On Bilayer Deformation Energetics With and Without Gramicidin A Channel

By

© 2017

Andrew Harrison Beaven

B.S., Indiana State University, 2011

Submitted to the graduate degree program in Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Co-Chair: Wonpil Im

_____

Co-Chair: Ward H. Thompson

_____

Eric J. Deeds

_____

Krzysztof Kuczera

_____

Mario Rivera

Date Defended: 12 May 2017

The dissertation committee for Andrew Harrison Beaven certifies that this
is the approved version of the following dissertation:

On Bilayer Deformation Energetics With and Without Gramicidin A Channel

_____

Co-Chair: Wonpil Im

_____

Co-Chair: Ward H. Thompson

Date Approved: 12 May 2017

Abstract

Lipid membranes are not simply passive barriers. Embedded proteins are coupled to the membrane and can deform the surrounding bilayer, which incurs an energetic penalty. To minimize these penalties, proteins are known to tilt, aggregate, and experience major conformation changes. The degree to which the protein is influenced by the bilayer is dependent on the bilayer material properties and protein-bilayer coupling strength, for example. In this dissertation, the effects of bilayer material properties and protein-bilayer coupling are detailed using gramicidin A channel. This simple channel experiences one major conformational change, its transmembrane dimerization, which produces a bilayer deformation if the bilayer and dimer do not have the same hydrophobic lengths. Herein, molecular dynamics simulations are used to describe bilayer material properties, channel-bilayer coupling, and general lipid energetics with and without gramicidin A.
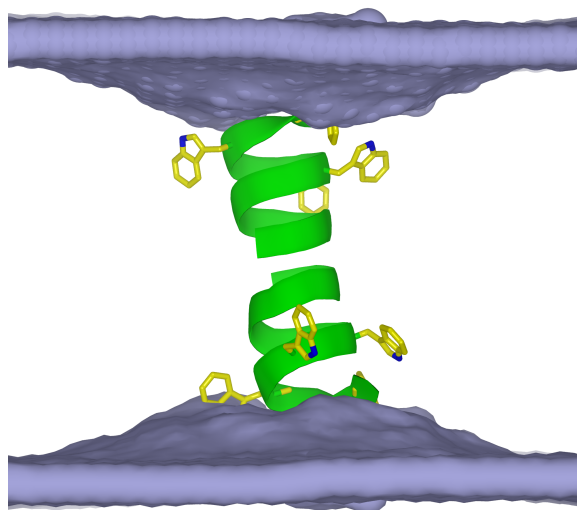
This work is dedicated to my parents, Lee W. and Clare A. Beaven. I cannot thank them enough.

Also, in honor/memory of my grandparents: Gordon L. and Dolena K. Beaven, and Lowell L. and Elizabeth (Betty) A. Kunce.

Acknowledgments

First, I thank my doctoral advisor, Dr. Wonpil Im, for his time, patience, knowledge, and caring. He is not only an outstanding researcher, but also has a deep appreciation for the importance of connectivity within science. With these connections, he provided me with more opportunities for academic and personal growth than I could have ever imagined, and likely ever deserved. With his guidance, I not only became a stronger and more competent scientist, but also a stronger and more competent human. For that, I will forever be indebted.

With Dr. Im's guidance, I have been fortunate to work with world-leading biophysicists: Dr. Olaf S. Andersen (Weill Cornell Medical College), Dr. Richard W. Pastor (NHLBI, NIH), Dr. Chaok Seok (Seoul National University), and Dr. Alexander J. Sodt (NICHD, NIH). Each of them has each been instrumental in my intellectual and personal growth, and their time invested in me has been invaluable. I also thank the Im Lab, particularly former members Dr. Sunhwan Jo and Dr. Kyu Il Lee, for helping me with both personal and academic matters.

I thank the members of my defense committee: Dr. Ward H. Thompson, Dr. Krzysztof Kuczera, Dr. Mario Rivera, and Dr. Eric J. Deeds, all of whom also served on my qualification committee in the spring of 2014, for their time and thoughtful comments and questions.

I have immense gratitude for my undergraduate advisor, Dr. Eric D. Glendening (Indiana State University), for introducing me to the world of science and scientific research, and also instilling a belief that I could find some success as a scientist.

Finally, I thank my family: my maternal grandparents, whose home has always been my second home; my brother for being a good distraction when I needed it; my father, who has *always* built my academic and personal confidence; and my mother, who has *always* been ready to hear me talk – whether it was in the middle of the day or the middle of night.

Table of Contents

List of Figures

List of Tables

1   Introduction

1.1   A brief historical background

The most obvious role of biological membranes is to serve as a barrier to define cell and organelle boundaries. Reinforcing this "simple barrier" mindset, older biophysical studies demonstrated that small molecules permeate biological membranes similarly as a ~30 Å thick sheet of liquid hydrocarbon (1, 2). This view of the membrane is both insightful and misleading. It is a correct view in that the hydrophobic region of the membrane is ~30 Å (typical hydrophobic thicknesses range from ~20–40 Å, e.g., Zhuang et al. (3)) and treating it like a bulk material has provided many useful insights (4–9). Additionally, biological membranes do serve as barriers between the inside and outside of cells and also compartmentalize organelles (for a nice review on basic lipid chemistry and organization, see Cooper (10)). However, these membranes are not simply barriers (see Andersen & Koeppe II for a review (11)), and the simplicity of viewing membranes as such comes with the cost of ignoring the incredible complexity that their components exhibit.

Biological membranes are assemblies of amphipathic molecules called lipids (10), which are composed of a hydrophilic head group and generally two or more hydrophobic tails. Given their chemical geometry, lipids aggregate with their head groups face the aqueous environment and their tails packed together to minimize water exposure. Therefore, there is a fairly distinct hydrophobic region encased by a region of "tumultuous chemical heterogeneity," where water begins interacting with the chemically diverse lipid head groups (12).

Lipids are a complex set of molecules, and the lipidome is thought to contain thousands of unique species (13, 14) – each of these species introducing possibly unique interactions with

other lipids and proteins, as well as possibly unique bulk properties (15–21). These interactions and properties depend on each species' unique chemistry.

1.2   Lipid packing and resulting properties

Lipids pack differently depending on their chemistry, ambient temperature, water/lipid ratio, and other factors (22–25), and many lipid types prefer non-bilayer phases (22). For example, an archetypical lipid type can be led through the various phases (**Figure 1.1**; from Brown (26)). At low temperature, there exists a gel state (high tail order and slow lateral diffusion) L_β, in which lipids tails have all *trans* bonds and are highly organized. Raising the temperature brings the lipids to a liquid-crystalline (liquid-disordered; low tail order and fast diffusion) state L_α, in which there are more *gauche* bonds and higher bilayer fluidity.
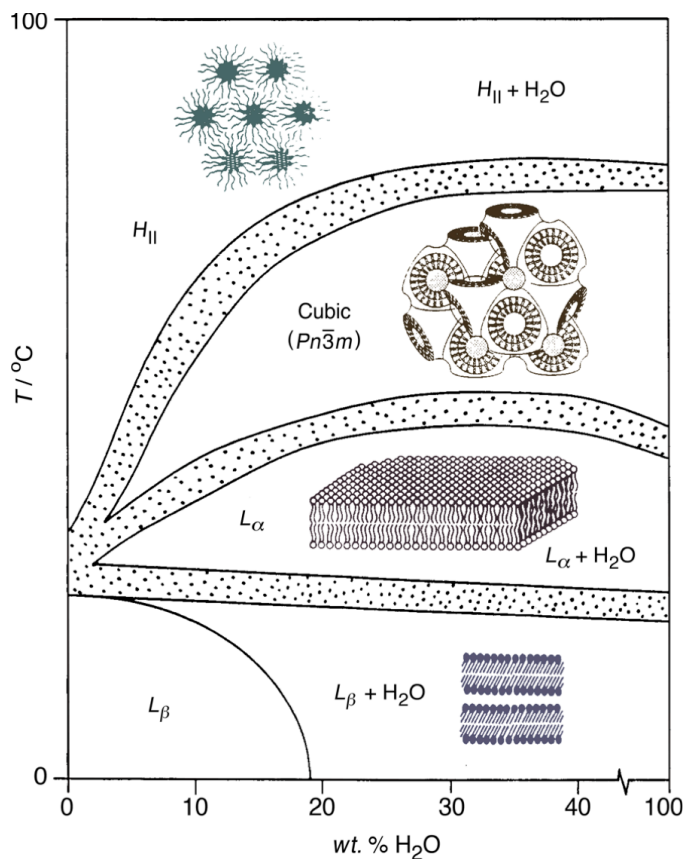
Figure 1.1: Representative phase diagram. Lipid head groups are represented by circles and acyl chains are lines. Reprinted with permission from Brown, M.F. Curvature Forces in Membrane Lipid–Protein Interactions. *Biochemistry*. 51:9782–9795. DOI:10.1021/bi301332v. Copyright 2012 American Chemical Society.

The exact characteristics of the phase diagram are strongly dependent on the chemical nature of the targeted lipid (or lipid mixture). A major factor in the phase equilibria is the curvature propensity of the lipid. A common, and perhaps elementary, method of thinking about lipid packing is by considering the molecule's "shape" (26). By this reasoning, a lipid with a large head group cross-sectional area (CSA) and small tail CSA would pack with so-called positive curvature (**Figure 1.2A**). Conversely, a lipid could have a relatively smaller head group CSA compared to its tail CSA. These lipids pack with so-called negative curvature (**Figure 1.2C**).

With some chemical intuition, one can predict the "shape" of a lipid. For two lipids with the same tails but different head groups, a lipid with a phosphatidylcholine ($-CH_2-N(CH_3)_3^+$; PC) head group will pack differently than a lipid with a phosphatidylethanolamine ($-CH_2-NH_3^+$; PE) head group (30–37). PE head groups can hydrogen bond with other PE head groups, meaning that it is favorable for them to be close together (negative intrinsic curvature; $R_0 = -29$ Å; (32, 37)). In contrast, a PC head group cannot hydrogen bond to other head groups, so their preferred packing is different than a PE lipid, but they still have negative intrinsic curvature; $R_0 = -87$ Å; (32)). Also reinforcing the "shape" concept, the addition of PC lysolipids (big head group CSA, small tail CSA) adds high positive curvature to leaflets, while PE lysolipids (small head group CSA, small tail CSA) is more cylindrical and adds little curvature (32). One can also imagine that tight packing of lipid head groups in a bilayer results in tight packing of the lipid tails (34–36). Indeed, the addition of PE lipids to PC bilayers increases the PC tail order (34).

Additionally, as we are taught in introductory biochemistry, lipids with long, saturated tails pack tighter than lipids with short, unsaturated tails (for overviews, see (38, 39)). The length of the lipid tail determines how many van der Waals (dispersion) interactions occur between different tails. If the tails are long, there are many favorable interactions, and the tails become

rigid. Similarly, unsaturated carbon atoms (typically *cis* double bonds) kill orderliness among lipid tails because the unsaturations introduce packing imperfections. These tail characteristics influence the phase diagram as well. For lipids with the same head group, those with high tail order have high melting (gel to liquid, i.e., L_β to L_α) transition temperatures. By the same reasoning, lipid order also affects the other phase transitions. Therefore, instead of considering a generic lipid "shape," it is better to think of a lipid's "chemical shape." That is, "What chemical properties does a lipid have that can lead to a certain preferred packing arrangement and material properties?"
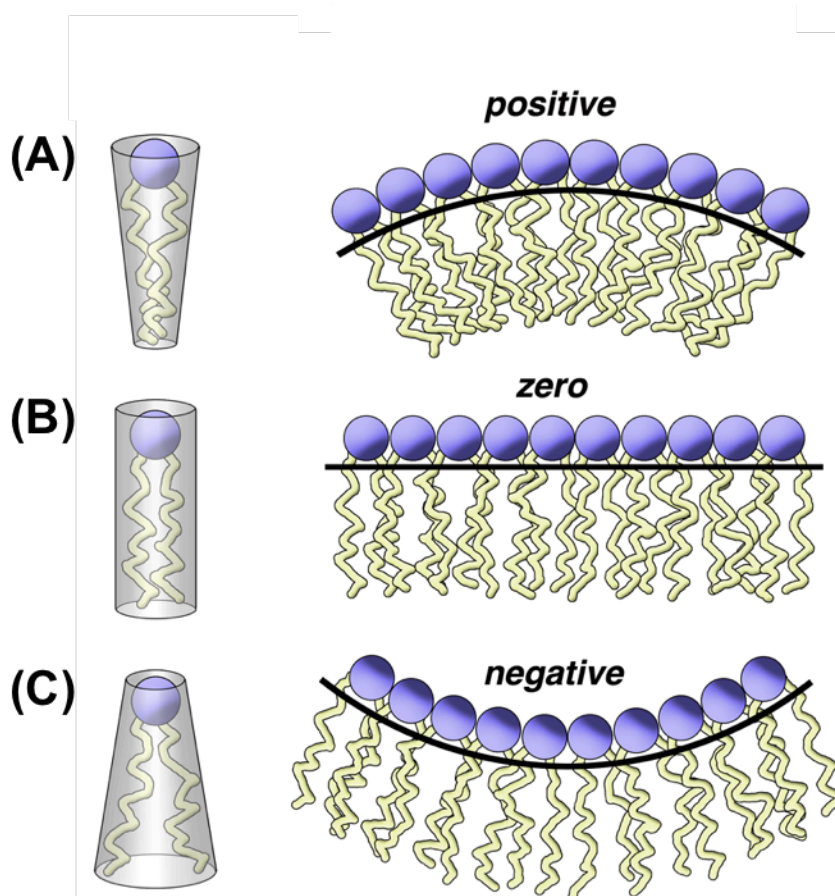
Figure 1.2: Lipid "shapes" and their effect on curvature propensity: (A) a lipid with a large head group (or small tails) creates positive curvature, (B) a cylindrical lipid creates zero curvature, and (C) a lipid with a small head group (or large tails) creates negative curvature. Lipid head groups are represented by circles and acyl chains are lines. Altered with permission from Brown, M.F. Curvature Forces in Membrane Lipid–Protein Interactions. *Biochemistry*. 51:9782–9795. DOI:10.1021/bi301332v. Copyright 2012 American Chemical Society.

For a single lipid leaflet or total bilayer, we might wonder how easy it is to bend and/or compress this leaflet/bilayer. All of the aforementioned chemical factors manifest themselves lipid material properties, which can be used to answer questions on membrane energetics. There are various material moduli that address how easily lipids can be bent, compressed, tilted, and twisted (31–33, 37, 40–46); each with an associated energy. This dissertation details the bilayer energetics around a cylindrical membrane channel, so lipid bending and compression energies are detailed (**Sections 1.3** and **1.4**). Lipid twist has not been deeply studied beyond theory, and a cylindrical channel likely would not have much effect on twisting energetics. Lipid tilt could be a factor with respect to how lipids minimize compression and bending penalties (or adjust to non-cylindrical proteins), and this could be studied in the future.

