

Machine Learning Assisted Phase Transition Temperatures From Generalized Replica Exchange Simulations of Dry Martini Lipid Bilayers

Zeke A. Piskulich and Qiang Cui*

Department of Chemistry, Boston University, Boston, MA, 02215, USA

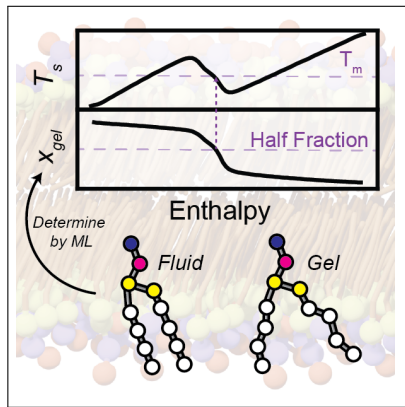
E-mail: qiangcui@bu.edu

June 29, 2022

TOC Graphic

Abstract

Accurate estimation of phase transition temperatures has been a longstanding challenge for molecular simulations. Recently, the generalized Replica Exchange technique for estimating phase transition temperatures has allowed for improved sampling of the phase transition; however, it requires a significant number of simultaneous replicas both inside and outside of the transition region leading to costly computational expense. In this work, the recently developed machine learning-assisted lipid phase analysis technique for learning the phase of individual lipids has been combined with generalized Replica Exchange Molecular Dynamics to reduce the overall computational expense of evaluating transition temperatures. This technique is then applied to eight different Dry Martini lipids to demonstrate its ability to describe transition temperatures as a function of chain length and tail saturation.



Upon cooling, lipids undergo a first order phase transition from the liquid-disordered fluid phase (L_α) to the liquid-ordered gel phase (L_β).¹⁻⁴ Across this transition, there are significant effects on local membrane structure, including straightening of the lipid tails, tighter lipid packing (and thus a reduced area per lipid), and increased bilayer thickness. As most lipids in biological membranes are found in the fluid phase, such transitions have recently been of interest due to occurrence of gel-phase lipid domains that may modulate the behavior of membrane proteins.⁵⁻¹⁰

As with any first order phase transition, the lipid fluid-gel transition is typically characterized by its transition temperature (T_m) and latent heat ($L_{f \rightarrow g}$). While these quantities have been well-characterized experimentally for a wide-range of lipids using Differential Scanning Calorimetry (DSC),^{11,12} results from molecular dynamics (MD) simulations have been inconsistent. Aside from known discrepancies inherent to the chosen description of the molecular interactions (*e.g.*, lipid forcefield, atomistic *vs.* coarse-grained), simulations using different techniques find different values of T_m while using the same lipid and forcefield. For instance, simulations with the Martini 2 coarse-grained forcefield on one of the most commonly simulated lipids, 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), have reported phase transition temperatures ranging from 286 K to 312 K.¹³⁻¹⁵

In part, these discrepancies originate from hysteresis linked to sampling limitations inherent in the techniques used to estimate the transition temperature. Indeed, the largest differences in transition temperature estimates tend to come from simulations where the transition temperature is estimated by simply heating or cooling the system. These types of simulations, which most closely mimic the experimental DSC measurements, unfortunately do not allow for simulation timescales long enough to overcome phase metastability leading to the formation of supercooled (or superheated) configurations in the case of cooling (or heating).

The presence of hysteresis has led to the design of enhanced sampling techniques that

can better sample the phase transition region, one such method is generalized Replica Exchange Molecular Dynamics (gREMD).¹⁶⁻¹⁹ Traditional temperature-based replica exchange methods are limited by the latent heat which causes poor energetic overlap in adjacent simulation windows in the transition region. gREMD avoids this limitation by simulating with H used as a control parameter, rather than T . In gREMD, simulation windows are controlled by linear effective temperature functions,

$$T_\alpha(H) = \lambda_\alpha + \eta(H - H_0), \quad (1)$$

where α is a label for individual replicas, η is a negative parameter that determines the slope of the function, H_0 is the reference enthalpy, and λ_α determines where this function intersects the true statistical temperature $T_S(H)$. This intersection occurs at exactly one place, thus avoids sampling multiple states like are sampled in traditional replica exchange. Functionally, the range and interval of λ_α s that are included determine the number of simulation windows required for a gREMD calculation.

The key output parameter from a gREMD calculation is the statistical temperature, $T_S(H)$, which is calculated using the Statistical Temperature Weighted Histogram Analysis Method (ST-WHAM).²⁰ From this profile, hereafter referred to as the “S-Loop” due to its shape, a Maxwell equal area construction can be used to obtain the transition temperature as well as the latent heat of the gel-fluid transition. In this construction, a horizontal line is placed so it intersects $1/T_S(H)$ such that the area bounded above and below the line are equal. The vertical location of this line is $1/T_m$, and the latent heat is found by calculating the enthalpy difference between the outer two intersections of this line with the S-Loop. We have shown this schematically in Figure 1b.

