Biophysical Journal, Volume 121

Supplemental information

Molecular mechanisms of spontaneous curvature and softening in complex lipid bilayer mixtures

Henry J. Lessen, Kayla C. Sapp, Andrew H. Beaven, Rana Ashkar, and Alexander J. Sodt

Supplemental Material for "Molecular mechanisms of spontaneous curvature and softening in complex lipid bilayer mixtures"

Henry J. Lessen^{1,*}, Kayla C. Sapp^{1,*}, Andrew H. Beaven^{1,2}, Rana Ashkar^{3,4}, and Alexander J. $Sodt^{1,**}$

¹ Eunice Kennedy Shriver National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892

²Postdoctoral Research Associate Program, National Institute of General Medical Sciences,

National Institutes of Health, Bethesda, MD 20892

³Department of Physics, Virginia Tech, Blacksburg, VA 24061

⁴Center for Soft Matter and Biological Physics, Virginia Tech, Blacksburg, VA 24061

*These two authors contributed equally.

**alexander.sodt@nih.gov

July 1, 2022

1 HMM state detection

A general overview of the HMM methodology used in this study will be provided. First, dimer conformations are extracted from MD trajectories from user defined parameters that will be described in detail below. These structures are clustered and subsequently divided into HMM observables that describe the energy landscape. The HMM observables are then used to define hidden states and optimize parameters that define the model. The states are then used in additional calculations to determine curvature preference of specific dimer subpopulations from their dynamic redistribution (curvature-coupled redistribution, CCR). A graphical summary of the steps in the HMM/CCR method is shown in Figure 1, and a list of terms is presented in Table S1.

Table S1: Terms used to describe the configurations and models for developing the HMM.

1.1 Using an HMM to extract states

For example, PSM approach complexes are identified using five headgroup atoms from each PSM molecule (10 atom positions from each dimer). The atoms are { OF, C1F, HNF, O3, HO3 }. For a complete list of atoms used to define approach complexes in each system analyzed, refer to Table S1. These atom positions are used to calculate the distance between molecules using the center of geometry. Pairs of lipids within $R_{\text{pair}} = 14 \text{ Å}$ were considered approach complexes. Atom positions are also used to calculate MSD between different approach complexes (used for scoring in K-medoid clustering).

We used approximately 15,000 approach complexes per simulation to define a pool from which the similarity centers were extracted. The set was extracted by traversing the trajectory uniformly, taking all complexes within R_{pair} . The quantity was limited by increasing the amount of trajectory skipped. Our software was able to efficiently determine similarity centers when the pool was less than 20,000.

Additional considerations must be made when identifying ion-mediated approach complexes. When processing simulations to extract this class of approach complex, an ion type and distance cutoff must be specified in the processing input file. When searching for the bridging ion, first the approach complex center (based on the specified atoms) is calculated. Then the distances from this point and the ions proximal to the approach complex are calculated to find the closest ion. If distance from the ion is then compared to the maximum radial distance (4 Å) , this ion is taken as the bridging ion.

Construction of the HMM

K-medoid clustering [1] was used to reduce the complexity of the extracted approach complexes and identify HMM observable states. This method uses both the mean-square-displacement (MSD) between approach complexes (following rotational and translational alignment) and a heuristic hydrogen bond matching scheme to rank similarity between approach complexes. The unweighted MSD between two approach complexes with a matching hydrogen bond were reduced by ϵ_+ \AA^2 for each bond, where ϵ_+ is a tunable parameter. If a hydrogen bond was present in one of the approach complexes but not the other, the MSD was penalized by $\epsilon_ \AA^2$. An exception was made for the amide-amide sphingolipid hydrogen bond. Penalties were applied that restricted members of a similarity cluster to have this bond (or the lack thereof) in common. The identity of one approach complex's 1 and 2 were swapped during the comparison, and the smaller MSD was selected, yielding a process independent of 1/2 ordering. The function applied by the K-medoid clustering to compare two pairs is called a similarity function, and any one function is denoted as S , with ϵ_{\pm} parameters indicated in the subscript. The final value of the similarity function is denoted $\sim \chi^2$. For each similarity function, the number of medoid centers was set to 52, defining similarity centers labeled $\{A-Z, a-z\}$.

