

Measurement of Something

by

Kiyotaka Akabori

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

at

Carnegie Mellon University

Department of Physics

Pittsburgh, Pennsylvania

Advised by Professor John F. Nagle
and Professor Stephanie Tristram-Nagle

June 23, 2014

Contents

1	Introduction	1
2	Materials and Methods	3
2.1	X-ray optics	3
2.2	Hydration Chamber	3
2.3	Sample Preparation	4
2.3.1	Stock Solutions	4
2.3.2	Thin Film Samples	4
2.4	CCD detector	5
3	Structural Perturbation of Lipid Bilayers Due to Tat Peptide	6
3.1	Introduction	6
3.2	Materials and Methods	9
3.2.1	Volume Measurements	9
3.2.2	Analysis of Diffuse Scattering	11
3.2.3	Modeling the Bilayer Structure	17
3.2.4	Molecular Dynamics Simulation	23
3.3	Analysis of Molecular Dynamics Simulation Data	26
3.3.1	SIMtoEXP program	26
3.3.2	Local Thinning of Membranes	26
3.3.3	Lateral Decay Length of Membrane Thinning	28
3.4	Results	30
3.4.1	Bending and Bulk Modulus	30
3.4.2	Volume results	33
3.4.3	Electron Density Profile Modeling	33
3.4.4	Hard Wall Constrain Fits	44

3.4.5	Summary of Electron Density Profile Modeling	45
3.4.6	Molecular Dynamics Simulations	47
3.5	Discussion	56
3.6	Conclusion	59
4	Ripple Phase	58
4.1	Introduction	58
4.2	Materials and Methods	61
4.2.1	Sample Preparation	61
4.2.2	Low Angle X-ray Scattering Experiment	62
4.2.3	Near Grazing Incidence Wide Angle X-ray Scattering Experiment	64
4.2.4	Transmission Wide Angle X-ray Scattering Experiment . . .	67
4.3	Some Theories	73
4.3.1	Lattice Structure	73
4.3.2	Sample q-space	74
4.3.3	Lorentz Correction	76
4.3.4	Absorption Correction for LAXS	81
4.3.5	Absorption Correction for WAXS	83
4.3.6	Effect of mosaic spread	83
4.4	Model	83
4.4.1	Contour Part of the Form Factor	83
4.4.2	Transbilayer Part of the Form Factor	84
4.5	Results	86
4.5.1	Data and Electron Density Profile	86
4.5.2	Near Grazing Incidence Wide Angle X-ray Scattering (NGI-WAXS)	86
4.5.3	Transmission WAXS	87
4.6	Discussion	87
4.7	Conclusion	90
Appendices		91
A Tat		92
A.1	Analysis of Fixed Angle Data using NFIT	92
A.1.1	Theory	92

A.1.2	Results	92
A.2	Proper Incorporation of Mosaic Spread to NFIT analysis	92
A.2.1	Mosaic Spread: Calculation	92
A.2.2	Mosaic Spread: Experiment	95
A.2.3	Results	96
A.3	Domain Size Distribution: Gaussian and Exponential	96
A.4	Hard Wall Constraints in SDP	96
A.5	Some More Details of Tat Stuff	96
B	Ripple Phase	97
B.1	Derivation of the contour part of the form factor	97
B.2	Rotation of a Two-Dimensional Function	99
B.3	Derivation of the transbilayer part of the form factor in the 2G hybrid model	100
B.4	Correction due to refractive index	102

List of Tables

3.1	Some Amino Acids Data	10
3.2	Number of electrons per lipid and volume per lipid.	22
3.3	Some structural parameters for each component. n_i^e is the number of electrons and ρ_i is the average electron density.	22
3.4	Tat basic structural parameters. The notations are the same as in Table 3.3. $x_{\text{Tat}} = \text{Tat}/(\text{Tat}+\text{Lipid})$	22
3.5	Volume results at 37 °C	33
3.6	Fitting Results for DOPC membranes for the THG (Tat in headgroup) model. $z_{\text{PC}} - z_{\text{CG}} = 3.1 \text{ \AA}$ and $z_{\text{CG}} - z_{\text{HC}} = 1.3 \text{ \AA}$ in all fits.	36
3.7	Fitting Results for DOPC:DOPE (3:1) membranes for the THG model. $z_{\text{PC}} - z_{\text{CG}} = 3.1 \text{ \AA}$ and $z_{\text{CG}} - z_{\text{HC}} = 1.3 \text{ \AA}$ in all fits.	36
3.8	(Numbers are wrong) Fitting Results for DOPC:DOPE (1:1) membranes for the THG model. $\Delta z_1 = z_{\text{PC}} - z_{\text{CG}}$ and $\Delta z_2 = z_{\text{CG}} - z_{\text{HC}}$	39
3.9	Fitting Results of the bound THG model for DOPC membranes. $\Delta z_1 = z_{\text{PC}} - z_{\text{CG}}$ and $\Delta z_2 = z_{\text{CG}} - z_{\text{HC}}$	45
3.10	Comparison of the simulated form factors to the experimental form factors.	52
3.11	Summary of simulation results. $\langle D_{\text{PP}} \rangle$, phosphorus-phosphorus distance averaged over all lipids; D_{PP} , Tat-perturbed phosphorus atoms; x , thickness away from Tat; Δt , $\langle D_{\text{PP}}^{\text{DOPC}} \rangle - D_{\text{PP}}$; H_{Tat} , Tat height; R_{Tat} , radius of Tat cylinder; R_2 , radius of the calculated in-plane Tat-perturbed region; R_3 , effective radius of the simulation box.	53
3.12	Summary of weighted average results. The caption is the same as Table 3.11.	53

4.1	Lattice constants for DMPC at $T = 18.0$ °C reported by Wack and Webb [1].	60
4.2	Definitions of Z_{CH_2} and Z_{W}	85
4.3	Lattice constants	86

List of Figures

3.1 LAXS of DOPC:DOPE (1:1) with $x_{\text{Tat}} = 0.034$ at 37 °C. White lobes of diffuse scattering intensity have large grey numbers, while lamellar orders and beam are shown to the left of the molybdenum beam attenuator (short, dark rectangle). q_z and q_r are the cylindrical coordinates of the sample q -space, where q_z -axis is along the bilayer normal and q_r -axis is along the in-plane direction. The lamellar repeat spacing was $D = 66.2 \text{ \AA}$	12
3.2 Schematic of an oriented stack of lipid bilayers. Thick green curves represent an instance of thermally fluctuating bilayers. The dashed lines show the thermally averaged positions $z = nD$ of the centers of each bilayer and $u_n(x, y)$ gives the instantaneous deviation from the average. Each bilayer extends in the $\mathbf{r} = (x, y)$ plane.	13
3.3 Expanded view of a fluctuating bilayer. Along the two black solid lines, the electron density profile is identical in an incompressible bilayer. Along the dashed line, the bilayer appears thicker by a factor $1/\cos \alpha$. This apparent thickness variation along the z direction is corrected by the undulation correction.	14
3.4 Schematic of DOPC showing each lipid component. The dash lines show where the lipid is divided into different components. The lipid headgroup is divided into two components, phosphate-choline (PC) and carbonyl-glycerol (CG). The hydrocarbon chain region is also divided into two components, methylene+methine (CH_2+CH) and terminal methyl groups (CH_3).	18
3.5 A model electron density profile for DOPC with Tat.	19

3.6 Our simple model to extract the local bilayer thickness from simulation trajectories. Tat is modeled as a cylinder with its height H_{Tat} and radius R_{Tat} . The local thickness is defined as $D'_{\text{phos-phos}}$. The thickness of the unperturbed DOPC bilayer is $D_{\text{phos-phos}}$. Blue highlighted lipids fall within the imaginary cylinder extended from the Tat. Unperturbed lipids are highlighted in green.	27
3.7 Simple model of the lateral decay of the membrane thickness perturbation due to Tat.	29
3.8 Bilayer bending modulus, K_c , vs. Tat mole fraction x_{Tat} . D -spacings for DOPC/Tat mixtures varied from 64 to 68 Å, for DOPC/DOPE/Tat mixtures from 64 to 69 Å, for DOPC/DOPS/Tat (3:1) mixtures from 57 Å to 100 Å (pure DOPS was unbound), and for nuclear mimic/Tat mixtures from unbound (nuclear mimic) to 64 Å. Estimated uncertainty in all values is about ± 2	31
3.9 Form factors of lipid mixtures (arbitrarily scaled and vertically displaced) with increasing Tat mole fractions x_{Tat} indicated on figure legends. Lipid mixtures: A. DOPC B. DOPC/DOPE (3:1) C. DOPC/DOPE (1:1) D. DOPC/DOPS (3:1) E. Nuclear mimic. The entire q_z range is shown in C, while others show partial ranges. Solid vertical lines indicate the q_z values where the form factors equal zero between the lobes of diffuse data.	32
3.10 The best fits to DOPC form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).	35
3.11 The best fits to DOPC:DOPE (3:1) form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).	37
3.12 The best fits to DOPC:DOPE (1:1) form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).	38
3.13 Modeling results for absolute electron density profiles and for the Tat location as a function of distance z along the bilayer normal. A. DOPC B. DOPC:DOPE (3:1), and C. DOPC:DOPE (1:1).	40

3.14 A. Bilayer thickness, D_{PP} ; B. Bilayer thickness, D_{HH} ; C. Area/lipid, A_L ; D. Twice the Tat location, $2z_{\text{Tat}}$: all plotted vs. Tat mole fraction x_{Tat} . Error bars are standard deviations from imposing Tat Gaussian widths, $\sigma_{\text{Tat}} = 2.5, 3.0$ or 3.5 \AA . Inverted blue triangles connected with dotted line are results from MD simulations, averaging the best fits to the X-ray data for each parameter, with standard deviations shown.	42
3.15 χ^2 as a function of z_{Tat} for DOPC, DOPC:DOPE (3:1), and DOPC:DOPE (1:1) (from left to right) with $x_{\text{Tat}} = 0.016, 0.034$, and 0.059 (from top to bottom). $\sigma_{\text{Tat}} = 3.0$. The THG model (black squares) and the THC model (red circles).	43
3.16 χ^2 as a function of z_{Tat} for DOPC with $x_{\text{Tat}} = 0.016, 0.034$, and 0.059 (from top to bottom). $\sigma_{\text{Tat}} = 3.0$. The bound THG model was used.	46
3.17 DPP graph with bound fits	47
3.18 DHH graph with bound fits	47
3.19 AL graph with bound fits	47
3.20 zTat graph with bound fits	47
3.21 MD simulated form factors for DOPC at $A_L = 68 \text{ \AA}^2$ (blue solid line), 70 \AA^2 (red solid line), and 72 \AA^2 (green solid line) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for 70 \AA^2	48
3.22 The simulated, symmetrized electron density profile for DOPC at $A_L = 70 \text{ \AA}^2$ as a function of the distance away from the bilayer center. Each component profile is labeled with its name: PC (phosphate-choline), CG (carbonyl-glycerol), CH ₂ +CH (methylene-methine combination), CH ₃ (terminal methyl). The sum of all the components is labeled as total.	49
3.23 MD simulated form factors for DOPC with $x_{\text{Tat}} = 0.015$ at $A_L = 72 \text{ \AA}^2$ (top) and 74 \AA^2 (bottom), with $z_{\text{Tat}} = 18 \text{ \AA}$ (red solid lines), 16 \AA (green solid lines), and 14 \AA (blue solid lines) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for $z_{\text{Tat}} = 18 \text{ \AA}$	50

3.24 MD simulated form factors for DOPC with $x_{\text{Tat}} = 0.030$ at $A_L = 74 \text{ \AA}^2$ (top) and 76 \AA^2 (bottom), with $z_{\text{Tat}} = 18 \text{ \AA}$ (red solid lines), 16 \AA (green solid lines), and 14 \AA (blue solid lines) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for $z_{\text{Tat}} = 18 \text{ \AA}$	51
3.25 Electron density profiles of guanidinium groups from the four best matched simulations for DOPC with $x_{\text{Tat}} = 0.015$ (one Tat on each leaflet). Tat on the lower and upper leaflets are shown on the left and right plots, respectively.	54
3.26 MD simulated form factors (red solid lines in A and C) of Tat/(DOPC+Tat), $x_{\text{Tat}}=0.030$, with Tat fixed at $z_{\text{Tat}}= 18 \text{ \AA}$ (panel A) and 5 \AA (panel C) from the bilayer center compared to experimental form factors (open circles) scaled vertically to provide the best fit to the simulations. Corresponding snapshots are shown in Panels B and D in which the lipid chains are represented as grey sticks on a white background, Tats are yellow, phosphate groups are red and water is blue.	55
3.27 Location of Tat in DOPC bilayer. Tat is represented as a cylinder, z is the distance from the bilayer center, and R is the in-plane distance from the center of Tat. The average z of the lipid phosphates as a function of R and the arginine guanidiniums are shown in red and blue, respectively.	56
4.1 Lattice structure of the DMPC ripple phase. Unit cells are shown in dash lines. Center of bilayers are shown by thick, solid lines. Notations in the figure are a and b : lattice unit vectors, D : D -spacing along z , D_s : stacking distance, λ_r : ripple wavelength, γ : oblique tilt angle, A : ripple amplitude, and x_0 : projected length of the major arm.	59
4.2 A picture of an annealing chamber. Need to take a picture	61
4.3 The horizontal profile of the beam used in the low resolution study.	62
4.4 The vertical profile of the beam used in the low resolution study.	63

4.5	1 second exposure (left) and 60 second exposure (right) of the low angle X-ray scattering from the DMPC ripple phase. The dark rectangle in the right image extending from $q_z = 0 \text{ \AA}^{-1}$ to 0.2 \AA^{-1} is the shadow cast by $100 \mu\text{m}$ thick molybdenum beam stop. $D = 57.8 \text{ \AA}$, $\lambda_r = 145.0 \text{ \AA}$, and $\gamma = 97.8^\circ$. The gray scales used are [0 100] (left) and [0 500] (right).	65
4.6	CCD images of X-ray scattering taken with (left) and without (right) a nominally $25 \mu\text{m}$ thick Mo attenuator. These data were taken at a fixed angle of incidence. The sample was an oriented film of DOPC:DOPE (3:1) in the fluid phase at 37°C . The wavelength was 1.175 \AA , the same as the one used for the ripple phase experiment. The same gray scale is used in both images.	66
4.7	Vertical p_z slices of X-ray images shown in Fig. 4.6 (left). The scattering intensity measured with the attenuator (red solid circles) was multiplied by a factor of 6.9 and compared to the intensity measured without the attenuator (black solid circles, right).	66
4.8	Beam used in the high resolution experiment. The horizontal beam profile (left) and vertical beam profile (right) are shown.	68
4.9	The geometric broadening due to the sample width and the beam width. A top view of the sample (gray), Si wafer (blue), and incoming and diffracted X-rays (bounded by red solid lines) are shown. The horizontal beam width is labelled as Δx_{beam} , the sample width as w_s , the total scattering angle for a lipid chain-chain correlation as 2θ , and the geometric broadening as Δx	68
4.10	Schematics of the sample holder showing a small piece I fabricated . .	69
4.11	Picture of the sample holder looking from the above. A lead tape was attached to the back of the sample holder to help reduce the background scattering, typically coming from the air gap between the flightpath snout and the mylar window of the chamber.	70
4.12	Schematics of the sample holder in the transmission mode.	70
4.13	Sample holder geometry.	70
4.14	Geometric broadening in TWAXS. The cross section of the incoming X-ray with the sample and the CCD detector are both shaded in red.	71

4.15	Top view of geometric broadening in TWAXS. The cross section of the incoming X-ray with the sample is shaded in red.	71
4.16	Side view of geometric broadening in TWAXS. The cross section of the outgoing X-ray with the CCD detector is shaded in red.	72
4.17	Projection of rectangular beam on the detector.	72
4.18	Experimental reflectivity geometry.	75
4.19	Ewald sphere construction for the ripple phase diffraction in the low angle regime. A ripple $k = 0$ peak is the solid, black circle on the q_z -axis. A ripple $k \neq 0$ ring is the black ring centered about the q_z -axis. The portion of the ring that is inside the Ewald sphere is shown as a red dashed line and the portion of the ring that is outside but behind the Ewald sphere is shown as a black dotted line. The magnitude of the total scattering angle is exaggerated. With a wavelength of 1.175 Å, the magnitude $ \mathbf{k}_{\text{in}} = 5.35 \text{ \AA}^{-1}$. For a $h = 5$ peak, $q_{z0}^{\text{z}} = 0.54 \text{ \AA}^{-1}$, one tenth of k_{in}	78
4.20	Side view of an arc of $k = 0$ peak shown as a light blue line.	79
4.21	q -space representations of Bragg peaks and Bragg rings for $h = 1$ and $k = 0, 1$, and 2 in q_{hk}^z planes. The shaded rectangles show cross sections of the rotating Ewald sphere along q_{hk}^z plane. The intersection between the Ewald sphere and a Bragg peak/ring is indicated in red. The observed intensity for the $k \neq 0$ orders is proportional to the fraction of the length of red arcs in the circumference. This fraction is equal to one for a $k = 0$ order. Because the orders are not in the same q_z plane, the range of q_y integration indicated by the height of the rectangle is different for different orders. The magnitude of curvature of arcs is exaggerated.	79
4.22	The path of X-rays within the sample. The incident angle is ω and the total scattering angle is 2θ . An X-ray with a penetration depth of z is shown. The total thickness of the sample is t	81
4.23	NGIWAXS of the DMPC ripple phase for $D = 59.2 \text{ \AA}$ (left) and 60.8 \AA (right). The angle of incidence ω was 0.2° . The black regions around the edge of each image are the q -space that was not probed. The distorted, non rectangular shape of the probed q -space signifies non-linear relation between the CCD space and sample q -space.	87

4.24 Enlarged view of the right image in Fig. 4.23. To show smaller features around the peak, a different contrast is used.	88
4.25 q_r swaths, each averaged over 0.02 \AA^{-1} . The center q_z value of a swatch is shown in the figure legends.	89

Chapter 3

Structural Perturbation of Lipid Bilayers Due to Tat Peptide

3.1 Introduction

The name cell-penetrating peptide (CPP) connotes a peptide that easily penetrates cell membranes (for Reviews see [2–4]).

This thesis focuses on the transactivator of translation, Tat, from the HIV-1 virus, which plays a role in AIDS progression. Earlier work showed that the HIV-Tat protein (86 amino acids) was efficiently taken up by cells, and concentrations as low as 1 nM were sufficient to transactivate a reporter gene expressed from the HIV-1 promoter [5, 6]. It has been reported that Tat protein uptake does not require ATP [7]. Studies using inhibitors of different types of endocytosis, including clathrin and caveolae-mediated, or receptor-independent macropinocytosis reached the same conclusion that ATP mediated endocytosis is not involved in Tat protein permeation [?, 8–10]. However, this issue is controversial, as other studies found evidence for endocytosis in Tat protein import [11–19]. Still other studies have concluded that an ATP requirement for Tat protein entry depends on the size of the cargo attached to Tat protein, or on the specific cell type [20–22]. The part of the Tat protein responsible for cellular uptake was assigned to a short region Tat (48–60), G₄RKKRRQRRRPPQ₆0, which is particularly rich in basic amino acids [7]. Deletion of three out of eight positive charges in this region caused loss of its ability to translocate [7]. In this chapter, short basic regions will be called Tat, while the entire

86 amino acid protein will be called Tat protein. Tat was shown to be responsible for the Tat proteins permeation into the cell nucleus and the nucleoli [7], and this was confirmed using live cell fluorescence in SVGA cells [23]. Tat (48-60) was shown to have little toxicity on HeLa cells at 100 μ M concentration [7], but the longer Tat protein (2-86) was toxic to rat brain glioma cells at 1-10 μ M [24]. Interestingly, no hemolytic activity was found when human erythrocytes were incubated with a highly neurotoxic concentration (40 μ M) of Tat (2-86) [24]. These results prompt the question, what is the mechanism of Tats translocation through membranes? To address this question, many biophysical studies have used simple models of biological membranes composed of a small number of lipid types. These studies are valuable because there is no possibility for ATP-dependent translocation, thus ruling out endocytosis if translocation occurs. For example, Mishra et al. reported that the rate of entry into giant unilamellar vesicles (GUVs) composed of PS/PC (1:4 mole ratio) lipids of rhodamine-tagged Tat is immeasurably slow, but it crosses a GUV composed of PS/PC/PE (1:2:1) lipids within 30 seconds [25]. This study suggests that negative curvature induced by the inclusion of PE facilitates translocation. In a subsequent study using much smaller unilamellar vesicles (LUVs), Tat did not release an encapsulated fluorescent probe in LUVs composed of lipids modeling the outer plasma membrane, PC/PE/SM/Chol (1:1:1:5), but did release the probe in LUVs composed of BMP/PC/PE (77:19:4) [26]; BMP (bis(monoacylglycero)-phosphate) is an anionic lipid specific to late endosomes. In that study [26], the inclusion of PE did not suffice to cause leaky fusion in LUVs in the absence of a negatively charged lipid. The contrasting results in these two experiments may also be due to the use of LUVs instead of GUVs since it was reported that Tat does not translocate across LUVS of PC/PG (3:2) but does translocate across GUVs of the same lipid composition [27]. In a similar experiment, Tat did not translocate into egg PC LUVs [28]. In another experiment confirming these results, Tat did not translocate into GUVs containing only PC with 20 mol% cholesterol, but when PS or PE was included with PC, then rapid translocation of Tat was observed [29]. These experiments demonstrate that the choice of lipids and model systems influences Tat translocation.