While gREMD has been shown to be successful for estimating phase transition temperatures for a variety of systems (including lipids),^{18,19} its implementation has been relatively restrictive. Many additional simultaneous simulation windows outside of the phase transition region are needed to meet the requirements of the

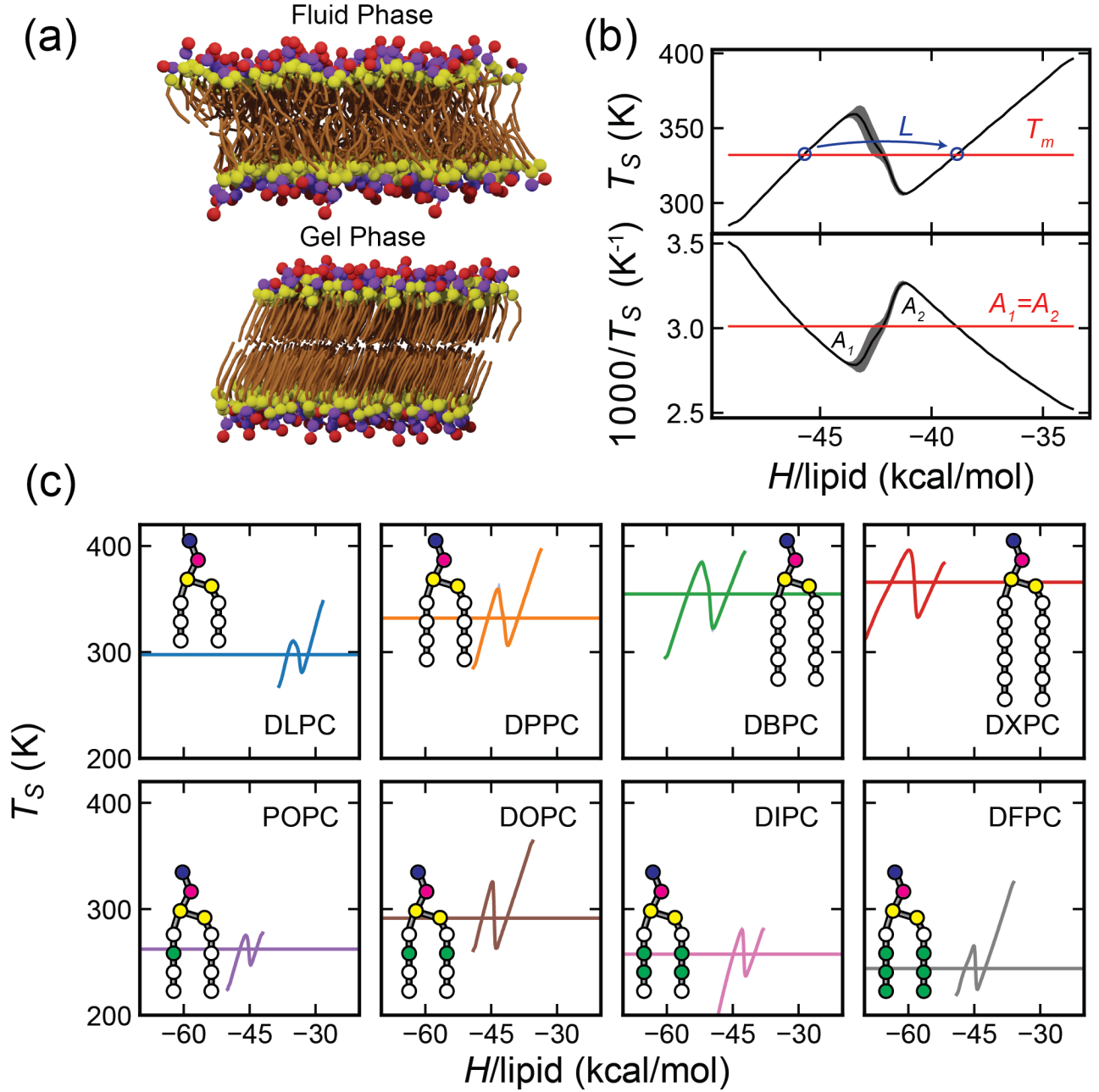


Figure 1: a) Simulation snapshots of a DPPC bilayer in the fluid and gel phases (periodic images of the x,y dimensions have been included for clarity). b) Upper Panel: Schematic diagram of the S-Loop structure for DPPC, the transition temperature, and the latent heat are included in black, red, and blue, respectively. Lower Panel: $1/T_S$ profile used for the Maxwell construction is shown in black, with the construction line shown in red. c) S-Loops for each lipid are presented on the same scale. Transition temperatures are included as horizontal lines in the same color. CG lipids are presented schematically, with NH_3 , PO_4 , glycerol, saturated tail, unsaturated tail beads represented in red, blue, yellow, white, and green, respectively.