1.3   The area compressibility modulus, K_A

Obtaining the correct surface area for a lipid type has been a focus of the CHARMM force field community from the beginning of its lipid parameter development. In the earlier days of the CHARMM lipid force field, surface tensions ($\gamma > 0$) were applied to enforce area per lipid values close to experiment (41, 47). Improvements in the force field have allowed bilayer simulations to be run under tensionless conditions ($\gamma = 0$), and agreement between simulation and experiment is generally good (35).

The proper average area per lipid and fluctuation around this average is energetically important as they define the area compressibility modulus, K_A (41):

$$K_A = \frac{k_B T}{\langle \delta A^2 \rangle} = A \left(\frac{d\gamma}{dA} \right)_T ,$$

(1.1)

where k_B is Boltzmann's constant, T is the temperature, $\langle A \rangle$ is the average total area, and $\langle \delta A^2 \rangle$ is the mean square area fluctuation. Thus, the fluctuations of a simulation's unit cell can be used to obtain K_A (**Chapter 2**). Strictly speaking, K_A relates to the ease of which a lipid (or group of

lipids) changes its xy area. However, if it is assumed that lipids compress with constant volume, a change in area (xy) is directly proportional to a change in height (z) (48). Throughout this dissertation, all lipids are aligned with their primary axis along z unless explicitly noted otherwise (**Chapter 5**). Therefore, we have a method to describe the energetics of lipid compression along the z-axis (E_{comp}):

$$E_{comp} = \frac{K_A}{2}\left(\frac{d_0 - d}{d_0}\right)^2 = \frac{K_A}{2}\left(1 - \frac{d}{d_0}\right)^2, \tag{1.2}$$

where d_0 and d are the unperturbed and perturbed bilayer thicknesses, respectively.

1.4   Per area free energy with respect to curvature, $\overline{F}'(0)$

As noted above, PC lipids have a relatively small, but negative intrinsic curvature (i.e., they would bend toward their head groups if unconstrained as a single leaflet). However, if two leaflets form a bilayer, they are constrained to be planar (and further constrained to be planar in simulations with periodic boundary conditions). This restriction to planarity introduces stress (internal forces from neighboring atoms), within the leaflets (**Figure 1.3**) (11, 26). The pressure within a leaflet is not directly obtainable from experiment, but can be calculated by the molecular dynamics program NAMD, for example (49–51). NAMD produces a lateral pressure profile, $p(z) = p_L(z) - p_N(z)$, where $p_L(z)$ and $p_N(z)$ are the lateral and normal components of the pressure tensor; i.e., $p_L(z) = \left[p_{xx}(z) + p_{yy}(z)\right]/2$ and $p_N(z) = p_{zz}(z)$.

Following **Figure 1.3** from $z = 0$ to $+\infty$, there are: i) relatively small positive pressures in the bilayer core due to tail entropy; ii) a large, negative pressure at the bilayer-water interface associated with tail hydrophobicity; iii) a large positive pressure associated with head group repulsions; and iv) in an equilibrated system, a pressure of 1 bar at large z, which is the applied system pressure (11, 26).
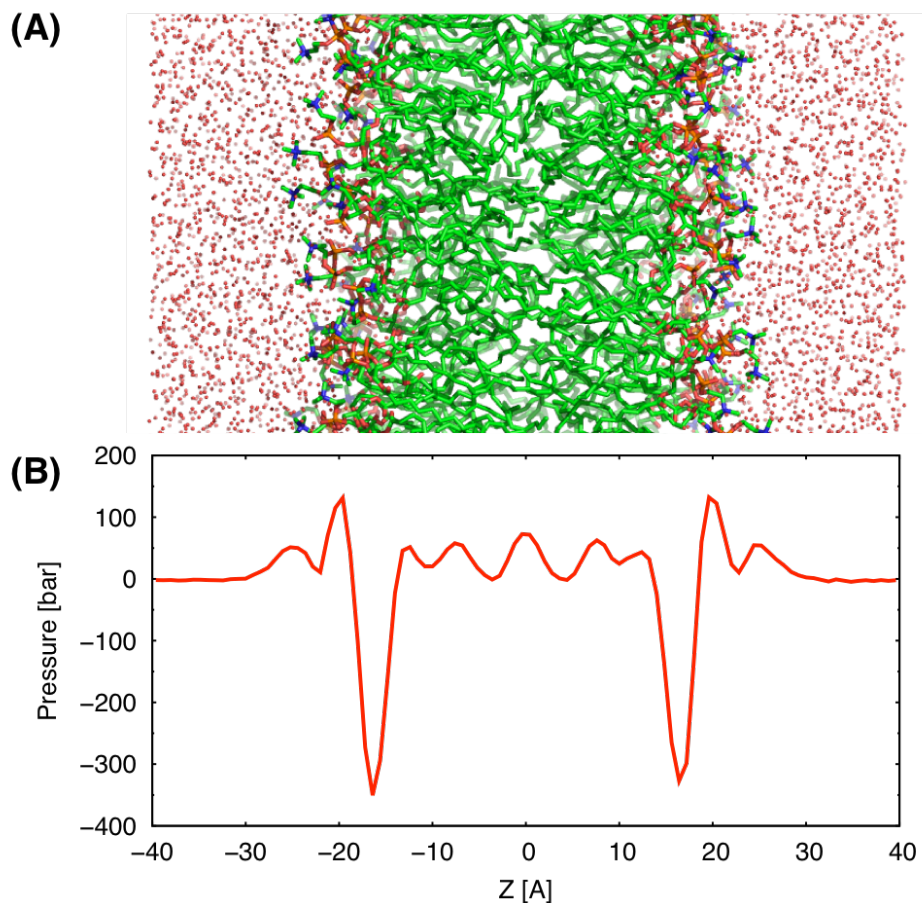
Figure 1.3: (A) DOPC bilayer that is enforced to be planar by hydrophobic and periodic boundary considerations. The individual leaflets would separate and bend toward their head groups if able. (B) A pressure profile where the **z**-axis is the membrane normal. Data from Sodt et al. (9)

The pressures within the bilayer have direct relation to physical quantities. These quantities are also discussed when applicable within each chapter. For a planar bilayer, the surface tension is the integral of the pressure profile (37, 52–54):

$$\gamma = -\int_{-\infty}^{\infty} p(z)\,\mathrm{d}z. \tag{1.4}$$

If a simulation is performed at $\gamma = 0$, the integral of the pressure profile should be zero. As discussed in **Chapters 2** and **3**, it is often more convenient to consider leaflet characteristics rather than total bilayer characteristics. In the case of the surface tension, the integration limits could be changed to $-\infty$ to 0 and 0 to ∞ (or, more appropriately for simulations with periodic

7

boundary conditions, $-L_z/2$ to 0 and 0 to $L_z/2$ where L_z is the unit cell size along z). Assuming that the leaflets are acting independently of each other, calculating leaflet quantities can effectively double the statistical sampling.

In model systems, the Helfrich Hamiltonian ($\bar{F}_H$) for the per area energy with respect to curvature is calculated by (55, 56):

$$\bar{F}_H(c_1, c_2) = \frac{k_c}{2}(R_1^{-1} + R_2^{-1} - R_0^{-1})^2 + k_G R_1^{-1} R_2^{-1}, \tag{1.5}$$

where k_c is the bending modulus, R_1^{-1} and R_2^{-1} are curvatures of two principle axes, R_0^{-1} is the intrinsic curvature, and k_G is the Gaussian bending modulus (or saddle splay modulus). Note that R indicates a radius of curvature, and that curvatures are also written as $c = R^{-1}$. It has been shown that Gaussian curvature is only important when there is a large change in membrane topology (e.g., pore formation or fission/fusion events; instances of high curvature structures) (summarized in (57)). This leads to a simplified expression of:

$$\bar{F}_H = \frac{k_c}{2}(R^{-1} - R_0^{-1})^2, \tag{1.6}$$

where R^{-1} has become the overall curvature ($R^{-1} = R_1^{-1} + R_2^{-1}$).

The derivative of the Helfrich Hamiltonian (Equation 1.6) with respect to curvature related to the first moment of the pressure profile (37, 52–54):

$$\bar{F}'(0) = \frac{d\bar{F}}{dR^{-1}}\Big|_{R^{-1}=0} = -k_c R_0^{-1} = -\int_{-L_z/2}^{L_z/2} zp(z)\,dz, \tag{1.7}$$

where the integral is written in terms of the unit cell size (L_z) centered at $z = 0$. It is of note that k_c and c_0 are obtainable from the lipid hexagonal phase in experiment, and more recently, simulation (see **Chapter 5**). Conversely, the pressure profile is accessible from simulation and not experiment. This derivative of the Helfrich Hamiltonian relates the curvature frustration within a bilayer (or leaflet, given the correct integration limits), and states what will occur to a

leaflet to minimize its energy by returning to R_0^{-1}, if it was able to do so. Since k_c is always positive, a positive value of $\bar{F}'(0)$ indicates that a lipid leaflet would bend toward its head groups if able (a negative curvature preference) and a negative $\bar{F}'(0)$ indicates that a lipid leaflet would bend toward its tails if able (a positive curvature preference). The magnitude of $\bar{F}'(0)$ indicates the degree of frustration that the leaflet is experiencing (i.e., $\bar{F}'(0) = 0$ denotes a system with no curvature frustration).

1.5   Protein-bilayer interactions

Though there are many lipid binding proteins in the proteome (58), the vast majority of proteins must interact with the bilayer by a more general means. These general interactions can be viewed in a continuum sense that a bilayer's bulk properties (hydrophobic thickness, intrinsic curvature, and moduli) affect protein behavior (11). What is intimately related is that the protein also affects the bilayer through deformations. The deformation of the bilayer can either stabilize or destabilize a given protein conformation.

In 1984, Mouritsen and Bloom published their seminal work describing the "mattress model" of matching between hydrophobic protein domains and the hydrophobic thickness of the bilayer (59). That is, bilayers that are too thick for a protein's hydrophobic domain must compress and bilayer that are too thin for a protein's hydrophobic domain must stretch. This model is very beneficial, but sometimes too simplistic because proteins are known to aggregate, tilt, and perform major conformational changes in response to hydrophobic mismatch (see Killian for a review (60)).

The energy involved with deforming the bilayer has a cost involving the material properties of the bilayer, and there is now a rich background of examples demonstrating how membrane proteins are regulated by differences in hydrophobic mismatch and lipid intrinsic curvature (see

Andersen and Koeppe II, for a review (11)). This dissertation details the bilayer energetics surrounding a semi-cylindrical channel, gramicidin A (gA) as a model for other embedded membrane proteins (**Figure 1.4**). Gramicidin is unique in that it is semi-cylindrical and forms conducting dimer channels (D) by a transmembrane dimerization of two non-conducting monomers (M), which float in opposite leaflets (61–63):

$$2M \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} D \tag{1.8}$$
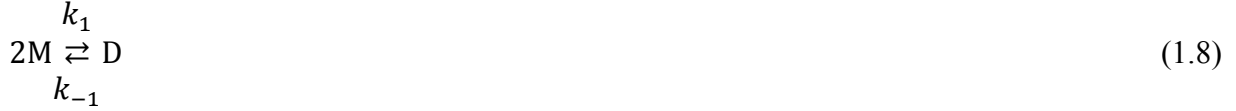
where k_1 and k_{-1} are the association and dissociation rate constants, respectively. The relative concentration between monomers and dimers is determined by the energy associated with forming/breaking the six inter-monomer hydrogen bonds ($\Delta G_{\text{protein}}^{\text{M}\to\text{D}}$) and the bilayer deformation is caused by the dimerization event ($\Delta G_{\text{bilayer}}^{\text{M}\to\text{D}}$) (18, 62, 64, 65):

$$\frac{[\text{D}]}{[\text{M}]^2} = \frac{k_1}{k_{-1}} = \exp\left\{-\frac{\Delta G_{\text{protein}}^{\text{M}\to\text{D}} + \Delta G_{\text{bilayer}}^{\text{M}\to\text{D}}}{k_{\text{B}}T}\right\}. \tag{1.9}$$

In general, the bilayer deformation energy caused by a semi-cylindrical gA is characterized by:

$$\Delta G_{\text{def}}^0 = \int_{r_0}^{\infty}\left[\frac{K_A}{2}\left(\frac{2u}{d_0}\right)^2 + \frac{k_c}{2}(\nabla^2 u - R_0^{-1})^2\right]\cdot 2\pi\cdot r\,dr - \int_{r_0}^{\infty}\left[\frac{k_c}{2}\cdot(R_0^{-1})^2\right]\cdot 2\pi\cdot r\,dr \tag{1.10}$$

where $u = (d_0 - d)/2$ with d_0 being the unperturbed bilayer thickness and d being the local, perturbed bilayer thickness (7, 9, 66, 67). Thus, u describes a single leaflet deformation, the second derivative of which describes the local curvature of the deformation ($\nabla^2 u = R^{-1}$). Even without a deformation, there is generally a curvature energy associated with forming the bilayer, and this is represented by the second integral. If $2u/d_0$ and $\nabla^2 u$ are zero, there is still some latent energy, $k_c \cdot (R_0^{-1})^2$, associated with the general frustration a leaflet feels due to being in a bilayer. The subtraction of this energy ensures that ΔG_{def}^0 is due to the deformation.
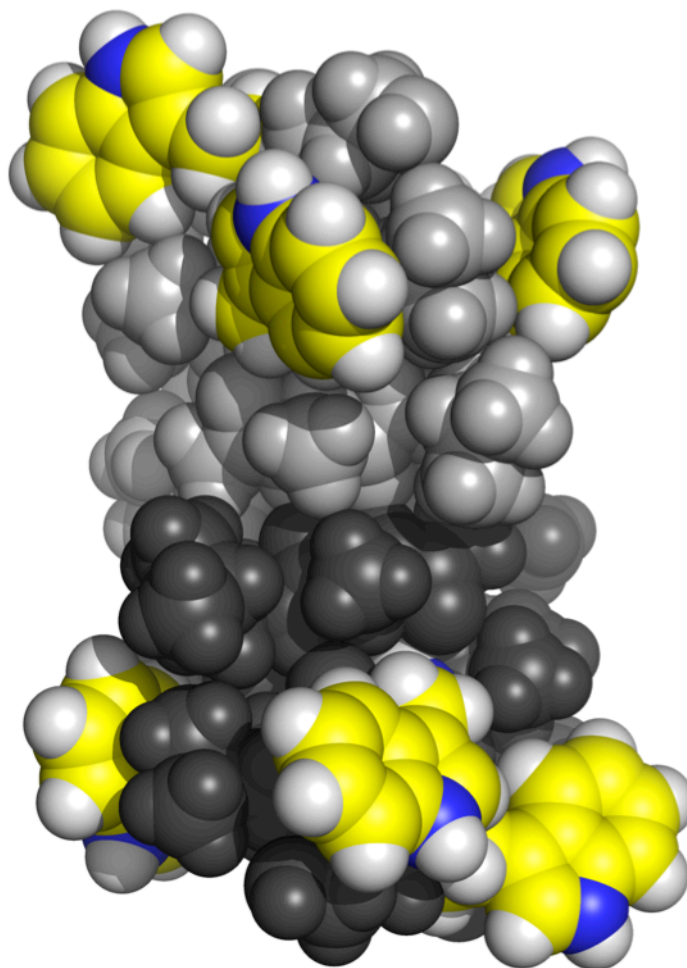
Figure 1.4: gA dimer with monomers colored light and dark grey. Tryptophan residues are yellow with indole nitrogen in blue.