Is a pore formed during Tat translocation? Although direct conductance measurements of Tat and lipid membranes have not been carried out, two studies measured conductance with the somewhat similar CPP oligoarginine R₉C peptide. Using single-channel conductance of gramicidin A in planar lipid membranes consisting of anionic,

neutral or positively charged lipids, R₉C did not increase conductance, even in anionic lipid membranes [30]. By contrast, in a similar experiment using planar lipid membranes, a current was induced by R₉C in PC/PG (3:1) membranes, with increasing destabilization over time [31]. Thus questions remain about pore formation of Tat in membranes. In the GUV experiment with Tat mentioned above [29], Ciobanasu *et al.*, using size exclusion methods, suggested a pore in the nanometer range, which could only be passed by small dye tracer molecules. Thus, if a true pore forms, it is likely to be small and transitory.

The secondary structure of Tat has been characterized by many researchers. Ref. [27] carried out Circular dichroism (CD) spectroscopy on a variation of Tat where the penultimate proline on Tat (48-60) was replaced by a tryptophan [27]. Their study found a random coil secondary structure in aqueous solution as well as when Tat was mixed with PC/PG/PE (65:35:5) LUVs. Ziegler *et al.* [10] obtained the same result using CD in PC/PG (3:1) vesicles. In addition, solid state NMR has identified a random coil structure of Tat in DMPC/DMPG (8:7 mole ratio) multibilayers [32]. In the larger Tat-(1-72)-protein NMR measurements at pH 4 have determined there is no secondary structure, with a dynamical basic region [33]. Similarly, NMR was used to study the full Tat protein and found a highly flexible basic region [34]. These previous studies indicate that an alpha helix is not required for Tats translocation ability.

Regarding the mechanism of translocation of this randomly structured, short basic peptide, many models have been proposed based on the conflicting results listed above. Molecular dynamics simulations offer some insight into the molecular details of translocation. Herce and Garcia simulated the translocation of Tat (Y₄₇GRKKRRQRRR₅₇) across DOPC at various lipid:peptide molar ratios [35]. Their simulations indicated that Tat binds to the phosphate headgroups, with 1 Tat binding with 14 lipids, each positive charge on Tat associated with nearly 2 phosphate groups [35]. Translocation involved a localized thinning, and snorkeling of arginine side chains through the hydrophobic layer to interact with phosphates on the other side of the membrane. This allowed some water molecules to penetrate the membrane along with Tat, forming a pore [35]. In this simulation, performed without inclusion of counterions, pore formation was only observed at high ratios of peptide:lipid (1:18) or at elevated temperature. However, a subsequent Gromacs simulation with counterions found no thinning and no pore formation when Tat was added to DOPC

membranes [36]. Instead it found a membrane invagination associated with a cluster of Tat peptides. From their findings, the authors suggested that micropinocytosis could be the model for Tat translocation across membranes [36].

In this thesis, I combine experimental low-angle X-ray scattering (LAXS) data with MD simulations to obtain the structure of fully hydrated, oriented lipid bilayers with Tat (47-57) added at several mole ratios. The lipid systems were DOPC, DOPC/DOPE (3:1 mole ratio), DOPC/DOPS (3:1), DOPC/DOPE (1:1) and a mimic of the nuclear membrane (POPC/POPE/POPS/SoyPI/Chol, 69:15:2:4:11).

3.2 Materials and Methods

3.2.1 Volume Measurements

Multilamellar vesicles (MLVs) were prepared by mixing dried lipid mixtures with MilliQ water to a final concentration of 2-5 wt% in nalgene vials and cycling three times between 20 °C and 60 °C for ten minutes at each temperature with vortexing. Pure Tat was dissolved in water at 0.4 wt%.

Volumes of lipid mixtures with and without peptides in fully hydrated multilamellar vesicles (MLV) were determined at 37 ± 0.01 °C using an Anton-Paar USA DMA5000M (Ashland, VA) vibrating tube densimeter. This instrument measures the average density of a solution and compares it to the density of air using $\rho_s - \rho_0 = k(\tau_s - \tau_0)^2$ where k is an instrumental ??? that depends on the atmospheric pressure.

The Tat peptide sequence used in X-ray experiments and MD simulations was Y₄₇GRKKRRQRRR₅₇. Table 3.1 lists the chemical formulas and molecular weights of these amino acids for convenience. The molecular weight of this sequence is $181.2 + 75.1 + 146.1 + 2 \times 146.2 + 6 \times 174.2 - 10 \times 18 = 1560$. The Tat peptides were synthesized in trifluoroacetic acid, which has the chemical formula CF₃CO₂H, and is made into a powder form by the freeze-dry method. Therefore, each positively charged amino acid such as an arginine and lysine was counter-balanced by a trifluoroacetate (TFA) (C₂F₃O₂). Since Tat has six arginines and two lysines, it came with eight trifluoroacetates. This complex has a molecular weight of $1560 + 113 \times 8 = 2464$. We used the molecular weight of this complex in order to calculate the molarity of Tat correctly. The same molecular weight was also used in preparing oriented samples.

The Tat volume V_{Tat} was calculated from the measured average density of a Tat-

Code	Amino acid	Chemical Formula	Molecular weight (g/mol)
K	Lysine	C ₆ H ₁₄ N ₂ O ₂	146.2
R	Arginine	C ₆ H ₁₄ N ₄ O ₂	174.2
G	Glycine	C ₂ H ₅ NO ₂	75.1
Y	Tyrosine	C ₉ H ₁₁ NO ₃	181.2
Q	Glutamine	C ₅ H ₁₀ N ₂ O ₃	146.1

Table 3.1: Some Amino Acids Data

water solution in the following way. Assuming that Tat molecules in water do not change the volume of water molecules, the density of Tat-water solution is equal to the mass of Tat-water solution divided by the sum of volumes of water and Tat,

$$\rho_{\text{sol}} = \frac{m_w + m_c}{V_w + V_c N_c}, \quad (3.1)$$

where m_w and m_c are the total masses of water and Tat-TFA complex, respectively, V_w is the total volume of water, V_c is the molecular volume of a Tat-TFA complex, and N_c is the total number of this complex in the solution. Denoting $V_w = m_w/\rho_w$ and $N_c = N_A m_c/W_c$, where W_c is the molecular weight of the complex, N_A is the Avogadro's number, and ρ_w is the density of water, we have

$$V_c = \frac{W_c}{\rho_{\text{sol}} N_A} \left(1 + \frac{m_w}{m_c} \left(1 - \frac{\rho_{\text{sol}}}{\rho_w} \right) \right), \quad (3.2)$$

which allows us to calculate the molecular volume of a Tat-TFA complex from the experimentally measured quantities. Assuming that the molecular volume scales with the molecular weight gives the volume of Tat, $V_{\text{Tat}} = 1560/2464 \times V_c \text{ \AA}^3$.

3.2.2 Analysis of Diffuse Scattering

Figure 3.1 shows our typical Low angle X-ray scattering (LAXS) data from oriented stacks of fluctuating bilayers in the fluid phase. Analysis of diffuse scattering intensity patterns like the one shown in Fig. 3.1 results in material parameters such as the bending modulus K_c and bulk modulus B as well as the absolute form factor $|F(q_z)|$. The form factor is the Fourier transform of the bilayer electron density profile $\rho(z)$ and gives us information about the internal structure of the bilayers interacting with Tat peptides.

The form factor $F(q_z)$ is obtained by realizing that the diffuse scattering intensity pattern $I(\mathbf{q})$ is a product of the structure factor $S(\mathbf{q})$ and the form factor; $I(\mathbf{q}) = S(\mathbf{q})|F(q_z)|^2$. This means that once the two dimensional structure factor map $S(\mathbf{q})$ is calculated from a model free energy for bilayer fluctuations, the form factor can simply be calculated by dividing the intensity by the structure factor. Getting the best fit of a model structure factor to the intensity results in the material parameters, K_c and B .

We used an analysis program called NFIT developed by Dr. Yufeng Liu [37–39]

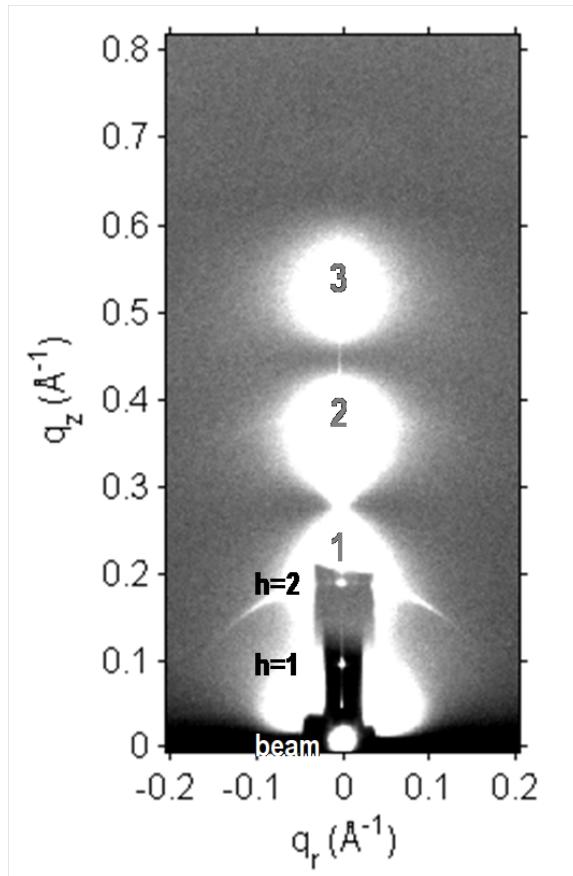


Figure 3.1: LAXS of DOPC:DOPE (1:1) with $x_{\text{Tat}} = 0.034$ at 37 °C. White lobes of diffuse scattering intensity have large grey numbers, while lamellar orders and beam are shown to the left of the molybdenum beam attenuator (short, dark rectangle). q_z and q_r are the cylindrical coordinates of the sample q -space, where q_z -axis is along the bilayer normal and q_r -axis is along the in-plane direction. The lamellar repeat spacing was $D = 66.2 \text{ \AA}$.

to analyze the diffuse scattering and obtain the bending modulus, bulk modulus, and form factor. The details of the analysis are found in Dr. Yufeng Liu's thesis [39]. This section outlines the method.

X-ray scattering essentially reflects the electron density distribution of the sample system. This includes not only the average electron density but also the disorder caused by fluctuations about the average. In our samples, the electron density distribution reflects the stacking of the bilayers, whose average structure can be described as a one-dimensional array of membranes, each of which is a two-dimensional in-plane fluid. A sketch of the membrane stack is shown in Fig. 3.2.

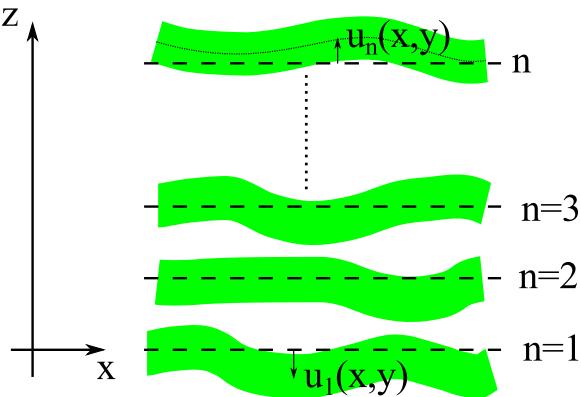


Figure 3.2: Schematic of an oriented stack of lipid bilayers. Thick green curves represent an instance of thermally fluctuating bilayers. The dashed lines show the thermally averaged positions $z = nD$ of the centers of each bilayer and $u_n(x, y)$ gives the instantaneous deviation from the average. Each bilayer extends in the $\mathbf{r} = (x, y)$ plane.

Fluctuations in the stack of the bilayers are described by the quantities $u_n(\mathbf{r})$, which are the spatial deviations of the center of the n -th bilayer from its average position in the z direction at the in-plane location $\mathbf{r} = (x, y)$. Given this description of structure of the lipid bilayer system, we write the electron density ρ_n of the n -th bilayer as

$$\rho_n(z, r) = \rho(z - nD - u_n(\mathbf{r})) \quad (3.3)$$

where $\rho(z)$ is the electron density profile of a single, flat (no fluctuations) bilayer centered at $z = 0$ with its normal in the z direction. Eq. (3.3) is not quite accurate for an incompressible fluid phase bilayers because it ignores the $\cos \alpha$ factor [39] (see Fig. 3.3). To correct for the absence of this factor, the undulation correction [40] was applied to all our form factors in this thesis. For a typical sample, this correction is

about 2% [39]. We then write the electron density of a stack of N bilayers,

$$\rho(\mathbf{R}) = \sum_{n=0}^{N-1} \rho_n(z, \mathbf{r}). \quad (3.4)$$

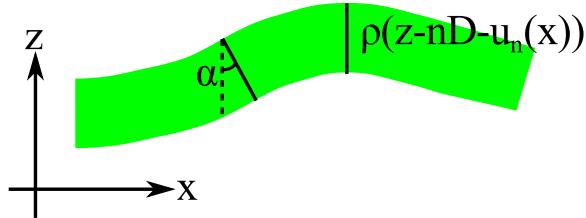


Figure 3.3: Expanded view of a fluctuating bilayer. Along the two black solid lines, the electron density profile is identical in an incompressible bilayer. Along the dashed line, the bilayer appears thicker by a factor $1/\cos \alpha$. This apparent thickness variation along the z direction is corrected by the undulation correction.

Basic X-ray scattering theory in the usual Born approximation gives

$$\begin{aligned} I(\mathbf{q}) &\propto \int d\mathbf{R} d\mathbf{R}' \rho(\mathbf{R}) \rho(\mathbf{R}') e^{i\mathbf{q} \cdot (\mathbf{R} - \mathbf{R}')} \\ &= \left| \int_V d^3\mathbf{R} \rho(\mathbf{R}) e^{i\mathbf{q} \cdot \mathbf{R}} \right|^2, \end{aligned} \quad (3.5)$$

where $\rho(\mathbf{R})$ is the electron density distribution function of a stack of bilayers. Because X-ray scattering measures integrated intensity over time, Eq. (3.5) requires thermal averaging indicated by angular brackets,

$$I(\mathbf{q}) = \left\langle \left| \int_V \rho(\mathbf{R}) e^{i\mathbf{q} \cdot \mathbf{R}} d^3\mathbf{R} \right|^2 \right\rangle. \quad (3.6)$$

Using Eq. (3.3) and (3.4), it can be shown that [39]

$$I(\mathbf{q}) = |F(q_z)|^2 S(\mathbf{q}) \quad (3.7)$$

where

$$S(\mathbf{q}) = \left\langle \left| \sum_{n=0}^{N-1} \int e^{inDq_z + iu_n(\mathbf{r})q_z + i\mathbf{q}_r \cdot \mathbf{r}} d^2\mathbf{r} \right|^2 \right\rangle, \quad (3.8)$$

and

$$\begin{aligned} F(q_z) &= \int_V \rho(z) e^{iq_z z} dz \\ &= \int \rho(z) \cos(q_z z) dz. \end{aligned} \quad (3.9)$$

The second equality in Eq. (3.9) follows for centro-symmetric lipid bilayers. Assuming that $u_n(\mathbf{r}) - u_m(\mathbf{r}')$ has a normal distribution, we write Eq. (3.8) as

$$S(\mathbf{q}) = \sum_{n,m=0}^{N-1} e^{iq_z(n-m)D} \int_V d^2\mathbf{r} d^2\mathbf{r}' e^{i\mathbf{q}_r \cdot (\mathbf{r}-\mathbf{r}')} G(r, r', n, m), \quad (3.10)$$

where

$$\begin{aligned} G &= \left\langle e^{iq_z(u_n(\mathbf{r}) - u_m(\mathbf{r}'))} \right\rangle \\ &\approx e^{-\frac{q_z^2 \langle [u_n(\mathbf{r}) - u_m(\mathbf{r}')]^2 \rangle}{2}}. \end{aligned} \quad (3.11)$$

G is called the scattering pair correlation function. The second approximation is obtained by employing the harmonic approximation [39]. Eq. (3.10) is a general expression for a system whose electron density distribution has the form given by Eq. (3.3) and (3.4).

The membrane fluctuations given by $\langle [u_n(\mathbf{r}) - u_m(\mathbf{r}')]^2 \rangle$ in Eq. (3.11) are calculated from the free energy functional for smectic liquid crystals. The basic degrees of freedom in this free energy include (1) bending of each membrane independently of others in the stack and (2) interactions between membranes in the stack. The bending free energy is proportional to the curvature squared with a bending modulus K_c as shown in the first term in the following free energy equation

$$F_U = \frac{1}{2} \int d\mathbf{r} \sum_{n=0}^{N-1} \left\{ K_c [\nabla_r^2 u_n(\mathbf{r})]^2 + B [u_{n+1}(\mathbf{r}) - u_n(\mathbf{r})]^2 \right\}. \quad (3.12)$$

The second term is a harmonic approximation to the interactions between membranes with a modulus B . From analysis of X-ray data, K_c and B are determined.

From the free energy functional, we calculate the pair correlation function G . We

consider the Fourier representation of the displacement variables u

$$u_n(r) = \sum_{\mathbf{Q}} U(\mathbf{Q}) e^{i\mathbf{Q}_r \cdot \mathbf{r} + iQ_z n D} \quad (3.13)$$

where Q_z takes the discrete value of $Q_z = \frac{2\pi m}{DN}$ ($m = -N/2 + 1, \dots, -1, 0, 1, \dots, N/2$) (where we use \mathbf{Q} to distinguish the Fourier space of the sample from the Fourier space \mathbf{q} of the scattering). The sum in Eq.(3.13) includes both an integration over \mathbf{Q}_r and a sum over Q_z .

Because the model is harmonic, Eq. (3.13) allows us to express the free energy in normal modes. Then, it can be shown that applying the equipartition theorem of statistical physics leads to

$$\langle U(\mathbf{Q}_r, Q_z) U(\mathbf{Q}'_r, Q'_z) \rangle = \frac{1}{A_r N} \frac{k_B T}{K_c Q_r^4 + 4B \sin^2(Q_z D/2)} \delta_{\mathbf{Q}_r + \mathbf{Q}'_r, 0} \delta_{Q_z + Q'_z, 0}. \quad (3.14)$$

$\langle [u_n(\mathbf{r}) - u_m(\mathbf{r}')]^2 \rangle$ can be written in terms of $\langle U(\mathbf{Q})U(\mathbf{Q}') \rangle$. After some lengthy calculus [39], we arrive at

$$\begin{aligned} & \left\langle [u_n(\mathbf{r}) - u_m(\mathbf{r}')]^2 \right\rangle \\ &= \frac{D^2 \eta}{2\pi^2} \int_0^\infty dx \frac{1 - J_0(\sqrt{2x}|r - r'|/\xi) (\sqrt{1+x^2} - x)^{2|n-m|}}{x\sqrt{1+x^2}}. \end{aligned} \quad (3.15)$$

where $\eta = \pi k_B T / (2D^2 \sqrt{K_c B})$, $\xi^4 = K_c/B$, and $J_0(x)$ is the first order Bessel function. Essentially, Eq. (3.7), (3.10), (3.11), and (3.15) relate the X-ray data to the material parameters, K_c and B . Through a non-linear least square fitting, the values of K_c and B that yield the best fit of the model to the data are obtained. From these values, the absolute form factor $|F(q_z)|$ is calculated.

3.2.3 Modeling the Bilayer Structure

In the case of X-rays, the features with the most contrast are the electron-dense headgroups, providing the head-head spacing D_{HH} , and also the terminal methyl groups in the bilayer center with the least electron density. Modeling of the bilayer structure was done similarly to the SDP model written by Dr. Norbert Kucerka when he was a postdoc in the Nagle/Tristram-Nagle lab [41].

Parsing of DOPC into lipid components is shown in Fig. 3.4. The phosphate/choline (PC) and carbonyl/glycerol (CG) components together make up the lipid headgroup whereas the hydrocarbon chain region (HC) is divided into two components, the methylene (CH_2) and methine (CH) group combination (denoted as CH_2+CH) and terminal methyl groups (CH_3). We combine methylene (CH_2) and methine groups (CH) in order to avoid proliferation of fitting parameters.

Functional forms

Our model for the electron density profile (EDP) of Tat/lipid bilayer system consists of five structural subgroups: PC, CG, CH_2+CH , CH_3 , and Tat (see Fig. 3.5). The volume probability distributions of components PC, CG, CH_3 , and Tat are described by Gaussian functions,

$$P_i(z) = \frac{c_i}{\sqrt{2\pi}} \left(\exp\left\{-\frac{(z+z_i)^2}{2\sigma_i^2}\right\} + \exp\left\{-\frac{(z-z_i)^2}{2\sigma_i^2}\right\} \right), \quad (3.16)$$

where c_i is an integrated area underneath the curve and the two parts of the expression describe the two bilayer leaflets.