Maxwell construction. Alternatively, peaks in explicit derivative of $1/T_S(H)$ with respect to E have been used to identify the transition location, but these require significant sampling

and can become complicated in systems with multiple transitions.²¹

In this letter, we demonstrate that gREMMD can be combined with the recently developed

Machine Learning-assisted Lipid Phase Analysis (ML-LPA) technique,²² and that this combination allows for accurate determination of the phase transition temperature with fewer simultaneous simulations. We have selected nine lipids with phosphatidylcholine headgroups and modeled them with the solvent-free coarse-grained Dry Martini model. These lipids cover a wide range of lipid chain lengths (DHPC, DLPC, DPPC, DBPC, and DXPC), and extent of tail saturation (POPC, DOPC, DIPC, DFPC).²³ The implicit solvent model for the lipids was selected for the following reasons: i) the phase behavior of this model (outside of DPPC) has not been well characterized previously, ii) the lack of explicit solvent avoids overlaying of solvent transitions and lipid transitions that complicates estimation of T_m , and iii) computational efficiency.

To begin, we built small model systems containing 32 lipids for each considered lipid type. We have chosen the present system size as the goal of this work is primarily to demonstrate the combination of the gREMD and ML-LPA methods. Furthermore, Stelter and Keyes demonstrated¹⁹ that increasing the number of DPPC lipids to 390 only decreased the transition temperature by 5 K, thus while the present results may not be quantitatively exact we expect no significant modification of the qualitative trends.

We then set up and ran gREMD calculations for the transition temperature for each lipid. For these simulations, we ran simulations over a range of values for λ_α to identify the approximate phase transition region. Based on this, we chose a reasonable bounds for λ_α and equilibrated each independently in the order from low enthalpy to high enthalpy, using the results of the previous calculation as the initial point for the new equilibration. Following this, a full gREMD calculation was run until reaching equilibrium, after which production simulations of 600 ns were run. Full details of this calculation, as well as the selected values of λ_α , H_0 , and η chosen for each lipid are included in the SI.

We have calculated the S-Loop for each of the PC lipids in the present study and eval-

uated both the phase transition temperature and latent heat for each of the Dry Martini lipids and have reported their values, as well as experimental measured values, in Table 1. Of the studied lipids, all but POPC and DHPC demonstrated a fluid/gel transition in the gREMD simulations. DHPC, which has by far the shortest chain length of any of the lipids studied (2 CG beads in each tail), likely is too small to observe a transition in the region accessible with gREMD, as simulations become unstable at very low effective temperatures. The lack of transition for POPC is more surprising as the corresponding wet model has been previously shown to undergo the fluid/gel transition. To correct this, we incorporated the angle modifications suggested by Daily and co-workers for POPC which have been shown to improve the wet model and gain closer agreement with the experimentally measured transition temperature.²⁴ Thus, we will include the results from the angle corrected model for completeness.

While gREMD is capable of calculating T_m , a significant number of simulation replicas were required outside of the transition region to meet the minimum criteria for the Maxwell construction. For the small model system described presently, this is not such an issue as the largest number of replicas used was 76 (for DBPC and DXPC). However, as simulation size increases the width of the transition region, ΔH , increases, meaning that for larger systems (or systems containing explicit solvent) additional replicas would be needed. For instance, in our DPPC simulations, of the 43 total replicas only 6 were located in the transition region. While the remaining 37 could be cut down slightly by exactly tuning the overlaps between adjacent replicas, such tuning would not lead to a significant reduction in the total number of replicas. Furthermore, it has been shown previously that adding another component that undergoes a phase transition (*e.g.*, water) can disrupt the shape of the S-Loop, making it harder (or potentially impossible) to identify the phase transition temperature through a Maxwell construction.¹⁹

