May 29, 2014

1 Modeling of Bilayer Structure

Modeling of the bilayer structure was done similarly to the SDP model [1]. The lipid component groups are shown in Fig. We assigned two Gaussian functions for a headgroup, one representing a sum of phosphate and choline groups and another representing a sum of carbonyl and glycerol groups. Another Gaussian was assigned to the terminal methyl group. The hydrocarbon was represented by an error function minus the terminal methyl Gaussian. To avoid having too many parameters, we combined methylece (CH_2) and methine (CH). Tat was represented by a Gaussian. In our modeling procedure, we assumed that the lipid headgroup is somewhat rigid, so that the distance between PC and CG groups is fixed as well as the distance between CG and the end of the hydrocarbon chains. These amounted to fixing dX_1 and dX_2. We allowed the chains to change its length freely, leading to free D_c. The center of the terminal methyls was fixed at the center of the bilayer. The constraints on component group volumes were taken from [2]. Table 1 shows all the constraints. Because our experiment was carried out at 37 ℃while [2]

	Braun et. al	DOPC	DOPC:DOPC (3:1)
R_{12}	141.7	57.94	98.4°
r_{12}	145	57.8	98.2°
r	?	?	?

Table 1: Lattice constants

is for 30 ℃, our measured volume for DOPC was higher. The difference was assigned equally to hydrocarbon chains and terminal methyls. This led to distributing the measured extra volume to 17 carbon groups. This distribution resulted in slightly different values for r_{12}. The headgroup volume was assumed to be the same phosphatidylcholine headgroup measured for the gel phase of DPPC [**?**].

Using the model described above, we first fitted the DOPC form factor. The result of the best fit is shown in Table. For fitting the data with Tat, we considered two cases: 1) free lipid component widths and 2) lipid component widths assumed to be greater than the corresponding pure DOPC values. The second case avoided increasing the number of free parameters. In this

case, we assumed that Tat decreases the thickness of the bilayer locally, which leads to smearing of the headgroups in z direction, which necessarily leads to the wider headgroup.

The width of Tat was fixed. We fitted with three different values to study the range of variation due to the Tat width. The choice was made based on MD simulation results.

We also investigated a possibility of Tat in the hydrocarbon region. For this, because Tat occupies some volume in the chain region, we subtracted Tat volume from the error function. Then, this modified error function was used as the hydrocarbon chains. Figure shows that χ^2 is equally good at Tat inside and outside the chain region for case 1. For case 2, the minimum χ^2 was obtained with Tat in the headgroup region. While modeling suggested that Tat could be in the tail region, this position seems energetically unfavorable as Tat is a hydrophilic molecule. With a help of MD simulations, we were able to discard the interior position as an artifact of our modeling. The simulation results are discussed in section X.

case 1 led to a deeper position of Tat inside the bilayer.

2 MD simulation

Three systems with different petpide concentrations (DOPC:Tat = 128:0, 128:2, and 128.4) were studied with Gromacs 4.6.1 package [3]. DOPCs were modeled by Slipid force field [4, 5] and HIV-Tats were modeled by Amber99SB [6]. The systems were simulated at 310 K with a constant area in the x-y plane. The z direction was coupled to 1 atm with constant pressure. The center of mass (COM) distance between each peptide and the bilayer was constrained by an umbrella potential with a force constant of 3000 kJ/mol/nm^2. Each system was explored 18 independent simulations as a combination of 3 different constant area and 6 different peptide insertion depths (except the pure DOPC system).

$$z_{\mathrm{cm}} = \frac{\sum_{i=1}^{N} m_i z_i}{\sum_{i=1}^{N} m_i} \tag{1}$$

The center of mass constraint was applied through an external force field, which derives from an added pontential energy of the system. The pontential

is like a spring, where it depends on the deviation of the distance between the center of mass of Tat and DOPC from a prefered value, z_0,

$$U(z_1^{\mathrm{Tat}}, \ldots, z_1^{\mathrm{DOPC}}, \ldots) = -\frac{1}{2}k\big(z_{\mathrm{cm}}^{\mathrm{Tat}} - z_{\mathrm{cm}}^{\mathrm{DOPC}} - z_0\big)^2 \tag{2}$$

Then, $-\partial U/\partial z_i$ is equal to the external force acting on atom, i. Before applying this constraint, Tats were attached to the bilayer from the water region. During the first 20 ns for pre-equilibration, Tats were allowed to change their configuration, which resulted in different configurations for each Tat when attached to the bilayer. It is possible that the Tats configuration inside the bilayer at the end of simulations was affected by this initial configuration of Tats. Instead of preparing many simulations with different initial Tat configurations, we avaraged over all Tats present in the system. We also performed many simulations with different A_L and z_{Tat} to investigate how robust some of the Tat structural features are. Many simulation results are shown in the appendix of this thesis.