The hydrocarbon chain region (HC) is represented by error functions,

$$P_{\text{HC}}(z) = \frac{1}{2} [\text{erf}(z, -z_{\text{HC}}, \sigma_{\text{HC}}) - \text{erf}(z, z_{\text{HC}}, \sigma_{\text{HC}})], \quad (3.17)$$

where

$$\text{erf}(z, z_i, \sigma_i) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{z-z_i}{\sqrt{2}\sigma}} dx e^{-x^2}. \quad (3.18)$$

The volume probability distribution for the methylene and methine group combination can then be expressed as

$$P_{\text{CH}_2+\text{CH}}(z) = P_{\text{HC}}(z) - P_{\text{CH}_3}(z). \quad (3.19)$$

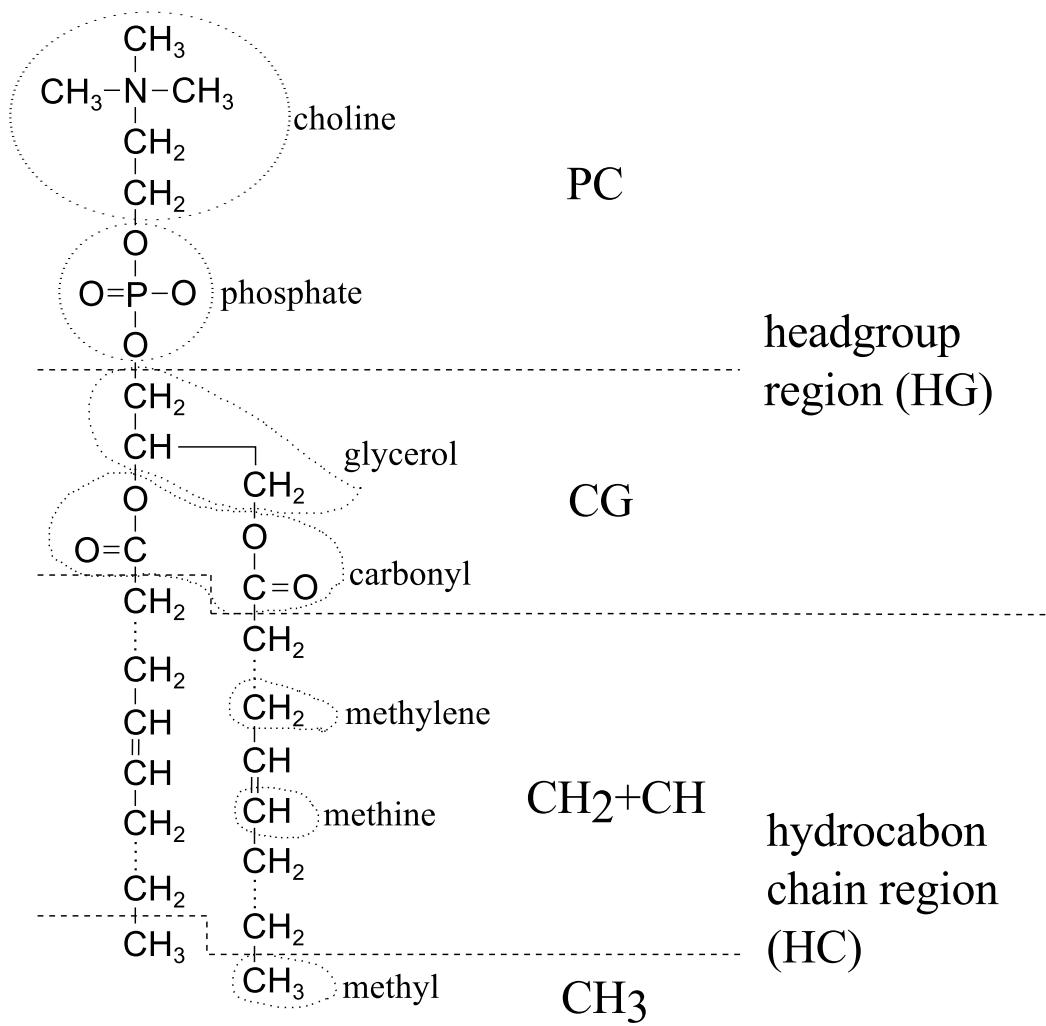


Figure 3.4: Schematic of DOPC showing each lipid component. The dash lines show where the lipid is divided into different components. The lipid headgroup is divided into two components, phosphate-choline (PC) and carbonyl-glycerol (CG). The hydrocarbon chain region is also divided into two components, methylene+methine (CH_2+CH) and terminal methyl groups (CH_3).

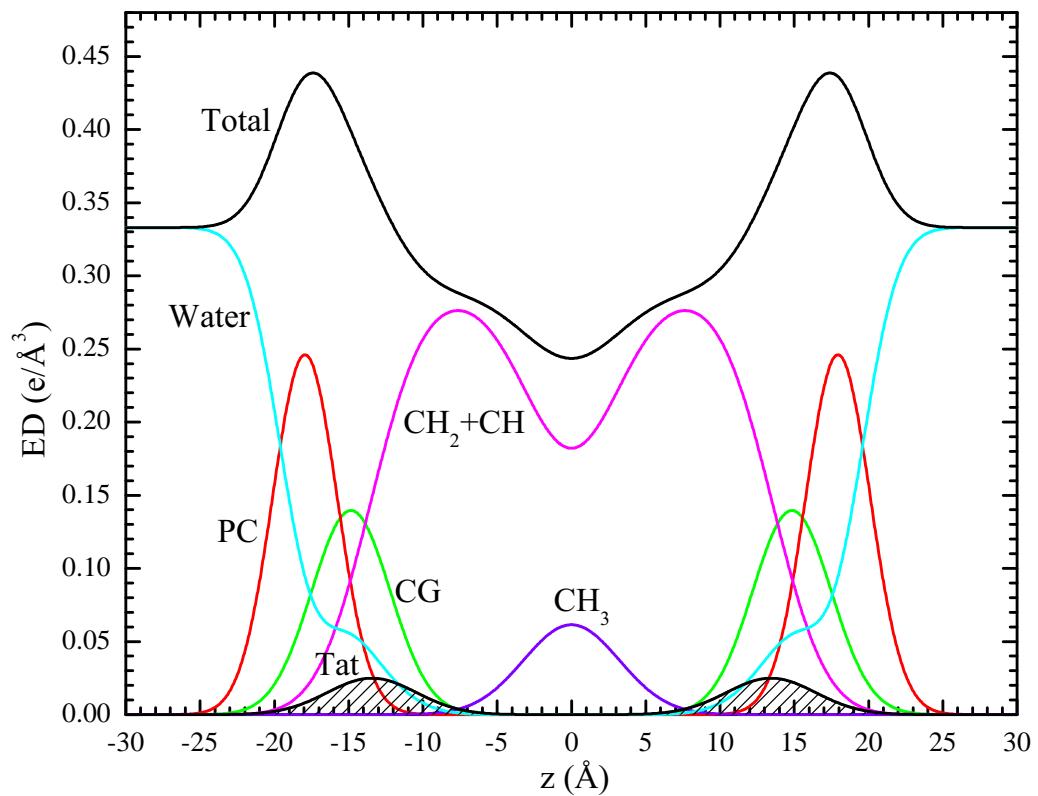


Figure 3.5: A model electron density profile for DOPC with Tat.

This definition enforces the total probability P_{HC} in the hydrocarbon chain region to equal one, which in turn means that placement of Tat in the chain region is prohibited. We call the model defined by Eq. (3.19) model A. To allow Tat to be placed inside the hydrocarbon chain region, we also consider an alternative definition,

$$P_{\text{CH}_2+\text{CH}}(z) = P_{\text{HC}}(z) - P_{\text{CH}_3}(z) - P_{\text{Tat}}(z), \quad (3.20)$$

where the volume probability of CH_2+CH combined component is reduced by the Tat volume probability distribution. We call this model B. The spatial conservation requires the water volume probability distribution to be

$$P_{\text{W}}(z) = 1 - P_{\text{PC}}(z) - P_{\text{CG}}(z) - P_{\text{Tat}}(z) - P_{\text{HC}}(z) \quad (3.21)$$

for model A and

$$P_{\text{W}}(z) = 1 - P_{\text{PC}}(z) - P_{\text{CG}}(z) - P_{\text{HC}}(z) \quad (3.22)$$

for model B.

Because X-rays measure the contrast between the bilayer and surrounding solvents, water, the experimental form factor is compared to the water subtracted model form factor,

$$F(q_z) = 2 \int_0^{\frac{D}{2}} dz \left(\sum_i (\rho_i - \rho_{\text{W}}) P_i(z) \right) \cos(q_z z), \quad (3.23)$$

where $i = \text{PC}, \text{CG}, \text{Tat}, \text{CH}+\text{CH}_2$, and CH_3 .

Constraints

The height of the hydrocarbon chain error function is fixed to one by imposing spatial conservation, whereas the mean position of the terminal methyls is fixed to $z_{\text{CH}_3} = 0$ by symmetry arguments. The total lipid volume V_L is fixed to the experimentally measured value. The headgroup volume V_{HL} was determined to be 331 \AA^3 for gel phase phosphatidylcholine (PC) bilayers [42], and we assume the same volume for the fluid phase PC bilayers. The volumes of PC and CG components satisfy

$$V_{\text{PC}} + V_{\text{CG}} = V_{\text{HL}}, \quad (3.24)$$

and the volumes of CH_3 and CH_2+CH components satisfy

$$2(16V_{\text{CH}_2+\text{CH}} + V_{\text{CH}_3}) = V_L - V_{\text{HL}}. \quad (3.25)$$

These component volumes constrain the height of the Gaussians as

$$c_{\text{PC}} = \frac{V_{\text{PC}}}{A_L \sigma_{\text{PC}}} \quad (3.26)$$

$$c_{\text{CG}} = \frac{V_{\text{CG}}}{A_L \sigma_{\text{CG}}} \quad (3.27)$$

$$c_{\text{CH}_3} = \frac{2V_{\text{CH}_3}}{A_L \sigma_{\text{CH}_3}} \quad (3.28)$$

$$c_{\text{Tat}} = \frac{V_{\text{Tat}}}{A_L \sigma_{\text{Tat}}} \quad (3.29)$$

where A_L is area per lipid.

The ratio of the carbonyl/glycerol volume to the headgroup volume V_{HL} was reported to be 0.41 [43], so we constrain the CG component volume to 135.7 \AA^3 and the PC component volume to 195.3 \AA^3 .

The most detailed structural study on DOPC to date was published by Braun *et al.* [43], and many of constraints on our model parameters can be derived from their study. However, in that work, the authors used the SDP model [41], which is specifically tailored for combined analysis of neutron and X-ray form factors. Therefore, we need to convert their structural results to the corresponding parameters in our simpler model. For example, from the reported values of the ratio of the volumes of the chain terminal methyl (CH_3) to the chain methylenes (CH_2) and the ratio of the volumes of the chain methines (CH) to the chain methylenes, we can calculate the ratio r_{CH_3} of the volumes of CH_3 to the CH_2 and CH combined component. Furthermore, the study by Braun *et al.* was at 30°C while our study was at 37°C , so our measured volume of DOPC was slightly higher.

At 30°C , the volume of DOPC was reported to be 1303 \AA^3 , so the volume of hydrocarbon chain region at the same temperature is $1303 - 331 = 972 \text{ \AA}^3$. The ratio r of the volumes of the chain terminal methyl (CH_3) to the chain methylenes (CH_2) was reported to be 1.95, and the ratio r_{12} of the volumes of the chain methines (CH) to the chain methylenes 0.91 at 30°C . Because there are 14 CH_2 groups, 2 CH groups, and 1 CH_3 group in each DOPC hydrocarbon chain, we have $2 \times (14V_{\text{CH}_2} + 2V_{\text{CH}} +$

$V_{\text{CH}_3}) = 972 \text{ \AA}^3$. Using $r = V_{\text{CH}_3}/V_{\text{CH}_2} = 1.95$ and $r_{12} = V_{\text{CH}}/V_{\text{CH}_2} = 0.91$, we get $V_{\text{CH}_2} = 27.3 \text{ \AA}^3$, $V_{\text{CH}} = 24.9 \text{ \AA}^3$, and $V_{\text{CH}_3} = 53.3 \text{ \AA}^3$. These calculated volumes lead to $V_{\text{CH}_3}/V_{\text{CH}_2+\text{CH}} = 1.97$ for 30 °C.

At 37 °C, the volume of DOPC was measured to be 1313.5 \AA^3 , so we have $2 \times (16V_{\text{CH}_2+\text{CH}} + V_{\text{CH}_3}) = 1313.5 - 331$. Assuming that the ratio $V_{\text{CH}_3}/V_{\text{CH}_2+\text{CH}}$ at 37 °C is the same as that at 30 °C gives $V_{\text{CH}_2+\text{CH}} = 27.3 \text{ \AA}^3$ and $V_{\text{CH}_3} = 53.9 \text{ \AA}^3$. We constrain the components for the hydrocarbon chain region in our model to these calculated values.

lipid	number of electrons	volume (\AA^3)
DOPC	434	1313.5
DOPE	410	1212.3
DOPC:DOPE (3:1)	428	1288.2

Table 3.2: Number of electrons per lipid and volume per lipid.

component	n_i^e	$V_i (\text{\AA}^3)$	$\rho_i (\text{e}/\text{\AA}^3)$
PC	97	195.3	0.497
PE	73	94.1	0.776
PC:PE (3:1)	91	170	0.535
CG	67	135.7	0.494
CH_2+CH	7.875	27.3	0.288
CH_3	9	53.9	0.167

Table 3.3: Some structural parameters for each component. n_i^e is the number of electrons and ρ_i is the average electron density.

number of electrons	838	mole fraction (x_{Tat})	n_{Tat}^e	$V_{\text{Tat}} (\text{\AA}^3)$
volume (\AA^3)	1877	0.016	13.6	30.5
$\rho_{\text{Tat}} (\text{e}/\text{\AA}^3)$	0.446	0.034	29.5	66.1
		0.059	53.0	118.8

Table 3.4: Tat basic structural parameters. The notations are the same as in Table 3.3. $x_{\text{Tat}} = \text{Tat}/(\text{Tat}+\text{Lipid})$.

Fits with Lower Bounds

Non-linear least squared fits with upper and lower bounds for the model parameters are implemented using an internal-external parameter transformation method. This

method is described in MINUIT User's Guide, section 1.3 [44]. This section briefly describes the method. The details can be found in the MINUIT website [45].

Basically, instead of a model parameter, which is also called the external variable, the minimization procedure varies a related variable called the internal variable. This internal variable can take any values between $-\infty$ to $+\infty$. At every χ^2 calculation, the internal variable is transformed to the external variable, which can take values only between the lower and upper bounds (a and b). This non-linear transformation allows an existing minimization algorithm that was developed for fits with no bounds to work for fits with bounds. This point was important because it allowed us to implement bound fits in the model fitting program called the SDP program, fully developed by Dr. Norbert Kucerka, without too many additional changes. Downsides of the transformation method include turning a linear problem into a non-linear one and some computational overhead, neither of which is particularly problematic in this study.

For variables with both lower and upper bounds (a and b , respectively), the transformation between the internal and external variables is

$$P_{\text{int}} = \arcsin\left(2\frac{P_{\text{ext}} - a}{b - a} - 1\right) \quad (3.30)$$

$$P_{\text{ext}} = a + \frac{b - a}{2}(\sin P_{\text{int}} + 1). \quad (3.31)$$

For variables with a lower bound a only, the transformation is

$$P_{\text{int}} = \sqrt{(P_{\text{ext}} - a + 1)^2 - 1} \quad (3.32)$$

$$P_{\text{ext}} = a - 1 + \sqrt{P_{\text{int}}^2 + 1}, \quad (3.33)$$

and for variables with an upper bound b only,

$$P_{\text{int}} = \sqrt{(b - P_{\text{ext}} + 1)^2 - 1} \quad (3.34)$$

$$P_{\text{ext}} = b + 1 - \sqrt{P_{\text{int}}^2 + 1}. \quad (3.35)$$

3.2.4 Molecular Dynamics Simulation

This section describes the MD simulations performed by Dr. Kun Huang, who was a graduate student of Prof. Angel Garcia at Rensselaer Polytechnic Institute.

Systems with different DOPC/Tat mole ratios (128:0, 128:2 and 128:4, corresponding to 0, 0.015 and 0.030 mole fractions) were simulated atomistically using the Gromacs 4.6.1 package [46]. DOPC was modeled by the Slipid force field [47, 48] and HIV Tat was modeled by Amber 99SB [49]. Tip3p water was used [50]. The number of Tats was divided equally on each side of the bilayer to mimic experimental conditions. All systems were simulated at 310 K with a constant area in the x - y plane and 1 atm constant pressure in the z direction. Each system was simulated for 100 ns and the last 50 ns was used as the production run. At each DOPC/Tat mole ratio, we studied systems with three different area/lipid (A_L). For the DOPC system, we fixed $A_L = 68, 70, 72 \text{ \AA}^2$; DOPC/Tat (128:2), we fixed the $A_L = 72, 74, 76 \text{ \AA}^2$; DOPC/Tat (128:4), we fixed the $A_L = 72, 74, 76 \text{ \AA}^2$. These values were based on the analysis of experimentally obtained form factors, which is discussed in Sec. 3.4.3. For each DOPC/Tat system at fixed A_L , we then conducted seven independent simulations with the center of mass (COM) of each Tat constrained at different bilayer depths from the bilayer center (18, 16, 14, 12, 10, 8 and 5 \AA). In total, 45 independent simulations were conducted. The goal of constrained simulations is to find the best match between experimental and MD simulation form factors. Comparison to the X-ray form factors was performed using the SIMtoEXP software written by Dr. Norbert Kucera [51].

All simulations were conducted with a 2 fs time integration step. SETTLE [52] was used to constrain water molecules and LINCS [53] was used to constrain all other bond lengths in the system. VdW interactions were truncated at 1.4 nm with a twin-range cutoff scheme and a dispersion correction was applied to both energy and pressure. Electrostatics interactions were treated with the particle-mesh Ewald (PME) method [54]. The direct term for electrostatics was evaluated within 1.0 nm cutoff and the Fourier term was evaluated with a 0.12 nm grid spacing and a 4th order interpolation. Each system was simulated at 310 K using the V-rescale algorithm [55] with a 0.2 ps time coupling constant. The semi-isotropic parrinello-rahman barostat [56] was used to couple the system at 1 atm in the z direction with a 5 ps time coupling constant, while the projected area at the x - y plane was fixed by setting the system compressibility to 0. We inserted the Tats into the system by initially turning off all interactions between Tats and the rest of the system, with Tats constrained at different depths. Then we slowly turned on the interactions to normal strength through thermodynamics integrations. We used umbrella potentials

to constrain Tats at desired depths with a force constant of 3000 kJ/mol/nm².

The center of mass (COM) distance between each peptide and the bilayer was constrained by an umbrella potential with a force constant k of 3000 kJ/mol/nm². Essentially, this potential acts as a spring, where its potential energy depends on the deviation of the distance between the center of mass of Tat and DOPC from a preferred value, z_0 ,

$$U(z_1^{\text{Tat}}, \dots, z_1^{\text{DOPC}}, \dots) = -\frac{1}{2}k(z_{\text{cm}}^{\text{Tat}} - z_{\text{cm}}^{\text{DOPC}} - z_0)^2.$$

Then, $-\partial U / \partial z_i$ is the external force acting on atom, i .

3.3 Analysis of Molecular Dynamics Simulation Data

3.3.1 SIMtoEXP program

This section briefly describes the SIMtoEXP program developed by Dr. Norbert Kucerka [51]. Essentially, for each snapshot, positional distribution of each atom averaged over the xy plane is calculated. Then, the distribution is averaged over snapshots. The product of this distribution and the average electron density gives the electron density profile of the atom. The sum over all the atoms provide the total electron density profile. This total electron density profile minus the average electron density of water is Fourier transformed to provide the X-ray form factor.

$$F^{\text{sim}}(q_z) = \int_0^\infty dz(\rho(z) - \rho_W) \cos(q_z z). \quad (3.36)$$

Electron density profiles were symmetrized and then X-ray form factors were calculated with $\rho_W = 0.326 \text{ e}/\text{\AA}^3$, which was the average electron density of water molecules in the MD simulations. Because $\rho(z)$ is equal to ρ_W outside the bilayer, the upper integration limit takes on a finite value.

Because the experimental form factor is on arbitrary units, it is scaled by a single constant a to produce the best fit to the simulated form factor through a linear least squared fit that minimizes the following goodness of fit

$$\chi^2 = \sum_i \left(\frac{1}{\sigma_i} (a|F_i^{\text{exp}}| - |F^{\text{sim}}(q_{z,i})|) \right)^2 \quad (3.37)$$

where σ_i is the input experimental uncertainties and F_i^{exp} is the experimental form factor measured at $q_z = q_{z,i}$. The SIMtoEXP program does not scale the input uncertainties, so the relative errors ($\sigma_i/|F_i^{\text{exp}}|$) depend on the value of the overall scaling factor a . Consequently, the χ^2 values calculated by the program had to be multiplied by $1/a^2$. These corrected χ^2 are reported in this chapter.

3.3.2 Local Thinning of Membranes

My contribution to the MD simulations was to help analyze the results.

The SIMtoEXP program only gives the average quantities for each leaflet. While our X-ray data are sensitive to the bilayer average electron density, local information

of Tat-bilayer interactions can be obtained from MD simulations. In this section, we discuss a method to extract a local membrane thickness around the Tat peptides from the MD simulation trajectories.

One of the expected effects of Tat interacting with a bilayer is compression of the lipid bilayer along the z -direction. It is reasonable to assume that this compression is greater near Tat and weaker far from Tat. If this is the case, the distance between phosphorus atoms in opposite leaflets near Tat should be different from the distance between phosphorus atoms away from Tat. For a small Tat concentration, $D_{\text{phos-phos}}$ is the same as that of pure DOPC if the distance from all Tats is large enough. For our experimental 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, we expect that the thickness near the Tats is smaller than the average thickness, so $D'_{\text{phos-phos}}$ is what we want to measure.