The data was then processed to describe the time evolution of the approach complexes using the similarity centers determined through K-medoid clustering.This is done by reiterating through the MD trajectory in order to assign individual approach complexes to specific similarity clusters throughout the simulation. When processing the MD trajectory to build the HMM, the same spacing used for extracting approach complexes was used. An approach complex was assigned the letter code of the medoid center with which it had the lowest $\sim \chi^2$. This yields many distinct series of similarity cluster observables, stored as simple character strings. These distinct series (called a pair trajectory) describe the evolution of independent approach complexes. The pair trajectories are then used to build the kinetic portion of the HMM.

Finally, we have to determine the "hidden" states used in our HMM, which will will refer to as HMM states. The model defines $N_s = 6$ states. There is no correct number of states no more than there is a correct model for any physical system. Rather, if "too many" states are selected, an interesting state may be split into two states that interchange rapidly, but which exchange slowly with the other states. If "too few" states are selected, features of interest can not sort cleanly into states. Quotes here indicate that these are qualitative judgments. Once states are determined, all approach complexes are then classified using the HMM states and used for further analysis, such as the curvature preference calculations described below.

2 PIP2: protonated on the 4- or 5-phosphate

Figure S2 shows the curvature coupling of PIP2 protonated on the 4-phosphate (left) and the 5-phosphate (right). While there are differences between the two states, results for each protonation state are consistent with the observation that PIP2-Ca²⁺-POPC induces strong negative curvature.

Figure S1: A flowchart showing the process of dimer discovery and assignment via a hidden Markov model (HMM). The training (left side) and analysis (right side) steps are distinguished by color.

Figure S2: PIP2 protonated on the 4-phosphate (left) and the 5-phosphate (right).

Figure S3: PIP2 spontaneous curvature, but now with interactions bridged by Mg^{2+} instead of Ca²⁺. PIP2 is protonated either on the 4-phosphate (left) or the 5-phosphate (right).

Table S2: Results from Students T-test comparison of experienced curvatures determined from the DOPE/DOPC/DOPS simulations.

3 PIP2: Calcium vs. magnesium ions

Figure S3 shows the same curvature-dependent dimers now coupled by Mg^{2+} instead of Ca²⁺. Results for Mg^{2+} are consistent with the mechanism that it is the cation-mediated interaction with neighboring generic phospholipid (here, POPC) that induces strong negative curvature.

Table S3: Results from Students T-test comparison of experienced curvatures determined from the PIP2/POPC simulations.

Table S4: Results from Students T-test comparison of experienced curvatures determined from the PSM simulations.

4 P values for comparisons of experienced curvatures

5 Softening theory for complex bilayers

With $c = c_1 + c_2$, where c_1 and c_2 are the two principal curvatures,

$$
H_{\rm HC} = \frac{\kappa_{\rm b}}{2} (c_1 + c_2 - c_0)^2 + \bar{\kappa} c_1 c_2, \tag{S1}
$$

the total elastic curvature energy is:

$$
E_{\rm HC} = \frac{\kappa_{\rm b}}{2} \int_A d\mathbf{S} H_{\rm HC},\tag{S2}
$$

To account for the preferred curvature $\Delta c_0 = c_{0,p} - c_{0,\text{background}}$ of a lipid $(c_{0,p})$ in a background spontaneous curvature $(c_{0,\text{background}})$, the energy density is modified by

$$
H_{\rm c} = \frac{\kappa_{\rm m} A_p}{2} \int d\mathbf{S} \left[\left(c - c_{0,\rm p} \right)^2 - \left(c - c_{0,\rm background} \right)^2 \right],\tag{S3}
$$

Here we do not include the Gaussian curvature modulus \bar{k} as we assume all lipids have equivalent \bar{k} , and the integral of the Gaussian curvature does not vary with fluctuations.