3 Local Thinning of Membranes

The SIMtoEXP program only gives the average quantities for each leaflet. While our x-ray data are sensitive to the avarage bilayer electron density, local information of Tat-DOPC interactions can be obtained from MD simulations. In this section, we discuss a method to extract a local membrane thickness.

The presence of Tat may result in compression of lipid bilayer along z-direction. If so, the phosphorus-phosphorus distance of the bilayer near Tat (denoted by DPP' in the figure below) may be different from that away from Tat (DPP). For small Tat concentration, DPP would be the same as that of pure DOPC if the distance from all Tats is large enough. Of course, for our concentrations, the thinning effect may extend throughout the bilayer because the lateral effect of Tat might have a larger lateral decay length than the distance between Tats. Whether that is the case or not, one would expect that the thickness near the Tats is smaller than the average thickness.

So, we want to measure DPP'. First, let us define what we mean by lipids close to Tat. As in Fig. 1, we imagine a cylinder around Tat and pick up all the phosphorus atoms within it. Approximating Tat as a cylinder with its height given by the FWHM of its electron density distribution, its radius $R_{\mathrm{Tat}} = 9$ Å comes from the volume of Tat $= 1876$ Å^3 and $h_{\mathrm{Tat}} = 7.6$ Å

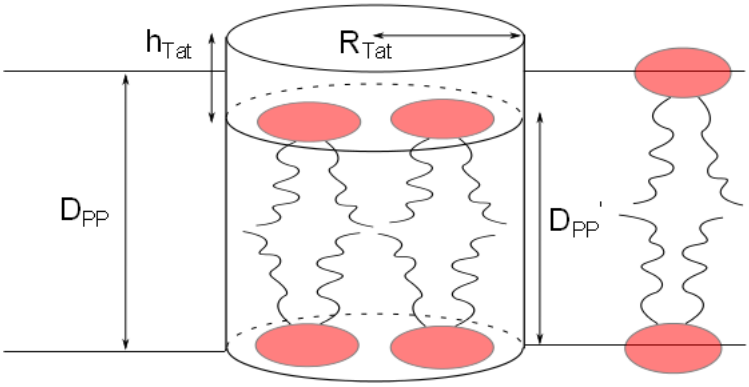

Figure 1: test

measured from one of the simulations. Let us define the lateral center of the cylinder in some way - a crude approximation would put it at the arginine in the middle of the amino acid sequence. Then let us define DPP' using only those lipids whose phosphorus atoms lie within these 9Å cylinders around the tats. Then $D_{\mathrm{PP}} = Z^{+}_{\mathrm{phos}} - Z^{-}_{\mathrm{phos}}$ where Z^{+}_{phos} and Z^{-}_{phos} are the average Z of the n1 (n2) lipids in the upper and lower monolayer, respectively. To be more precise, assume that the arginine in the middle of the amino acid sequence is at the center of the cylinder. For a refined method, we could find the center of mass of each Tat and use them as the lateral center of cylinders (instead of a particular carbon atom in an arginine).

The algorithm for doing this is straightforward. For each time frame (snapshot), the positions (X_i ,Y_i, Z_i) of each Tat, i, are listed. We choose phosphorus atoms whose (X, Y) lateral position lies within 9 Å of any one of the Tat's lateral position. Then, Z position of the chosen phosphorus atoms are placed in a list. Then, Z_{phos} are calculated from the list. The number of selected phosphorus atoms in each monolayer was also recorded. This value gives local lateral depletion if the Tat cylinders are assumed not to overlap. After averaging over many snapshots, there will hopefully be decent statistics for local thinning and local difference between Tat and phosphorus Z positions, and maybe local lateral depletion if overlaps are taken into account.