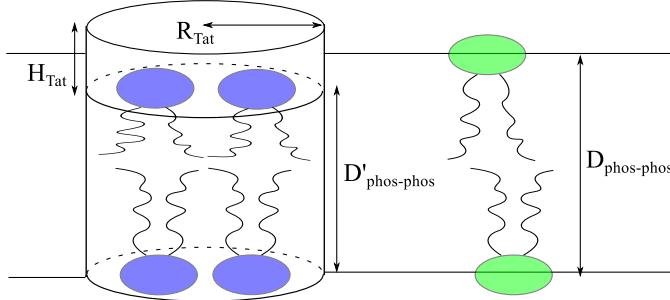


Figure 3.6: Our simple model to extract the local bilayer thickness from simulation trajectories. Tat is modeled as a cylinder with its height H_{Tat} and radius R_{Tat} . The local thickness is defined as $D'_{\text{phos-phos}}$. The thickness of the unperturbed DOPC bilayer is $D_{\text{phos-phos}}$. Blue highlighted lipids fall within the imaginary cylinder extended from the Tat. Unperturbed lipids are highlighted in green.

First, let us define what we mean by lipids close to Tat. As in Fig. 3.6, we imagine a cylinder around Tat and find 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 $R_{\text{Tat}} = 9 \text{ \AA}$ comes from the experimentally determined volume $V_{\text{Tat}} = 1876 \text{ \AA}^3$ and $H_{\text{Tat}} = 7.6 \text{ \AA}$ measured from one of the simulations (see Sec. 3.4.6). Let us define the lateral center of the cylinder as the center of mass of each Tat. Then we define $D'_{\text{phos-phos}}$ using only those lipids whose phosphorus atoms lie within these 9 \AA

cylinders around the Tats. Then $D_{\text{phos-phos}} = z_{\text{phos}}^+ - z_{\text{phos}}^-$ where z_{phos}^+ and z_{phos}^- are the average z of the n_1 (n_2) lipids in the upper and lower monolayer, respectively.

The algorithm for doing the above was straightforward. For each time frame, the positions (x_i, y_i, z_i) of each Tat, i , are listed. We chose phosphorus atoms whose (x, y) lateral position lied within 9 Å of any one of the Tat's lateral position. Then, z positions of the chosen phosphorus atoms were placed in a list. Then, z_{phos} were calculated from the list. We averaged over many snapshots to gain better statistics.

3.3.3 Lateral Decay Length of Membrane Thinning

This section describes a method to measure the lateral decay length of membrane thinning due to Tat-lipid interactions. As in the previous section, Tat is modeled here 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 as in Fig. 3.7. The two leaflets are assumed to be decoupled. In our model, lipids are separated into three regions: suppressed, boundary, and unperturbed region. The suppressed region extends from $r = 0$ to R_1 and is directly beneath (above) Tat in the top (bottom) leaflet. In this region, lipids are uniformly compressed by Tat toward the center of the bilayer, so that $h(r)$ is a constant equal to z_{phos} . From $r = R_1$ to R_2 is the boundary region, where $h(r)$ is assumed to linearly increase with the lateral distance r . The lateral decay length of membrane thinning is given by $R_2 - R_1$. In the unperturbed region ($r > R_3$), lipids do not interact with Tat, behaving identically to DOPC, so the phosphorus position is the same as that of DOPC. A continuous $h(r)$ that satisfies the above criteria is

$$h(r) = \begin{cases} z_{\text{phos}} & \text{if } 0 \leq r < R_1 \\ mr + b & \text{if } R_1 \leq r < R_2 \\ z_{\text{phos}}^0 & \text{if } R_2 \leq r < R_3 \end{cases} \quad (3.38)$$

with $m = (z_{\text{phos}} - z_{\text{phos}}^0)/(R_1 - R_2)$ and $b = (z_{\text{phos}}^0 R_1 - z_{\text{phos}} R_2)/(R_1 - R_2)$. Approximating the simulation box as a cylinder gives $R_3 = \sqrt{NA_{\text{L}}/\pi}$, where N is the number of lipids in a leaflet. z_{phos} can be measured directly from simulation trajectories. z_{phos}^0 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 .

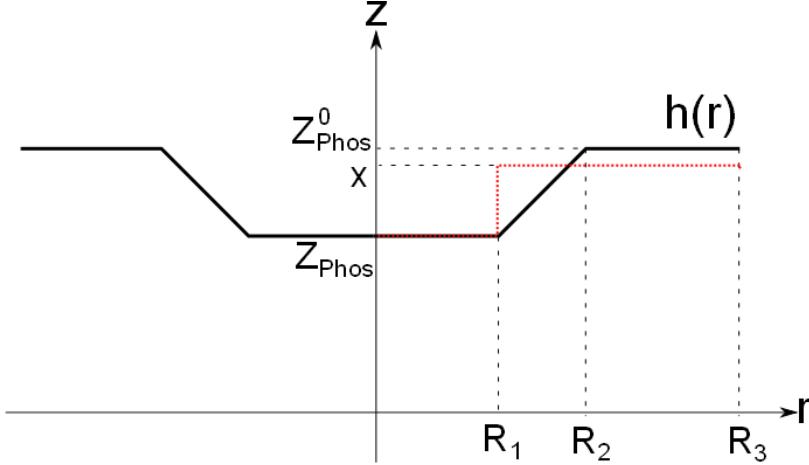


Figure 3.7: Simple model of the lateral decay of the membrane thickness perturbation due to Tat.

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

$$\langle h(r) \rangle = \frac{1}{\pi R_3^2} \int_0^{2\pi} d\phi \int_0^{R_3} dr r h(r) \quad (3.39)$$

The ϕ integration is trivial. The r integration is

$$\begin{aligned} & \int_0^{R_3} dr r h(r) \\ &= \int_0^{R_1} dr z_{\text{phos}} r + \int_{R_1}^{R_2} dr (mr + b)r + \int_{R_2}^{R_3} dr z_{\text{phos}}^0 r \\ &= \frac{1}{2} [z_{\text{phos}} R_1^2 + z_{\text{phos}}^0 (R_3^2 - R_2^2)] + \frac{1}{3} m (R_2^3 - R_1^3) + \frac{1}{2} b (R_2^2 - R_1^2) \\ &= \frac{1}{2} [z_{\text{phos}} R_1^2 + z_{\text{phos}}^0 (R_3^2 - R_2^2)] + \frac{1}{3} (z_{\text{phos}}^0 - z_{\text{phos}}) (R_2^2 + R_1 R_2 + R_1^2) \\ &\quad + \frac{1}{2} (z_{\text{phos}} R_2 - z_{\text{phos}}^0 R_1) (R_1 + R_2) \end{aligned} \quad (3.40)$$

Using Eq. (3.40), we get

$$\langle h(r) \rangle = \frac{(z_{\text{phos}} - z_{\text{phos}}^0) (R_1^2 + R_1 R_2 + R_2^2) + 3z_{\text{phos}}^0 R_3^2}{3R_3^2} \quad (3.41)$$

Eq. 3.41 is a quadratic equation in terms of R_2 . Solving for R_2 gives

$$R_2 = \frac{-R_1 + \sqrt{R_1^2 + 4C}}{2} \quad (3.42)$$

with

$$C = \frac{3R_3^2(z_{\text{phos}}^0 - \langle h(r) \rangle)}{z_{\text{phos}}^0 - z_{\text{phos}}} - R_1^2 \quad (3.43)$$

3.4 Results

3.4.1 Bending and Bulk Modulus

(Under construction) Show X-ray data. Show fitting boxes. Show the Kc values. Also, show the resultant form factors, which qualitatively show the membrane thinning. Also describe how I got error bars.

Fig. 3.1 shows the scattering intensity pattern from DOPC/DOPE (1:1) with mole fraction $x_{\text{Tat}} = 0.034$. The diffuse lobes are due to equilibrium fluctuations that occur in these fully hydrated, oriented lipid/peptide samples. The intensity $I(\mathbf{q})$ in the diffuse patterns provide the absolute values of the form factors $F(q_z)$, which are the Fourier transforms of the electron density profile, through the relation $I(\mathbf{q}) = S(\mathbf{q})|F(q_z)|^2/q_z$, where $\mathbf{q} = (q_r, q_z)$, $S(q)$ is the structure interference factor, and q_z^1 is the usual LAXS approximation to the Lorentz factor [57–59]. The first step in the analysis takes advantage of the q_r dependence of the scattering to obtain the bending modulus K_c with results shown in Fig. 3.8. As positively charged Tat concentration was increased, the lamellar repeat spacing D generally increased in neutral lipid bilayers and decreased in negatively charged bilayers, consistent with changes in electrostatic repulsive interactions. With few exceptions, the water space between bilayers exceeded 20 Å.

The analysis that obtains K_c also obtains the structure factor $S(\mathbf{q})$ and then the unsigned form factors $|F(q_z)|$ are obtained from the intensity $I(\mathbf{q})$ by division. Results for five different membrane mimics are shown in Fig. 3.9. Vertical lines indicate the zero position between the lobes of diffuse data where $F(q_z)$ change sign. In every sample, the zero positions shift to larger q_z , indicating a thinning of the membranes.

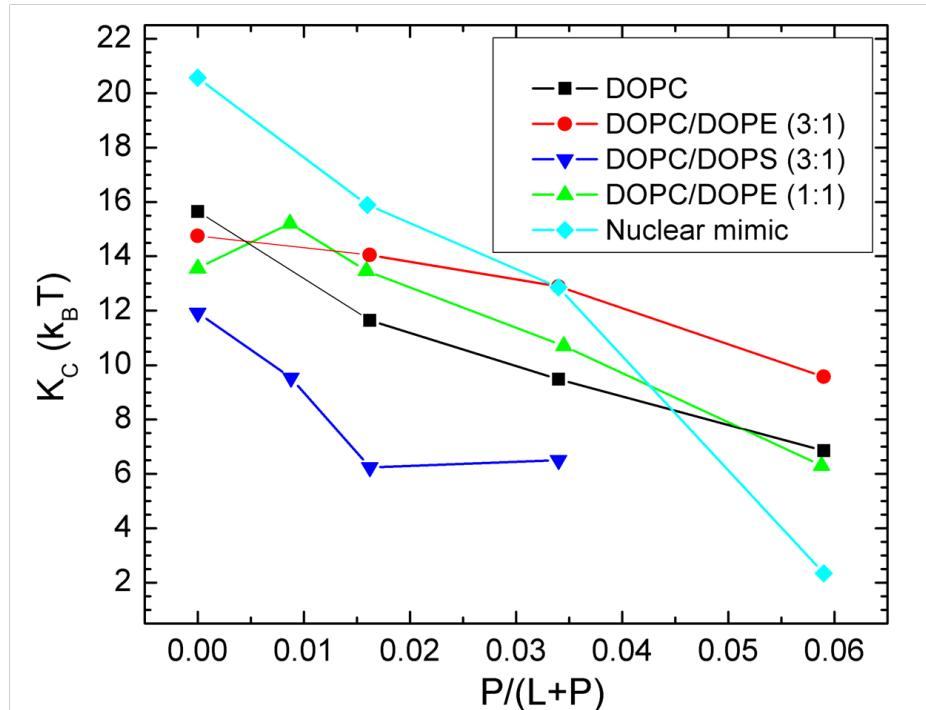


Figure 3.8: Bilayer bending modulus, K_c , vs. Tat mole fraction x_{Tat} . D -spacings for DOPC/Tat mixtures varied from 64 to 68 Å, for DOPC/DOPE/Tat mixtures from 64 to 69 Å, for DOPC/DOPS/Tat (3:1) mixtures from 57 Å to 100 Å (pure DOPS was unbound), and for nuclear mimic/Tat mixtures from unbound (nuclear mimic) to 64 Å. Estimated uncertainty in all values is about ± 2 .

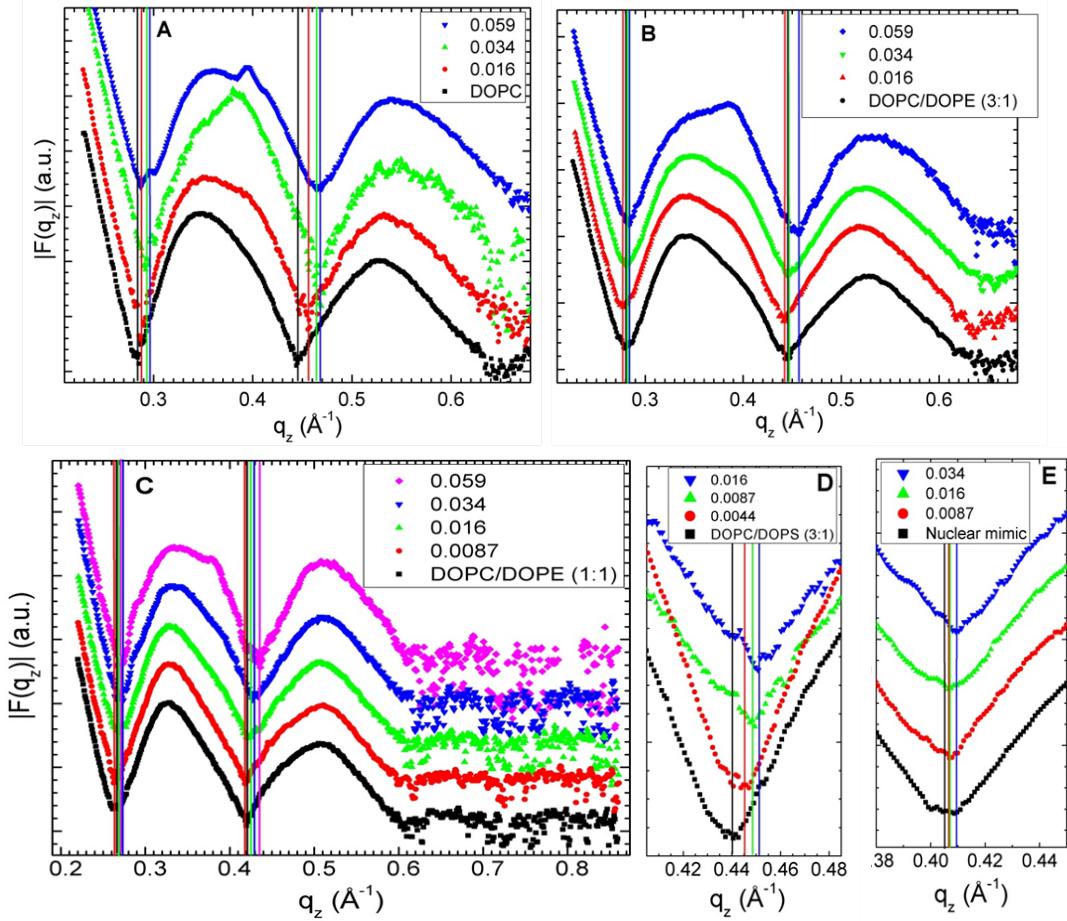


Figure 3.9: Form factors of lipid mixtures (arbitrarily scaled and vertically displaced) with increasing Tat mole fractions x_{Tat} indicated on figure legends. Lipid mixtures: A. DOPC B. DOPC/DOPE (3:1) C. DOPC/DOPE (1:1) D. DOPC/DOPS (3:1) E. Nuclear mimic. The entire q_z range is shown in C, while others show partial ranges. Solid vertical lines indicate the q_z values where the form factors equal zero between the lobes of diffuse data.

3.4.2 Volume results

Experimental and simulated volumes are given in Table 3.5. The simulated volume was obtained using the volume app in the SIMtoEXP program. The experimental Tat volume was calculated from the measured density assuming that the lipid volume was the same as with no Tat. In general, there may be an interaction volume between the peptide and the lipid membrane as previously reported for bacteriorhodopsin [60]. As lipid was present in excess to Tat, the partial molecular volume of the lipid should be the same as with no Tat, so this way of calculating includes all the interaction volume in V_{Tat} . Comparison of V_{Tat} in water with the result for 5:1 Lipid:Tat suggests that the interaction volume may be negative, consistent with a net attractive interaction with lipid. Understandably, values of V_{Tat} were unreliable for small mole ratios of Tat:Lipid. Therefore we used simple additivity for those mimics not shown in Table 3.5 for the volumes used in the electron density profile modeling. All volumes obtained from the Gromacs MD simulations were somewhat smaller than the measured volumes, but it supports the Tat volume being closer to 1877 Å³ than the outlying values obtained experimentally at small Tat concentrations. The measured volume was in a good agreement with the value calculated from a peptide calculator website [61], which gave 1888 Å³.

Experiments			
Tat in:	V_{lipid} (Å ³)	Lipid:Tat	V_{Tat} (Å ³)
water			1877
DOPC:DOPE (3:1)	1288	5:1	1822
DOPC	1314	39.6:1	676
DOPC:DOPS (3:1)	1298	39.6:1	2613

Simulations			
Tat in:	V_{lipid} (Å ³)	Lipid:Tat	V_{Tat} (Å ³)
DOPC	1283	128:2	1694
DOPC	1294	128.4	1699

Table 3.5: Volume results at 37 °C

3.4.3 Electron Density Profile Modeling

(Under construction) Using the model described in section 3.2.3, we fitted our measured X-ray form factors. In all fits, the positions of component groups were free

parameters, but we assumed that the lipid headgroup is somewhat rigid so that it cannot compress or expand. This assumption led to fixing the distance $z_{\text{PC}} - z_{\text{CG}}$ between the PC and CG components as well as the distance $z_{\text{CG}} - z_{\text{HC}}$ between the CG component and the Gibbs dividing surface for the hydrocarbon chains. We also constrained the width of Tat Gaussian. We fitted with three different values of widths, 2.5, 3.0, and 3.5, to study the range of variation due to the Tat width. The choice was made based on MD simulation results. (Check this again) We constrained the Tat width because we found that this parameter tended to become very small when it was free. This tendency to a unphysically small value was due to lack of higher q_z data points. A very narrow feature in an electron density profile led to large magnitude of the form factor at larger q_z . Because data points at larger q_z were not available, this narrow feature did not get penalized. In wide angle X-ray scattering, which probes much larger q_z than LAXS does, we did not observe much diffuse scattering for $q_z > 1.0 \text{ \AA}^{-1}$ (data not shown). Also, from a stereochemical point of view, a peptide width cannot be too small. These arguments allowed us to disregard the best fits with a too small value of σ_{Tat} and led to fixing the Tat width.

Figure 3.10 shows the results for DOPC with Tat. As shown, the fits were generally very good. Table 3.6 shows the best fit parameters for DOPC bilayers. It is clear that the area per lipid A_L increased as the Tat concentration was increased. An increase in A_L implies thinning of the bilayer which is an incompressible membrane. The results for DOPC:DOPE (3:1) are shown in Fig. 3.11 and Table 3.7, and the results for DOPC:DOPE (1:1) in Fig. 3.12 and Table 3.8. Figure 3.13 summarizes these results. For the results shown in Fig. 3.13, a consistent trend is that Tat moved away from the bilayer center as concentration increased.

