributions are non-Gaussian and heavy-tailed (see Figure 2).

- Heavily sampled states fit well to Lévy stable distributions.
- The rarer states have less data, but we still assume Lévy stable distributions.

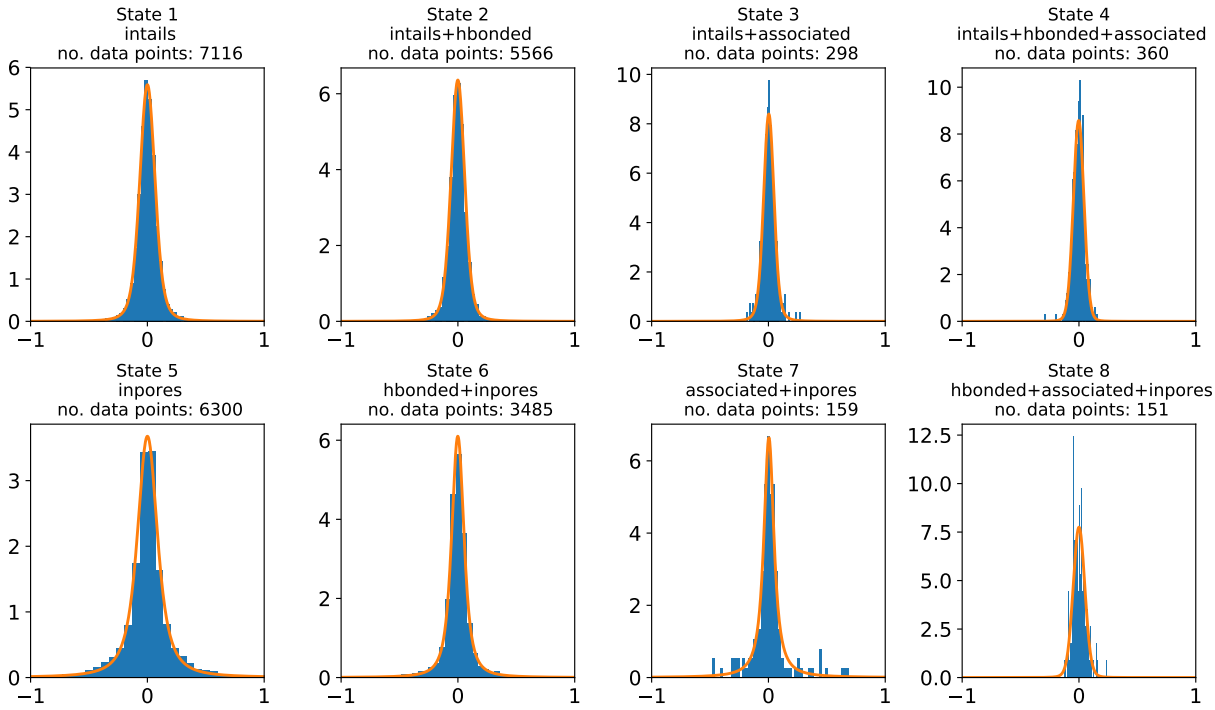

Figure 2

Motion within states is anticorrelated.

- We calculated the self-similarity parameter for motion within each state.

The emissions resulting from transitions between states are anticorrelated.

- We simulated these emissions by combining all transition emissions into a single distribution with a single self-similarity parameter.
- We are unaware of techniques for simulating correlated draws from Lévy distributions with varying shapes.

We projected the long timescale behavior using this model.

4 Conclusion

We have tested two different mathematical frameworks for describing solute motion in an H_{II} phase LLC membrane.

- Subordinated fractional Brownian motion has a nice theoretical foundation in the anomalous diffusion literature. A two mode model that describes dynamics based on whether a solute is in or out of the pore region leads to MSDs fairly consistent with MD simulated trajectories.
- Markov state modeling with predefined states gives a nice description of transitions between observed states as well as the type of stochastic behavior shown in each state. However, it doesn't accurately portray correlated time series behavior leading to overpredicted MSDs.