In the remainder of this Letter, we will

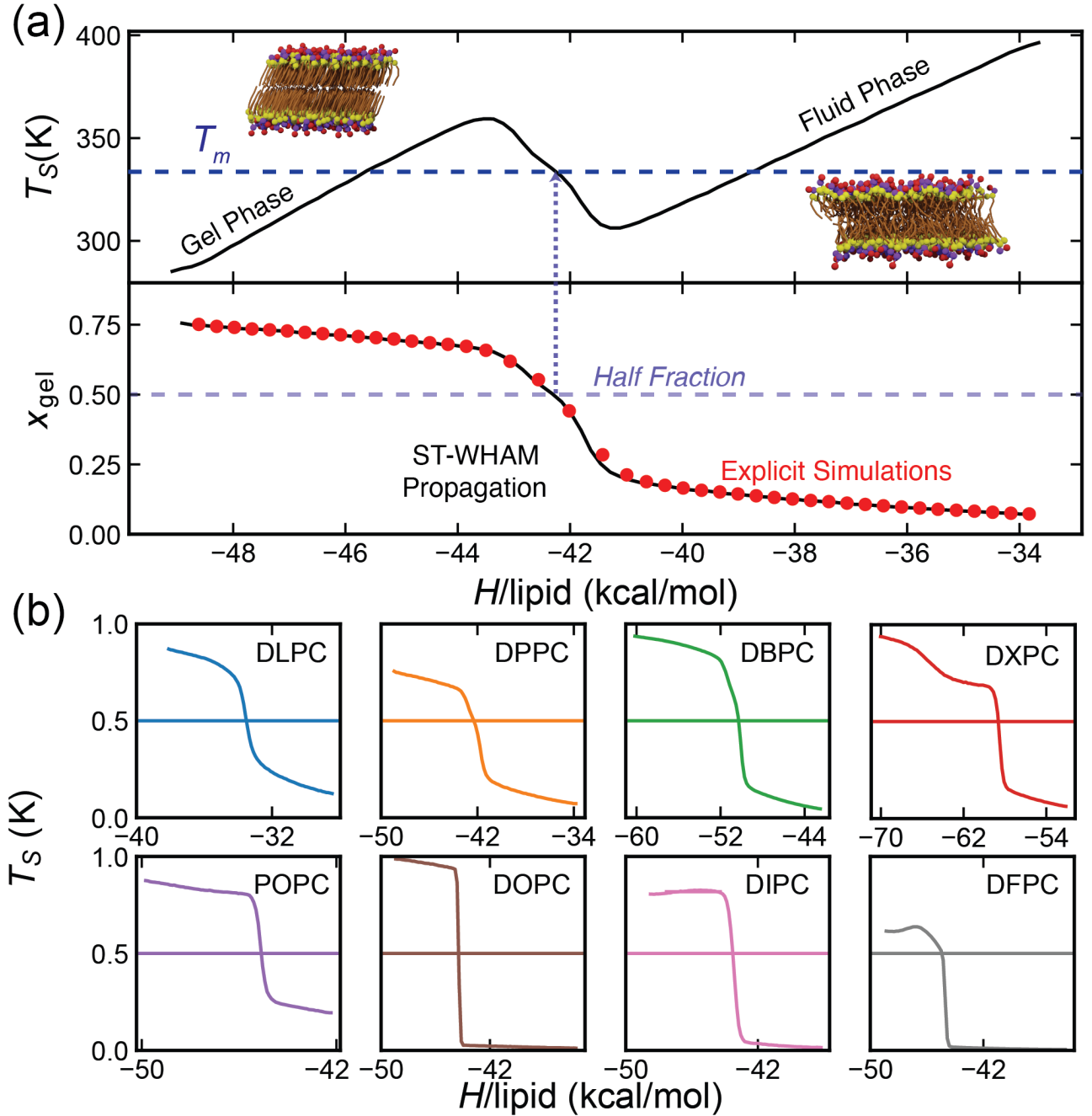


Figure 2: a) Schematic diagram of ML-LPA approach, mapping from gel half fraction (bottom) to S-Loop (top) (example shown: DPPC). b) Gel fraction plots with half fraction marked for each of the eight considered lipids.

demonstrate an alternative technique to calculating the phase transition temperature from a gREM simulation that does not require the use of an equal area construction. Instead, we will use the ML-LPA model recently developed by Walter and co-workers to identify individual lipid phases and then use this information in tandem with the S-Loop to evaluate the transition temperature.²²

In this work, we have adapted the publicly available ML-LPA code^{22,25} for use with LAMMPS.²⁶ This code uses a combination of Naive-Bayes (NB), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) to learn (and classify) the features that go into individual lipids as either gel or fluid phase.²⁷ These classification algorithms are applied to a combination of both coordinates and distance

Table 1: Tabulated transition temperatures and latent heats from Expt., gREMD, and ML-LPA. Uncertainties in the final digit(s) are reported as subscripts. For the transition region calculations, DPPC-Tr42 has 42 replicas in the transition region, and DPPC-Tr12 has 12 replicas.

Lipid	CG Beads		Expt.		gREMD		ML-LPA
	n_{tail}	n_{unsat}	$L/\text{Lipid}^{4,a}$ (kcal/mol)	T_m^{exp} (K)	L/Lipid (kcal/mol)	T_m (K)	T_m (K)
DLPC	6	0	2.1 _{1.1}	271.15	4.69 ₄	297.6 ₈	287.7 ₈
DPPC	8	0	8.2 _{1.4}	314.15	6.85 ₅	332.1 _{1.8}	335.0 _{2.3}
DPPC-Tr42	8	0	8.2 _{1.4}	314.15	-	-	334.0 ₈
DPPC-Tr12	8	0	8.2 _{1.4}	314.15	-	-	337.4 _{3.5}
DBPC	10	0	13.3 _{2.5}	339.15	9.07 ₅	354.8 _{1.2}	353.0 _{1.4}
DXPC	12	0	15.4 _{1.6}	353.45	9.88 ₁₀	365.8 ₆	365.4 _{1.6}
POPC[‡]	8	1	-	271.15	3.76 ₆	262.2 ₄	254.9 _{1.0}
DOPC	8	2	-	256.15	5.66 ₃	291.4 ₂	291.5 _{2.6}
DIPC	8	4	-	216.15	4.72 ₄	257.5 ₄	258.1 _{1.1}
DFPC	8	6	-	213.15	4.61 ₄	244.0 ₁	264.5 ₆

[‡] POPC includes angle corrections not included in any of the other studied models.

