Black: Actual Text  
Purple : Discussion  
  
Dear Dr. Grossfield,  
  
The reviews of your manuscript have been received. Based on these reviews, and on my own reading, your paper cannot be accepted for publication in its present form. Although there is no guarantee, a revised manuscript may become acceptable if it adequately addresses the concerns described below. Biophysical Journal normally permits only one major revision.   
  
As you will see, the referees find your work interesting, but are of mixed opinion on the overall significance of the work. Referee #3 does not support publication for that reason. However, in my opinion this recommendation is not well justified, and I would welcome a revised version of your manuscript. In particular, please pay attention to the following points: (i) try to provide a better rationale for why certain CVs work better than others, to what extent they are generalizable to other membrane systems, and how they can distinguish between local domain formation and macroscopic phase separation, (ii) try to connect your results better to the thermodynamics of phase separation, including improvement on the statistical quality of the analysis.  
  
When you submit your revised version of the paper, it will be sent back for re-evaluation to at least one of the original referees to help me determine if it can be accepted for publication. Please find the referee comments at the end of this email.  
  
The changes that you make in the revision must be clearly and explicitly explained. When submitting your revision, please prepare a response to reviewer and editor comments and provide this upon uploading your submission. If the changes are reasonably localized, your changes in the manuscript should be specifically marked (i.e., using colored font, underlined text, highlighted background, or italicized font).  
  
Although Biophysical Journal accepts format-neutral manuscripts on first submission, note that you must follow formatting guidelines for your revised manuscript. You may visit the Author Guidelines at [https://www.cell.com/pb-assets/journals/society/biophysj/PDFs/author-guidelines.pdf](https://urldefense.proofpoint.com/v2/url?u=https-3A__www.cell.com_pb-2Dassets_journals_society_biophysj_PDFs_author-2Dguidelines.pdf&d=DwMGaQ&c=4sF48jRmVAe_CH-k9mXYXEGfSnM3bY53YSKuLUQRxhCwrmQ9ZDKmVaAI6sBPw4Jo&r=49qnaP-kgQR_zujl5kbj_PmvQeXyz1NAoiLoIzsc27zuRX32UDM2oX8NQCaAsZzH&m=CBmb8WmZqKSoaP7cJpYtJRv5VIhGfBU8yI6fv35788jknt4wrB4k1JQrAfQD60bS&s=lgZgUyGcrqoPaLtqsHQ1Uxybs4u7BV7ur56akm95QX0&e=)  
  
Your revised manuscript is due by Jul 18, 2023. A revised manuscript received after this date will be treated as a new submission.  
  
Regards,  
Siewert Jan Marrink  
Editor, Biophysical Journal

Reviewers' Comments to the Author

Reviewer 1: The manuscript "Free energy of phase separation in lipid bilayers from molecular dynamics" by Poruthoor and Grossfield reports an investigation of three different collective variables (CVs) used in conjunction with the weighted ensemble (WE) algorithm to sample lipid mixing degrees of freedom. This is an important problem and a well-premised choice of methods, and the technical aspects of the results will be useful to researchers in the field. However, the manuscript only really offers some technical recommendations for such simulations, and falls short of delivering on the promise in the title. This is due mainly because the results are never really connected to fundamental aspects of lipid phase separation thermodynamic, as described below in the first two "major criticisms."   
  
May be changing the title “Thermodynamic” to Free energy  
  
Major criticisms.  
  
1. Figures 5, 6, and 7, and accompanying discussion. The discussion below Figure 5 is confusing. The DPPC-DAPC-chol plot shows double well behaviour at lower temperatures, and (I think) single well behaviour at higher temperatures. (It is hard to say, because the curves are all piled on top of each other.) But the text says that "both basins correspond to relatively high FLC and appear to be phase separated based on visual inspection." It is also mentioned that at the higher temperature "60% of the lipids prefer to be in clusters." Is this system phase separated? Normally when a free energy curve like this moves from a double well to a single well as a function of some thermodynamic control parameter, one well corresponds to one phase, and the other well corresponds to the other phase. Is that the case here?  
  