1.6   Dissertation outline

This dissertation balances the two views of lipids: i) that they exist in a bilayer as a bulk material and ii) that individual lipids interact specifically with embedded proteins. The focus of **Chapters 2, 3,** and **4** is the interplay between gA dimers and monomers and the bilayer. In each instance, it is demonstrated that bulk properties as well as individual lipid-channel interactions are important. **Chapter 5** details building inverse hexagonal lipid systems to obtain the bending

modulus and intrinsic radius of curvature. In all cases, molecular dynamics simulations were used to provide atomistic insight into the complexity of protein-bilayer interactions.

Specifically, **Chapter 2** describes how two-component bilayers rearrange to minimize hydrophobic exposure and bilayer frustration. The results of long timescale (3.5 µs) simulations are compared to experimental results and a simple compression-only energy model. **Chapter 3** is the companion piece to **Chapter 2** and further details the energetics of bilayer deformations due to gA. **Chapter 4** is a study of how gA mutations affect bilayer energetics and how these different interactions affect explain experimentally measured channel lifetimes. In **Chapter 5**, a methodology for quickly and reproducibly building the lipid inverse hexagonal phase is developed and justified. This module in CHARMM-GUI (http://charmm-gui.org) will be made available to the public and hopefully allow more researchers to investigate this useful, but infrequently used lipid phase. Finally, **Chapter 6** provides some overarching conclusions and thoughts for future research based on the results from this dissertation.

2   Gramicidin A Channel Formation Induces Local Lipid Redistribution: Experiment and Simulation[a]

2.1   Introduction

As mentioned in **Chapter 1.1**, cell membrane lipid composition is diverse and tightly regulated (68, 69), with the cellular lipidome containing 1,000+ unique species (13, 14). Each of these species has distinct physical properties that contribute to the organization of the bilayer and the energetics of membrane protein-bilayer interactions. Studies using single-component bilayers have demonstrated how changes in a given lipid characteristic, such as acyl chain length, may alter protein function (15–21) and highlighted the importance of the hydrophobic coupling between a protein and its host lipid bilayer (11). Mismatch between a protein's hydrophobic domain and the bilayer's hydrophobic core is highly unfavorable, as exposing hydrophobic residues to water incurs a 25–75 cal/(mol·Å^2) energetic penalty (11, 70). Deforming the bilayer to match the protein's hydrophobic domain also has an associated energetic cost, which includes compression and bending energy contributions (7, 59, 66, 71, 72). These bilayer contributions to the change in free energy are known to affect the conformational preference of many diverse membrane proteins (see Andersen and Koeppe (11) for a review).

In a multicomponent bilayer, these energetic penalties may be reduced if a particular lipid species redistributes laterally to be preferentially enriched near a given protein. This lipid sorting has been studied experimentally with Ca^{2+}-ATPase (SERCA) (15, 73, 74) and bacteriorhodopsin (BR) (75), which showed limited lipid preference in liquid-crystalline bilayers. Both SERCA and BR, however, aggregate when embedded in bilayers that are either very thin or thick (76, 77), such that the protein/bilayer boundary can vary as the bilayer thickness is varied, which may

---

obscure any lipid sorting. KcsA (K^+ channel) (78), however, exhibited clear selectivity among lipids with different chain lengths (a three-fold change in the preference for the protein/bilayer interface between $dC_{10:1}$ and $dC_{22:1}$).

Less is known about single membrane-spanning α-helices, but the available evidence shows little, if any, acyl chain-length-dependent preference among phospholipids (79). This could reflect that the imperfect hydrophobic matching between single bilayer-spanning α-helices and their host bilayers (80) or the perfect matching by helix tilting (81). In the case of gramicidin channels (showing only relatively small tilting), Fahsel et al. (82) found a preference for the thinner di-myristoyl-phosphatidylcholine ($dC_{14:0}$, DMPC) in studies on $dC_{14:0}$-$dC_{18:0}$ mixtures in both fluid-gel and gel-gel coexistence.

The lateral distribution of lipids also has been explored in simulation studies. Mouritsen and colleagues (75, 83) used lattice Monte Carlo lipid models and a smooth cylindrical membrane protein to show that lateral lipid sorting should be feasible, although there was little evidence for such sorting in the experiments on BR (75). Later Monte Carlo-molecular simulations on gramicidin channels or the OmpA protein embedded in $dC_{14:0}$-$dC_{18:0}$ mixtures in the liquid-crystalline state similarly found little enrichment of either lipid adjacent to the channel or membrane protein (84, 85). Coarse-grain simulations, however, have more convincingly demonstrated lipid redistribution as a function of hydrophobic mismatch (85–87).

This study explores the redistribution of liquid-crystalline lipid species with different acyl chain lengths adjacent to a bilayer-spanning channel, [Val[1]]gramicidin (gA), which has been extensively characterized by electrophysiology (18, 88–90), spectroscopy (91–95), and molecular dynamics (MD) simulations (20, 96–98). Gramicidin channels are small and cylindrical (allowing for lipid radial properties to be easily calculated) with ample evidence for

hydrophobic matching between the channel and the host bilayer (99), and it can be studied at single-molecular resolution (where there is no lateral aggregation). Hence, gA is an excellent tool for probing the physics underlying lipid redistribution.
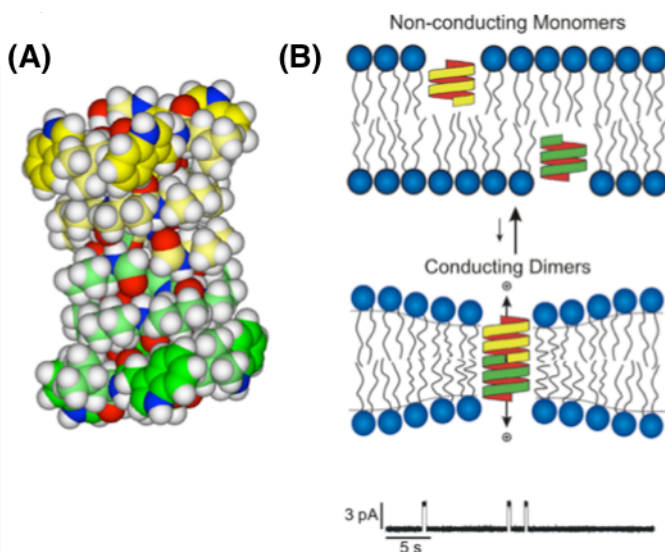
Figure 2.1: (A) Space-filling model of [Val^1]gramicidin; the two β-helical subunits are indicated by yellow and green coloring of the carbon atoms. (B) Gramicidin channels form by the transmembrane dimerization of two non-conducting subunits. Channel formation and dimer dissociation are "visible" as rectangular current transitions. Channel formation is associated with a local decrease in bilayer thickness, with an associated energetic cost, which gives rise to a disjoining force that the bilayer imposes on the channel. Thus, the average channel lifetimes depend on the channel-bilayer hydrophobic mismatch.

gA channels have only one major conformational transition, which is their formation by transmembrane dimerization of two nearly cylindrical monomers residing in opposite leaflets with their axes aligned to form a channel across the bilayer (**Figure 2.1**). That gA has only a single major conformational change is important, in that it mainly adjusts to its environment through changes in the monomer↔dimer equilibrium, with minimal changes in other properties, such as the helical pitch (100) and single-channel conductance (101). Additionally, the single-leaflet monomer most likely produces minimal bilayer deformations (20). Thus, the bilayer must deform to accommodate the dimer (and the dimer-to-monomer transition state), meaning that any

local enrichment of a preferred lipid in the vicinity of a gramicidin dimer must be reflected in the channel lifetime.

First, single-channel experiments demonstrate that gA single-channel lifetimes are increased (relative to a single component reference bilayer) in bilayers formed from lipid mixtures, where the lipids can redistribute to minimize the total bilayer deformation energy. The tentative conclusion from the electrophysiological experiments, that lipid redistribution occurs, was then substantiated using long-timescale MD simulations to determine the lipid redistribution as a function of time and distance from the channel. The relevant timescales and concentrations are discussed, and a comparison between experimental and MD results with a discussion of redistribution energetics follows. Because the lipid environment can modulate any embedded protein's function (11, 102), results and conclusions here should be applicable to a wide range of membrane-embedded proteins.

It is important to note that although redistribution is an important feature by itself, the magnitude of redistribution is related to channel-bilayer hydrophobic mismatch, and so it can confirm model energetics that in turn affect protein function. When considered with the results of the second part of this work (seen in **Chapter 3**, which briefly develops and analyzes a continuum model of hydrophobic mismatch), a cross-validated base is built for interpreting the effect of protein-lipid coupling on function and compositional heterogeneity. The connections among experiment, simulation, and theory are briefly summarized at the conclusion of this article to emphasize how our results provide, to our knowledge, new insight into the physics of lipid redistribution.

2.2   Methods

2.2.1   Experimental methods

Di-palmitoleoyl-phosphocholine, $(16:1c\Delta^9)_2PC$ $(dC_{16:1})$, di-oleoyl-phosphocholine, $(18:1c\Delta^9)_2PC$ $(dC_{18:1})$, di-eicosenoyl-phosphocholine, $(20:1c\Delta^{11})_2PC$ $(dC_{20:1})$, di-erucoyl-phosphocholine, $(22:1c\Delta^{13})_2PC$ $(dC_{22:1})$ and di-nervonoyl-phosphocholine, and $(24:1c\Delta^{15})_2PC$ $(dC_{24:1})$, were from Avanti Polar Lipids (Alabaster, AL). n-Decane (99.9% pure) was from Wiley Organics (Columbus, OH). gA was a gift from Drs. Roger E. Koeppe II and Denise V. Greathouse (Dept. of Chemistry and Biochemistry, University of Arkansas). gA was dissolved in ethanol (200 proof) or DMSO (Sigma-Aldrich, St. Louis, MO). NaCl was Purissimum grade from Sigma-Aldrich. Electrolyte solutions were prepared daily using deionized Milli-Q water (Millipore Corp., Bedford, MA).

Planar lipid bilayers were painted with a solution of lipid/n-decane across a 0.8–1.2 mm diameter hole in a Teflon partition separating two Teflon chambers each containing 5 ml of solution (unbuffered 1.0 M NaCl). Single channel experiments were done using the bilayer punch (88). All experiments were done at 25 ± 1°C, with a membrane potential of ±200 mV. The current signal was amplified using an AxoPatch 1B (Axon Instruments, Foster City, CA). The signal was filtered at 500–1000 Hz, digitized, and sampled by a PC/AT compatible computer. Single-channel average lifetimes (τ) were determined by fitting single exponential distributions to the lifetime distributions, $N(t)/N(0) = \exp(-t/\tau)$, where $N(t)$ is the number of channels with a duration longer than t. All experiments were done at very low gA/lipid molar ratios ($\sim 1/10^5$) to minimize any uncertainties associated with lateral interactions among gA channels.

2.2.2 Simulation methods

2.2.2.1 System setup

The simulation systems were built using the *Membrane Builder* module (103, 104) in CHARMM-GUI (www.charmm-gui.org) (105). A single gA dimer (PDB:1JNO) (94) was

inserted into three different bilayers ($dC_{16:1}+dC_{24:1}$, $dC_{18:1}+dC_{22:1}$, and $dC_{20:1}$) (**Figure S2.1**). Each bilayer has 90 lipids per leaflet, which has been shown to have at least three lipid shells around the protein and extend into the effective bulk in previous gA simulations (20). 0.15 M KCl was used for all simulations. Including water and ions, each of the three systems contains ~63,000 atoms (see **Supplemental information Table S2.1** for system information).

Each system was initially minimized and equilibrated using CHARMM (106). After this short procedure, a further 60-ns equilibration was performed using NAMD (51) in the isothermal-isobaric (NPT) ensemble at 303.15 K and 1.0 atm. Langevin dynamics were used to maintain constant temperature with a Langevin coupling coefficient of 1 ps^{-1}. A Nosé-Hoover Langevin piston (107, 108) maintained constant pressure with a piston period of 50 fs and a piston decay of 25 fs. All simulations were run using P1 periodic boundary conditions. The CHARMM all-atom C22 protein force field (109) including dCMAP (110, 111) was used together with the C36 lipid force field (35) and a TIP3P water model (112). A 2-fs time-step was used along with the SHAKE algorithm (113). Electrostatic interactions were calculated using the particle-mesh Ewald method (114) (mesh size ~1 Å, $\kappa = 0.34$ $Å^{-1}$, and sixth-order B-spline interpolation), and van der Waals interactions were smoothly switched off between 10–12 Å by a force-switching function (115).

The coordinates from the end of the NAMD simulations were used to initiate 3.5-μs simulations on Anton, a special-purpose supercomputer designed for long timescale MD simulations (116), using the same force fields and water model as above. The NPT ensemble was employed in these Anton simulations with the pressure and temperature held constant at 1 bar and 303.15 K using Berendsen's coupling scheme (117). The lengths of all bonds involving hydrogen atoms were constrained using M-SHAKE (118), and the cutoff distance of the van der

Waals and short-range electrostatic interactions was set to 10.05 Å. Long-range electrostatic interactions were evaluated with the k-space Gaussian split Ewald method (119) using a 64×64×64 mesh. The r-RESPA integration method (120) was employed with a time-step of 2 fs, and the long-range electrostatic interactions were evaluated every 6 fs. Due to the length of the simulation, atomic coordinates were saved every 0.24 ns.

In addition to the gA-bilayer systems, single- and two-component bilayer-only simulations were performed to gather equilibrium bilayer properties (thickness, area compressibility, and per lipid area) to better quantify how lipids adapt to the channel and each other in the mixed bilayers, as well as to use the computed bilayer properties for continuum model calculations. Single-component bilayers ($dC_{16:1}$, $dC_{18:1}$, $dC_{20:1}$, $dC_{22:1}$, and $dC_{24:1}$) were built with 50 lipids per leaflet. Two-component bilayers ($dC_{18:1}+dC_{22:1}$ and $dC_{16:1}+dC_{24:1}$) were built with 90 lipids per leaflet (see **Table S2.1** for system information). Each bilayer was simulated for at least 100 ns with the same simulation parameters and protocols described above for the gA-bilayer NAMD simulations.