^a Experimental latent heats taken from Table 1 of Koyanova *et al.*⁴

vectors between pairs of beads separated by at least four bonds. We then trained individual ML-LPA models for each lipid on 3- microsecond NPT simulations of two gel temperatures and two fluid temperatures, selected based on prior knowledge of the transition region for a total of 4,000 configurations. Full training details, as well as model validation accuracy, are included in the SI.

These learned models were then applied to every gREMD configuration and used to evaluate the gel fraction,

$$x_{\text{gel}} = \frac{N_{\text{gel}}}{N_{\text{lipids}}} \quad (2)$$

where N_{gel} is the number of gel lipids identified, and N_{lipids} is the total number of lipids. We then propagated x_{gel} through ST-WHAM to obtain the profile of $x_{\text{gel}}(H)$. This profile is similar for each lipid, it is close to one at low enthalpies (gel phase) and decreases sharply across the phase transition, until it becomes near zero at high enthalpies (fluid phase). By combining $x_{\text{gel}}(H)$ and $T_s(H)$ we access the overarching profile $T_s(x_{\text{gel}})$ which provides information about the statistical temperature as a function of gel fraction. In practice, we access

the overarching profile by using a cubic spline interpolation on both profiles individually in order to map both more precisely to enthalpy.

If we define the phase transition as the location where there is an equal proportion of lipids in the fluid and gel phase, we can identify the transition temperature from the statistical temperature profile as $T_m = T_s(x_{\text{gel}} = 0.5)$. Thus, the ML-LPA approach provides an alternative to the Maxwell Construction by providing otherwise unavailable access to x_{gel} .

In summary, the typical procedure the combined calculation works as follows:

- Run two simulations, one at low temperature and one at high temperature to estimate the enthalpies. Use the procedure described by Kim *et al.* to extract values of H_0 and η .¹⁷
- Run consecutive simulations one after another from low- λ_α to high- λ_α using the previous final configuration as the starting point for the simulation. From each simulation calculate the temperature and use this information to obtain the range of λ_α needed to encompass the transition region.

- Run a gREMD calculation using these values of λ_α .
- Train the ML-LPA code on two temperatures for each phase using at least a few hundred configurations for each.
- Sort the gREMD configurations by value of λ_α and apply the ML-LPA model to each configuration to evaluate x_{gel} . Use ST-WHAM to stitch these, as well as the statistical temperature, together into $x_{\text{gel}}(H)$ and $T_s(H)$, respectively. Evaluate the transition temperature by locating the point on $T_s(H)$ where $x_{\text{gel}}(H) = 0.5$.

We have applied this approach, which is shown schematically in Figure 2, to each of the lipids and used it to evaluate the transition temperature. These transition temperatures are listed in the eighth column of Table 1.

In Figure 3, we have plotted values of all the transition temperatures calculated by both methods. In general, we find that the ML-LPA approach successfully reproduces the values of the transition temperature calculated by gREMD. For instance, for DPPC, the value of T_m extracted using the Maxwell construction approach is $332.0 \text{ K} \pm 0.8 \text{ K}$, whereas the ML-LPA approach gives a value of $335.0 \text{ K} \pm 2.3 \text{ K}$, which is in agreement within uncertainty. Similarly good agreement is seen for the vast majority of the lipids.

One exception to this good agreement is in the case of the DFPC bilayer, which contains three unsaturated tail beads. From the Maxwell construction, we obtained a value of $244.0 \text{ K} \pm 0.1 \text{ K}$, whereas we obtained using the present ML-LPA approach a value of $264.5 \text{ K} \pm 0.6 \text{ K}$, which are significantly different. Importantly, by looking at the $x_{\text{gel}}(H)$ profile for this lipid in Figure 2B, it becomes clear that the highest gel fraction it finds for this lipid is about 60%, or about 20 lipids in the gel phase. Attempts at further training to improve this gel fraction in the low enthalpy regime were not fruitful, and did not provide meaningful modifications to the ML-LPA predicted transition temperature.

We believe that the highly kinked conformation of the DFPC lipid tails leads to the ML al-

gorithm to incorrectly identify gel phase lipids as fluid, leading to the failure to obtain the correct transition temperature. In the Supporting Information, we include calculations of the area per lipid and tail order parameter calculated for the training data. These parameters are most similar between the fluid and gel phases for DFPC, which suggests that the phases may not be distinct enough for classification through ML-LPA. This could potentially be resolved in the future by improving the features used for training and classification.