We largely agree with this assessment, and have rewritten the discussion accordingly.

We don’t know where the phase separation is: Potential energy fluctuations as a function of FLC (Phase transition point?) – Heat Capacity  
Try from the conventional MD as well by omitting the first 3/4th of the vanilla sims – last 1us may be?  
Next pass would be to get it from the WE  
One as a function of temperature and other as FLC (at a given temperature)

One of the struggle is the location shift and the DAPC is inherently phase separating even at high temperatures – and this is being reflected in the curves: Stressing the actual value of FLC than where the minima are  
  
More broadly, the results are never clearly connected to the basic thermodynamics of phase separation. Some attempts are made in this direction in the section entitled "Computing the free energy of separation" (did the authors mean to write "phase separation here?"), but the presented analysis falls short of the mark, in the reviewer's opinion. The difficulty seems to be that it is not straightforward to transform the FLC collective variable into a useful order parameter (in the Landau sense, not the SCD sense), with clearly separated values that discriminate the two phases. Perhaps this is not surprising, since the thermodynamic variable that is varied is the temperature. As temperature is varied the composition of the two phases also changes; meanwhile, the FLC is a local composition variable. Thus one expects that the values of the FLC CV that defines the two phases will change with temperature. I expect that this approach would be cleaner if instead of temperature composition were varied, in a such a way that the system follows a tie-line through the coexistence region.

We agree that FLC is not especially intuitive as an order parameter, and were surprised that quantities much closer to classic order parameters (e.g. CEI, which is an analog of density) do not perform well. To improve this discussion, we added a new coordinate, FLC\_opt, which does a better job of discriminating between the two states (see modified Fig 6 and a new Fig 8), and discussed why we think it discriminates more easily. We further agree that performing simulations along a tie-line rather than as a function of would yield more easily interpretable results; however, to our knowledge, there is no easy way to find the tie lines short of truly exhaustive sampling of composition space, which is prohibitively costly.

Alternatively, one might consider an analysis that is expected to give a clean signal as a function of temperature — the heat capacity. This has the advantage of obtaining an estimate of the transition temperature for each system (at least for the finite size simulation), at the location of the peak in the heat capacity, which ties the results back to the thermodynamics of phase separation. The POPC system is also useful here, as one expects no signature of a phase transition, based on the longer, standard CG simulations.

This is an excellent idea, but unfortunately the heat capacities are unusably noisy. In a sense this is not surprising, since the fluctuations of the potential energy for systems of this size cover thousands of kcal/mol, while the expected signal is perhaps 2 orders of magnitude smaller.  
  
2. The presented work falls short of delivering data of high statistical quality as promised in the introduction. The work is premised by stating correctly that it is challenging to obtain simulation results for lipid phase separation of good statistical quality. If one could obtain such data it would be of great value. Simulation models could be rigorously tested, and experimental results for more complex mixtures rationalized. However, no such demonstrations are presented. No phase boundaries are reported, no transition temperatures reported, the compositions of the two phases are nowhere obtained. The data points in Fig. 7 are presented without error bars. I agree with the authors that the goal is a worthy and impactful one, but as it stands the work seems to get only halfway to the finish line. A new sampling protocol is presented that looks promising, but the proof that it really works for calculating key observables is not provided.

In the original submission, we do demonstrate one way to calculate the Tm (Figure 7) given the restriction of model and the FLC cutoff choice. In the revised manuscript, the addition of FLC\_opt as a coordinate gave us an additional option, since nearly all systems have 2 wells, which allows us to use the free energy curves themselves to estimate ΔΔG.

Figure S7 shows the variation of the free energy curves from individual replicas, which are generally quite consistent (usually less than 0.5 kcal/mol difference). Computing error bars for free energy curves unambiguously is actually a very challenging problem, since the error bars vary with the choice of the arbitrary additive constant. That, plus a desire to show curves for all temperatures on the same graph in an effort to keep the number of figures manageable, led us to choose not to put error bars on the free energy curves in the main manuscript.  
  