As shown in Fig. 3.10, the membrane thickness can be defined as the distance D_{PP} between the PC components in the opposing leaflets or the distance D_{HH} between the maxima in the opposing leaflets. D_{HH} is more reliable than D_{PP} because it is a property of the total electron density of a bilayer and, therefore, does not depend strongly on the specific model employed for fitting the data. Indeed, the total electron density profile can be determined independently of a bilayer model by writing the electron density profile in terms of Fourier series, Fourier transforming the profile, and fitting the resulting model-independent form factor to the data. On the other hand, D_{PP} is a property that depends on lipid components, which are influenced by how the lipid is parsed and what assumptions and constraints go into the specific

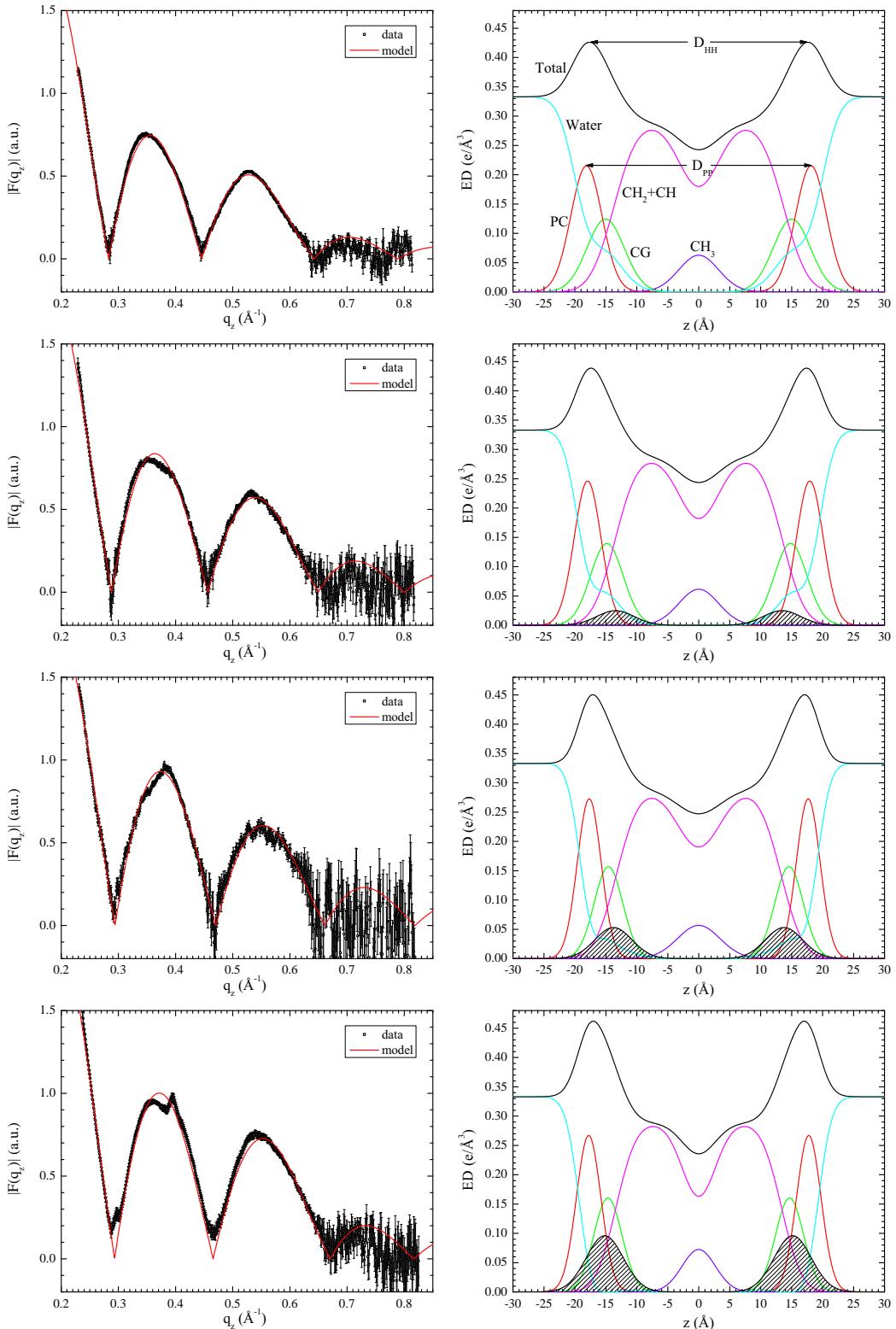


Figure 3.10: The best fits to DOPC form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).

x_{Tat}	0	0.016	0.016	0.016	0.034	0.034	0.034	0.059	0.059	0.059
χ^2	2961	1554	1570	1581	1563	1587	1607	2342	2338	2363
z_{PC}	18.1	18.0	17.9	17.9	17.8	17.7	17.6	17.8	17.8	17.7
σ_{PC}	2.52	2.14	2.17	2.18	1.86	1.92	1.93	2.02	1.97	1.93
z_{CG}	15.0	14.9	14.8	14.8	14.7	14.6	14.5	14.7	14.7	14.6
σ_{CG}	3.00	2.62	2.64	2.66	2.22	2.30	2.31	2.58	2.27	2.14
z_{HC}	13.7	13.6	13.5	13.5	13.4	13.3	13.2	13.4	13.4	13.3
σ_{HC}	3.00	2.69	2.84	2.95	2.65	2.82	3.01	2.47	2.58	2.83
σ_{CH_3}	3.20	3.19	3.22	3.24	3.37	3.43	3.47	2.70	2.70	2.74
z_{Tat}	NA	12.9	13.4	14.2	13.1	13.8	14.4	15.2	15.2	15.7
σ_{Tat}	NA	2.5	3.0	3.5	2.5	3.0	3.5	2.5	3.0	3.5
A_L	71.5	72.4	72.5	72.7	73.6	74.0	74.4	73.6	73.5	73.9

Table 3.6: Fitting Results for DOPC membranes for the THG (Tat in headgroup) model. $z_{\text{PC}} - z_{\text{CG}} = 3.1 \text{ \AA}$ and $z_{\text{CG}} - z_{\text{HC}} = 1.3 \text{ \AA}$ in all fits.

x_{Tat}	0	0.016	0.016	0.016	0.034	0.034	0.034	0.059	0.059	0.059
χ^2	924.5	4972	4985	4994	6758	6826	6863	2293	2280	2296
z_{PC}	18.3	18.5	18.5	18.4	18.5	18.4	18.3	18.2	18.2	18.1
σ_{PC}	2.66	2.23	2.26	2.27	2.25	2.31	2.34	2.31	2.19	2.11
z_{CG}	15.2	15.4	15.4	15.3	15.4	15.3	15.2	15.1	15.1	15.0
σ_{CG}	2.92	2.63	2.65	2.69	2.52	2.58	2.63	2.40	2.20	2.01
z_{HC}	13.9	14.1	14.1	14.0	14.1	14.0	13.9	13.8	13.8	13.7
σ_{HC}	2.73	2.70	2.83	2.91	2.86	2.79	2.84	2.25	2.38	2.60
σ_{CH_3}	3.24	2.94	2.97	2.98	2.87	2.90	2.91	2.63	2.61	2.65
z_{Tat}	NA	13.5	14.0	15.0	14.3	14.9	16.0	16.3	16.4	16.9
σ_{Tat}	NA	2.5	3.0	3.5	2.5	3.0	3.5	2.5	3.0	3.5
A_L	70.9	69.8	69.9	70.1	69.5	70.0	70.6	71.3	71.4	71.7

Table 3.7: Fitting Results for DOPC:DOPE (3:1) membranes for the THG model. $z_{\text{PC}} - z_{\text{CG}} = 3.1 \text{ \AA}$ and $z_{\text{CG}} - z_{\text{HC}} = 1.3 \text{ \AA}$ in all fits.

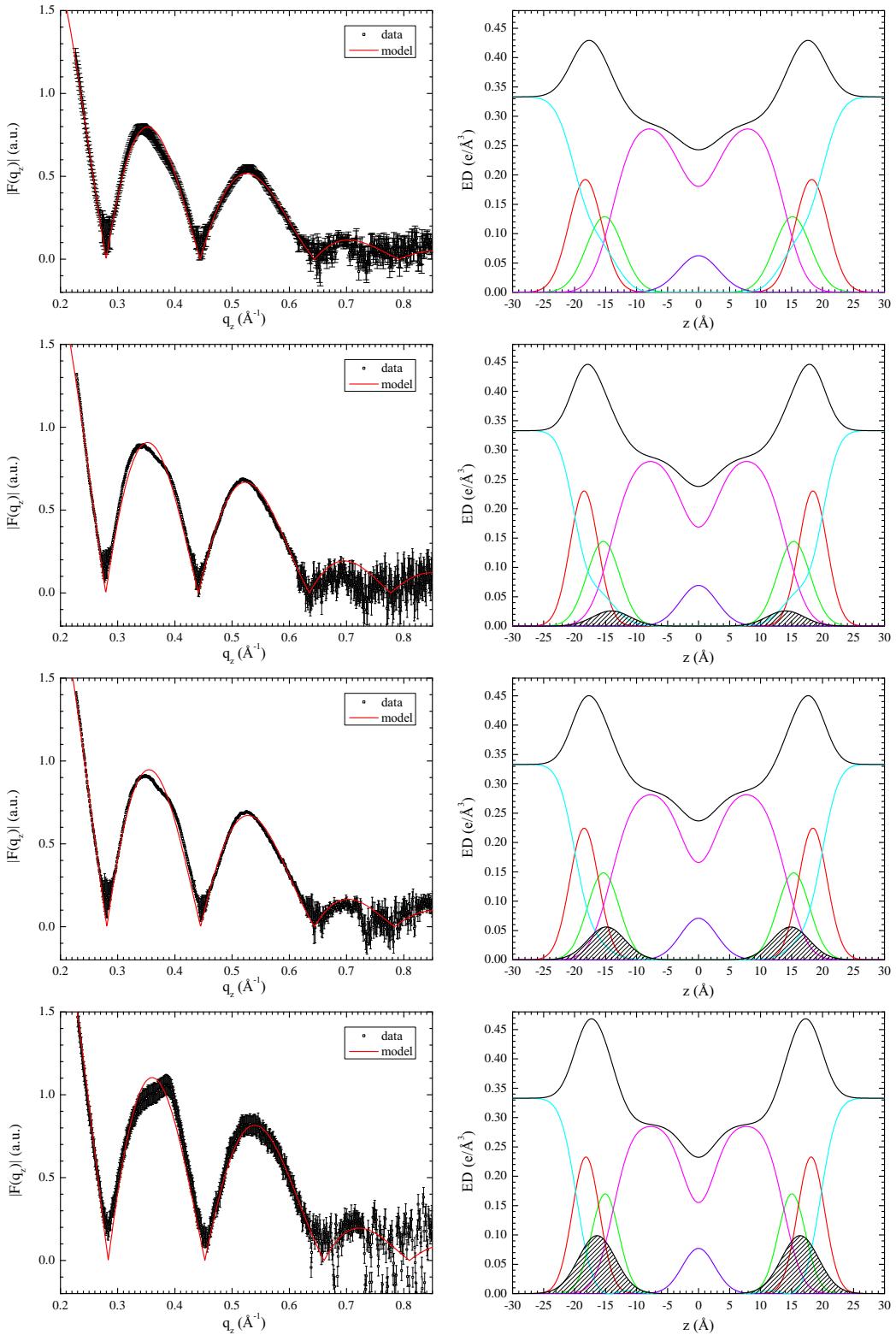


Figure 3.11: The best fits to DOPC:DOPE (3:1) form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).

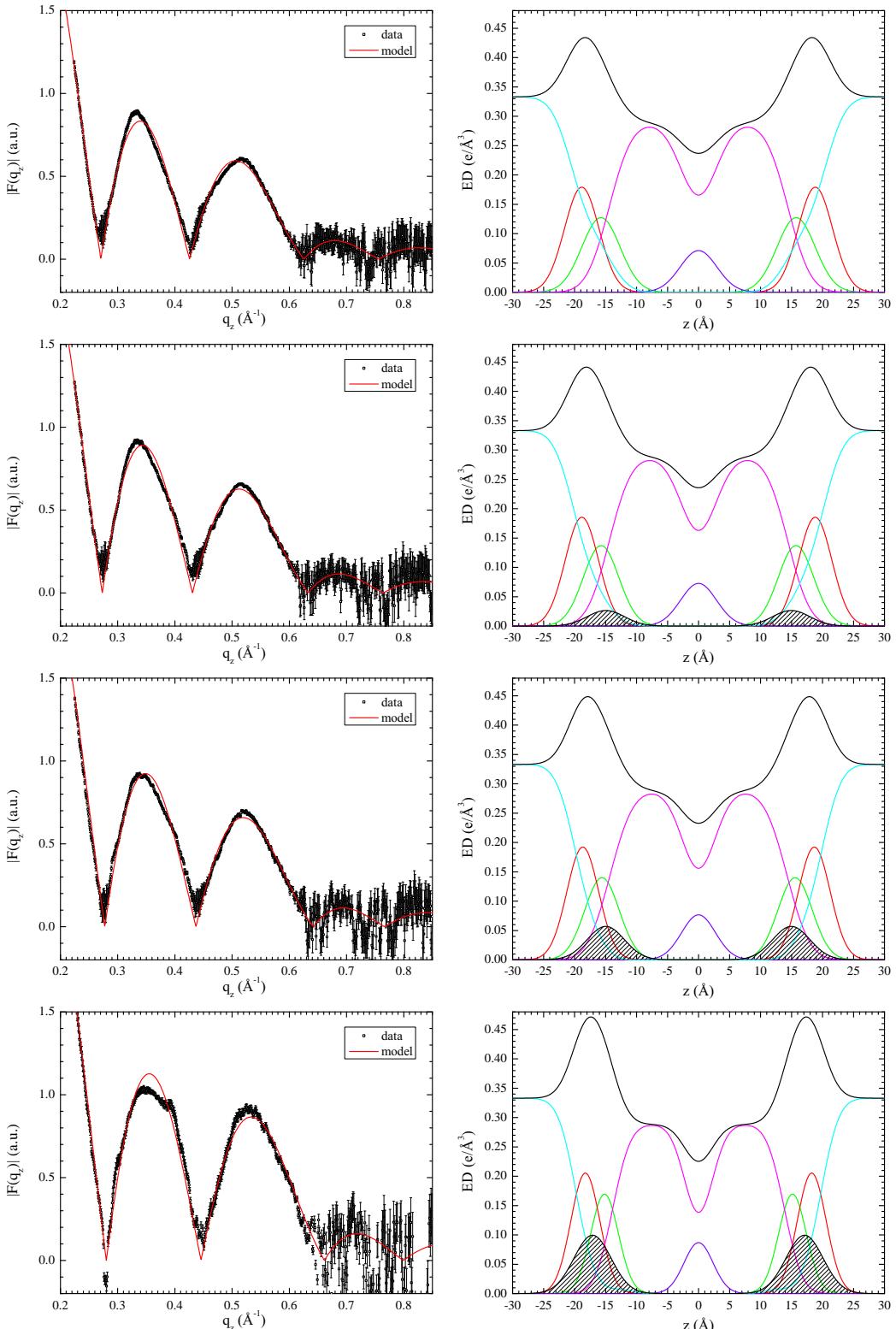


Figure 3.12: The best fits to DOPC:DOPE (1:1) form factors (left) and the corresponding electron density profiles (right) with $x_{\text{Tat}} = 0, 0.016, 0.034$, and 0.059 (from top to bottom).

x_{Tat}	0	0.016	0.016	0.016	0.034	0.034	0.034	0.059	0.059	0.059
χ^2	2961	1554	1570	1581	1563	1587	1607	2342	2338	2363
z_{PC}	18.1	18.0	17.9	17.9	17.8	17.7	17.6	17.8	17.8	17.7
σ_{PC}	2.52	2.14	2.17	2.18	1.86	1.92	1.93	2.02	1.97	1.93
z_{CG}	15.0	14.9	14.8	14.8	14.7	14.6	14.5	14.7	14.7	14.6
σ_{CG}	3.00	2.62	2.64	2.66	2.22	2.30	2.31	2.58	2.27	2.14
z_{HC}	13.7	13.6	13.5	13.5	13.4	13.3	13.2	13.4	13.4	13.3
σ_{HC}	3.00	2.69	2.84	2.95	2.65	2.82	3.01	2.47	2.58	2.83
σ_{CH_3}	3.20	3.19	3.22	3.24	3.37	3.43	3.47	2.70	2.70	2.74
z_{Tat}	NA	12.9	13.4	14.2	13.1	13.8	14.4	15.2	15.2	15.7
σ_{Tat}	NA	2.5	3.0	3.5	2.5	3.0	3.5	2.5	3.0	3.5
A_{L}	71.5	72.4	72.5	72.7	73.6	74.0	74.4	73.6	73.5	73.9

Table 3.8: (Numbers are wrong) Fitting Results for DOPC:DOPE (1:1) membranes for the THG model. $\Delta z_1 = z_{\text{PC}} - z_{\text{CG}}$ and $\Delta z_2 = z_{\text{CG}} - z_{\text{HC}}$.

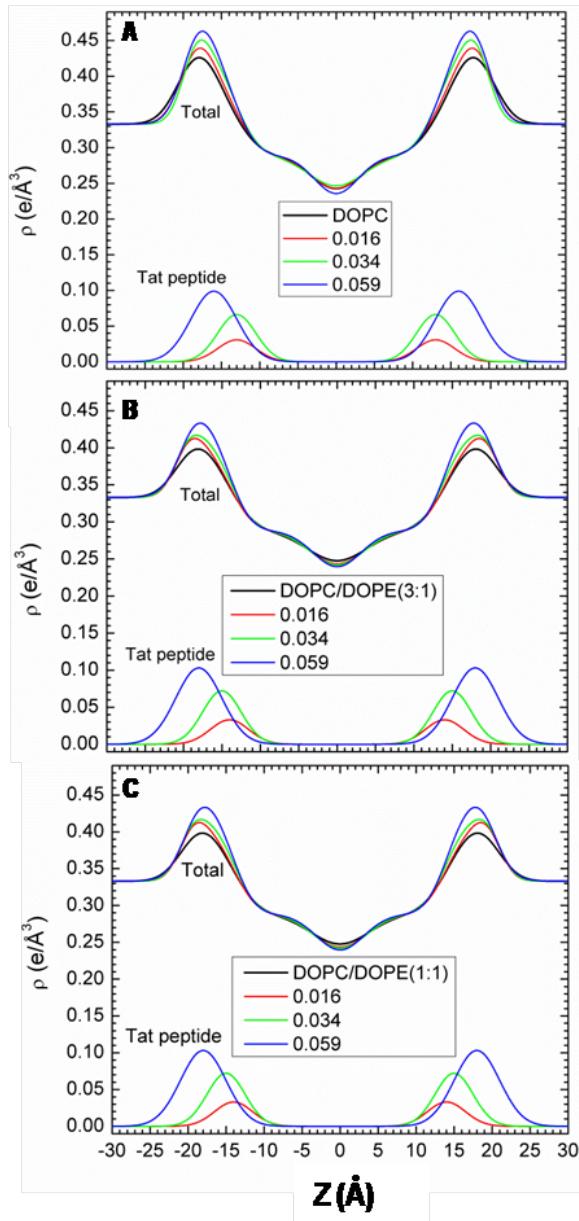


Figure 3.13: Modeling results for absolute electron density profiles and for the Tat location as a function of distance z along the bilayer normal. A. DOPC B. DOPC:DOPE (3:1), and C. DOPC:DOPE (1:1).

model. A disadvantage of using D_{HH} as a measure of the membrane thickness is that D_{HH} is influenced by the electron density of Tat because the total electron density profile includes a contribution from the electron density of Tat. Especially when the mole fraction of Tat in a system becomes large, the Tat electron density contributes significantly to the total electron density profile. If the Tat resided slightly outside of the PC component, the apparent membrane thickness measured by D_{HH} would be larger than D_{PP} . Then, even if the actual bilayer thickness defined by D_{PP} were reduced by the presence of Tat, the effect of thinning might not be obvious. With the above caveat in mind, we report both quantities in what follows since they can be easily calculated from the model.

More structural detail from the modeling is shown in Fig. 3.14. Figs. 3.14A and 3.14B show that both D_{PP} and D_{HH} tended to decrease with increasing Tat mole fraction x_{Tat} , showing that Tat thins membranes, increasingly so as its concentration is increased, even though both simulation and modeling suggested that Tat moves further from the membrane center with increasing concentration as shown in Fig. 3.14D. Figure 3.14C shows that the area per lipid A_{L} usually increased with increasing mole fraction of Tat, similar to the findings from MD simulations (Sec. 3.4.6), as would be expected.

We also investigated how the goodness of fits varied as the position of the Tat Gaussian was varied. Figure 3.15 plots χ^2 as a function of the fixed Tat position z_{Tat} . We found that the two models, THG (Tat in headgroup region) and THC (Tat in hydrocarbon chain region), resulted in similar electron density profiles, yielding similar χ^2 values when Tat was placed near the hydrocarbon-water interface region. In the THC model, the error function representing the hydrocarbon chain region became wider as Tat was placed near the interface region but further from the bilayer center. The subtraction of the Tat component from the hydrocarbon chain error function resulted in a smooth error function-like profile with an apparent smaller value of σ such that the total profile calculated from the THC model was very similar to that calculated from the THG model.

In general, while the total electron density profile is well determined by our modeling procedures, the values of the parameters for the components are not as well determined as the agreement of the fit to the data may suggest. In many cases, we found multiple local minima in the fitting landscape, including one with Tat closer to the center of the bilayer as shown in Fig. 3.15. χ^2 calculated at these local minima

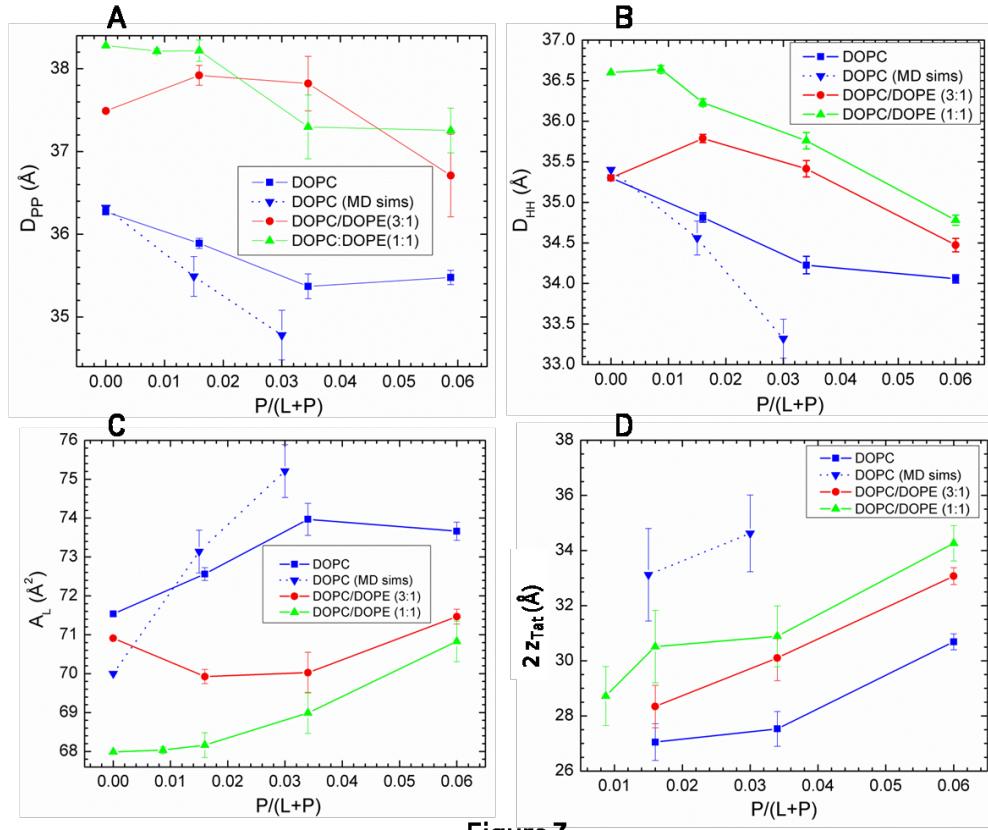


Figure 7

Figure 3.14: A. Bilayer thickness, D_{PP} ; B. Bilayer thickness, D_{HH} ; C. Area/lipid, A_L ; D. Twice the Tat location, $2z_{Tat}$: all plotted vs. Tat mole fraction x_{Tat} . Error bars are standard deviations from imposing Tat Gaussian widths, $\sigma_{Tat} = 2.5, 3.0$ or 3.5 \AA . Inverted blue triangles connected with dotted line are results from MD simulations, averaging the best fits to the X-ray data for each parameter, with standard deviations shown.