There are a few important advantages to the presently developed approach. Firstly, the calculation of the full S-Loop is not required to extract the transition temperature. To demonstrate this we re-ran DPPC using 12 windows located primarily in the transition region. Using ML-LPA, we find the transition temperature from this calculation to be $337.4 \text{ K} \pm 3.5 \text{ K}$, in good agreement with the above results calculated from the entire S-Loop. Likewise, we also ran a calculation with 42 replicas in over the same region, which provided a value of $334.0 \pm 0.8 \text{ K}$ for the transition temperature. While having many replicas in the transition region (as in the 42 replica calculation) improves the accuracy of the estimated transition region, it is clear from the 12 replica calculation that few replicas are needed to estimate the transition temperature with reasonable accuracy.

As more windows are required by gREMD as the system size increases, reducing the simultaneous effort to only the transition region empowers scaling to larger systems than would otherwise be possible. Secondly, in general both considered methods give comparatively better agreement with one another than other reported approaches have in the past. Even slight variations between gREMD and ML-LPA (*e.g.*, the 10 K difference for DLPC) are smaller than the range of reported simulation results from other approaches. Lastly, because this method effectively maps the statistical temperature to the gel fraction directly, it allows for estimation of transition temperatures in scenarios where the S-Loop takes on a complicated structure. Even a binary system (*e.g.*, wet Martini DPPC with water beads) has a significantly more com-

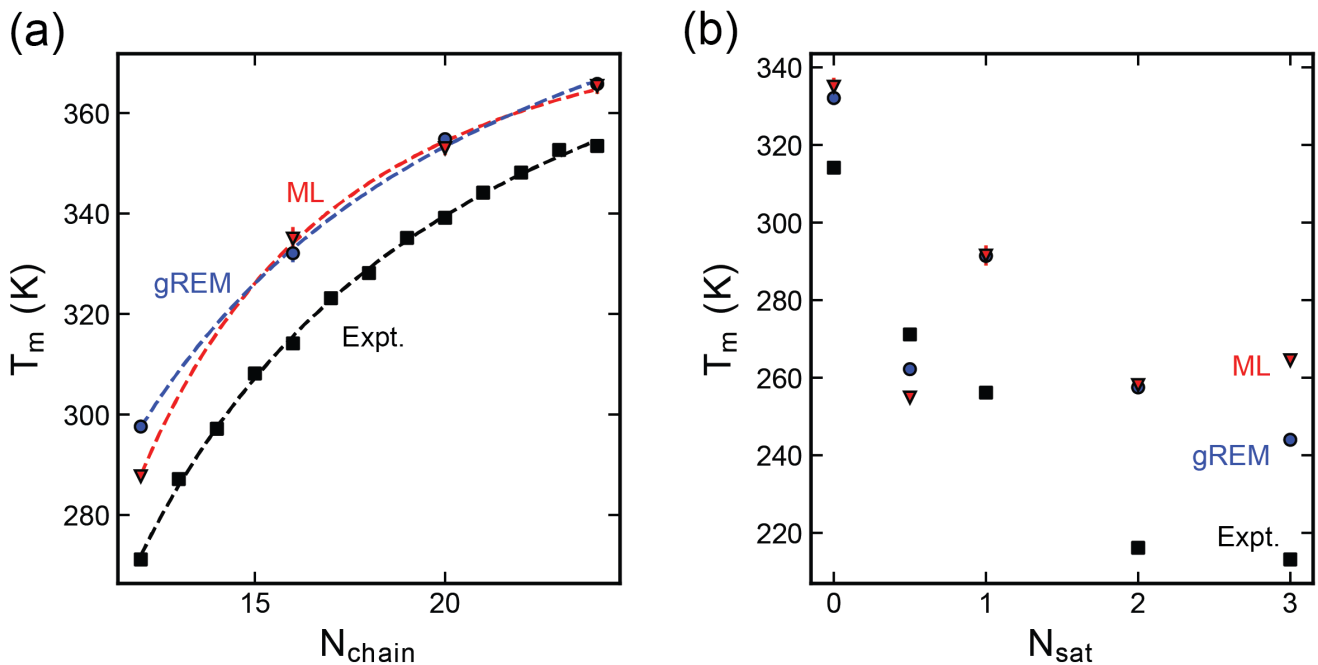


Figure 3: Phase transition temperatures are presented as a function of a) chain length and b) tail unsaturation. Maxwell construction and ML transition temperatures are shown as blue circles and red triangles, respectively. Experimental data is included as filled black squares. Fits to Cevc’s Eq. 1 are included in the same color as dashed lines to guide the eye.²⁸

plicated S-Loop¹⁹ that can present a challenge to the Maxwell Construction approach.

Both methods of calculating the transition temperature demonstrate that the Dry Martini model is able to qualitatively describe lipid phase behavior as a function of both chain length and tail saturation, with the exception of POPC, for which the unmodified Dry model did not demonstrate a phase transition and the angle-corrected model matches better with experiment than the other Dry lipids.

In summary, combining ML-LPA and gREMD greatly reduces the expense of estimating the phase transition temperature. Instead of requiring simulation windows far outside of the transition region in order to create a Maxwell Construction, the combined method requires only simulations in the S-Loop transition region, greatly reducing the cost of gREMD calculations. This decrease is significant enough that it will make gREMD calculations accessible for larger system sizes.