2.2.2.2 Lipid shell definition

Lipids that are approximately within the same radial distance from a protein define a lipid shell around the protein, and lipids in the same lipid shell generally have similar bilayer properties (e.g., thickness, area per lipid, and compressibility). Recent studies using estimations of shell locations based on one-dimensional (1D) radial and two-dimensional (2D) density distribution functions have yielded some success (20, 121, 122). However, lipid radial or density distribution functions often do not allow for high-resolution data of the location, population, and bilayer properties of lipid shells beyond the first shell.

In this study, an alternative approach was developed to define high-resolution lipid shells, based on marking lipid locations by a 2D Voronoi tessellation, in which the gA Cα atoms were used to represent the gA structure and the lipid locations were defined by the center of mass of each lipid projected onto the Z = 0 plane (i.e., the membrane plane). 2D Voronoi tessellations are particularly useful in this context as they divide the bilayer into regions associated with individual lipids on the bilayer plane. These regions are divided by tessellation borders, which represent the spatial boundaries between individual lipid areas. After defining the 2D areas for each lipid, it is possible to determine the neighbors (adjacent areas belonging to either a lipid or the channel) of a specific lipid site. Because tessellation borders between lipids rapidly fluctuate (and may be quite small) due to the nature of stochastic lipid diffusion and conformational change, a threshold (empirically defined as 4 Å) can be used to determine where lipids that share a border with another lipid or the channel are true neighbors. Thus, a target site and a neighbor are only considered to be true neighbors if the tessellation border between them is above the threshold. In **Figure 2.2A**, lipid site "1" is in the first shell since it shares at least a 4 Å tessellation border with the channel (in orange). Lipid site "2" is in the second shell because it shares a border with a lipid in the first shell (i.e., lipid site "1") and does not share a border with the channel. A lipid in the third shell would therefore neighbor lipids in shell 2, etc.

To determine instantaneous shell identities for all lipids throughout the trajectory, a 2D Voronoi tessellation was obtained for the first snapshot of the MD trajectory. Using this tessellation, lipids were examined iteratively to determine the first shell lipids. Then, other lipids in contact with these first shell lipids were assigned to the second shell. In other words, the lipids were continuously assigned, first defining and filling the first shell, then defining and filling the second shell, etc. The assignment stopped when all shell identities were unchanged from the

20

previous iteration. For each following snapshot of the trajectory, nearly the same procedure was applied for assigning neighbor relationships.
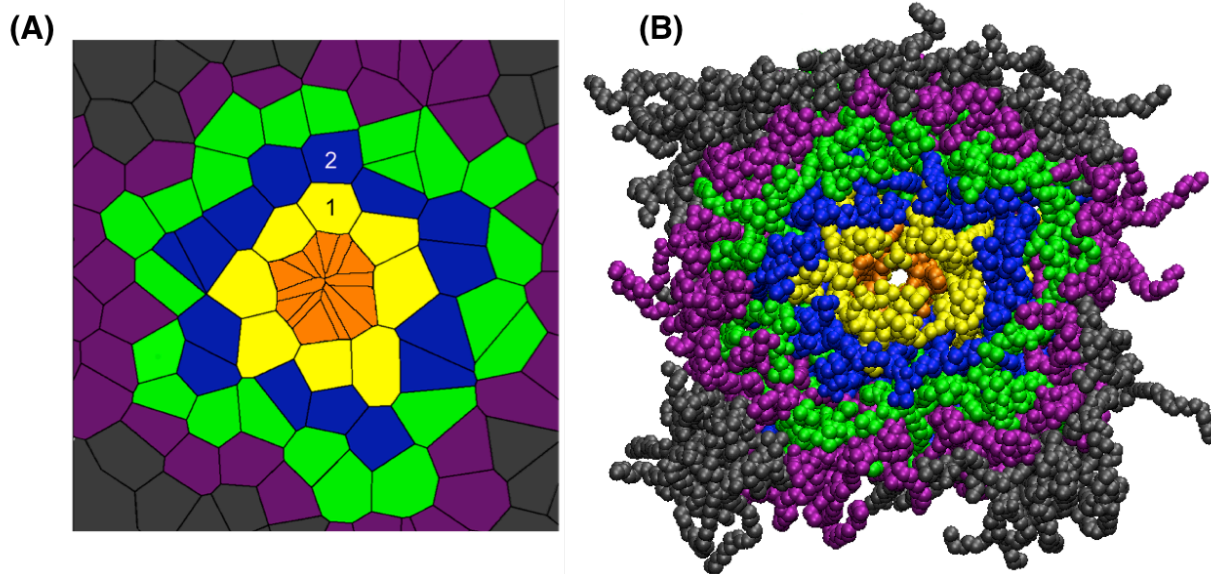
Figure 2.2: (A) Lipids shown with associated Voronoi area. The first four lipid shells surrounding the gA channel (orange) are defined by yellow, blue, green, and purple, respectively, with all lipids past these shells in grey. (B) Individual lipids around gA based on the same shell definition in (A).

The fast border fluctuations mean that a lipid's shell assignment could change nearly instantaneously (every snapshot or 0.24 ns). For example, a border could fluctuate between 3.8 Å and 4.2 Å throughout the simulation (i.e., crossing the threshold that defines a true lipid neighbor). To reduce these rapid, unphysical transitions, an additional step was added to the algorithm for all snapshots after the first. Once a border between two lipids reached the 4 Å threshold, the algorithm backtracked through time to the point of initial contact between the target lipid and the neighbor to mark the border as "effective". Thus, two lipids were considered to be neighbors from the point of first contact until the border between them went to 0 Å (complete separation). By using this scheme of reaching a defined threshold, backtracking, and then relaxing the threshold, the strong dependence of lipid shell identity on the threshold can be removed. Ultimately, the lipid shells can be defined with high resolution throughout the

simulation trajectory. Once a lipid's shell is determined, any of its characteristics (e.g., hydrophobic thickness, which is the average position of the C22 and C32 atoms (i.e., the carbon atom next to the carbonyl carbon in each acyl chain) along the membrane normal (i.e., the z-axis)) can be calculated. Subsequently, any characteristic can be averaged for all lipids in a particular shell.

2.2.2.3 First shell lipid residence time determination

The average amount of time that a single lipid spends in the first shell before moving to the second is called the "expected *residence time*" in this study. Characterizing the residence time is important to properly judge equilibration of the system (i.e., the simulation must be run many times longer than the residence time), and to build a model of how quickly the lipid environment around a channel relaxes. Reliable characterization of the residence time from a simulation trajectory is a challenging task, however, due to the co-existence of different lipid motions in and out of the first shell. Despite the shell definition using 2D Voronoi tessellations, fast lipid motions near the boundary of the first and second shells can occur due to difficulties in first-shell assignments, and the relaxation of such fast and physically meaningless events is unrelated to lipid diffusion for the residence time. Likewise, it is possible that lipids under strong interactions (e.g., hydrogen bonding) with the channel have much longer residence times. In this study, we used statistical modeling of the observed residence times and its technical procedure is described at the beginning of the **Supplemental information**. Briefly, two timescales are observed, the longer of which, typically ~50–100 ns, corresponds to a residence time that is consistent with lipid diffusion (**Table S2.2**).

2.3   Results and discussion

2.3.1   Single-channel lifetime measurements support lipid redistribution

Transmembrane association of gA subunits to form the conducting channel is described by:

$$2M \underset{k_{dis}}{\overset{k_{assn}}{\rightleftarrows}} D \text{ and } [D] = \frac{k_{assn}}{k_{dis}} \cdot [M]^2 , \tag{2.1}$$

where k_{assn} and k_{dis} are the rate constants for association and dissociation, respectively. The dissociation rate constant can be modeled as an activated process along a reaction coordinate that separates the subunits normal to the bilayer. Assuming that a free monomer subunit imposes no deformation of the membrane, only a dimerized gA has a contribution to the deformation energy. Thus, k_{dis} is increased and the channel lifetimes decrease as the deformation energy increases. gA dimerization in a lipid environment X (with a hydrophobic thickness that is longer than the channel's hydrophobic length) causes a generic bilayer deformation with an associated energetic cost (7, 71). Because the intrinsic curvature of phosphatidylcholines is close to zero (123), we can approximate the deformation energy as:

$$E_{def,dimer} = H (h_X - l)^2, \tag{2.2}$$

where H is a phenomenological spring constant, h_X is the unperturbed hydrophobic bilayer thickness of lipid environment X, and l is the channel length. The magnitude of H is determined by the shape of the protein and the physical properties of the bilayer, such as the bulk area compressibility and bending moduli and the channel-bilayer boundary condition (7). At the transition state (where the channel subunits have moved apart, breaking some inter-subunit hydrogen bonds), the channel length is increased by δ:

$$E_{def,t.s.} = H (h_X - (l + \delta))^2, \tag{2.3}$$

and the bilayer contribution to the activation energy (dissociation barrier) becomes, $\Delta E_{A,X} = E_{def,t.s.} - E_{def,dimer}$. For such a process, with uncorrelated events occurring at an average rate, the distribution of channel lifetimes will be determined by a parameter τ, the average lifetime, as

$$N(t)/N(0) = \exp\{-t/\tau\},\tag{2.4}$$

where $N(t)$ is the number of events occurring after some time t. The difference in activation energy between lipid environments X and Y that differ in thickness (h_X and h_Y, respectively), but have the same H, thus becomes (124):

$$\Delta E_{A,X-Y} = \Delta E_{A,X} - \Delta E_{A,Y} = 2H\delta(h_Y - h_X),\tag{2.5}$$

and the logarithm of the ratio of τ in two different lipid environments becomes:

$$\ln(\tau_X) - \ln(\tau_Y) = \frac{2H\delta(h_Y - h_X)}{RT}.\tag{2.6}$$

That is, the logarithm of the lifetime varies as a linear function of bilayer thickness with a slope of $2H\delta(RT)^{-1}$.

In the experiments reported here, τ was measured for three single-component bilayers ($dC_{16:1}$, $dC_{18:1}$, and $dC_{20:1}$, which form stable, though decane-containing bilayers) to obtain the phenomenological product, $2H\delta(RT)^{-1}$, for this lipid chemistry (phosphoglycerol lipids, two acyl chains with one double bond each). **Table 2.1** and **Figure 2.3A** show results obtained in planar lipid bilayers formed by $dC_{16:1}$, $dC_{18:1}$, and $dC_{20:1}$ (there was an ~10-fold change in τ when the phospholipid acyl chains were shortened or extended by two methylene groups).

A key issue is what to use as an appropriate bilayer thickness. The single-channel experiments were done in hydrocarbon-containing membranes, which are thicker than the hydrocarbon-free membranes used in the simulations. The hydrophobic thickness of $dC_{18:1}$ membranes formed using decane is ~40 Å (125), whereas the thickness of hydrocarbon-free $dC_{18:1}$ bilayers is ~27 Å (Refs. (76, 126) and **Table S4**). Although the hydrocarbon-containing and hydrocarbon-free thicknesses are different, Helfrich and Jakobsson (5) noted that the membrane thinning associated with channel formation is likely to first involve the expulsion of the hydrocarbon solvent in the bilayer adjacent to the channel followed by the compression of

the bilayer leaflets. The key variable of interest, however, is the difference in thickness when the acyl chain length is varied, which is similar in the two systems, varying ~1.6 Å per CH_2 residue ($dC_{18:1}$ to $dC_{24:1}$) in solvent membranes (76), vs. ~1.1 Å per CH_2 residue ($dC_{16:1}$ to $dC_{24:1}$) in hydrocarbon-containing planar bilayers (127). For this analysis, we therefore chose to use the bilayer thickness from MD simulations (see below and **Table S2.4**), which agrees well with the thickness determined by hydrocarbon-free experiments (76, 126). If the thicknesses of hydrocarbon-containing bilayers had been used instead, the slope of the $\ln(\tau)$-h relation would have been steeper, and we would have deduced a larger enrichment of the shorter lipid in the vicinity of the channel.

We thus find that the slope of the $\ln(\tau)$-h relation in single-component bilayers, Equation 2.6, is 0.622 Å^{-1} ± 0.021 Å^{-1} (fit using uncertainties found in **Tables 2.1** and **S2.4**). It is not necessary to individually specify the values of H and δ because it is their product that is important in the calibration of timescale and bilayer thickness. Nevertheless, δ can be reasonably assumed to be on the Å length scale necessary to break the hydrogen bonds between gA subunits. A δ value of 1.6 Å is justified in the literature (125, 128–130), which yields $H = 0.117 \pm 0.004$ kcal/(mol·Å^2) (49.0 ± 1.6 kJ/(mol·nm^2)) from the data in this work, in reasonable agreement with estimates obtained using different gA analogues (125). This value is in acceptable agreement with the estimate computed independently in the **Supplemental information** of **Chapter 3**, i.e., $H = 0.0853$ kcal/(mol·Å^2) (35.7 kJ/(mol·nm^2)).

Using the H estimate from gA dimer lifetime measurements in single-component bilayers, we can define an effective bilayer thickness around the channel (i.e., the thickness that the channel experiences) to interpret the lifetime measurements in two-component bilayers:

$$h_0^{\text{eff}} = h_{\text{dC}_{20:1}} - \frac{RT}{2\delta H} \cdot \ln\left(\frac{\tau_{\text{mix}}}{\tau_{\text{dC}_{20:1}}}\right) = h_{\text{l}} + n_{\text{s}}^{\text{eff}}(h_{\text{s}} - h_{\text{l}}), \tag{2.7}$$

which can be written in terms of the change in effective thickness, as:

$$\Delta h_0^{\text{eff}} = h_0^{\text{eff}} - \frac{1}{2}(h_{\text{s}} + h_{\text{l}}), \tag{2.8}$$

where $h_{\text{dC}_{20:1}}$ is the unperturbed thickness of the reference $\text{dC}_{20:1}$ bilayer, τ_{mix} is the gA single-channel lifetime in a mixed bilayer, $\tau_{\text{dC}_{20:1}}$ is the gA lifetime in $\text{dC}_{20:1}$ lipid (channel lifetime results are listed in **Figure 2.3B** and **Table 2.1**), and $n_{\text{s}}^{\text{eff}}$ is the fraction of short lipids in direct contact with gA (first lipid shell). Here, the subscript s refers to the "short" and the subscript l refers to "long" lipids in the mixture. Assuming that the thickness of a lipid mixture is the weighted average of the thicknesses of its constituents, i.e., $h_0^{\text{eff}} = n_{\text{s}}^{\text{eff}}h_{\text{s}} + n_{\text{l}}^{\text{eff}}h_{\text{l}}$ (where $n_{\text{l}}^{\text{eff}} = 1 - n_{\text{s}}^{\text{eff}}$), the effective fraction of short-chained lipids in the vicinity of the channel, $n_{\text{s}}^{\text{eff}}$, is then determined by:

$$n_{\text{s}}^{\text{eff}} = \frac{1}{2} + \frac{\Delta h_0^{\text{eff}}}{(h_{\text{s}} - h_{\text{l}})}. \tag{2.9}$$

If Δh_0^{eff} is zero, the lipids are randomly mixed in the vicinity of the channel ($n_{\text{s}}^{\text{eff}} = \frac{1}{2}$). The work of collaborators in **Chapter 3** shows that the major contribution to the deformation energy comes from the lipids in the first shell, justifying the interpretation of $n_{\text{s}}^{\text{eff}}$ as the first shell lipid enrichment. Estimates of $n_{\text{s}}^{\text{eff}}$ are shown in **Table 2.1**. The measured lifetimes thus correspond to bilayers composed of a mole-fraction 0.58 ± 0.04 $\text{dC}_{18:1}$ in the $\text{dC}_{18:1}+\text{dC}_{22:1}$ bilayer and 0.66 ± 0.02 $\text{dC}_{16:1}$ in the $\text{dC}_{16:1}+\text{dC}_{24:1}$ bilayer. The single-channel experiments thus provide strong evidence for lipid redistribution in the vicinity of the channel for $\text{dC}_{16:1}+\text{dC}_{24:1}$.