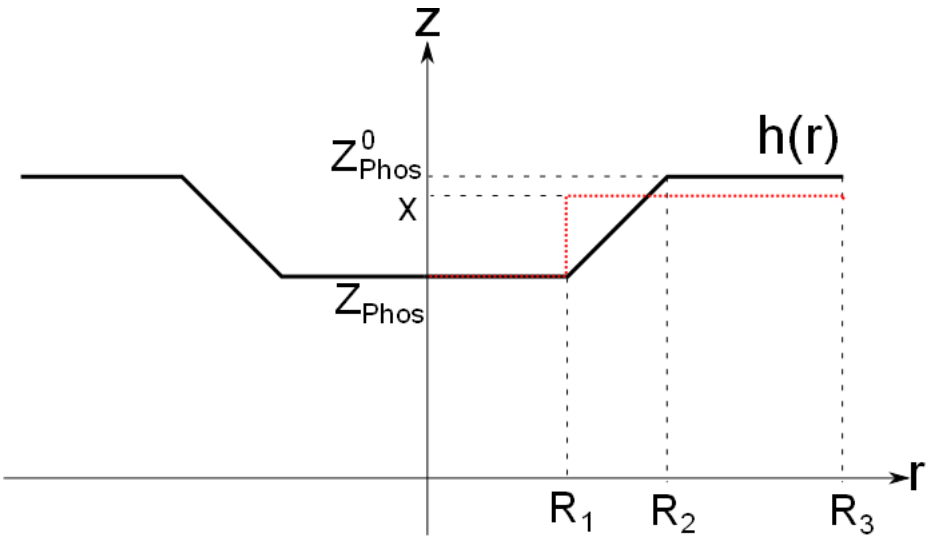

Figure 2: test

4 Linear Model

As described in the previous section, a Tat is modeled as a cylinder with its radius equal to R_1, height h_{Tat}, and volume V_{Tat} such that $R_1 = \sqrt{V_{\text{Tat}}/(\pi h_{\text{Tat}})}$. Let $h(r)$ represent the phosphorus height profile of a leaflet. The two leaflets are assumed to be decoupled. In the linear model, lipids are seprated into three regions: suppressed, boundary, and unperturbed (see Fig. for the definitions of various quantities). In the suppressed region, lipids are uniformly compressed by Tat toward the center of the bilayer; $h(r)$ is a constant, Z_{phos}. In the boundary region, $h(r)$ is assumed to linearly increase with the distance r from the center of the Tat. In the unperturbed region, lipids do not interact with Tat, behaving identially to DOPC, so the phosphorus position is the same as that of DOPC. A continuous $h(r)$ that satisfies the above criteria is

$$h(r) = \begin{cases} Z_{\text{phos}} & \text{if} \quad 0 \le r < R_1 \\ mr + b & \text{if} \quad R_1 \le r < R_2 \\ Z^0_{\text{phos}} & \text{if} \quad R_2 \le r < R_3 \end{cases} \tag{3}$$

with $m = (Z_{\text{phos}} - Z^0_{\text{phos}})/(R_1 - R_2)$ and $b = (Z^0_{\text{phos}} R_1 - Z_{\text{phos}} R_2)/(R_1 - R_2)$. Assuming that the simulation box is a cylinder, we have $R_3 = \sqrt{N A_{\text{L}}/\pi}$. Z_{phos} can be measured directly from simulation trajectories. Z^0_{phos} is a half of the average phosphorus-phosphorus distance in a DOPC simulation, which

can be easily obtained from the SIMtoEXP program. The average height profile over the monolayer, $\langle h(r)\rangle$, can be also obtained from the program in the same manner. The only unknown is R_2, which is the relaxation length of DOPC-Tat interaction.

Let us calculate $\langle h(r)\rangle$. In the cylindrical coodinates,

$$\langle h(r)\rangle = \frac{1}{\pi R_3^2}\int_0^{2\pi} d\phi \int_0^{R_3} rh(r)dr \tag{4}$$

The ϕ integration is trivial. The r integration is

$$\begin{aligned}
&\int_0^{R_3} rh(r)dr \\
&= \int_0^{R_1} Z_{\rm P} r dr + \int_{R_1}^{R_2} (mr+b) r dr + \int_{R_2}^{R_3} Z_{\rm P}^0 r dr \\
&= \frac{1}{2}\left[Z_{\rm P}R_1^2 + Z_{\rm P}^0(R_3^2 - R_2^2)\right] + \frac{1}{3}m\left(R_2^3 - R_1^3\right) + \frac{1}{2}b\left(R_2^2 - R_1^2\right) \\
&= \frac{1}{2}\left[Z_{\rm P}R_1^2 + Z_{\rm P}^0(R_3^2 - R_2^2)\right] + \frac{1}{3}\left(Z_{\rm P}^0 - Z_{\rm P}\right)\left(R_2^2 + R_1R_2 + R_1^2\right) \\
&+ \frac{1}{2}\left(Z_{\rm P}R_2 - Z_{\rm P}^0 R_1\right)(R_1 + R_2)
\end{aligned} \tag{5}$$