Our results demonstrate that for a wide range of Dry Martini lipids the combined method is able to determine the fluid/gel phase transi-

tion temperature within a reasonable accuracy for most lipids. The addition of another training feature, such as area per lipid (calculated through a Voronoi diagram), could potentially be used to further improve this agreement in the future. Additionally, the combined method provides a promising alternative to the Maxwell construction for systems that may not have an ideal S-Loop. For instance, lipids with explicit water have multiple phase transitions that have been shown previously to modify the shape of the S-Loop significantly, complicating the calculation of a transition temperature for these systems. It is likely that such methodology could also prove useful for systems with lipid mixtures. Lastly, the present work demonstrates that Dry Martini lipids are generally able to reproduce qualitative trends in the transition temperature as a function of chain length and tail saturation when compared with experiment.

Methods

For all simulations, a test set of phosphatidylcholine lipids were considered such that trends with respect to chain length (DLPC, DPPC, DBPC, DXPC) and chain saturation (POPC, DOPC, DIPC, DFPC) could be considered. A corrected Dry Martini force field file was generated from the publicly available Martini ITP files for use with the Moltemplate software,^{23,29} which was then used to generate systems of 32 lipids. This system size was selected to match the size used by Stelter and co-workers on DPPC bilayers.¹⁹ In that work, it was demonstrated that while finite size effects are present for such system sizes these effects are smaller than the uncertainty caused due to thermal hysteresis. After energy minimization, systems were equilibrated for 300 ns, while gradually increasing the timestep from 1.0 fs to the production value of 30 fs (the thermostat and barostat damping parameters were set to be 100 and 1000 times the timestep as it was scaled).

For gREMD simulations, a walkdown procedure from lowest to highest enthalpy replica was applied for 15 ns on each replica. gREMD simulations were run until block averaging of the S-Loops no longer showed a block dependence, anywhere between 1.5 μ s to 3 μ s, after which production runs of 600 ns were run from which reported values were calculated. ML-LPA training simulations were initialized from equilibrated gel configurations and allowed to relax for 300 ns in the NpT ensemble, these simulations were run for a further 3 μ s from which configurations were output every 300 ps. Models were trained on every fifth configuration to balance training expense and accuracy. For all calculated values, statistical uncertainties are estimated using block averaging over 5 blocks and are reported as a 95-% confidence interval according to Student's t -distribution.³⁰

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant

No. CHE-2001611, the NSF Center for Sustainable Nanotechnology. The CSN is part of the Centers for Chemical Innovation Program. The authors are pleased to acknowledge that the computational work reported on in this paper was performed on the Shared Computing Cluster which is administered by Boston University's Research Computing Services. The authors would furthermore like to thank Profs. John Straub and Tom Keyes, as well as Dr. David Stelter, for useful conversations about the generalized replica exchange method.

Supporting Information

Supporting information is available online. Included are gREMD simulation details, ML-LPA training details, calculated latent heats, comparisons of the S-Loop and gel fraction profiles, and lipid order parameters.

References

- (1) Marrink, S. J.; Corradi, V.; Souza, P. C.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119*, 6184–6226.
- (2) Kučerka, N.; Nagle, J. F.; Sachs, J. N.; Feller, S. E.; Pencier, J.; Jackson, A.; Katsaras, J. Lipid Bilayer Structure Determined by the Simultaneous Analysis of Neutron and X-ray Scattering Data. *Biophys. J.* **2008**, *95*, 2356–2367.
- (3) Tristram-Nagle, S.; Nagle, J. F. Lipid Bilayers: Thermodynamics, Structure, Fluctuations, and Interactions. *Chem. Phys. Lipids* **2004**, *127*, 3–14.
- (4) Koynova, R.; Caffrey, M. Phases and Phase Transitions of the Phosphatidylcholines. *Biochim. Biophys. Acta, Rev. Biomembr.* **1998**, *1376*, 91–145.
- (5) Simons, K.; Ikonen, E. Functional Rafts in Cell Membranes. *Nature* **1997**, *387*, 569–572.