For simplicity, the membrane surface, and thus its form in Eqs. S2 and S3, is parameterized in the linearized Monge gauge, a simplified surface parameterization capable of describing nearly planar configurations. The parameterization is $\mathbf{r}(x, y) = \{x, y, h(x, y)\}\$ and linearization imposes a quadratic order cut-off [2, 3] that is accurate for thermal (here, weak) fluctuations:

$$
E_{\rm H}\left[h(\boldsymbol{r})\right] = \int_A d\boldsymbol{S} \, \frac{\kappa_{\rm b}}{2} \left[\boldsymbol{\nabla}^2 h(\boldsymbol{r})\right]^2 \tag{S4}
$$

In Fourier space:

$$
h(\mathbf{r}) = \frac{1}{A} \sum_{\mathbf{q}} h_{\mathbf{q}} e^{i \mathbf{q} \cdot \mathbf{r}} \tag{S5}
$$

$$
h_{q} = \int_{A} dr h(r) e^{-i q \cdot r}.
$$
 (S6)

For a real space function $f(r)$ the relation $f_q = f^*_{-q}$ applies to the Fourier transform coefficients. The coefficients are not independent. Furthermore, terms that must be real are real only if q and $-q$ are considered simultaneously. In recognition of this we sum over q using the shorthand $\{q > 0\}$,

$$
\{\boldsymbol{q} > 0\} \equiv \{q_x, q_y\} \tag{S7}
$$

,

such that

$$
\begin{cases} 0 < q_x < q_{\text{max}} & \text{for } q_y = 0\\ -q_{\text{max}} < q_x < q_{\text{max}} & \text{for } q_y > 0 \end{cases}
$$

yielding two independent functions, $f_{q,a}$ and $f_{q,b}$:

$$
f_{q,a}(r) = 2\operatorname{Re} f_q \cos(q \cdot r)
$$

\n
$$
f_{q,b}(r) = 2\operatorname{Im} f_q \sin(q \cdot r)
$$
\n(S8)

Henceforth we use the cos function's coefficient $f_{q,a} = 2\text{Re } f_q$ as the independent variable. Implicitly, the same derivations apply to the b-component, $h_{q,b}$.

The Fourier coefficients $\{h_{q,a}\}$ of the height (with $h_{q,b}$ defined analogously to $f_{q,b}$) are energetically uncoupled:

$$
E_{\rm H}(\{h_{\mathbf{q},a}\}) = \frac{\kappa_{\rm b}}{2A} \sum_{\{\mathbf{q} > 0\}} q^4 h_{\mathbf{q},a}^2 \tag{S9}
$$

with all other cross-terms energetically uncoupled at this order. Given a bilayer with a mole fraction χ of one lipid and $1 - \chi$ background lipids, the energy of a density fluctuation is given by

$$
E_{\rho}(\{p_{q,a}\}) = \sum_{\{q>0\}} \frac{k_{\rm B} T(p_{q,a}^2)}{2A\rho_0 (1 - \chi)}
$$

=
$$
\sum_{\{q>0\}} \frac{A_p k_{\rm B} T(p_{q,a}^2)}{4A\chi (1 - \chi)}
$$
(S10)

Where $p_{q,a}$ are Fourier coefficients for the distribution of particles (with a and b subscripts analogous to $f_{q,a}/f_{q,b}$ above), and $\rho_0 = \frac{2\chi}{A_p}$, with the factor of two representing the contribution from each leaflet. See derivation below.

The impact of the distribution of lipids with spontaneous curvature difference $\Delta c_{0,p}$ is now introduced using the lipid density $\rho(\mathbf{r})$ which is convoluted with the spatial extent $w(\mathbf{r})$ to account for the change in curvature energy density in Eq. S3:

$$
E_{c}[h(\mathbf{r}), \rho(\mathbf{r})] = \frac{\kappa_{\rm m}}{2} \int_{L^{2}} d\mathbf{r}_{p} \int_{L^{2}} d\mathbf{r} \rho(\mathbf{r}_{p}) w(\mathbf{r} - \mathbf{r}_{p}) \times
$$