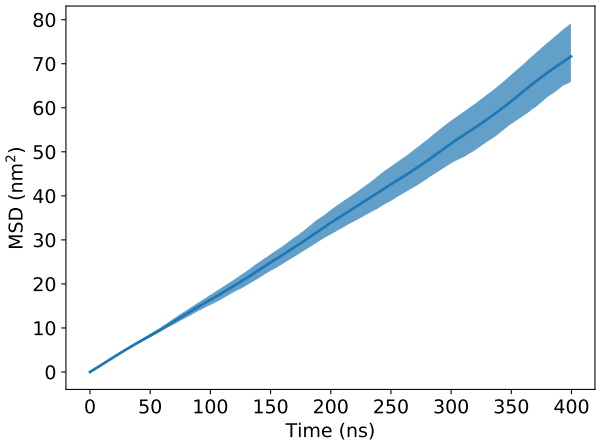

Figure 3

Supporting Information

Detailed explanations and expansions upon the results and procedures mentioned in the main text are described in the Supporting Information. This information is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgements

Molecular simulations were performed using the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. Specifically, it used the Bridges system, which is supported by NSF award number ACI-1445606, at the Pittsburgh Supercomputing Center (PSC). This work also utilized the RMACC Summit supercomputer, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. The Summit supercomputer is a joint effort of the University of Colorado Boulder and Colorado State University.

References

[1] Y. Meroz and I. M. Sokolov, "A Toolbox for Determining Subdiffusive Mechanisms," *Phys. Rep.*, vol. 573, pp. 1–29, Apr. 2015.

[2] E. W. Montroll and G. H. Weiss, "Random Walks on Lattices. II," *Journal of Mathematical Physics*, vol. 6, pp. 167–181, Feb. 1965.

[3] G. T. Morrin and D. K. Schwartz, "Three Regimes of Polymer Surface Dynamics under Crowded Conditions," *Macromolecules*, vol. 51, pp. 1207–1214, Feb. 2018.

[4] B. Mandelbrot and J. Van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Rev.*, vol. 10, pp. 422–437, Oct. 1968.

[5] J.-H. Jeon and R. Metzler, "Fractional Brownian Motion and Motion Governed by the Fractional Langevin Equation in Confined Geometries," *Phys. Rev. E*, vol. 81, p. 021103, Feb. 2010.

[6] D. S. Banks and C. Fradin, "Anomalous Diffusion of Proteins Due to Molecular Crowding," *Biophys. J.*, vol. 89, pp. 2960–2971, Nov. 2005.

[7] T. Neusius, I. Daidone, I. M. Sokolov, and J. C. Smith, "Subdiffusion in Peptides Originates from the Fractal-Like Structure of Configuration Space," *Phys. Rev. Lett.*, vol. 100, p. 188103, May 2008.

[8] C. D. Snow, E. J. Sorin, Y. M. Rhee, and V. S. Pande, "How Well Can Simulation Predict Protein Folding Kinetics and Thermodynamics?," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 34, pp. 43–69, May 2005.

[9] J. D. Chodera, N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope, "Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics," *J. Chem. Phys.*, vol. 126, p. 155101, Apr. 2007.

[10] B. J. Coscia and M. R. Shirts, "Chemically Selective Transport in a Cross-Linked HII Phase Lyotropic Liquid Crystal Membrane," *J. Phys. Chem. B*, June 2019.

[11] V. S. Pande, K. Beauchamp, and G. R. Bowman, "Everything you wanted to know about Markov State Models but were afraid to ask," *Methods*, vol. 52, pp. 99–105, Sept. 2010.

[12] C. Wehmeyer, M. K. Scherer, T. Hempel, B. E. Husic, S. Olsson, and F. No, "Introduction to Markov state modeling with the PyEMMA software [Article v1.0]," *Living Journal of Computational Molecular Science*, vol. 1, no. 1, pp. 5965–, 2018.

[13] K. A. Beauchamp, G. R. Bowman, T. J. Lane, L. Maibaum, I. S. Haque, and V. S. Pande, "MSM-Builder2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale," *J. Chem. Theory Comput.*, vol. 7, pp. 3412–3419, Oct. 2011.