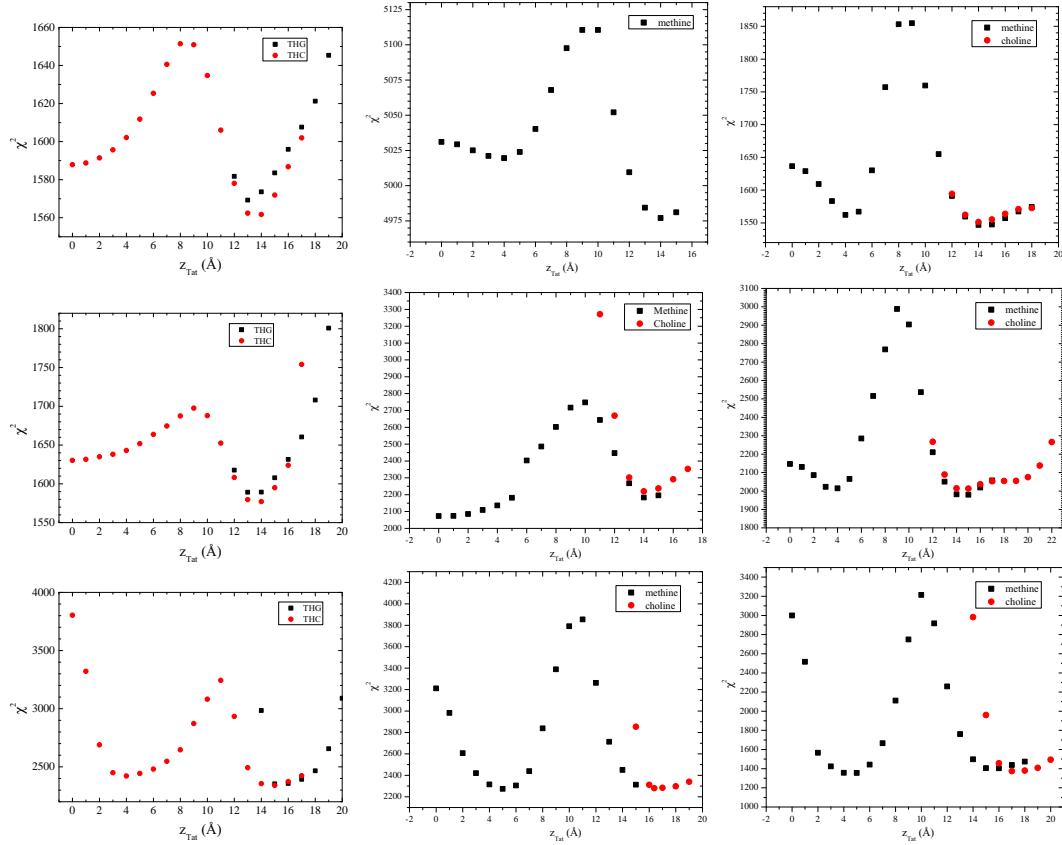


Figure 3.15: χ^2 as a function of z_{Tat} for DOPC, DOPC:DOPE (3:1), and DOPC:DOPE (1:1) (from left to right) with $x_{\text{Tat}} = 0.016, 0.034$, and 0.059 (from top to bottom). $\sigma_{\text{Tat}} = 3.0$. The THG model (black squares) and the THC model (red circles).

tended to be smaller for larger concentration of Tat. We also found that χ^2 with z_{Tat} in the hydrocarbon chain region and headgroup region was almost equal for the smallest value of x_{Tat} for DOPC:DOPE (1:1) bilayer. While Fig. 3.15 shows trends for the χ^2 minima with Tat in the hydrocarbon chain region, this position seemed energetically unfavorable as Tat is a hydrophilic molecule. Also, if Tat favored to be inserted deep in a membrane, it would be difficult for Tat to leave the membrane. This difficulty seems inefficient in terms of the HIV virus infection because Tat passing through the nuclear membrane and binding to the viral integrated DNA is crucial for proliferation of HIV infected cells. These considerations suggested that the local minima in the chain region were artifact of our models. The MD simulations performed by Dr. Kun Huang suggested that the interior positions of Tat were artifacts of our model. The simulation results are found in section 3.4.6.

Electron density profiles for DOPC/DOPS (3:1) and the nuclear membrane mimic were not successful, due to loss of diffuse scattering by Tats charge neutralization of these negatively charged membranes as described in section 3.4.1.

3.4.4 Hard Wall Constraint Fits

(Under construction) As seen from Table 3.6, 3.7, and 3.8, the widths of the headgroup components became smaller as Tat concentration increased in all membranes. These decreases seemed somewhat unreasonable; if Tat causes a bilayer to locally become thinner near where it is bound, we would expect that the headgroup components to become wider. Therefore, we also fitted the model with hard wall constraints on these headgroup widths. Namely, the minimum values of the widths of the headgroup components, PC and CG, were limited to the corresponding values for pure bilayers without Tat.

Table 3.9 shows results from fitting the data with lower bounds on the widths of the headgroup components for DOPC:Tat. In all cases, both headgroup widths, σ_{PC} and σ_{CG} , resulted in the same value as the value of their corresponding lower bounds. Figure 3.16 shows χ^2 landscape as a function of z_{Tat} similarly to Fig. 3.15. The χ^2 minima observed for $z_{\text{Tat}} > 25 \text{ \AA}$ were artifact; Tat are essentially in the water region while the bilayer structure was significantly perturbed. This action-at-distance seemed unreasonable, so these minima were considered as artifact of the model. Indeed, when we fixed lipid component parameters in these fits to be identical

to those of the DOPC model, we did not observe any minima with Tat in the water region. Although we did not note in the previous section, we observed similar minima in the unbound model as well.

x_{Tat}	0	0.016	0.016	0.16	0.034	0.034	0.034
χ^2	2961	1853	1979	2118	2398	2893	3414
z_{PC}	18.1	17.8	17.8	17.8	17.4	17.4	17.4
σ_{PC}	2.5	2.5	2.5	2.5	2.5	2.5	2.5
z_{CG}	15.0	14.7	14.7	14.7	14.3	14.3	14.3
σ_{CG}	3.0	3.0	3.0	3.0	3.0	3.0	3.0
z_{HC}	13.7	13.4	13.4	13.4	13.0	13.0	13.0
σ_{HC}	3.0	2.7	2.7	2.7	2.7	2.7	2.7
σ_{CH_3}	3.2	3.1	3.1	3.1	3.6	3.6	3.7
z_{Tat}		16.9	16.8	17.0	16.4	16.5	16.7
σ_{Tat}		2.5	3.0	3.5	2.5	3.0	3.5
Δz_1	3.1	3.1	3.1	3.1	3.1	3.1	3.1
Δz_2	1.3	1.3	1.3	1.3	1.3	1.3	1.3
A_L	71.5	73.5	73.5	73.5			
x_{Tat}	0.059	0.059	0.059				
χ^2	3160	4298	5539				
z_{PC}	17.5	17.4	17.3				
σ_{PC}	2.5	2.5	2.5				
z_{CG}	14.4	14.4	14.3				
σ_{CG}	3.0	3.0	3.0				
z_{HC}	13.1	13.0	12.9				
σ_{HC}	2.7	2.7	2.7				
σ_{CH_3}	2.6	2.6	2.5				
z_{Tat}	16.3	16.6	17.1				
σ_{Tat}	2.5	3.0	3.5				
Δz_1	3.1	3.1	3.1				
Δz_2	1.3	1.3	1.3				
A_L							

Table 3.9: Fitting Results of the bound THG model for DOPC membranes. $\Delta z_1 = z_{\text{PC}} - z_{\text{CG}}$ and $\Delta z_2 = z_{\text{CG}} - z_{\text{HC}}$.

3.4.5 Summary of Electron Density Profile Modeling

(Under construction) Figure 3.19 shows that the area per lipid A_L as defined by $(V_L - V_{\text{HL}})/D_C$ decreased as the mole fraction of DOPE in DOPC:DOPE membranes

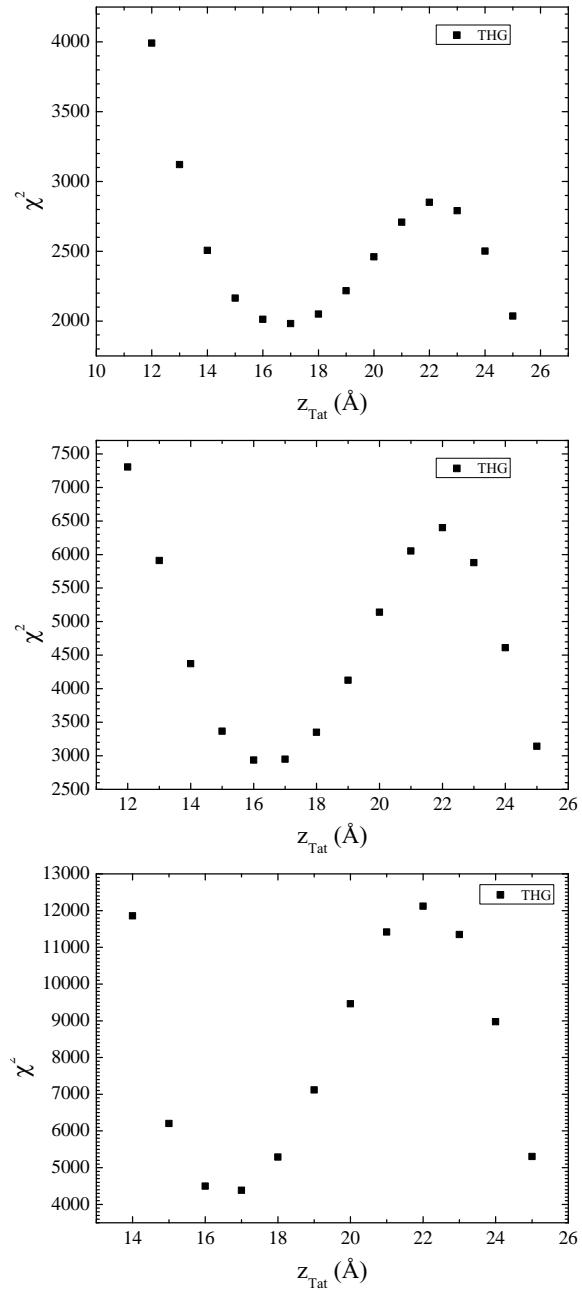


Figure 3.16: χ^2 as a function of z_{Tat} for DOPC with $x_{\text{Tat}} = 0.016, 0.034$, and 0.059 (from top to bottom). $\sigma_{\text{Tat}} = 3.0$. The bound THG model was used.

Figure 3.17: DPP graph with bound fits

Figure 3.18: DHH graph with bound fits

Figure 3.19: AL graph with bound fits

Figure 3.20: zTat graph with bound fits

increased. This decrease of A_L is qualitatively consistent with previous studies which attributed the decrease to the small size of PE head group (references?). Because DOPE has a smaller headgroup than DOPC, lipids in DOPC:DOPE bilayers pack more compactly than in DOPC bilayers. Then, more compact packing of lipids leads to a smaller A_L .

Figure 3.20 shows that Tat is located further out from the bilayer center with higher content of PE lipids. This trend is consistent with a potential mean force calculated from MD simulations (ref?), which has shown that arginine insertion costs more energy in a PE membrane than PC membrane. The higher energy cost of arginine insertion has been suggested to be due to more possible hydrogen bonding between PE groups and arginines. (After I wrote this paragraph, I realized that the argument presented does not make much sense. z_{Tat} could be larger for PE simply because DOPE membranes are thicker than DOPC membrane. z_{Tat} must be measured with respect to, say, the hydrocarbon interface. Let's do this later.)

3.4.6 Molecular Dynamics Simulations

(Under construction) Due to the slow relaxation in lipid bilayers and limited accuracy of the force field, a good agreement between experimental and MD simulation calculated form factors may be difficult to reach. Consequently, we carried out several constrained simulations at various A_L and z_{Tat} as described in Sec. ???. We then compared the simulated form factor $F(q_z)$ with the experimentally measured one. Figure 3.21 shows such comparison for a DOPC bilayer. As discussed earlier, the simulated form factor shifted to larger q_z as the area per lipid was increased. From this comparison, we found the simulation at $A_L = 70 \text{ \AA}^2$ to be the best match with the experimental form factor, yielding the lowest χ^2 . However, the form factor for

$A_L = 72 \text{ \AA}^2$ matched the experiment better than that for 70 \AA^2 near $q_z = 0.3 \text{ \AA}^{-1}$, which suggests that a better match might lie between 70 and 72 \AA^2 . This case was not investigated further. The electron density profile for the best fit is shown in Fig. 3.22. The comparison for DOPC with $x_{\text{Tat}} = 0.015$ where there is one Tat in each monolayer is shown in Fig. 3.23. The same comparison for DOPC with $x_{\text{Tat}} = 0.03$ is shown in Fig. 3.24.

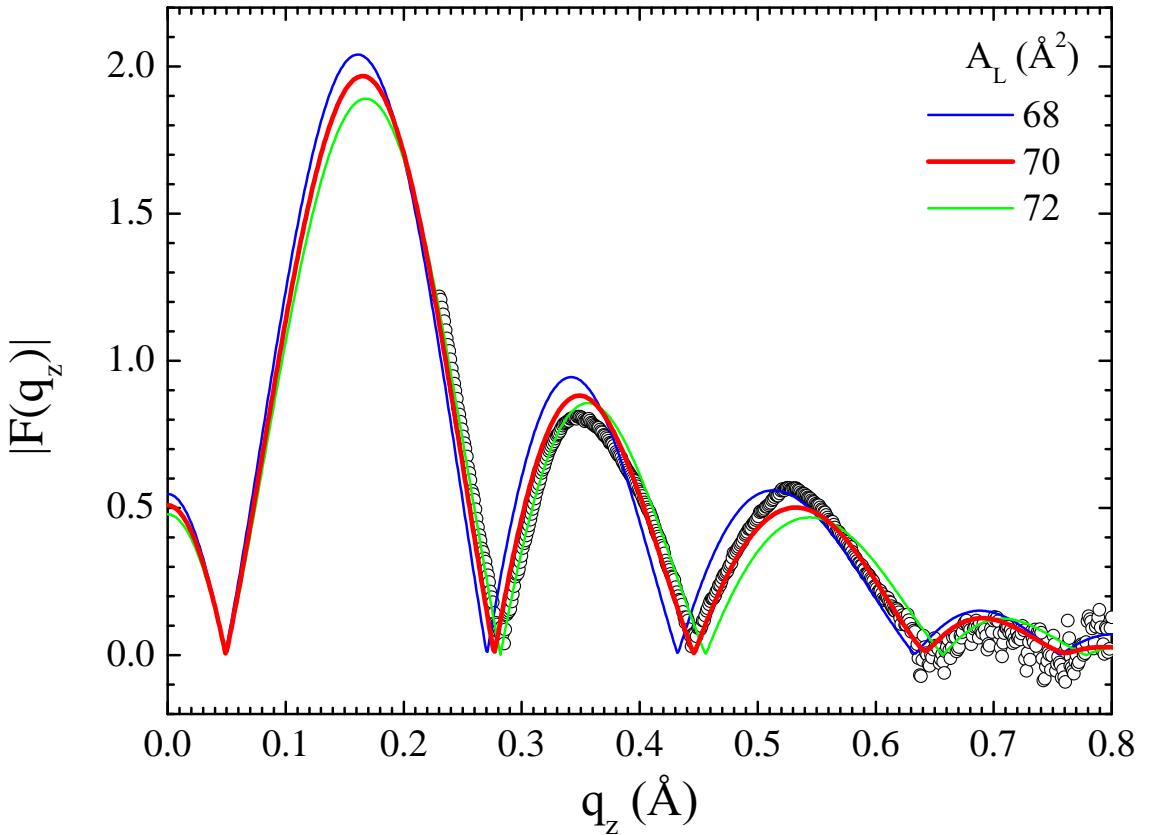


Figure 3.21: MD simulated form factors for DOPC at $A_L = 68 \text{ \AA}^2$ (blue solid line), 70 \AA^2 (red solid line), and 72 \AA^2 (green solid line) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for 70 \AA^2 .

The best match for DOPC/Tat (128:4) was found when the Tats were constrained at 18 \AA away from the bilayer center (Fig. 3.26). The other best fit results were: DOPC $A_L = 70 \text{ \AA}^2$ and DOPC/Tat(128:2) $A_L = 72 \text{ \AA}^2$, $z_{\text{Tat}} = 18 \text{ \AA}$. It clearly indicates that with increasing Tat concentration, A_L increases. The agreement worsened as Tat was constrained to be closer to the center of the bilayer. When Tats were constrained at 5 \AA away from the bilayer center, we observed a spontaneous formation of water pores

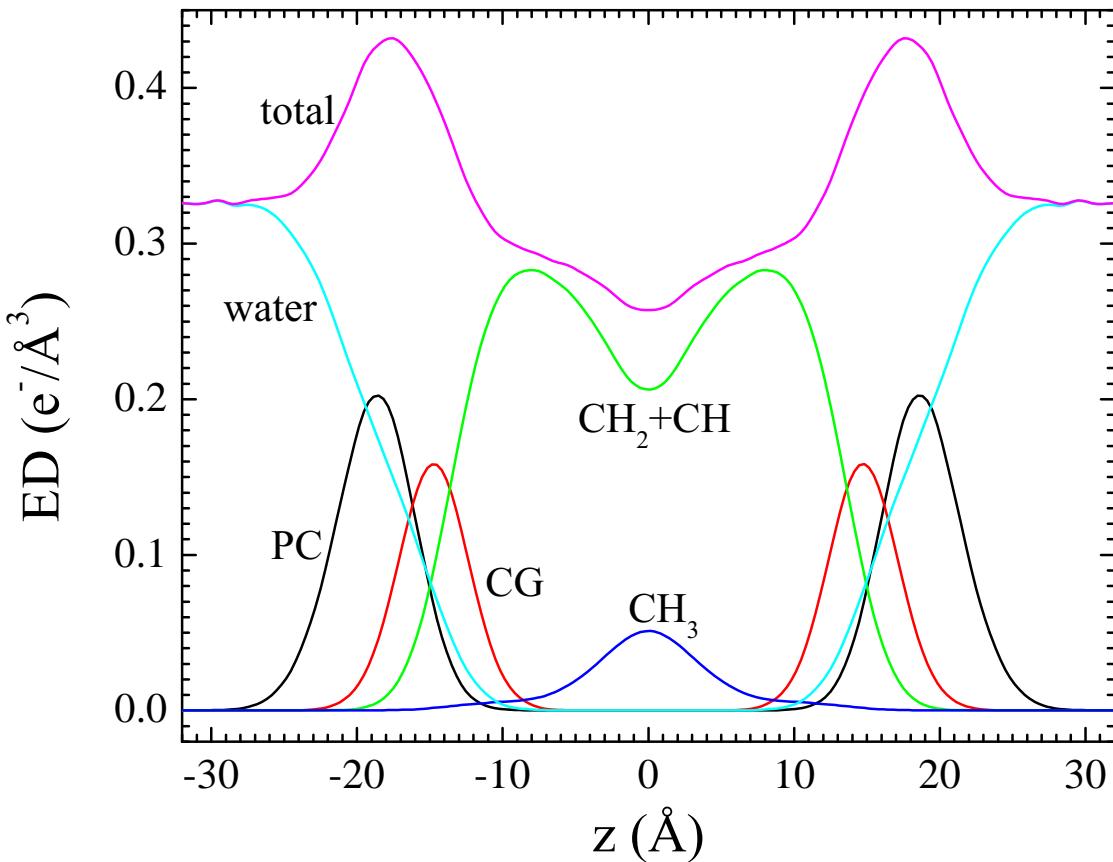


Figure 3.22: The simulated, symmetrized electron density profile for DOPC at $A_L = 70 \text{ \AA}^2$ as a function of the distance away from the bilayer center. Each component profile is labeled with its name: PC (phosphate-choline), CG (carbonyl-glycerol), CH_2+CH (methylene-methine combination), CH_3 (terminal methyl). The sum of all the components is labeled as total.