\n
$$
\left[\left(\nabla^{2} h(\mathbf{r}) - c_{0, p} \right)^{2} - \left(\nabla^{2} h(\mathbf{r}) \right)^{2} - c_{0, \text{background}} \right]
$$

\n
$$
= \frac{A_{p} \kappa_{\rm m}}{2} \int_{L^{2}} d\mathbf{r} \rho(\mathbf{r}) \times
$$

\n
$$
\left[\left(\nabla^{2} h(\mathbf{r}) - c_{0, p} \right)^{2} - \left(\nabla^{2} h(\mathbf{r}) \right)^{2} - c_{0, \text{background}} \right]
$$
\n(S11)

Here Eq. S12 uses $w(r - r_p) = A_p \delta(r - r_p)$ as justified by the observed locality of PC, PE, and PS [4].

Performing the Fourier transform of Eq. S11 yields the particle-distribution/undulation coupling:

$$
E_{c}(\lbrace h_{q,a}\rbrace, \lbrace p_{q,a}\rbrace) = -\sum_{\lbrace q>0\rbrace} \frac{A_p \Delta c_{0,p} \kappa_m q^2 p_{q,a} h_{q,a}}{A}.
$$
 (S12)

with $p_{q,a}$ the Fourier coefficient for the density of lipid with spontaneous curvature differ $\Delta c_{0,p} = c_{0,p} - c_{0,\text{background}}$. Combining Eqs. S9, S10, and S12, the total energy is

$$
E_a = E_H(\{h_{\mathbf{q},a}\}) + E_\rho(\{p_{\mathbf{q},a}\}) + E_c(\{h_{\mathbf{q},a}\}, \{p_{\mathbf{q},a}\}).
$$
\n(S13)

The expectation of $h_{q,a}^2$ is determined by

$$
\langle h_{\mathbf{q},a}^2 \rangle = Z^{-1} \int \mathrm{d}h_{\mathbf{q},a} \mathrm{d}p_{\mathbf{q},a} h_{\mathbf{q},a}^2 e^{-\beta E_a}, \tag{S14}
$$

where

$$
Z = \int \mathrm{d}h_{\mathbf{q},a} \mathrm{d}p_{\mathbf{q},a} \mathrm{e}^{-\beta E_a}.\tag{S15}
$$

For a single-component membrane ($\chi = 0$ or $\Delta c_{0,p} = 0$), integration leads to

$$
\langle h_{\mathbf{q},a}^2 \rangle = \frac{Ak_{\rm B}T}{\kappa_{\rm b}q^4} \quad \text{(single component)},\tag{S16}
$$

from which the bending rigidity can be determined. For the two-component mixture, the expectation of $h_{q,a}^2$ is now:

$$
\langle h_{\mathbf{q},a}^2 \rangle = \frac{Ak_{\rm B}T}{\kappa_{\rm b}q^4} (1 + \frac{\Delta c_{0,\rm p}^2 A_p \kappa_{\rm b} \chi \left(1 - \chi\right)}{2k_{\rm B}T} + \mathcal{O}[\Delta c_{0,\rm p}^3]) \tag{S17}
$$

Thus, a bilayer with a symmetric or asymmetric mixture of lipids (with unequal spontaneous curvatures) will experience apparent softening according to

$$
\kappa_{\text{apparent}} = \kappa_{\text{b}} \left(1 - \frac{\Delta c_{0,\text{p}}^2 A_p \kappa_{\text{b}} \chi \left(1 - \chi \right)}{2k_{\text{B}} T} \right)
$$
\n
$$
= \kappa_{\text{b}} \left(1 - \alpha \right). \tag{S18}
$$

where α is the same as in Eq. 2 of the main text.

5.1 Fourier amplitudes of variations in lipid density

The derivation for the free-energy/entropic variation of the Fourier amplitudes of the density modes follows below.