Table 2.1: Experimental gA dissociation timescales as well as Δh_0^{eff} and n_s^{eff} for the mixtures.

Bilayer	$\tau^{\#}$ (ms)	$\Delta h_0^{\text{eff}\,\$}$ (Å)	$n_s^{\text{eff}\,\%}$
$dC_{16:1}$	6820 ± 510		
$dC_{18:1}$	530 ± 22		
$dC_{20:1}$	64 ± 10		
$dC_{18:1}+dC_{22:1}$	95 ± 10	-0.72 ± 0.3	0.58 ± 0.04
$dC_{16:1}+dC_{24:1}$	195 ± 20	-2.9 ± 0.3	0.66 ± 0.02

$^{\#}$ Data for $dC_{20:1}$ and the mixtures are shown in **Figure 3**.
$^{\$}$ Computed from **Equation 7**.
$^{\%}$ Computed from **Equation 9**.
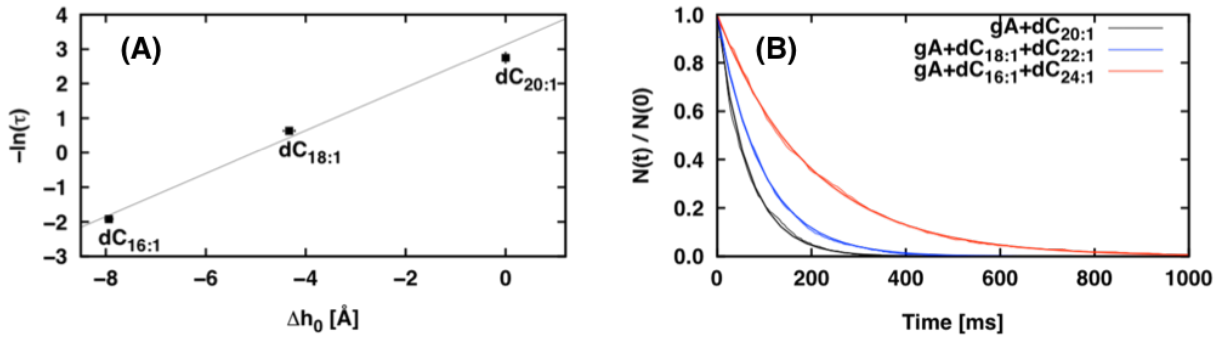
Figure 2.3: (A) The natural log of the average gA single-channel lifetimes as a function of the difference in bilayer thickness for $dC_{16:1}$, $dC_{18:1}$, and $dC_{20:1}$, relative to $dC_{20:1}$. The points show the averages with associated uncertainty. The black line is a linear fit to the data, with a slope of 0.622 Å^{-1} $\pm$ 0.021 Å^{-1}. (B) Single-channel lifetime distributions in $dC_{20:1}$, $dC_{18:1}+dC_{22:1}$, and $dC_{16:1}+dC_{24:1}$ membranes. The distributions (thinner line) were fit with single exponential distributions (thicker line) and the average lifetimes are listed in Table 1.

2.3.2 Microsecond simulations show ~100-ns first shell lipid residence times

In the case of semi-cylindrical integral membrane proteins such as gA, it is convenient to define lipid shells and other bilayer properties as a function of radial distance from the protein. In this context, the lipid shell locations from gA, defined using the 2D Voronoi tessellation described in the **Section 2.2.2.2**, correspond to the center of each shell. By averaging shell locations, a typical shell boundary can be approximated (**Table S2.3**), e.g., the first lipid shell boundary is the average of the first and second shell locations. The shell locations (black vertical lines, in **Figure 2.4**) coincide well with the first peak in the 1D lipid radial distribution function (RDF) in gA+$dC_{16:1}$+$dC_{24:1}$; the results for gA+$dC_{18:1}$+$dC_{22:1}$ and gA+$dC_{20:1}$ are similar and not shown. Slight discrepancies arise because the lipid shell location represents an average location within each shell, which is not necessarily the most populated location in the 1D-RDF peaks. When the shell boundaries (dashed red vertical lines) are considered, this shell method clearly provides much finer resolution of each lipid shell compared to the 1D-RDF. In fact, it is difficult to define even the second lipid shell based solely on the 1D-RDF. The Voronoi tessellation, however, provides an unambiguous lipid shell definition that allows for a high-quality definition of the first shell residence time and other shell-dependent bilayer properties.
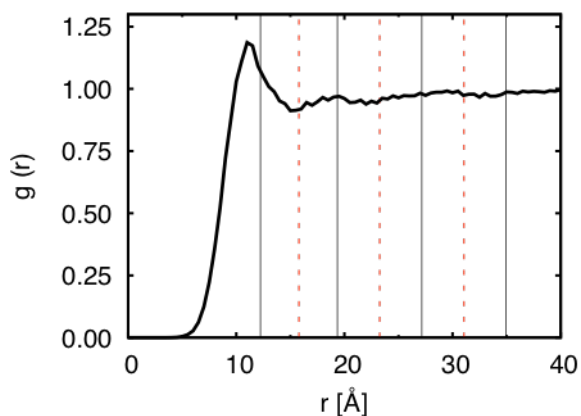
Figure 2.4: 1D lipid radial distribution function in gA+$dC_{16:1}$+$dC_{24:1}$. The vertical black lines represent the calculated locations of the radial lipid shells by the 2D Voronoi tessellation, and the red dashed lines denote each shell boundary.

Using the lipid shells, as well as a fitting procedure that properly weights the transition time series in the longer time regime (see **Section 2.2.2.3** and **Figure S2.2**), the first shell residence times were calculated; the results are summarized in **Table S2.2**. Not unexpected, the fit for the residence time requires a double exponential distribution. Rabinowitch in the 1930s showed that in a condensed phase, a particle must wait some amount of time (proportional to the density) to find a new (unoccupied) lattice hole (131). During the wait, the particle rattles in an effective cage created by its neighbors. That is, with respect to lipid hops, a lipid molecule has a short timescale motion that is associated with being confined (this short timescale is not related to the timescale associated with the Voronoi glitches, see **Section 2.2.2.2**), and a long timescale motion that is associated with the full hop into a hole in the lattice.

In $gA+dC_{16:1}+dC_{24:1}$ and $gA+dC_{20:1}$, the mean first shell residence times are ~70–75 ns; $gA+dC_{18:1}+dC_{22:1}$ has longer residence times (~90–110 ns). A simple estimate using the Einstein-Smoluchowski equation suggests that the time required for a lipid to diffuse 7 Å (the approximate distance between lipid shells) in one dimension with a diffusion coefficient of ~5×10^{-8} cm^2/s is ~50 ns, in agreement with the above values. The longer mean residence time in $gA+dC_{18:1}+dC_{22:1}$ is due mainly to a few very long lipid residence times; six lipid residence times are longer than 500 ns, the longest being ~760 ns. In comparison, there is only one lipid residence time longer than 500 ns in $gA+dC_{16:1}+dC_{24:1}$. These very long lipid residence times (and their fitting) cause the difference in mean values between $gA+dC_{18:1}+dC_{22:1}$ and the other two systems. Our analysis demonstrates that: first, measuring lipid redistribution near a membrane protein using all-atom MD simulations remains a difficult task due to the long timescales involved; and second, a simulation time of 3.5 μs in each system is sufficient to allow for ~3 μs sampling at equilibrium (assuming that the system is in equilibrium after ~5 residence

times). Additionally, the relaxation time for lipid redistribution around the gA is ~500 ns, which is less than 1/100,000 of the lifetime of gA in $dC_{20:1}$ (64 ± 10 ms). So, relative to gA dimer average lifetimes, lipid redistribution is nearly instantaneous, and the measured experimental lifetimes can truly be considered a result of the channels residing in an enriched environment.

Whereas the observed residence times are much shorter than the total duration of the simulation, the decorrelation time of the first shell enrichment of a lipid requires multiple lipid hops. In the **Supplemental information** of **Chapter 2**, a model created by coworkers is briefly described. They used simulation-extracted timescale (75 ns) as well as the membrane energetic model to probe the expected enrichment and standard deviation in the simulation results. This model predicts first shell enrichments in good agreement with both experiment and MD (model predictions in **Chapter 3**: 0.59 ± 0.03 $dC_{18:1}$ in $dC_{18:1}+dC_{22:1}$ and 0.65 ± 0.03 $dC_{16:1}$ in $dC_{16:1}+dC_{24:1}$).

2.3.3   Simulated lipid distributions agree qualitatively with experiment

Assuming a 500-ns equilibration (deduced from the residence time analysis), the average first shell $dC_{18:1}$ mole-fraction in gA+$dC_{18:1}$+$dC_{22:1}$ is 0.50 ± 0.04 the average first shell $dC_{16:1}$ mole-fraction in gA+$dC_{16:1}$+$dC_{24:1}$ is 0.66 ± 0.02. Standard errors were calculated by dividing the last 3 µs into 6 consecutive 500-ns blocks. **Table 2.2** summarizes the mole-fractions of $dC_{18:1}$ and $dC_{16:1}$ in the first three lipid shells. The mole-fraction of short lipid per shell in the two leaflets is similar, indicating satisfactory convergence. The 95% confidence interval (CI) shows that redistribution is statistically significant for $dC_{16:1}+dC_{24:1}$, but not for $dC_{18:1}+dC_{22:1}$; i.e., the fluctuations in the latter obscure the results. Hence, the simulation results for both bilayers agree with experiment (**Table 2.1**), but the CIs suggest that 3 µs is still a short timescale for this type of biophysical problem.

Table 2.2: Summary of $dC_{18:1}$ and $dC_{16:1}$ shell mole-fractions in their respective bilayers. Means and standard errors are reported for the individual leaflets, and the average of both leaflets. The 95% CI for the shell averages are the means ± the values in parentheses.

	gA+dC$_{18:1}$+dC$_{22:1}$				gA+dC$_{16:1}$+dC$_{24:1}$		
Shell	Top	Bottom	Average		Top	Bottom	Average
1	0.52 ± 0.07	0.48 ± 0.02	0.50 ± 0.04 (0.10)		0.64 ± 0.04	0.69 ± 0.03	0.66 ± 0.02 (0.06)
2	0.53 ± 0.03	0.48 ± 0.02	0.51 ± 0.02 (0.05)		0.53 ± 0.02	0.57 ± 0.01	0.55 ± 0.01 (0.04)
3	0.51 ± 0.01	0.52 ± 0.01	0.52 ± 0.01 (0.02)		0.49 ± 0.01	0.49 ± 0.01	0.49 ± 0.01 (0.02)
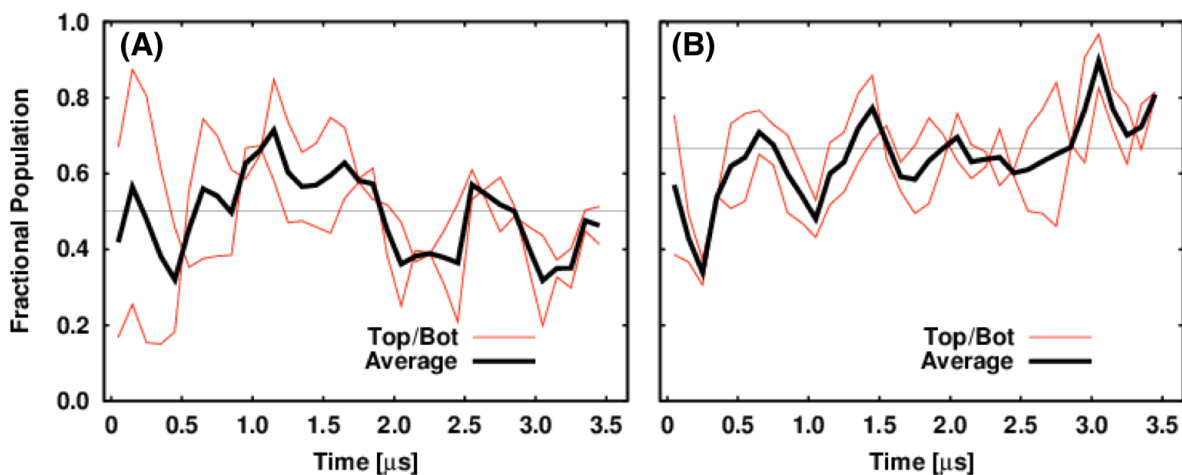
Figure 2.5: First shell mole-fractions of (A) $dC_{18:1}$ in gA+dC$_{18:1}$+dC$_{22:1}$ and (B) $dC_{16:1}$ in gA+dC$_{16:1}$+dC$_{24:1}$. In both plots, the red curves are the top and bottom leaflet concentration, the thick black curve is the average of the leaflets at each time, and the thin black line is the overall average concentration after leaving off the first 500 ns. The plotted mole-fractions denote 100-ns block averages of the trajectory.

The first shell mole-fractions of $dC_{18:1}$ and $dC_{16:1}$ are plotted as functions of time in **Figure 2.5**. Both systems are dynamic, and the evolution of the mole-fractions show lateral lipid redistribution in the first and second shells in $gA+dC_{16:1}+dC_{24:1}$, with the second shell's enrichment being lesser in magnitude (**Table 2**). Although the simulations should have been long enough to ensure equilibration, the first shell concentrations were still changing rapidly up to the end of the 3.5 μs. Moreover, because there were on average only seven lipids in the first shell per leaflet (**Table S2.3**), it is difficult to assign an equilibrium concentration because the fluctuation of a single lipid can drastically alter the instantaneous value.

The disjoining force that a bilayer exerts on the channel is mostly due to the strain on the first shell lipids (Nielsen et al. (7) and **Chapter 2**), which allows for comparison of the first shell lipid enrichment in the experiments (**Table 2.1**) and simulations (**Table 2.2**). These results are in qualitative agreement. There is less first-shell enrichment in the $gA+dC_{18:1}+dC_{22:1}$ simulation than in the experiment, but, given the sampling difficulties, we attach the greater significance to the $gA+dC_{16:1}+dC_{24:1}$.