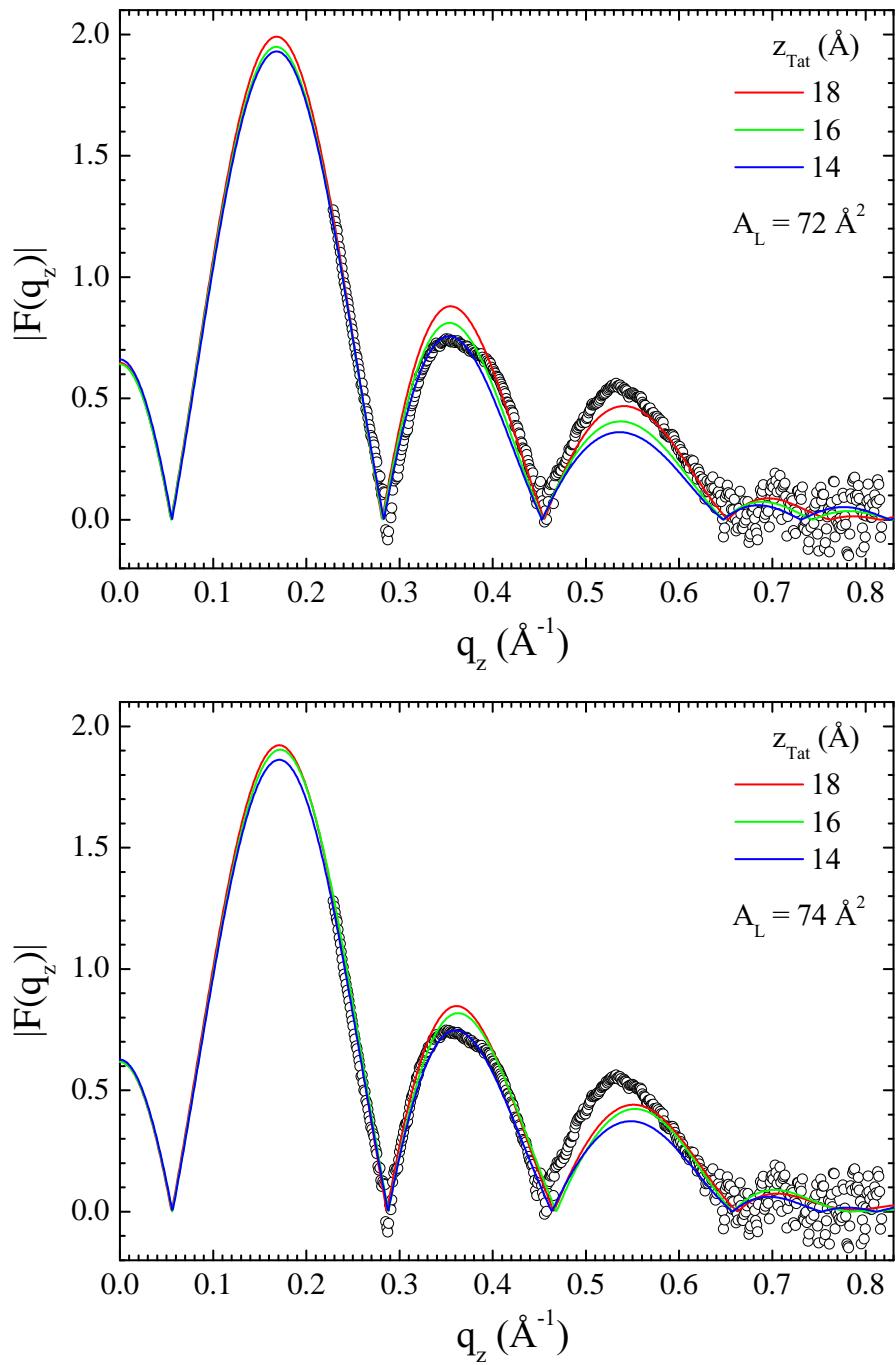


Figure 3.23: MD simulated form factors for DOPC with $x_{\text{Tat}} = 0.015$ at $A_L = 72 \text{ \AA}^2$ (top) and 74 \AA^2 (bottom), with $z_{\text{Tat}} = 18 \text{ \AA}$ (red solid lines), 16 \AA (green solid lines), and 14 \AA (blue solid lines) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for $z_{\text{Tat}} = 18 \text{ \AA}$.

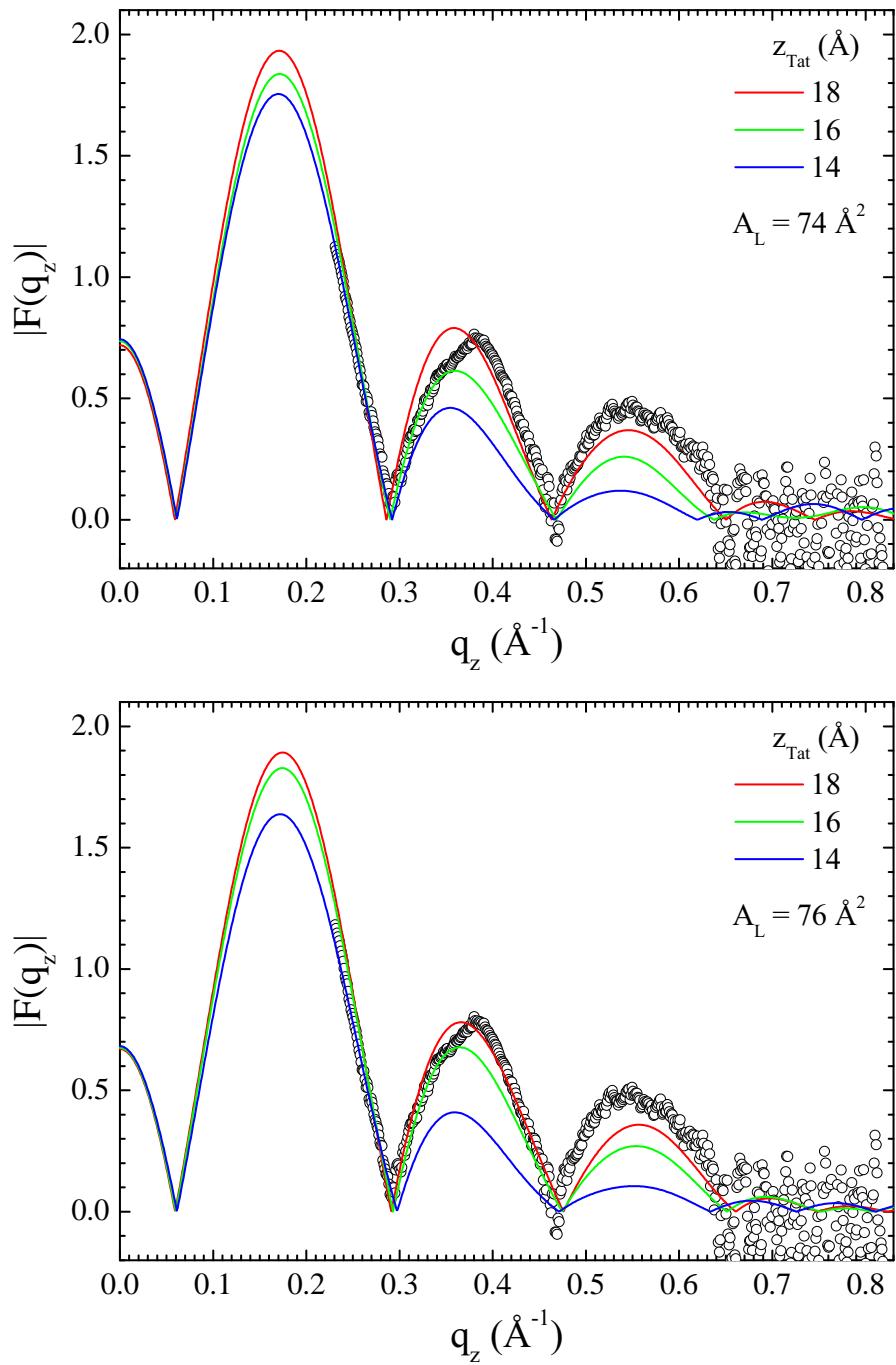


Figure 3.24: MD simulated form factors for DOPC with $x_{\text{Tat}} = 0.030$ at $A_L = 74 \text{ \AA}^2$ (top) and 76 \AA^2 (bottom), with $z_{\text{Tat}} = 18 \text{ \AA}$ (red solid lines), 16 \AA (green solid lines), and 14 \AA (blue solid lines) compared to the experimental form factor (open circles) scaled vertically to best match the form factor for $z_{\text{Tat}} = 18 \text{ \AA}$.

$x_{\text{Tat}} = 0.015$				$x_{\text{Tat}} = 0.030$			
$A_L (\text{\AA}^2)$	$z_{\text{Tat}} (\text{\AA})$	a	χ^2	$A_L (\text{\AA}^2)$	$z_{\text{Tat}} (\text{\AA})$	a	χ^2
70	18	0.621	60.1	70	18	0.621	60.1
70	16	0.568	69.1	70	16	0.568	69.1
70	14	0.439	131	70	14	0.439	131
70	12	0.285	391	70	12	0.285	391
70	10	0.199	440	70	10	0.199	440
70	8	0.196	374	70	8	0.196	374
70	5	0.159	527	70	5	0.159	527
72	18	0.72	18.0	72	18	0.72	18.0
72	16	0.65	24.9	72	16	0.65	24.9
72	14	0.6	31.4	72	14	0.6	31.4
72	12	0.426	104	72	12	0.426	104
72	10	0.219	443	72	10	0.219	443
72	8	0.205	336	72	8	0.205	336
72	5	0.165	448	72	5	0.165	448
74	18	0.722	21.3	74	18	0.722	21.3
74	16	0.704	25.9	74	16	0.704	25.9
74	14	0.631	24.7	74	14	0.631	24.7
74	12	0.412	81.9	74	12	0.412	81.9
74	10	0.312	194	74	10	0.312	194
74	8	0.246	351	74	8	0.246	351
74	5	0.177	427	74	5	0.177	427

Table 3.10: Comparison of the simulated form factors to the experimental form factors.

x_{Tat}	A_L	z_{Tat}	$\langle D_{\text{PP}} \rangle$	D_{PP}	x	Δt	H_{Tat}	R_{Tat}	R_2	z_{phos}	z_{guan}	χ^2
0	70		36.3									
0.015	72	18	35.6	32.8	35.8	3.5	9.2	8.1	15.0	14.7	15.5	18
0.015	72	16	36.1	33.0	36.3	3.3	9.4	8.0	9.0	14.9	14.5	24.9
0.015	74	18	35.0	33.0	35.1	3.3	8.6	8.3	23.9	14.9	16.5	21.3
0.015	74	16	35.0	32.1	35.2	4.2	7.6	8.9	20.4	14.0	13.5	25.9
0.030	74	18	35.3	32.6	NA	3.7	7.6	8.9	NA	14.5	15.5	24.3
0.030	74	16	35.3	31.2	NA	5.1	7.7	8.8	NA	13.1	13.5	40.1
0.030	76	18	34.2	32.0	NA	4.3	7.6	8.9	NA	13.9	16.5	14.8
0.030	76	16	34.9	31.4	NA	4.9	7.8	8.7	NA	13.3	14.5	30.4

Table 3.11: Summary of simulation results. $\langle D_{\text{PP}} \rangle$, phosphorus-phosphorus distance averaged over all lipids; D_{PP} , Tat-perturbed phosphorus atoms; x , thickness away from Tat; Δt , $\langle D_{\text{PP}}^{\text{DOPC}} \rangle - D_{\text{PP}}$; H_{Tat} , Tat height; R_{Tat} , radius of Tat cylinder; R_2 , radius of the calculated in-plane Tat-perturbed region; R_3 , effective radius of the simulation box.

x_{Tat}	A_L	z_{Tat}	$\langle D_{\text{PP}} \rangle$	D_{PP}	Δt	H_{Tat}	R_{Tat}	R_2	z_{phos}	z_{guan}
0.015	72.9	17.1	35.4	32.7	3.6	8.7	8.3	17.1	14.6	15.1
0.030	75.2	17.3	34.8	31.9	4.4	7.7	8.8	NA	13.8	15.4

Table 3.12: Summary of weighted average results. The caption is the same as Table 3.11.

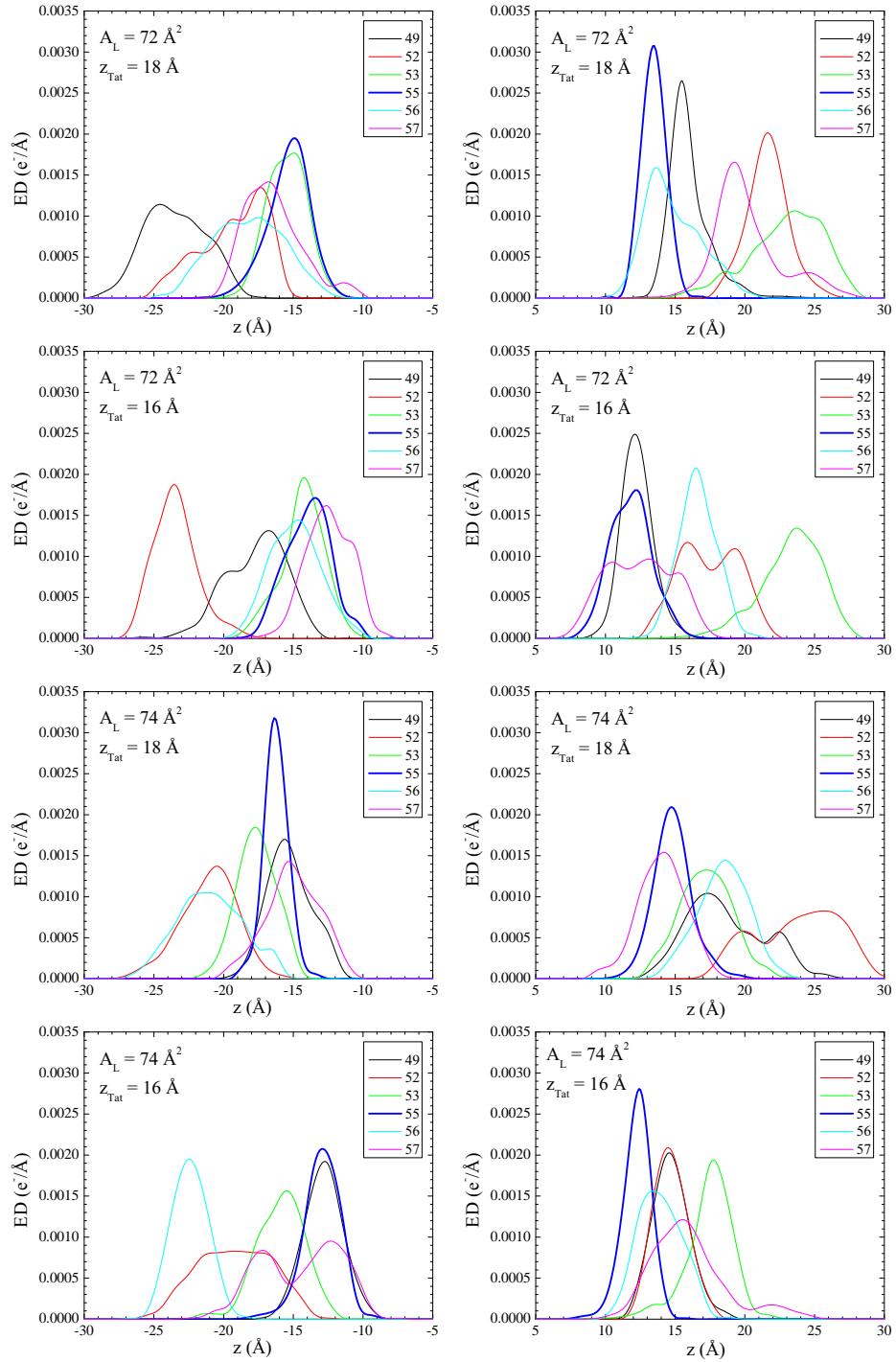


Figure 3.25: Electron density profiles of guanidinium groups from the four best matched simulations for DOPC with $x_{\text{Tat}} = 0.015$ (one Tat on each leaflet). Tat on the lower and upper leaflets are shown on the left and right plots, respectively.

in the MD simulation. However, as shown in Fig. 3.26 the corresponding form factor calculated from MD simulations does not match well with experiments.

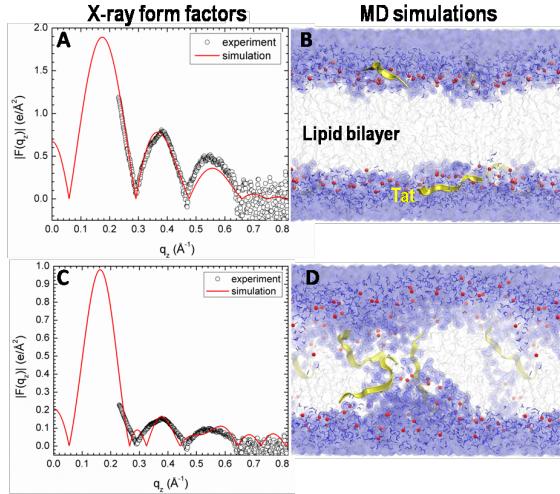


Figure 3.26: MD simulated form factors (red solid lines in A and C) of Tat/(DOPC+Tat), $x_{\text{Tat}}=0.030$, with Tat fixed at $z_{\text{Tat}}=18 \text{\AA}$ (panel A) and 5\AA (panel C) from the bilayer center compared to experimental form factors (open circles) scaled vertically to provide the best fit to the simulations. Corresponding snapshots are shown in Panels B and D in which the lipid chains are represented as grey sticks on a white background, Tats are yellow, phosphate groups are red and water is blue.

We summarize our results for how Tat affects the lipid bilayer in Fig. 3.27. The height of Tat, $H_{\text{Tat}} = 8.7 \text{\AA}$, was the full width at half maximum of the Tat electron density profiles obtained from simulations and the cylindrical radius, $R_{\text{Tat}} = 8.3 \text{\AA}$, was calculated to give the measured volume. The Z distances from the center of the bilayer were derived from weighted averages of four MD simulations of Tat:DOPC 2:128. The χ^2 obtained by comparison to experiment indicated that the best Z_{Tat} lay between the simulated values of 16 \AA and 18 \AA and the best area/lipid A_L lay between the simulated values of 72\AA^2 and 74\AA^2 , so averages were obtained from these four combinations of Z_{Tat} and A_L , weighted inversely with their χ^2 . The average positions, z'_{phos} , of phosphates situated underneath the Tats were calculated by averaging over the phosphates whose in-plane distance, R , from the center of Tat is smaller than R_{Tat} . The simulation cell extended to 38 \AA , far enough to ensure that z_{phos} for most of the lipids is the same as for DOPC. Assuming a simple linear ramp in z_{phos} , Fig. 3.27 then indicates a ring of boundary lipids that extends twice as far in R as Tat itself. Although the guanidinium electron density profile was broad (Fig. not yet

included), indicating that some were pointing away from the bilayer relative to the center of Tat, more were pointing towards the bilayer center as indicated in Fig. 3.27.

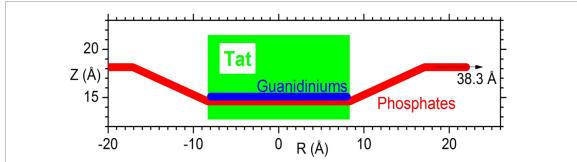


Figure 3.27: Location of Tat in DOPC bilayer. Tat is represented as a cylinder, z is the distance from the bilayer center, and R is the in-plane distance from the center of Tat. The average z of the lipid phosphates as a function of R and the arginine guanidiniums are shown in red and blue, respectively.

3.5 Discussion

Given that 8 of the 11 amino acids in Tat (47-57) are arginines and lysines, one would have suggested 20 years ago that highly charged Tat would partition strongly into solution rather than being associated with lipid bilayers. By contrast, but in agreement with more recent perspectives on arginine partitioning into the interfacial region [62], we find that Tat interacts with lipid bilayers, even with neutral DOPC and DOPC/DOPE mixtures, as well as with negatively charged DOPC/DOPS and nuclear membrane mimic lipid mixtures. This paper presents multiple lines of evidence for a Tat/membrane interaction. Fig. ?? shows that Tat decreases the bending modulus. Although one could argue that such a decrease is only apparent and could instead be due to local changes in membrane spontaneous curvature [63], either interpretation supports a Tat-bilayer interaction. The changes with increasing Tat concentration in the X-ray membrane form factors in Fig. ?? prove that Tat affects membrane structure, and the shift of the zero positions to higher q_z suggests thinning. Thinning is substantiated by quantitative analysis of the X-ray data and by MD simulations. Fig. 7A shows that the average membrane thickness, as measured by the distance D_{PP} between phosphocholines on opposite surfaces, decreases with increasing Tat concentration. Similar thinning is shown in Fig. 7B for the distance D_{HH} between the maxima in the electron density profiles of opposite surfaces. Compared to D_{PP} , D_{HH} is pulled towards both the carbonyl/glycerol groups and Tat because both have electron densities ($0.4 \text{ e}/\text{\AA}^3$) greater than water ($0.33 \text{ e}/\text{\AA}^3$) or hydrocarbon (0.3

$\text{e}/\text{\AA}^3$). Although the thinning shown in Figs. 7A and 7B is not large, it obviously requires interaction of Tat with the bilayers. Fig. 7C shows that A_L increases with increasing Tat concentration, by both model fitting and MD simulations.

It is of considerable interest to learn where Tat resides, on average, in the membrane, as this would establish a base position from which translocation would be initiated. We have combined our two main methods, MD simulations and X-ray scattering, to address this question. In general, Tats locate at the bilayer/water interface as indicated in Section 3.2, and they are close to the phosphocholine headgroup region by comparing the simulated 2ZTat in Fig. 7.D with 7.A. Although the SDP modeling of the X-ray data obtains excellent fits to the experimental form factors for a model with Tat deep in the hydrocarbon interior (see Fig. S5), the corresponding MD simulation (shown in Fig. 4.C) eliminates this spurious result. Fig. 7D also shows that modeling gives smaller values for z_{Tat} than the simulation. The modeling result is supportive of the original simulation result of Herce and Garcia that Tat resides closer to the bilayer center than do the phosphocholine groups [35]. That is a base position that would be a possibly important precursor to translocation, as would the larger A_L .