Consider a bilayer with mol fraction χ of lipid A and $1-\chi$ of background lipids. Break the bilayer into M narrow strips that are long in y with length L_y and narrow in x with width d. There are N lipids in each strip. The expected number of lipid A, $\langle n_A \rangle$ in each narrow strip is χN . Given random fluctuations, the variance of n_A is $\chi(1-\chi)N$ according to the binomial distribution. For convenience define the density fluctuation function $\rho_A(i)$ as:

$$
\Delta n_A(i) = n_A(i) - \chi N \tag{S19}
$$

If the number of lipids in bin i, $n_A(i)$ is independent of $n_A(j)$, the autocorrelation function of $\Delta n_A(i)$, $r_{nn}(j-i)$ is

$$
r_{nn}(j-i) = \begin{cases} \chi(1-\chi)N & \text{: } j=i\\ 0 & \text{: } j\neq i \end{cases} \tag{S20}
$$

By the Wiener-Kninchin theorem, the power spectral density $S(f)$ is

$$
S(f) = \sum_{k=-\infty}^{\infty} r_{nn}(k) e^{-i2\pi f k}
$$

= $\chi(1-\chi)N,$ (S21)

that is, $S(f)$ is independent of f; it is "white noise." Given the power spectral density $S(f)$ in Eq. S21, discrete variables can be translated to the continuous case of interest: the Fourier transform of the lateral distribution of lipid A density, p_q . Dividing by the bin area dL_y , transforming from f to q with $f = \frac{qL_x}{2\pi}$, and multiplying by 2π (by the convention in which the factor of $(2\pi)^{-1}$ is in the *inverse* transform) yields:

$$
S(q) = (2\pi) \frac{L_x}{2\pi} \frac{1}{dL_y} S(f) \tag{S22}
$$

$$
=L_x \rho_A (1-\chi) \tag{S23}
$$

Where here $\rho_A = \frac{N\chi}{dL_y}$ is the number of lipids per unit area. Instead of restricting analysis to x, performing the two-dimensional Fourier analysis with $q = \{q_x, q_y\}$ yields

$$
S(\mathbf{q}) = (1 - \chi)A\rho_A \tag{S24}
$$

Given a sufficiently large bilayer, we now substitute a normal distribution with matching variance. The probability $pr(p_q)$ of observing density fluctuation p_q is given by

$$
\text{pr}(p_q) \propto \exp\left(-\frac{p_{q,a}^2 + p_{q,b}^2}{(1-\chi)A\rho_A}\right) \tag{S25}
$$

where A_p is the area of the lipid, $p_{q,a}$ and $p_{q,b}$ are defined as for $h_{q,a}/h_{q,b}$ above. The expectation value $\langle p_q^2 \rangle$:

$$
\langle |p_{\mathbf{q}}|^2 \rangle = Z^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dp_{\mathbf{q},a} dp_{\mathbf{q},b} (p_{\mathbf{q},a}^2 + p_{\mathbf{q},b}^2) e^{-\frac{p_{\mathbf{q},a}^2 + p_{\mathbf{q},b}^2}{(1 - \chi)A\rho_A}}
$$
(S26)

$$
= (1 - \chi)A\rho_A \tag{S27}
$$

with

$$
Z = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dp_{\mathbf{q},a} dp_{\mathbf{q},b} e^{-\frac{p_{\mathbf{q},a}^2 + p_{\mathbf{q},b}^2}{(1-\chi)A\rho_A}}
$$
(S28)

has yielded the correct variance.

6 Tabulated simulation details

Table S5: Key descriptors of the simulations analyzed in this work.

References

- [1] Kaufman L. and Rousseeuw P. Clustering by means of Medoids. In Statistical Data Analysis Based on the L1 Norm and Related Methods, pages 405–416, 1987.
- [2] Frank L.H. Brown. Elastic modeling of biomembranes and lipid bilayers. Annual Review of Physical Chemistry, 59:685–712, 2008.
- [3] Markus Deserno. Fluid lipid membranes: From differential geometry to curvature stresses. Chemistry and Physics of Lipids, 185:11–45, 2015.
- [4] Kayla C. Sapp, Andrew H. Beaven, and Alexander J. Sodt. Spatial extent of a single lipid's influence on bilayer mechanics. *Physical Review E*, 103(4), 2021.