2.3.4   Lipids redistribute to better match the channel and each other

Individual lipids adjust to the channel as well as to the other lipids in the bilayer. Even in the case of a bilayer with large disparity in the constituent chain lengths (e.g., ~9 Å difference in preferred monolayer height in $dC_{16:1}+dC_{24:1}$, **Table S2.4**), the lipids are observed to adopt configurations that lead to a nearly smooth surface (**Figure 2.6**) and the thickness of the mixed bilayer is near the average of the two constituents' thicknesses (**Table 2.3**).
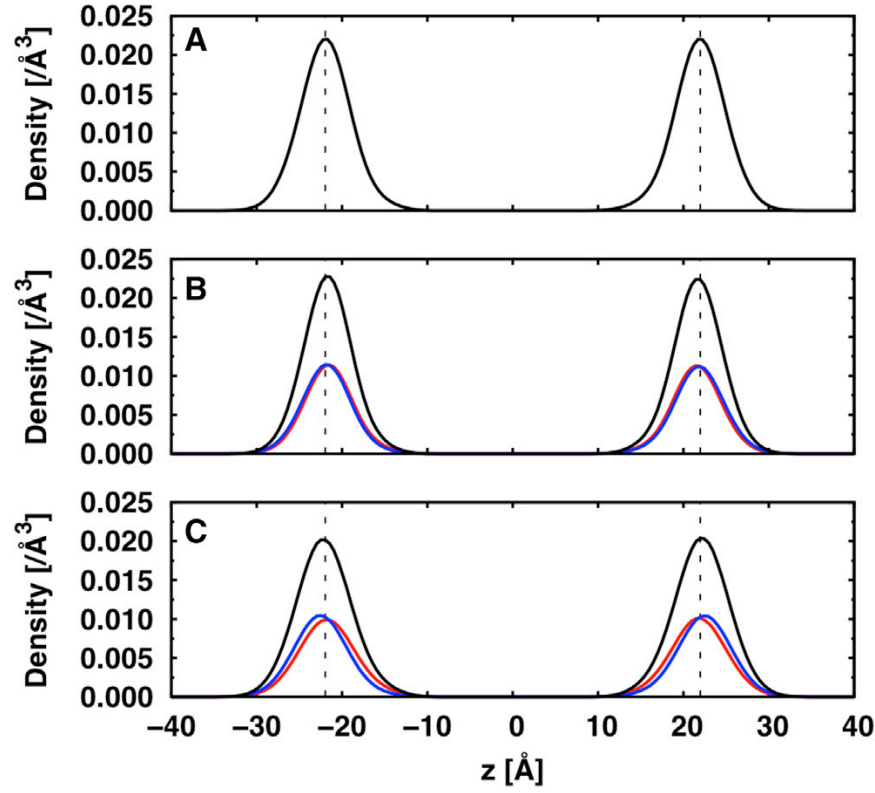
Figure 2.6: Number density plots of lipid head groups, defined as the phosphocholine groups: (A) $dC_{20:1}$; (B) $dC_{18:1}$ (red), $dC_{22:1}$ (blue), and $dC_{18:1}+dC_{22:1}$ (black); and (C) $dC_{16:1}$ (red), $dC_{24:1}$ (blue), and $dC_{16:1}+dC_{24:1}$ (black). The decomposition of head group locations in (B) and (C) show that head groups match very closely in an equimolar mixed lipid bilayer.

Table 2.3: Lipid hydrophobic thickness by shell (in Å) with associated standard errors.[$]

	gA+$dC_{18:1}$+$dC_{22:1}$			gA+$dC_{16:1}$+$dC_{24:1}$			gA+$dC_{20:1}$
Shell	$dC_{18:1}$	$dC_{22:1}$	$dC_{18:1}+dC_{22:1}$	$dC_{16:1}$	$dC_{24:1}$	$dC_{16:1}+dC_{24:1}$	$dC_{20:1}$
1	25.94 ± 0.12	26.24 ± 0.13	26.05 ± 0.07	25.28 ± 0.11	25.86 ± 0.15	25.46 ± 0.08	26.11 ± 0.04
2	29.88 ± 0.09	30.25 ± 0.06	30.05 ± 0.06	29.44 ± 0.10	30.22 ± 0.09	29.79 ± 0.08	30.30 ± 0.02
3	31.81 ± 0.05	32.04 ± 0.05	31.91 ± 0.04	31.55 ± 0.05	32.48 ± 0.06	32.01 ± 0.06	32.04 ± 0.02
4	32.41 ± 0.03	32.66 ± 0.03	32.54 ± 0.02	33.47 ± 0.08	33.50 ± 0.04	33.10 ± 0.03	32.57 ± 0.02

[$]Standard errors were calculated by dividing the trajectory into 350-ns blocks.

The adjustment of the acyl chains is most visible in gA+dC$_{16:1}$+dC$_{24:1}$, where the dC$_{16:1}$ acyl chains do not strongly interdigitate across the bilayer center ($z = 0$), whereas the dC$_{24:1}$ acyl chains are highly interdigitated in all lipid shells (**Figure S2.3)**. The same acyl chain changes also were observed, though to a lesser amount, for gA+dC$_{18:1}$+dC$_{22:1}$ (data not shown). The lipid head matching thus occurs at the cost of deforming the chains.

The overall changes in lipid order due to lipid-lipid and lipid-channel hydrophobic matching can be further illustrated using acyl chain lipid order parameters (S$_{CD}$) calculated for the neat bilayers, as well as for each lipid type as a function of shell in gA-containing systems:

$$S_{CD} = \left| \frac{1}{2} \left\langle 3 \cos^2 \theta_{CH} - 1 \right\rangle \right|, \tag{2.10}$$

where θ_{CH} is the angle between the CH bond vector and the bilayer normal. Changes in these lipid order parameters represent how much an acyl chain has deformed relative to its bulk state. For dC$_{16:1}$ in gA+dC$_{16:1}$+dC$_{24:1}$, as shown in **Figure S2.4**, chain order increases as the radial distance from the protein increases (i.e., shell number increases), indicating that dC$_{16:1}$ in the bulk becomes more ordered. The dC$_{16:1}$ lipid order in the first shell, in contrast, is similar to that of the single-component dC$_{16:1}$ system, indicating a good match to the channel. As expected, dC$_{24:1}$ shows a drastically different order parameter profile when comparing the first shell and the dC$_{24:1}$-only bilayer. To match the length of the channel, the dC$_{24:1}$ acyl chains must adjust their configurations (bend or compress), which decreases the chain order. Consequently, the dC$_{24:1}$ acyl chain order increases as the shell number increases, but it does not become as ordered as in the single-component dC$_{24:1}$ system because it is still neighbored by some dC$_{16:1}$. These profiles suggest that: (i) dC$_{16:1}$ is a good match to the channel; (ii) dC$_{16:1}$ becomes more ordered to accommodate head group matching in the bulk of gA+dC$_{16:1}$+dC$_{24:1}$ and in dC$_{16:1}$+dC$_{24:1}$; (iii) dC$_{24:1}$ deforms near the channel; and (iv), dC$_{24:1}$ has more order in the bulk (and binary lipid-only

systems), but must compress to match the shorter lipid. Although $dC_{24:1}$ is near its melting temperature (26.7 °C; Lewis et al. (132)), the order parameters indicate that the acyl chains are in the fluid phase in each shell, including the effective bulk. Therefore, the simulations show that lateral lipid redistribution occurs in completely fluid two-component lipid systems.

The $dC_{18:1}$ and $dC_{22:1}$ order parameters also explain why there is less lateral lipid redistribution in gA+$dC_{18:1}$+$dC_{22:1}$. **Figure S2.5** shows that both $dC_{18:1}$ and $dC_{22:1}$ adjust their conformations to be near the channel, and the order parameters of the two lipid types in the fourth lipid shell are not drastically different from those of their respective pure bilayers. That is, there is no strong penalty for having them mixed in the bilayer (i.e., they need not stretch or compress dramatically to coexist).

2.3.5   A simple continuum model can explain lipid redistribution

Lipid redistribution around a membrane protein will be affected by general and specific mechanisms. General mechanisms refer to redistribution that depends on the material properties of the membrane, e.g., thickness, spontaneous curvature, compressibility, and bending modulus, whereas specific mechanisms refer to occupancies that cannot be simply correlated to material properties (and are likely to involve residence times typical of a strongly bound lipid, e.g., much longer than 100 ns). Therefore, this study pertains to the general mechanism, and the discussion below refers only to this case.

The thicknesses of the $dC_{18:1}$+$dC_{22:1}$, $dC_{16:1}$+$dC_{24:1}$, and $dC_{20:1}$ systems are similar, and importantly, all three have nearly the same thickness profiles when perturbed by gA (**Figure 2.7**). All show local thinning near the channel, which has been demonstrated for model and biological systems (see Kim et al. (20) and Mitra et al. (133)). The small differences between the profiles are explained by the redistribution discussed above; for $dC_{16:1}$+$dC_{24:1}$, the thickness is

less in the first shell that is enriched in the short lipid, which results in an enrichment of the long lipid in the outer shells.

The total bilayer deformation energy includes compression and curvature penalties. Assessing the simulated curvature in the first shell is difficult because curvature is highly sensitive to the definition of the complex hydrophobic surface (**Figure 2.7**). Also, the calculation is dependent on first (tangential curvature) and second (radial curvature) derivatives with respect to the hydrophobic surface profile (see **Chapter 3**). Lipid compression, in contrast, can be simply estimated from the radial (per shell) thicknesses from simulation. Additionally, it is shown in **Chapter 3** that CEMs produced by coworkers predict the compression energy is the major component governing redistribution (only ~25% of the deformation energy is from curvature, similar to the ~35% for model bilayers examined by Nielsen et al. (7)). Therefore, the analysis in this article uses a simple compression-only model (neglecting curvature contributions that are difficult to obtain by simulation and are shown to be smaller than the compression contributions) that is based on the simulation results. Nonetheless, the compression-only model provides insight into the redistribution and also highlights the importance of curvature in reducing the predicted lipid redistribution.
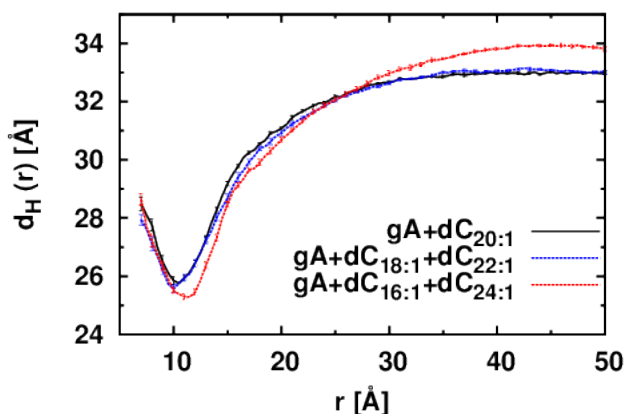
Figure 2.7: Radial hydrophobic thickness profiles around gA in gA+dC$_{20:1}$, gA+dC$_{18:1}$+dC$_{22:1}$, and gA+dC$_{16:1}$+dC$_{24:1}$. Hydrophobic thickness is the average z-distance between the C22+C32 atoms (i.e., the first carbon atoms after the carbonyl of each fatty acid tail) in each leaflet.

Compression penalties are incurred when the lipid length is perturbed from its equilibrium length (h_0) along the membrane normal. Assuming that a lipid can be compressed like a spring with constant density (i.e., a change in height is directly correlated to a change in lateral area, $V_{lipid} = Ah = A_0 h_0$), the amount of energy required to stretch/compress a lipid is given by:

$$E_{area} = \frac{1}{2} K_{A,m} A_0 \left(\frac{A}{A_0} - 1\right)^2 = \frac{1}{2} K_{A,m} A_0 \left(\frac{h_0}{h} - 1\right)^2$$
$$= \frac{1}{2} K_{A,m} A_0 \left(\frac{h}{h_0} - 1\right)^2 + \mathcal{O}[(h - h_0)^3], \tag{2.11}$$

where $K_{A,m}$ is the monolayer lipid area compressibility ($K_{A,m}$ is assumed to be one half of the bilayer K_A and has units of energy per area), A_0 is the unperturbed area per lipid, and A and h are the perturbed area and height, respectively. **Equation 2.11** is a simple quadratic expansion of the energy about the compression energy minimum (i.e., when the lipid is unperturbed).

Using simulation values, the compression energy is estimated by subtracting the energy necessary to compress an unperturbed long lipid (h_{pure}, the lipid thickness in a single-component bilayer) to the thickness of the bulk mixture (h_{mixed}) from the energy necessary to compress an unperturbed lipid to the thickness of the channel (l):

$$\Delta E = \frac{1}{2} K_{A,m} A_0 \left[\left(\frac{l}{h_{pure}} - 1\right)^2 - \left(\frac{h_{mixed}}{h_{pure}} - 1\right)^2\right]. \tag{2.12}$$

The resulting ΔE is the energy difference between a lipid residing in the mixed bulk and a lipid residing in the first lipid shell near to the channel. Values for $K_{A,m}$ and h_0 (**Table S2.4**) are taken to be the two-component values (naturally with the exception of dC$_{20:1}$). The situation is clear for long lipids – they must compress to be in the mixed bilayer, and also must compress to be near the channel. Short lipids must *stretch* to be in the mixed bulk membrane and little energy will be required for them to be near the channel. For simplicity, we assume their energy to be in the mixed bulk also to be zero because the conformational space accessible to the lipid is

assumedly unhindered by having excess space in the opposing leaflet. This is justified by considering the terminal methyl density in **Figure S2.3A**, which is *shifted* in Z, rather than stretched to fit the full leaflet profile. Order parameters in **Figure S2.4A** agree qualitatively, where the order is unchanged for the first 10 carbons as the lipid is placed in the outer shells. These compression-only ΔE values can be compared to those calculated by the 3D continuum elastic model (3D-CEM) created by coworkers in **Chapter 3** (**Table 2.4**).

The canonical hydrophobic length of gA is 22 Å, as deduced from both experiment and theory (66, 99, 134), but some lipids are still able to protrude above the channel, i.e., matching is non-ideal, and so the definition of the hydrophobic length is ambiguous even given the molecular detail of a simulation. In **Figure 2.7**, the minima of the hydrophobic thickness profiles only reach ~26 Å due to lipid protrusions (e.g., Kim et al. (20)). Therefore, for the analysis in **Table 2.4**, we decided to use channel lengths of both 22 Å and 26 Å for illustrative purposes.

Table 2.4: First shell compression energy for each lipid type (kcal/mol), dependent on the assumed length of the channel (either 22 or 26 Å). 3D-CEM values are from coworkers and described in **Chapter 3** and assume l = 26 Å. Dashes indicate situations that are presumed to have negligible compression penalty.