Several groups have carried out calculations and MD simulations showing that the cost of moving an arginine group from water to the bilayer center is 12-26 kcal/mol [62, 64–66] or 6-7 kcal/mol if side-chain snorkeling to the surface is taken into account [67]. This is not inconsistent with our result that Tat interacts with the membrane because, as is well known, the bilayer is not just a hydrocarbon slab, but has interfacial headgroup regions where Tat can reside. It has been suggested that the free energy cost for charged amino acids entering the headgroup region is similar to that for partitioning into octanol, about an order of magnitude smaller free energy cost than partitioning into cyclohexane [68–70]. Simulations suggest that the free energy is smaller for an arginine residing in the interfacial region than in water, roughly by 3 kcal/mole, depending upon the lipid [62, 70]. Our results therefore appear energetically reasonable.

One concern with diffraction experiments on samples consisting of adjacent bilayers in a stack or in a multilamellar vesicle is that the samples have to be partially dried to obtain conventional diffraction data. But then there is no pure water layer between adjacent bilayers, so a hydrophilic peptide is forced into the interfacial, partially hydrophilic region of the lipid bilayer. In contrast, by using diffuse scattering,

we obtained structure from experimental samples that had a range of lamellar D spacings (see Fig. 2 caption) that were considerably larger than the thickness of the bilayer in Fig. 7A, thereby providing an ample pure water space, typically greater than 20Å. The result that $2z_{\text{Tat}}$ shown in Fig. 7D is so much smaller than our repeat spacings shows that Tat preferentially associates with the membrane rather than dissociating into water.

Tat also increases the mosaic spread observed by X-ray and neutron scattering as shown in Figs. S1-3; this is a much larger scale disordering of the stack of bilayers.

We analyzed the secondary structures of Tats from MD simulations using the Define Secondary Structure of Proteins (DSSP) program [71]. Data from the MD simulation which has the best fit to experimental X-ray form factors show that Tat contains neither α - nor β -helix structures. It appears that the membrane does not influence the conformation of solubilized Tat.

Given our structural and elastic moduli results, we now compare to other experiments in the literature. In 2008, the Wong group implicated Tats ability to induce saddle-splay curvature with a potential role of bidentate hydrogen bonding as key [25]. Rhodamine-tagged Tat only entered GUVs when the PE headgroup was included with PS and PC lipids (PS/PC/PE, 20:40:40), indicating that hydrogen-bonding, and/or curvature-promoting lipids are required for Tat translocation. In PS/PE (20:80) lipids, they found Tat caused a highly curved cubic phase using X-ray diffraction [25]. In our experiments, there was little effect of adding DOPE to DOPC at either a 3:1 or 1:1 mole ratio on decrease in the bending modulus, bilayer thinning, or Tats outward movement with increasing concentration. Our two results are not inconsistent, however, since curvature-promotion appears not to be required for Tats ability to lower the energy required to bend nor to locate Tat in the bilayer, both of which may be important for Tat translocation. Yet Tat does translocate across membranes in their experiments only with PE in the membrane, so the ability to induce saddle-splay curvature may also be required for Tats translocation. An X-ray, neutron and AFM study reported thickening upon initial Tat binding, in contradiction to our result in Fig. 7B that shows thinning [72]. We suggest that this difference was caused by their using stiff gel phase DPPC lipid that did not allow bound Tat to perturb the bilayer. Using a variety of techniques, including high sensitivity isothermal titration calorimetry and ^2H - and ^{31}P -NMR, Ziegler *et al.* [?] presented evidence that the lipid bilayer remains intact upon Tat binding and our results confirm this.

Finally, we compare our structural results to those obtained by solid state NMR, although at a lower hydration level than in our sample. Su *et al.* [32] found that Tat lies parallel to the bilayer surface in the headgroup region of DMPC/DMPG (8:7) bilayers, similar to our cartoon in Fig. 9.

3.6 Conclusion

Although a recent MD simulation using umbrella sampling [73] found that the free energy required for R₉C to traverse a membrane was smaller if a water pore was present, we could not directly test the existence of a transient water pore from our X-ray scattering experiment. This is because, even with a water pore, the translocation process still requires crossing a free energy barrier which is a non-equilibrium process. X-ray form factors measure an equilibrium state. If the form factors obtained from water pore structures agreed well with experiments, it would indicate that the pore structure was thermodynamically stable. This may be the case for some antimicrobial peptides, but certainly not for cell-penetrating peptides. Finding a kinetically competent pathway for the interesting phenomenon of translocation of highly charged Tat through hydrophobic membranes is difficult. An energetically passive translocation likely occurs very seldom on an MD simulation time scale, and it probably happens quickly, so it would not significantly change the average structure of the membrane in which it occurs. Although our results in this paper do not reveal a kinetically competent pathway, they do show that Tat is drawn to the surface of the membrane, and is therefore ready for translocation at a region of local thinning. And they show that these interactions tend to soften (Fig. 2) the membrane and increase the area per lipid A_L , thereby likely reducing the energy barrier for passive translocation.

Appendices

Bibliography

- [1] Daniel C. Wack and Watt W. Webb. Synchrotron x-ray study of the modulated lamellar phase $p\beta'$ in the lecithin-water system. *Phys. Rev. A*, 40:2712–2730, Sep 1989.
- [2] Rainer Fischer, Mariola Fotin-Mleczek, Hansjrg Hufnagel, and Roland Brock. Break on through to the other sidebiophysics and cell biology shed light on cell-penetrating peptides. *ChemBioChem*, 6(12):2126–2142, 2005.
- [3] Alain Joliot and Alain Prochiantz. Transduction peptides: from technology to physiology. *Nat Cell Biol*, 6(3), 2004.
- [4] Maria Lindgren, Mattias Hllbrink, Alain Prochiantz, and lo Langel. Cell-penetrating peptides. *Trends in Pharmacological Sciences*, 21(3):99 – 103, 2000.
- [5] Alan D. Frankel and Carl O. Pabo. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*, 55(6):1189 – 1193, 1988.
- [6] Maurice Green and Paul M. Loewenstein. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell*, 55(6):1179 – 1188, 1988.
- [7] Eric Vifs, Priscille Brodin, and Bernard Lebleu. Hiv-1 tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *Journal of Biological Chemistry*, 272(25):16010–16017, 1997.
- [8] Gohar Ter-Avetisyan, Gisela Tnnemann, Danny Nowak, Matthias Nitschke, Andreas Herrmann, Marek Drab, and M. Cristina Cardoso. Cell entry of arginine-rich peptides is independent of endocytosis. *Journal of Biological Chemistry*, 284(6):3370–3378, 2009.

- [9] Gisela Tnnemann, Robert M. Martin, Simone Haupt, Christoph Patsch, Frank Edenhofer, and M. Cristina Cardoso. Cargo-dependent mode of uptake and bioavailability of tat-containing proteins and peptides in living cells. *The FASEB Journal*, 20(11):1775–1784, 2006.
- [10] Andr Ziegler, Pierluigi Nervi, Markus Drrenberger, and Joachim Seelig. The cationic cell-penetrating peptide cpptat derived from the hiv-1 protein tat is rapidly transported into living fibroblasts: optical, biophysical, and metabolic evidence. *Biochemistry*, 44(1):138–148, 2005. PMID: 15628854.
- [11] J. S. Wadia, R. V. Stan, and S. F. Dowdy. Transducible tat-ha fusogenic peptide enhances escape of tat-fusion proteins after lipid raft macropinocytosis. *Nature Medicine*, 10(3):310–315, 2004.
- [12] I. M. Kaplan, J. S. Wadia, and S. F. Dowdy. Cationic tat peptide transduction domain enters cells by macropinocytosis. *Journal of Controlled Release*, 102(1):247–253, 2005.
- [13] David A Mann and Alan D Frankel. Endocytosis and targeting of exogenous hiv-1 tat protein. *The EMBO journal*, 10(7):1733, 1991.
- [14] Jean Philippe Richard, Kamran Melikov, Hilary Brooks, Paul Prevot, Bernard Lebleu, and Leonid V Chernomordik. Cellular uptake of unconjugated tat peptide involves clathrin-dependent endocytosis and heparan sulfate receptors. *Journal of Biological Chemistry*, 280(15):15300–15306, 2005.
- [15] Simon W Jones, Richard Christison, Ken Bundell, Catherine J Voyce, Sarah Brockbank, Peter Newham, and Mark A Lindsay. Characterisation of cell-penetrating peptide-mediated peptide delivery. *British journal of pharmacology*, 145(8):1093–1102, 2005.
- [16] Agnès Vendeville, Fabienne Rayne, Anne Bonhoure, Nadir Bettache, Philippe Montcourier, and Bruno Beaumelle. Hiv-1 tat enters t cells using coated pits before translocating from acidified endosomes and eliciting biological responses. *Molecular biology of the cell*, 15(5):2347–2360, 2004.
- [17] Christina Foerg, Urs Ziegler, Jimena Fernandez-Carneado, Ernest Giralt, Robert Rennert, Annette G Beck-Sickinger, and Hans P Merkle. Decoding the entry of

- two novel cell-penetrating peptides in hela cells: lipid raft-mediated endocytosis and endosomal escape. *Biochemistry*, 44(1):72–81, 2005.
- [18] Antonio Fittipaldi and Mauro Giacca. Transcellular protein transduction using the tat protein of hiv-1. *Advanced drug delivery reviews*, 57(4):597–608, 2005.
 - [19] Ying Liu, Melina Jones, Cynthia M Hingtgen, Guojun Bu, Nick Laribee, Rudolph E Tanzi, Robert D Moir, Avindra Nath, and Johnny J He. Uptake of hiv-1 tat protein mediated by low-density lipoprotein receptor-related protein disrupts the neuronal metabolic balance of the receptor ligands. *Nature medicine*, 6(12):1380–1387, 2000.
 - [20] Vladimir P Torchilin, Ram Rammohan, Volkmar Weissig, and Tatyana S Levchenko. Tat peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors. *Proceedings of the National Academy of Sciences*, 98(15):8786–8791, 2001.
 - [21] Vladimir P Torchilin, Tatyana S Levchenko, Ram Rammohan, Natalia Volodina, Brigitte Papahadjopoulos-Sternberg, and Gerard GM D’Souza. Cell transfection in vitro and in vivo with nontoxic tat peptide-liposome–dna complexes. *Proceedings of the National Academy of Sciences*, 100(4):1972–1977, 2003.
 - [22] Carsten Rudolph, Christian Plank, James Lausier, Ulrike Schillinger, Rainer H Müller, and Joseph Rosenecker. Oligomers of the arginine-rich motif of the hiv-1 tat protein are capable of transferring plasmid dna into cells. *Journal of Biological Chemistry*, 278(13):11411–11418, 2003.
 - [23] Ashok Chauhan, Akshay Tikoo, Arvinder K Kapur, and Mahavir Singh. The taming of the cell penetrating domain of the hiv tat: myths and realities. *Journal of Controlled Release*, 117(2):148–162, 2007.
 - [24] JM Sabatier, E Vives, K Mabrouk, ABDELAZIZ Benjouad, H Rochat, A Duval, B Hue, and ELMOSTAFA Bahraoui. Evidence for neurotoxic activity of tat from human immunodeficiency virus type 1. *Journal of virology*, 65(2):961–967, 1991.
 - [25] A. Mishra, V. D. Gordon, L. H. Yang, R. Coridan, and G. C. L. Wong. Hiv tat forms pores in membranes by inducing saddle-splay curvature: Potential

- role of bidentate hydrogen bonding. *Angewandte Chemie-International Edition*, 47(16):2986–2989, 2008.
- [26] S. T. Yang, E. Zaitseva, L. V. Chernomordik, and K. Melikov. Cell-penetrating peptide induces leaky fusion of liposomes containing late endosome-specific anionic lipid. *Biophysical Journal*, 99(8):2525–2533, 2010.
 - [27] P. E. G. Thoren, D. Persson, E. K. Esbjorner, M. Goksor, P. Lincoln, and B. Norden. Membrane binding and translocation of cell-penetrating peptides. *Biochemistry*, 43(12):3471–3489, 2004.
 - [28] SD Krämer and H Wunderli-Allenspach. No entry for tat (44–57) into liposomes and intact mdck cells: novel approach to study membrane permeation of cell-penetrating peptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1609(2):161–169, 2003.
 - [29] C. Ciobanasu, J. P. Siebrasse, and U. Kubitscheck. Cell-penetrating hiv1 tat peptides can generate pores in model membranes. *Biophysical Journal*, 99(1):153–62, 2010.
 - [30] Philip A Gurnev, Sung-Tae Yang, Kamran C Melikov, Leonid V Chernomordik, and Sergey M Bezrukov. Cationic cell-penetrating peptide binds to planar lipid bilayers containing negatively charged lipids but does not induce conductive pores. *Biophysical journal*, 104(9):1933–1939, 2013.
 - [31] H. D. Herce, A. E. Garcia, J. Litt, R. S. Kane, P. Martin, N. Enrique, A. Rebollo, and V. Milesi. Arginine-rich peptides destabilize the plasma membrane, consistent with a pore formation translocation mechanism of cell-penetrating peptides. *Biophysical Journal*, 97(7):1917–1925, 2009.
 - [32] Y. C. Su, A. J. Waring, P. Ruchala, and M. Hong. Membrane-bound dynamic structure of an arginine-rich cell-penetrating peptide, the protein transduction domain of hiv tat, from solid-state nmr. *Biochemistry*, 49(29):6009–6020, 2010.
 - [33] S. Shojania and J. D. O’Neil. Hiv-1 tat is a natively unfolded protein - the solution conformation and dynamics of reduced hiv-1 tat-(1-72) by nmr spectroscopy. *Journal of Biological Chemistry*, 281(13):8347–8356, 2006.

- [34] P. Bayer, M. Kraft, A. Ejchart, M. Westendorp, R. Frank, and P. Rosch. Structural studies of hiv-1 tat protein. *Journal of Molecular Biology*, 247(4):529–535, 1995.
- [35] H. D. Herce and A. E. Garcia. Molecular dynamics simulations suggest a mechanism for translocation of the hiv-1 tat peptide across lipid membranes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(52):20805–20810, 2007.
- [36] S. Yesylevskyy, S. J. Marrink, and A. E. Mark. Alternative mechanisms for the interaction of the cell-penetrating peptides penetratin and the tat peptide with lipid bilayers. *Biophysical Journal*, 97(1):40–49, 2009.
- [37] Y. Lyatskaya, Y. F. Liu, S. Tristram-Nagle, J. Katsaras, and J. F. Nagle. Method for obtaining structure and interactions from oriented lipid bilayers. *Physical Review E*, 63(1):0119071–0119079, 2001.
- [38] Y. F. Liu and J. F. Nagle. Diffuse scattering provides material parameters and electron density profiles of biomembranes. *Physical Review E*, 69(4):040901–040904(R), 2004.
- [39] Yufeng Liu. *NEW METHOD TO OBTAIN STRUCTURE OF BIOMEMBRANES USING DIFFUSE -RAY SCATTERING: APPLICATION TO FLUID PHASE DOPC LIPID BILAYERS*. PhD thesis, Carnegie Mellon University, 2003.
- [40] John F Nagle and Stephanie Tristram-Nagle. Structure of lipid bilayers. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1469(3):159–195, 2000.
- [41] 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.
- [42] Stephanie Tristram-Nagle, Yufeng Liu, Justin Legleiter, and John F. Nagle. Structure of gel phase DMPC determined by x-ray diffraction. *Biophysical Journal*, 83(6):3324 – 3335, 2002.

- [43] 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.
- [44] <http://seal.web.cern.ch/seal/documents/minuit/mnusersguide.pdf>.
- [45] <http://lcgapp.cern.ch/project/cls/work-packages/mathlibs/minuit/index.html>.
- [46] 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.
- [47] 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.
- [48] 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.
- [49] 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.
- [50] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79(2):926–935, 1983.
- [51] Norbert Kuerka, John Katsaras, and JohnF. Nagle. Comparing membrane simulations to scattering experiments: Introducing the simtoexp software. *Journal of Membrane Biology*, 235(1):43–50, 2010.
- [52] Shuichi Miyamoto and Peter A Kollman. Settle: an analytical version of the shake and rattle algorithm for rigid water models. *Journal of computational chemistry*, 13(8):952–962, 1992.

- [53] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. Lincs: A linear constraint solver for molecular simulations. *J Comput Chem*, 18(12):1463–1472, 1997.
- [54] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An $n \log(n)$ method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.
- [55] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.
- [56] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.
- [57] Norbert Kučerka, Yufeng Liu, Nanjun Chu, Horia I Petrache, Stephanie Tristram-Nagle, and John F Nagle. Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-ray scattering from oriented multilamellar arrays and from unilamellar vesicles. *Biophysical journal*, 88(4):2626–2637, 2005.
- [58] Norbert Kučerka, Stephanie Tristram-Nagle, and John F Nagle. Closer look at structure of fully hydrated fluid phase dppc bilayers. *Biophysical journal*, 90(11):L83–L85, 2006.
- [59] Norbert Kučerka, Stephanie Tristram-Nagle, and John F Nagle. Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *The Journal of membrane biology*, 208(3):193–202, 2005.
- [60] Stephanie Tristram-Nagle, Chao-Ping Yang, and John F Nagle. Thermodynamic studies of purple membrane. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 854(1):58–66, 1986.
- [61] <http://www.basic.northwestern.edu/biotools/proteincalc.html>.
- [62] A. C. V. Johansson and E. Lindahl. The role of lipid composition for insertion and stabilization of amino acids in membranes. *Journal of Chemical Physics*, 130(18), 2009.

- [63] S. Tristram-Nagle and J. F. Nagle. Hiv-1 fusion peptide decreases bending energy and promotes curved fusion intermediates. *Biophysical Journal*, 93(6):2048–2055, 2007.
- [64] L. B. Li, I. Vorobyov, and T. W. Allen. Potential of mean force and pk(a) profile calculation for a lipid membrane-exposed arginine side chain. *Journal of Physical Chemistry B*, 112(32):9574–9587, 2008.
- [65] I. Vorobyov, L. B. Li, and T. W. Allen. Assessing atomistic and coarse-grained force fields for protein-lipid interactions: The formidable challenge of an ionizable side chain in a membrane. *Journal of Physical Chemistry B*, 112(32):9588–9602, 2008.
- [66] J. L. MacCallum, W. F. D. Bennett, and D. P. Tieleman. Distribution of amino acids in a lipid bilayer from computer simulations. *Biophysical Journal*, 94(9):3393–3404, 2008.
- [67] E. V. Schow, J. A. Freites, P. Cheng, A. Bernsel, G. von Heijne, S. H. White, and D. J. Tobias. Arginine in membranes: The connection between molecular dynamics simulations and translocon-mediated insertion experiments. *Journal of Membrane Biology*, 239(1-2):35–48, 2011.
- [68] W. C. Wimley, T. P. Creamer, and S. H. White. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry*, 35(16):5109–5124, 1996.
- [69] W. C. Wimley and S. H. White. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Structural Biology*, 3(10):842–848, 1996.
- [70] B. Roux. Lonely arginine seeks friendly environment. *Journal of General Physiology*, 130(2):233–236, 2007.
- [71] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–637, 1983.
- [72] D. Choi, J. H. Moon, H. Kim, B. J. Sung, M. W. Kim, G. Y. Tae, S. K. Satija, B. Akgun, C. J. Yu, H. W. Lee, D. R. Lee, J. M. Henderson, J. W. Kwong,

- K. L. Lam, K. Y. C. Lee, and K. Shin. Insertion mechanism of cell-penetrating peptides into supported phospholipid membranes revealed by x-ray and neutron reflection. *Soft Matter*, 8(32):8294–8297, 2012.
- [73] K. Huang and A. E. Garcia. Free energy of translocating an arginine-rich cell-penetrating peptide across a lipid bilayer suggests pore formation. *Biophysical Journal*, 104(2):412–420, 2013.
- [74] A. Tardieu, Vittorio Luzzati, and F.C. Reman. Structure and polymorphism of the hydrocarbon chains of lipids: A study of lecithin-water phases. *Journal of Molecular Biology*, 75(4):711 – 733, 1973.
- [75] JA Zasadzinski, J Schneir, J Gurley, V Elings, and PK Hansma. Scanning tunneling microscopy of freeze-fracture replicas of biomembranes. *Science*, 239(4843):1013–1015, 1988.
- [76] W J Sun, S Tristram-Nagle, R M Suter, and J F Nagle. Structure of the ripple phase in lecithin bilayers. *Proceedings of the National Academy of Sciences*, 93(14):7008–7012, 1996.
- [77] M. P. Hentschel and F. Rustichelli. Structure of the ripple phase P'_β in hydrated phosphatidylcholine multimembranes. *Phys. Rev. Lett.*, 66:903–906, Feb 1991.
- [78] Alex H. de Vries, Serge Yefimov, Alan E. Mark, and Siewert J. Marrink. Molecular structure of the lecithin ripple phase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15):5392–5396, 2005.
- [79] http://henke.lbl.gov/optical_constants.
- [80] M.C. Wiener, R.M. Suter, and J.F. Nagle. Structure of the fully hydrated gel phase of dipalmitoylphosphatidylcholine. *Biophysical Journal*, 55(2):315 – 325, 1989.