After grinding through algebra,

$$\langle h(r)\rangle = \frac{\left(Z_{\rm P} - Z_{\rm P}^0\right)\left(R_1^2 + R_1R_2 + R_2^2\right) + 3Z_{\rm P}^0 R_3^2}{3R_3^2} \tag{6}$$

Eq. 6 is a quadratic equation in terms of R_2. Solving for R_2 gives

$$R_2 = \frac{-R_1 + \sqrt{R_1^2 + 4C}}{2} \tag{7}$$

with

$$C = \frac{3R_3^2\left(Z_{\rm P}^0 - \langle h(r)\rangle\right)}{Z_{\rm P}^0 - Z_{\rm P}} - R_1^2 \tag{8}$$

The model quantities calculated from the MD simulations are shown in Table.

5 Some Lipid Bilayer Data

We also estimated the structure by fitting the experimental form factors to a model using the SDP method with the component groups identified in Fig.

(what The positions of these groups were free parameters and the agreement with the experimental form factors was excellent (see Fig. S.M. 5). Absolute total electron density profiles and the Tat profiles are shown for many samples in Fig. 6 (A-C). It must be emphasized, however, that, while the total EDP is well determined by this fitting procedure. Indeed, there are local minima in the fitting landscape, including one with Tat closer to the center of the bilayer as shown in Fig. S.M. 5. The simulations help to discard that result. For the results shown in Fig. 6, a consistent trend is that Tat moves away from the bilayer center as concentration increases.

Figure shows that A_L as defined by $(V_L - V_{HL})/D_c$ decreases as the amount of DOPE increases for systems without Tat. This is consistent with the previous studies (or predictions?) and attributed to the small size of PE head group. Because DOPE has smaller head group than DOPC, lipids in DOPC/DOPE bilayers pack more compactly than DOPC bilayers do, leading to a smaller A_L. Consequently, bilayers composed of DOPC and DOPE tend to have a higher order parameter than DOPC alone. (NO THEY DON'T. WHAT'S GOING ON HERE?). Similarly, the thickness of bilayers is larger at higher PE content.

Figure shows that Tat is located further out from the bilayer center with higher content of PE lipids. This is also consistent with MD simulation PMF, which showed that arginine insertion cost more energy for PE membrane than PC membrane, the result of which was attributed to more possible hydrogen bonding between PE group and arginines.

More structural detail from the modeling and from the simulations is shown in Fig. 7. The bilayer thickness can be described as DHH, which is the distance between the maxima in the electron density profile, or as DPP, which is the distance between the phosphocholines on the opposing monolayers. Figs. 7A and 7B show that both these quantities decrease with increasing Tat mole fraction (P/(L+P)), showing that Tat thins membranes, increasingly so as its concentration is increased, even though both simulation and modeling suggest that Tat moves further from the membrane center with increasing concentration as shown in Fig. 7D. Fig. 7C shows that the area per lipid AL usually increases with increasing mole fraction of Tat, as would be expected from consideration of conservation of lipid volume. Interestingly, the bilayer thickness did not increase for DOPC/DOPE (3:1) bilayers with x less than 0.03.

Figure 8 shows that the Sxray orientational order parameters generally decreases with increasing concentration of Tat for most of the membrane

mimics studied. These decreases in membrane chain order are compatible with the increase in softening of membranes by Tat observed by a decrease in the bending energy, KC, in Fig. 2.

References

[1] Norbert Kuerka, John F. Nagle, Jonathan N. Sachs, Scott E. Feller, Jeremy Pencer, Andrew Jackson, and John Katsaras. Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data. *Biophysical Journal*, 95(5):2356 – 2367, 2008.

[2] Anthony R. Braun, Jonathan N. Sachs, and John F. Nagle. Comparing simulations of lipid bilayers to scattering data: The gromos 43a1-s3 force field. *The Journal of Physical Chemistry B*, 117(17):5065–5072, 2013.

[3] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008.

[4] Joakim P. M. Jmbeck and Alexander P. Lyubartsev. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *The Journal of Physical Chemistry B*, 116(10):3164–3179, 2012.

[5] Joakim P. M. Jmbeck and Alexander P. Lyubartsev. An extension and further validation of an all-atomistic force field for biological membranes. *Journal of Chemical Theory and Computation*, 8(8):2938–2948, 2012.

[6] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.