	$h = l$ = 22 Å (kcal/mol)	$h = l$ = 26 Å (kcal/mol)	3D-CEM (kcal/mol)
dC$_{16:1}$	–	–	–
dC$_{18:1}$	0.20	–	–
dC$_{20:1}$	0.69	0.24	0.07
dC$_{22:1}$	1.1	0.48	0.27
dC$_{24:1}$	1.4	0.69	0.43

This simple compression-only calculation serves as a starting point for understanding redistribution around the channel. As described in in **Chapter 3**, these first shell compression penalties are reduced at the expense of curvature penalties in a way that is sensitive to the protein-lipid boundary. The high estimate of the compression energy from the simple compression-only model is typical of the majority of previous estimates (e.g., Huang (66)),

which have concluded that the lipid meets the channel with nearly zero slope, a conclusion that is inconsistent with simulation thickness profiles (**Figure 2.7**; see also Kim et al. (20) and Lee et al. (71)). The 3D-CEM from our coworkers in **Chapter 3** naturally incorporates curvature (i.e., nonzero slope boundary condition), which reduces the compression energies, which then reduces the expected redistribution compared to the compression-only model.

From the density distributions of the terminal carbons (**Figure S2.3**) and order parameters (**Figure S2.4**) we find that $dC_{18:1}$ and $dC_{16:1}$ configurations are not as strongly perturbed as their longer counterparts in the first shell, meaning that the short lipids do not contribute much to the compression energy when they are near the channel. We also find that lipids can extend over the top of the channel and interdigitate into the opposite leaflet. These forms of lipid compression relaxation (protrusion, interdigitation, and curvature, which are discussed above) allow for more long lipids to reside in the first lipid shell than would be expected from simple compression-only modeling.

2.4   Conclusions

It has been hypothesized that integral membrane proteins can introduce local order and nanoscale heterogeneity in the distribution of lipids adjacent to the protein. Theoretical models have predicted that lipids with better hydrophobic match to a protein would be preferentially enriched around the protein, but there has been little experimental evidence supporting this conjecture.

The present work intertwines experiment, long timescale all-atom MD, and simple compression theory to describe lipid redistribution driven by hydrophobic mismatch. Single-channel lifetime experiments provided insight into the energetics within a bilayer: the larger the disjoining force, the lower the lifetime. By using two-component bilayers, where one lipid

species provided better hydrophobic match to the channel, the mean channel lifetime was increased relative to a control bilayer, which led to the calculation of an effective bilayer concentration based on the estimated composition of the bilayer in the first shell surrounding the channel. We thus predicted the mole-fractions of $dC_{18:1}$ and $dC_{16:1}$ to be 0.58 ± 0.04 and 0.66 ± 0.02 in the channel's vicinity in the $gA+dC_{18:1}+dC_{22:1}$ and $gA+dC_{16:1}+dC_{24:1}$, respectively.

All-atom MD was then used to explore the timescales and details of lipid redistribution. A lipid shell definition was built using a 2D Voronoi tessellation that allowed for a well-defined and high-resolution method for tracking and calculating bilayer properties around the channel. Using the shell definition, the timescale for lipid redistribution in the first shell could be calculated, and was observed to be ~100 ns, demonstrating that long timescale simulations are required for a multi-component bilayer to reach an equilibrium state around an embedded protein. After 500-ns simulation, which allowed for first shell equilibration, concentrations were 0.50 ± 0.04 $dC_{18:1}$ and 0.66 ± 0.02 $dC_{16:1}$ in the $gA+dC_{18:1}+dC_{22:1}$ and $gA+dC_{16:1}+dC_{24:1}$, respectively. These values are in agreement with the concentrations inferred from experiment (0.58 ± 0.04 and 0.66 ± 0.02, respectively) and theory in **Chapter 3** (0.59 ± 0.03 and 0.65 ± 0.03, respectively). The large uncertainty in the $dC_{18:1}+dC_{22:1}$ bilayer indicates that long timescale simulations are necessary to compute equilibrium distributions of lipids around an embedded channel. Nevertheless, the modeling in **Chapter 3** is consistent with the all-atom model, and thus predicts that given longer simulations, enrichment of $dC_{22:1}$ would emerge. Testing this will require longer simulations to reduce stochastic uncertainty. Finally, lipid configurations as a function of shell were determined with atomic resolution. It was found that well-matched short lipids exist in near-native conformations, while long lipids must strongly compress to match the hydrophobic length of the channel.

Supported by the configuration analysis, a simple compression-only model was used to demonstrate that the lipid redistribution in this study could be mainly attributed to reducing the bilayer deformation energy (decreasing the disjoining force that the bilayer imposes on the channel) by general interactions between lipids and the channel. It is now well established that stresses within the bilayer influence conformation and dynamics of embedded proteins. One of the contributors to bilayer stress is compression, which can be alleviated by a protein conformational change, protein diffusion into a better-matched lipid patch, or lipid redistribution, so that better-matched lipids are near the protein. This study focused on the latter option, where experiment and simulation both predicted lateral lipid redistribution. The bridges in this work between experiment, all-atom MD, and theory allow for cross-validation of all methods and deeper understanding from the atomistic to the continuum levels.

2.5   Supplemental information

2.5.1   First shell lipid residence time determination

In principle, the *average* residence time could be calculated as the mean of residence times of individual lipids, computed in a simulation by simply counting how long an individual lipid remains in the first shell. Unfortunately, however, this cannot take into account different (fast and slow) mechanisms of relaxation. To separate different mechanisms, a single residence time is modeled as if it is sampled from some probability density function (PDF) as a function of time. For a memory-less process (i.e., a process for which the probability of an event occurring does not depend on the history of the system), the PDF is an exponentially decaying function by definition. The memory-less assumption is valid for a "jump" mechanism of diffusion, in which diffusion is controlled by the rare appearance of transient "vacancies" near the channel. The additional fast transition events can be fit using an additional exponential (with faster decay)

added to the PDF. Summation of the individual exponentials is appropriate if the decays are uncoupled, as expected if such fast events only apply to lipids near the boundary of the first and second shells. The form of the overall PDF, $p_0(t)$, is thus:

$$p_0(t) \approx A_{\text{fast}} \, e^{\frac{-t}{\tau_{\text{fast}}}} + A_{\text{slow}} \, e^{\frac{-t}{\tau_{\text{slow}}}} \tag{1}$$

where the fast and slow components of the lipid residence times in the first shell have different amplitudes (A) and timescales (τ). The units of the amplitudes are probability per unit time; integrating the normalized PDF over a histogram interval i (here, 4.8 ns or 20 snapshots of the trajectory) yields the total probability of an event happening during that period, p_i. Within a histogram interval i, the number of shell transitions observed is n_i. Given a PDF with probabilities p_i for each bin i, the probability of observing n_i for a given bin of the histogram can be computed from the binomial distribution:

$$\Pr(n_i | p_i) = \binom{N}{n_i} p_i^{n_i} (1 - p_i)^{N - n_i} \tag{2}$$

where N is the total number of Bernoulli trials (here, the total number of observed exits from the first shell), with an event either being in the target interval or not. Given the PDF, the total probability of observing the data is then:

$$\Pr(\{n_i\} | p(t)) = \prod_i \Pr(n_i | p_i) \tag{3}$$

The parameters of the PDF (i.e., amplitudes (A) and timescales (τ) in Eq. 1) are then varied to maximize this likelihood, just as in a least-squares procedure. Applied to the dC$_{16:1}$ lipids of the gA+dC$_{16:1}$+dC$_{24:1}$ simulation (**Supplemental information Figure S2**), the procedure yields τ_{fast} of 4.76 ± 0.3 ns and τ_{slow} of 75 ± 7 ns (see Results and Discussion for more information). According to the PDF, the fast timescale process is responsible for 74% of the observed residence times but only 18% of the ensemble when the event is weighted by the amount of time spent in the shell. The likelihood of the observed data (from Eq. 3) was compared with the

likelihood of randomly generated data to estimate a *p*-value; the observed data was more likely

than 22% of the randomly generated samplings, indicating an acceptable model.

2.5.2 Supplemental tables

Table S2.1: System information.

System	Lipids/Leaflet	System Size$^\$$	Water Molecules	Total Atoms
gA+DC$_{16:1}$+dC$_{24:1}$	90	$78 \times 78 \times 100$	11,734	62,823
gA+dC$_{18:1}$+dC$_{22:1}$	90	$78 \times 78 \times 99$	11,808	63,045
gA+dC$_{20:1}$	90	$78 \times 78 \times 99$	11,703	62,736
dC$_{16:1}$	50	$59 \times 59 \times 72$	4,530	26,219
dC$_{18:1}$	50	$59 \times 59 \times 80$	4,855	28,392
dC$_{20:1}$	50	$58 \times 58 \times 85$	4,753	29,292
dC$_{22:1}$	50	$57 \times 57 \times 88$	4,631	30,128
dC$_{24:1}$	50	$55 \times 55 \times 99$	4,543	31,066
dC$_{16:1}$+dC$_{24:1}$	90	$78 \times 78 \times 86$	9,152	54,519
dC$_{18:1}$+dC$_{22:1}$	90	$78 \times 78 \times 83$	8,580	52,795

[1]The system size is given in terms of $L_X \times L_Y \times L_Z$, which correspond to each system's X, Y, and Z dimensions.

Table S2.2: First shell residence times with associated uncertainty of gA in $dC_{16:1}+dC_{24:1}$, $dC_{18:1}+dC_{22:1}$, and $dC_{20:1}$ bilayers, respectively. Single lipid residence times (e.g., $dC_{16:1}$ and $dC_{24:1}$ in the $dC_{16:1}+dC_{24:1}$ bilayer), as well as the pooled residence time (e.g., $dC_{16:1}+dC_{24:1}$), are shown.

Bilayer	Residence Time (ns)
$dC_{16:1}$	75 ± 7
$dC_{24:1}$	74 ± 11
$dC_{16:1}+dC_{24:1}$	73 ± 11
$dC_{18:1}$	113 ± 15
$dC_{22:1}$	91 ± 12
$dC_{18:1}+dC_{22:1}$	103 ± 10
$dC_{20:1}$	70 ± 5

Table S2.3: Radial lipid shell locations (in Å) and the number of lipids in each shell per leaflet (in parentheses) for the first four shells with standard errors.[$]

System	First Shell	Second Shell	Third Shell	Fourth Shell
gA+DC$_{16:1}$PC+DC$_{24:1}$PC	12.23 ± 0.08 Å (7.27 ± 0.04)	19.35 ± 0.15 Å (13.84 ± 0.09)	27.16 ± 0.09 Å (20.28 ± 0.10)	34.95 ± 0.09 Å (26.01 ± 0.06)
gA+DC$_{18:1}$PC+DC$_{22:1}$PC	12.52 ± 0.17 Å (7.31 ± 0.07)	19.64 ± 0.10 Å (14.00 ± 0.12)	27.49 ± 0.10 Å (20.76 ± 0.15)	35.37 ± 0.05 Å (26.43 ± 0.08)
gA+DC$_{20:1}$PC	12.32 ± 0.05 Å (7.17 ± 0.03)	19.43 ± 0.08 Å (13.82 ± 0.07)	27.23 ± 0.08 Å (20.32 ± 0.08)	35.06 ± 0.07 Å (26.15 ± 0.06)

[$]Standard errors were calculated by dividing the trajectory into 100-ns blocks.

Table S2.4: Bilayer properties calculated from bilayer-only systems with associated standard errors.[$]

Bilayer	Area/Lipid (Å^2)	Hydrophobic Thickness[†] (Å)	Lateral Compressibility[‡] (dyn/cm)
dC$_{16:1}$	68.81 ± 0.14	23.87 ± 0.06	306 ± 30
dC$_{18:1}$	68.66 ± 0.20	27.48 ± 0.11	303 ± 34
dC$_{20:1}$	66.25 ± 0.13	31.81 ± 0.05	306 ± 48
dC$_{22:1}$	64.49 ± 0.44	36.27 ± 0.23	363 ± 39
dC$_{24:1}$	62.03 ± 0.30	41.79 ± 0.21	516 ± 39
dC$_{18:1}$+dC$_{22:1}$	67.10 ± 0.17	31.03 ± 0.09	348 ± 78
dC$_{16:1}$+dC$_{24:1}$	66.82 ± 0.17	31.67 ± 0.07	367 ± 53

[$]Standard errors were calculated by dividing the trajectory into 100 ns blocks. The calculated values are generally in good accord with experimental values (76, 126).
[†]The hydrophobic thickness of the lipids was estimated by the average Z positions of the C22 and C32 atoms (i.e., the acyl carbon atoms next to the carbonyl carbons) in each leaflet.
[‡]The lateral compressibility was calculated from fluctuations of the entire simulation cell.

2.5.3 Supplemental figures

Figure S2.1: Lipid components in (A) gA+dC$_{16:1}$+dC$_{24:1}$, (B) gA+dC$_{20:1}$ (as the control), and (C) gA+dC$_{18:1}$+dC$_{22:1}$.

Figure S2.2: The histogram probability distribution functions of the first shell transition time and the double exponential fits for total bilayer and constituents of gA in (A) $dC_{16:1}+dC_{24:1}$, (B) $dC_{18:1}+dC_{22:1}$, and (C) $dC_{20:1}$. The histograms for the total bilayer are in black, the shorter lipids ($dC_{16:1}$ and $dC_{18:1}$, respectively) in red, the longer lipids ($dC_{24:1}$ and $dC_{22:1}$, respectively) in blue, and dotted black curves for the fit. Insets for each of the plots zoom on the tail of the double exponentials to show the quality of fit at long time scales.

Figure S2.3: Terminal carbon densities decomposed by shell location. Dotted lines are densities for lipids with head groups above Z = 0 Å. Densities for (A) $dC_{16:1}$ and (B) $dC_{24:1}$ in the $dC_{16:1}+dC_{24:1}$ mixture. The terminal carbon atoms of $dC_{16:1}$ show small amounts of interdigitation, particularly in the first shell. $dC_{24:1}$ densities show high amounts of interdigitation (the peak terminal density of the lipids with head groups above Z = 0 Å is actually near Z = –3 Å to Z = –5 Å), which is only slightly relaxed in the outer shells.

Figure S2.4: Average acyl chain order parameters of (A) $dC_{16:1}$ and (B) $dC_{24:1}$ in the first four radial lipid shells in $gA+dC_{16:1}+dC_{24:1}$. For comparison, the order parameters from $dC_{16:1}$, $dC_{24:1}$, and $dC_{16:1}+dC_{24:1}$ are also shown. Uncertainties are less than or equal to the marker size. Reported order parameters are an average of the *sn*-1 and *sn*-2 tails.

Figure S2.5: Average acyl chain order parameters of (A) $dC_{18:1}$ and (B) $dC_{22:1}$ in the first four radial lipid shells in gA+$dC_{18:1}$+$dC_{22:1}$. For comparison, the order parameters from $dC_{18:1}$, $dC_{22:1}$, and $dC_{18:1}$+$dC_{22:1}$ are also shown. Uncertainties are less than or equal to the marker size. Reported order parameters are an average of the *sn*-1 and *sn*-2 tails.