3. The introduction states that there is functional phase separation in the living cell membrane, but the evidence supports (at best) instead localized domains of different composition. It is also claimed that this phase separation exists across all domains of life. Both statements are an overreach, in the reviewer's opinion.   
  
We respectfully disagree. We have added a reference to recent work from Sarah Keller’s lab, showing melting curves for phase separation in live yeast.  
  
Minor comments.  
  
In a few places nice sounding turns of phrase are used, but with unclear meanings: for example, "molecular grammar" and "curated ensemble." I'm not sure what either of these phrases mean.   
  
We need to specify what we mean by that : molecular grammar: which physical characteristic has which thermodynamic effects/signature  
Curated ensemble 🡪 ensemble  
  
The explanation of the CVs (Eqs 1,2,3) is not very clear.  
  
More discussion on CVs  
  
Reviewer 2: In this manuscript, Poruthoor and Grossfield describe a weighted ensemble (WE) approach to study the free energy landscape of phase separation in model membranes using molecular dynamics (MD) simulations. The proposed method and protocols provide a useful means by which to probe the thermodynamics of lipid phase separation using MD simulations; this is important but has been missing in the field. The work is rigorous and the manuscript is well written overall. However there are a few issues that I would like to encourage the authors to revisit. One is that they did not provided a sufficient rationale for why two of the three collective variables failed in their WE analysis while the third, which is related to the other two or at east not fundamental unique, works so well. At the minumum a hypothesis/hypotheses could be offered in order to complete the discussion, even if they may not test the hypothesis here.

Add some speculations Refer the paper by Ma et al about the alanine-dipeptide(How you don’t need psi – how some coordinates couples and some don’t)  
Make a comment on : It would be interesting apply ML to derive and discriminate more effective coordinates – Mention how CEI and SI tracks similar phenomena and FLC tracks something different but much more transient.

Another limitation is the rather terse Discussion section. This could be expanded through, for example, a more thorough discussion of the possible physical basis and hence different performances of the three reaction coordinates, especially given the data in Fig 3.

Expand the discussion more detailed.

Fig 7 should also be interpreted more. For example, what is the basis for the different temperature dependence of deltadetaG for the two bilayer models?  
  
The phase separation propensity is different for DAPC and DIPC system.  
We are not particularly confident about the 298K data point in the DIPC system as the FLC is struggling to get the multiphasic behavior – can be addressed by optimizing the FLC parameters  
More speculations   
  
Overall, this is an important piece of work from a group with a strong track record in free energy methods, with a potentially broad application but can be strengthened a bit more, as noted above.  
  
Add a thank you note  
Change the xrange for the S8 Fig

Reviewer 3: The paper presents results of coarse-grained (CG) MARTINI simulations of three ternary mixtures, two of which are known to exhibit phase separation and one not. For the simulations, the authors use Weighted Ensemble (WE) techniques, which supposedly expedite the phase segregation of the mixtures when starting from a uniform distribution of lipids. The work is purely methodological and contains no new biophysics.

We agreed but we explained why this protocol can be used for future method dev as well as applications.

This is not to suggest that papers presenting new methods or improvements of existing methods are not interesting and publishable, but it is far from obvious whether the WE implementation presented herein is truly a big step forward: WE methods are not new and the generalization of the results to simulations of other membranes is not clear. The most interesting result in the work is the demonstration that collective variables can serve as "reaction coordinates" for WE simulations, but (as shown in the work) some collective variables perform poorly, and it is not obvious which collective variable works efficiently for a given membrane under investigation.

As far as we can tell this CV is general to membranes – the obvious choice don’t work to drive a WE and non-obvious choice does work. We show this for three different lipid bilayer system thus showing its generalization  
  
I find the choice to apply the WE methods to CG simulations problematic, for two reasons: First, the problem of slow phase separation in mainly a problem in atomistic simulations.