- (6) Lingwood, D.; Simons, K. Lipid Rafts as a Membrane-Organizing Principle. *Science* **2010**, *327*, 46–50.
- (7) Starr, F. W.; Hartmann, B.; Douglas, J. F. Dynamical Clustering and a Mechanism for Raft-like Structures in a Model Lipid Membrane. *Soft Matter* **2014**, *10*, 3036–3047.
- (8) Barnoud, J.; Rossi, G.; Marrink, S. J.; Monticelli, L. Hydrophobic Compounds Reshape Membrane Domains. *PLoS Comput. Biol.* **2014**, *10*, 1–9.
- (9) Levental, I.; Levental, K. R.; Heberle, F. A. Lipid Rafts: Controversies Resolved, Mysteries Remain. *Trends Cell Biol.* **2020**, *30*, 341–353.
- (10) Unguryan, V. V.; Golysheva, E. A.; Dzuba, S. A. Double Electron–Electron Resonance of Spin-Labeled Cholestane in Model Membranes: Evidence for Substructures inside the Lipid Rafts. *J. Phys. Chem. B* **2021**, *125*, 9557–9563.
- (11) McElhaney, R. N. The Use of Differential Scanning Calorimetry and Differential Thermal Analysis in Studies of Model and Biological Membranes. *Chem. Phys. Lipids* **1982**, *30*, 229–259.
- (12) Keough, K. M.; Giffin, B.; Kariel, N. The Influence of Unsaturation on the Phase Transition Temperatures of a Series of Heteroacid Phosphatidylcholines Containing Twenty-Carbon Chains. *Biochim. Biophys. Acta* **1987**, *902*, 1–10.
- (13) Rodgers, J. M.; Sørensen, J.; De Meyer, F. J.; Schiøtt, B.; Smit, B. Understanding the Phase Behavior of Coarse-Grained Model Lipid Bilayers through Computational Calorimetry. *J. Phys. Chem. B* **2012**, *116*, 1551–1569.
- (14) Marrink, S. J.; Risselada, J.; Mark, A. E. Simulation of Gel Phase Formation and Melting in Lipid Bilayers Using a Coarse Grained Model. *Chem. Phys. Lipids* **2005**, *135*, 223–244.
- (15) Yang, L.; Kindt, J. T. Line Tension Assists Membrane Permeation at the Transition Temperature in Mixed-Phase Lipid Bilayers. *J. Phys. Chem. B* **2016**, *120*, 11740–11750.
- (16) Okamoto, Y. Generalized-Ensemble Algorithms: Enhanced Sampling Techniques for Monte Carlo and Molecular Dynamics Simulations. *J. Mol. Graphics Modell.* **2004**, *22*, 425–439.
- (17) Kim, J.; Keyes, T.; Straub, J. E. Generalized Replica Exchange Method. *J. Chem. Phys.* **2010**, *132*, 1–10.
- (18) Lu, Q.; Kim, J.; Straub, J. E. Exploring the Solid-Liquid Phase Change of an Adapted Dzugutov Model Using Generalized Replica Exchange Method. *J. Phys. Chem. B* **2012**, *116*, 8654–8661.
- (19) Stelter, D.; Keyes, T. Enhanced Sampling of Phase Transitions in Coarse-Grained Lipid Bilayers. *J. Phys. Chem. B* **2017**, *121*, 5770–5780.
- (20) Kim, J.; Keyes, T.; Straub, J. E. Communication: Iteration-free, Weighted Histogram Analysis Method in Terms of Intensive Variables. *J. Chem. Phys.* **2011**, *135*, 1–4.
- (21) Bachmann, M. *Thermodynamics and Statistical Mechanics of Macromolecular Systems*; Cambridge University Press: Cambridge, UK, 2014; pp 31–66.
- (22) Walter, V.; Ruscher, C.; Benzerara, O.; Marques, C. M.; Thalmann, F. A Machine Learning Study of the Two States Model for Lipid Bilayer Phase Transitions. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19147–19154.
- (23) Arnarez, C.; Uusitalo, J. J.; Masman, M. F.; Ingólfsson, H. I.; De Jong, D. H.; Melo, M. N.; Periole, X.; De Vries, A. H.; Marrink, S. J. Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations

with Implicit Solvent. *J. Chem. theor. Comput.* **2015**, *11*, 260–275.

- (24) Daily, M. D.; Olsen, B. N.; Schlesinger, P. H.; Ory, D. S.; Baker, N. A. Improved Coarse-Grained Modeling of Cholesterol-Containing Lipid Bilayers. *J. Chem. Theor. Comput.* **2014**, *10*, 2137–2150.
- (25) Walter, V.; Ruscher, C.; Benzerara, O.; Thalmann, F. MLLPA: A Machine Learning-assisted Python Module to Study Phase-Specific Events in Lipid Membranes. *J. Comput. Chem.* **2021**, *42*, 930–943.
- (26) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.
- (27) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. et al. Scikit-Learn: Machine Learning in {P}ython. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- (28) Cevc, G. How Membrane Chain-Melting Phase-Transition Temperature Is Affected by the Lipid Chain Asymmetry and Degree of Unsaturation: An Effective Chain-Length Model. *Biochemistry* **1991**, *30*, 7186–7193.
- (29) Jewett, A. I.; Stelter, D.; Lambert, J.; Saladi, S. M.; Roscioni, O. M.; Ricci, M.; Autin, L.; Maritan, M.; Bashusqeh, S. M.; Keyes, T. et al. Moltemplate: A Tool for Coarse-Grained Modeling of Complex Biological Matter and Soft Condensed Matter Physics. *J. Mol. Biol.* **2021**, *433*, 166841–166841.
- (30) Shoemaker, D. P.; Garland, C. W.; Nibler, J. W. *Experiments in Physical Chemistry*; McGraw-Hill: New York, 1989.