[14] M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Prez-Hernndez, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, and F. No, "PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models," *J. Chem. Theory Comput.*, vol. 11, pp. 5525–5542, Nov. 2015.

[15] R. Klages, G. Radons, and I. M. Sokolov, *Anomalous Transport: Foundations and Applications*. John Wiley & Sons, Sept. 2008. Google-Books-ID: TDch5DfNgWoC.

[16] O. E. BarndorffNielsen and N. Shephard, "Non-Gaussian OrnsteinUhlenbeck-based models and some of their uses in financial economics," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 167–241, 2001.

[17] E. Taufer and N. Leonenko, "Simulation of Lvy-driven OrnsteinUhlenbeck processes with given marginal distribution," *Computational Statistics & Data Analysis*, vol. 53, pp. 2427–2437, Apr. 2009.

[18] F. Thiel and I. M. Sokolov, "Weak Ergodicity Breaking in an Anomalous Diffusion Process of Mixed Origins," *Phys. Rev. E*, vol. 89, p. 012136, Jan. 2014.

[19] Y. Meroz, I. M. Sokolov, and J. Klafter, "Subdiffusion of Mixed Origins: When Ergodicity and Nonergodicity Coexist," *Phys. Rev. E*, vol. 81, p. 010101, Jan. 2010.

[20] C. Truong, L. Oudre, and N. Vayatis, "Ruptures: Change Point Detection in Python," *arXiv*, vol. arXiv:1801.00826, Jan. 2018.

[21] R. Metzler and J. Klafter, "The Random Walk's Guide to Anomalous Diffusion: A Fractional Dynamics Approach," *Phys. Rep.*, vol. 339, pp. 1–77, Dec. 2000.

[22] R. Metzler, J.-H. Jeon, A. G. Cherstvy, and E. Barkai, "Anomalous Diffusion Models and Their Properties: Non-Stationarity, Non-Ergodicity, and Ageing at the Centenary of Single Particle Tracking," *Phys. Chem. Chem. Phys.*, vol. 16, pp. 24128–24164, Oct. 2014.

[23] T. Neusius, I. M. Sokolov, and J. C. Smith, "Subdiffusion in Time-Averaged, Confined Random Walks," *Phys. Rev. E*, vol. 80, p. 011109, July 2009.

[24] A. Clauset, C. Shalizi, and M. Newman, "Power-Law Distributions in Empirical Data," *SIAM Rev.*, vol. 51, pp. 661–703, Nov. 2009.

[25] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. CRC Press: Boca Raton, May 1994.

[26] M. J. Beal, Z. Ghahramani, and C. E. Rasmussen, "The infinite hidden Markov model," in *Advances in Neural Information Processing Systems 14: Proceedings of the 2001 Neural Information Processing Systems (NIPS) Conference*, vol. 1, (Cambridge, MA, US), pp. 577–585, MIT Press, Sept. 2002.

[27] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "The Sticky HDP-HMM: Bayesian Nonparametric Hidden Markov Models with Persistent States," *MIT Laboratory for Information and Decision Systems*, p. 60, Nov. 2007.

[28] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet Processes," *Journal of the American Statistical Association*, vol. 101, pp. 1566–1581, Dec. 2006.

[29] J. Van Gael, Y. Saatci, Y. W. Teh, and Z. Ghahramani, "Beam sampling for the infinite hidden Markov model," in *Proceedings of the 25th international conference on Machine learning - ICML '08*, (Helsinki, Finland), pp. 1088–1095, ACM Press, 2008.

[30] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "Nonparametric Bayesian Learning of Switching Linear Dynamical Systems," in *Advances in Neural Information Processing Systems 21* (D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, eds.), pp. 457–464, Curran Associates, Inc., 2009.

[31] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "Bayesian Nonparametric Methods for Learning Markov Switching Processes," *IEEE Signal Processing Magazine*, vol. 27, pp. 43–54, Nov. 2010.

TOC Graphic