Yes, you can get P.S in both but the problem is that such events are not reversible in MD timescale even for CG which we and others showed.

Second, calculations of free energy differences from CG simulations is always a risky business because of the loss of degrees of freedom in the coarse-grained picture which causes the entropic component of the free energy to be replaced by enthalpy. Therefore, the thermodynamics of phase separation captured by MARTINI simulations is not necessarily correct.  
  
Yes, you are right but we are proposing a method that works and the use of CG is to determine whether the method works. There’s nothing here that is hardwired just for CG. Imagine the same analysis done by AA sims and the computational resources need for that just to optimize a first iteration of a method.  
  
Some additional comments:  
  
1. The authors state, regarding computer simulations, that "as a rule they cannot quantify the underlying thermodynamics" (page 2). This is probably correct for atomistic and CG molecular simulations, but miss the contribution of lattice simulations to the field (see, e.g., the works of Almeida and others).  
  
Yes we agree. Included additional references to reflect this. ACTION ITEM: NEED TO CHANGE THE TEXT  
  
2. I find it disappointing that the authors do not explain the essence of WE methods, and only give the technical details for running WE simulations in GROMACS. Readers who are not familiar with the concept of WE simulations will have hard time to follow the work.  
  
A few more sentences to explain how WE works  
  
3. The collective variables proposed in the paper are quantities that are sensitive to local clustering of molecules with themselves. The logic behind this choice is clear, but it neglects the fact that ternary mixtures not exhibiting thermodynamic phase separation still show formation of domains (termed microscopic or local phase separation in some papers). This is clearly seen in the DPPC-POPC-Chol simulation results in Fig. 1. The authors should discuss how the collective variables distinguish between macroscopic phase separation and formation of local domains.

See above discussion, the current CV – not being a Landau CV, can’t distinguish between phase separated or not.   
  
It would be also nice to include a general discussion on the phase behavior of ternary mixtures.  
  
Add references to other reviews (Not the scope of the current work)  
  
To conclude, the paper presents an application of WE methods for simulations of ternary mixtures of saturated and unsaturated lipids with cholesterol. The results may be promising, but the generalization of the work to other mixtures is not clear, let alone to more complex biological membranes. The paper presents no new physics and its value is limited to readers who are experts in these kind of simulations. I think that the work fits better to a more computational-oriented journal and recommends the authors to elaborate on the concept of WE simulation methods and on the formation of liquid ordered domains in mixtures not showing phase separation.  
  
Thanks a lot for the suggestion. “the formation of liquid ordered domains in mixtures not showing phase separation” is an oxymoron.

Reviewers' Responses to Questions

**Research transparency**

Does this paper meet the standards of research transparency? If no, please explain why.

Reviewer #1: yes

Reviewer #2: yes

Reviewer #3: Yes

**Importance of research/value to the field**

Reviewer #1: 3 - Strong

Reviewer #2: 3 - Strong

Reviewer #3: 2 - Average

**Originality**

Reviewer #1: 4 - Excellent

Reviewer #2: 3 - Strong

Reviewer #3: 3 - Strong

**Technical Rigor**

Reviewer #1: 2 - Average

Reviewer #2: 4 - Excellent

Reviewer #3: 3 - Strong

**Proper citation of previous work**

Reviewer #1: 3 - Strong

Reviewer #2: 4 - Excellent

Reviewer #3: 2 - Average

**Clarity of writing/English usage**

Reviewer #1: 4 - Excellent

Reviewer #2: 4 - Excellent

Reviewer #3: 4 - Excellent

**Quality of figures**

Reviewer #1: 3 - Strong

Reviewer #2: 3 - Strong

Reviewer #3: 3 - Strong

**ACTION ITEMS :   
  
1. Potential Energy Fluctuation Calculation.**

**Error estimation? (Low priority)**

**Eigen space separation of CVs based on the SGOOP Method? (Even low priority)  
2. Add further detailed discussion.  
 Discussion of DeldelG** **Why FLC is better (Speculations) and how we can make FLC less arbitrary.**