3   Curvature Stress Induced by Gramicidin A Channel Formation and its Application to Continuum Elastic Models[b][c]

3.1   Introduction

3.1.1   Background

The lipid environment must adapt to the specific conformation of the protein it surrounds. The same strong hydrophobic interaction that localizes proteins to the membrane also ensures that the thickness of the protein and membrane are matched (1,2). The soft lipid leaflets compress (3) and bend (4–8), and their lipids tilt (9–11) to accommodate the protein, including its cavities and convexities (12). Thus a conformation can be resisted or promoted by its compatibility with the surrounding bilayer (12–15).

The membrane itself has been successfully treated as a material, albeit one with significant vertical inhomogeneity and thus stress (16,17). This has been accomplished through continuum elastic models (CEMs), described by elastic moduli (18), local stresses, and with boundary conditions enforcing the strong hydrophobic matching as a constraint. A fundamental assumption of an elastic model of the membrane is that all the complexity of the protein can be summarized at the protein-membrane boundary condition, to which the membrane adapts across all protein classes. In the material view, two problems emerge: how lipid composition affects the local elastic parameters (19–22) and, once a model is determined, what force a local lipid composition exerts on the protein's conformation (23–26) and vice versa (27,28).

These questions tend to have been pursued using a two-dimensional (2D) CEM to describe the membrane deformation around the protein. The model begins with the Helfrich Hamiltonian

(4,5) for (lateral) curvature energetics, and then adds a term modeling the vertical (or normal) compression of the material that occurs to achieve hydrophobic matching at the protein. The model is attractive because the energy function, E (Equation 3.2), is parameterized by experimentally available quantities. However, at the crucial interface between the three-dimensional (3D) protein and 2D bilayer, a choice must be made for the connection. Without independent guidance, alternate (29,30), but conceptually reasonable boundary conditions can quadruple the deformation energy for a particular mismatch.

As demonstrated in **Chapter 2**, compression energies can be estimated from all-atom simulation. However, calculating the bending energetics is more complicated. In this work, all-atom simulations were used to calculate the curvature stress in each lipid leaflet of the bilayer. Leaflet curvature stress is explicitly available from simulation (32) and implicitly from experiments on the response of the protein to lipid-compositional curvature stress (13,23,33). These stresses were then compared to stresses with 2D- and 3D-CEMs built by collaborators.

A basic tool for this and similar studies is the use of gramicidin A (gA) as a "force probe" (34,35). The transmembrane channel is formed from two monomers in opposite leaflets. The activation energy for dissociation of the channel is modulated by the energy of the material deformation around the complete channel, and this directly impacts the channel lifetime. This system can thus be used to validate a material model of the bilayer usable for more complex systems. Additionally, gA has approximate rotational symmetry, which simplifies numerical simulations and analysis.

Although discussed in **Chapter 2**, gA simulation results are validated by single-channel lifetime experiments. For clarity, some material is repeated here, but now from the perspective of calculating lipid bending energetics and inferring the peptide-bilayer boundary condition. The

deformation energy, E, is a quadratic function of the per-leaflet mismatch at the peptide-bilayer boundary (u_0), given by Andersen & coworkers (34):

$$E = H(2h_0 - l)^2 = H(2u_0)^2, \tag{3.1}$$

where h_0 is the unperturbed thickness of a lipid leaflet ($2h_0 = d_0$, where d_0 is the bilayer thickness), l is length of the channel, and H is a phenomenological spring constant reflecting bilayer deformation energetics. All contributions to the energy, including the effect of the boundary conditions, are gathered into H. The energy difference between the transition state gA and dimer is $H(4u_0\delta - \delta^2)$, and in a transition state model explains the difference in mean channel lifetimes of gA in different environments. Two unknowns associated with this treatment are: the slope of a lipid leaflet at the peptide-bilayer interface (which influences H) and the difference in *effective* hydrophobic thickness between the TS and dimer gA (δ). Neither of these quantities is directly available from experiment.

The choice for a boundary condition has been debated since the theory of elastic bilayer deformations was applied in 1986 (Huang (30)) to analyze the observed relation between gramicidin single-channel lifetimes and lipid bilayer thickness (37). This analysis showed that the s value, the slope of the leaflet at the channel, deduced was quite different, in fact negative, from the value that would minimize the deformation energy (29); a similar conclusion was reached by Lundbæk and Andersen (38). It is in this context important that the choice of s effectively lumps all the uncertainties about the model (and the molecular origins of lipid packing adjacent to the channel, which cannot be included in the 2D-CEM) into a single adjustable parameter.

This work is focused on the use of simulation to resolve the ambiguity of the peptide-bilayer boundary and describe curvature stress caused by gA. All-atom simulations provide molecular

details of the boundary as well as information on the curvature stress, which is highly correlated to the boundary. In this way, the simulation is treated as if it were an experiment, i.e., that with no adjustable parameter it has predictive power. Although the lipid force field used here, CHARMM 36 (C36) (39), reproduces assorted mechanical properties quite well (40), the accuracy of any force field is open to question (41). Unfortunately, how lipids at the gA boundary adapt is difficult to validate directly by experiment, and is another source of concern. These issues highlight the importance of the indirect comparison with experiment detailed in **Chapter 2**, which estimates lateral lipid redistribution through H and an estimatation of the monomerization transition state. Finally, although not discussed in depth within this dissertation, collaborators used the information presented herein to build and validate a 3D-CEM, which provided support to data collected from simulation in this chapter and **Chapter 2** (discussed briefly in the **Supplemental information**).

Characterizing the protein-lipid interface continues to be a challenge. For example, Yoo and Cui (135) have computed spatially resolved local normal and tangential stresses around the gA channel to characterize the bilayer deformation beyond using structural information. In a review of continuum membrane modeling, Argudo et al. (136) present a method for performing hybrid continuum-atomistic simulations with a continuum membrane that provides a close match to atomistic modeling. This work is complementary to these studies, as it develops a technique for characterizing and validating continuum peptide-lipid boundaries against simulation.

The general effectiveness of this model beyond gA will be determined by the limit of its applicability; the first place to look is in the first shell of lipids around the protein, where the lipid material ends and the complexity of the protein begins. As part of the introductory material for this work, the theoretical background for first shell lipid properties, including compression

and curvature stress, will be discussed.

3.1.2 Leaflet bending and compression

For a system with an inclusion (like gA) whose hydrophobic thickness does not match the surrounding lipid environment, the soft leaflet must deform to match the length of the inclusion. This article deals with the case where leaflets are thicker than the inclusion, and so the language will follow this case; given the strong energetics of hydrophobic matching, lipids in the first solvation shell must be compressed in thickness to match the channel. Following the deformation radially outward, the leaflet must eventually bend up to match the thickness of the lipid environment far from the inclusion. However, the precise behavior of this deformation in the interim must be modeled or inferred from experiment (4, 7, 67).

The widely used 2D-CEM (4, 7, 67, 125, 136) combines the Helfrich Hamiltonian (55, 56) with the modulus for lipid area compressibility to account for curvature and compression, respectively, through the function describing the leaflet hydrophobic surface, i.e., the leaflet thickness at a radial position r, $h(r)$:

$$E = \int \left[\frac{k_{c,m}}{2} (-\nabla^2 h(r) - R_0^{-1})^2 + \frac{K_{A,m}}{2} \left(\frac{h(r) - h_0}{h_0} \right)^2 \right] dA, \tag{3.2}$$

where E is the deformation energy, $k_{c,m}$ is the leaflet (monolayer) bending modulus (137), $K_{A,m}$ is the area compressibility (138, 139), and R_0^{-1} is the intrinsic curvature (31). Experimentally, $K_{A,b}$ ($K_{A,b} = 2K_{A,m}$) can be determined by the pipette aspiration technique to be between 265–290 mN/m (40, 139) depending somewhat on the application of tension, and these values are generally replicated by simulation (see **Chapter 2**, for example). There are generally large disagreements on the values of $k_{c,m}$, as they appear to depend on methodology, and even buffer conditions (37, 123, 137, 140–142). The values of $k_{c,m}$ were taken from Venable et al. (142).

Previous modeling has pointed out that the total energetics of the deformation depend

critically on the behavior of the deformation in the first shell, frequently described using the slope of the deformation as it meets the inclusion (7, 20, 71). As briefly discussed in this chapter, it is noted that the total curvature stress of the leaflet is proportional to the value of the slope (s). Second, the compression of lipids in the first shell is very large for small slopes, where the leaflet is not allowed to immediately deflect away from the inclusion in a way that would minimize the compression. This first shell compression will lead to a large energetic penalty for long lipids to occupy the first shell. These effects are illustrated in **Figure 3.1**.

Figure 3.1: The consequences of the slope of the membrane surface intersecting the inclusion that relate to the first shell compression and curvature stress. If the slope is constrained to be zero, there is zero net curvature, but large compression (red). If the slope is unconstrained, there is net positive curvature (blue).

Through these two metrics, the simulations and analysis justify a model that considers the 3D deformation of the leaflet around the inclusion, rather than simply a 2D model of the hydrophobic surface. An important benefit of the 3D-CEM, with information from molecular simulation, is the removal of the ambiguity of the boundary condition. As shown, although not

quantitatively as accurate for curvature frustration, the properties of the 2D-CEM(*s* free) (i.e., where the slope at the protein-lipid boundary condition is not fixed) are comparable to those of the 3D-CEM. However, without molecular justification for the free boundary condition, extension to larger proteins is questionable.

The blue and red arrows in **Figure 3.1** illustrate the slope near the peptide for positive and zero slopes, respectively. In this figure, the zero slope condition in red shows the correlation with increased compression. The relation between slope, compression, and curvature is summarized in the figure by color.

These experimental and simulation observables (curvature frustration and compression) are critical to assessing model quality. Three models are presented below. The first is the all-atom model, which derives its predictive power from the molecular detail. For the all-atom model, the deformation energy itself is ambiguous, yet the curvature stress can be directly calculated. Second, two continuum models are presented that reduce the effect of the bilayer to that of a surrounding material. These models are useful for rapidly predicting the effect of the surrounding lipids on the inclusion. The continuum models were constructed by coworkers (7, 9).

As discussed, curvature stress is directly related to the amount of first shell compression. The Results section is structured to first validate the curvature stress of the 3D-CEM against that of the all-atom model, and then to use the 3D-CEM to predict the degree of lipid redistribution around the channel (**Supplemental information**).

3.2   Methods

3.2.1   All-atom methods

3.2.1.1 All-atom model description

The molecular model of the leaflet deformation does not explicitly account for curvature and

compression, rather, they are present naturally as a result of the simulation. The lipids diffuse on the surface as it fluctuates according to the simulated ensemble, and the deformation can be computed according to the average value of a particular site in the model chosen to correspond to the hydrophobic surface. In theory, the molecular model contains all of the relevant energetics of the surface. However, they cannot be extracted trivially. In particular, the energetics are not a local function of the height or material deformation, as they are in the continuum models, and so a local energy density as in Equation 3.2 cannot be computed.

Even with the quantity $h(r)$ computed from a histogram of atomic positions, the inclusion-lipid boundary condition is still ambiguous, as can be seen from **Figure 2.7** (**Chapter 2**). The quantity of $h(r)$ is defined even above the peptide, where the surface is not behaving as a continuum membrane. These must represent fluctuations of the hydrophobic atom above the peptide. For this case, Equation 3.2 cannot be applied to $h(r)$ in the vicinity of the peptide. Interpreted directly, they indicate that depending on where the slope is radially evaluated, it may be zero or even negative (71). The trace of the lipid atom positions, representing the local director of the lipid traced along its molecular path, is an alternative representation of the boundary condition and shows a dramatically different value of the slope (i.e., the slope at a point in z is defined as a radial vector orthogonal to the trace) that is positive for reasonable choices of the hydrophobic surface (**Section 3.2.1.2**).

In **Chapter 2**, extensive (multiple microseconds) simulations of gA in mixed lipid bilayers were reported. In these simulations, short acyl-chain lipids are shown to reside preferentially in the first shell around the inclusion, but with a fraction indicating only moderate lipid compression. In addition, the **Supplemental information** in this chapter briefly describes how a CEM is that uses the occupancy and timescales reported in **Chapter 2** can assess model

energetics.

The initial coordinates for all systems were obtained from the *Membrane Builder* module in CHARMM-GUI (charmm-gui.org) (63–65), and the gA structure was from PDB: 1JNO (66). Pure bilayer simulations were conducted of phosphatidylcholine (PC) lipids with varying tail length, $dC_{16:1}$, $dC_{18:1}$, $dC_{20:1}$, $dC_{22:1}$, and $dC_{24:1}$, with 50 lipids/leaflet, 45–50 waters/lipid, and 0.15 M KCl. Two sets of gA+bilayer systems were made for each lipid type, a set with two gA monomers and a set with a single gA dimer. Systems containing gA were built with 90 lipids/leaflet, ~45 waters/lipid, and 0.15 M KCl.

To eliminate differences in gA structure due to fluctuations, gA in all monomer and dimer systems were restrained by a harmonic root mean square deviation (RMSD) restraint with a force constant of 5.0 kcal/(mol·Å^2) relative to the minimized PDB: 1JNO structure. In all cases, the gA was parameterized using the C22 protein parameter set with the ''correction-map'' (dCMAP) addition (67), and lipids were parameterized using the C36 parameter set (39). Bonds with hydrogen atoms were constrained using the SHAKE algorithm (68). To reduce possibly long equilibration times, bilayers of systems including a gA dimer were deformed to an idealized surface from a Helfrich continuum model. The surface itself was determined by a deformation taken from the approximated length and radius of gA. By doing this, lipid head groups could be placed close to the hypothetical equilibrium conformation before any minimization was performed. Monomer gA systems contained one monomer per leaflet at maximal separation (to minimize transmembrane monomer-monomer interactions) and with the z position taken from previous results (53). Monomer x and y positions were held by a harmonic potential with a force constant of 1.1 kcal/(mol·Å^2), and were unconstrained in z. No positional restraints were placed on the dimer.

All simulations (pure, dimer+bilayer, and monomer+bilayer) were run with a 1 fs time step for 120 ns using NAMD. Constant pressure of 1 bar was regulated by Nosé-Hoover Langevin piston pressure control (piston period of 50.0 fs and piston decay of 25 fs), and constant temperature of 310.15 K was maintained by Langevin dynamics (coupling coefficient of 1 ps^{-1}). Electrostatic interactions were calculated using the particle-mesh Ewald method (69) (mesh size = 1 Å; κ = 0.34 Å^{-1}; and sixth-order B-spline interpolation), van der Waals interactions were switched off between 10 and 12 Å by a force-switching function (70), and a pair list distance cutoff of 16 Å was imposed. All by the PME contribution to the pressure profile was calculated during simulation using NAMD's pressure profile capability (71,72) with 100 equally spaced slabs and the profile calculated every 100 fs. The PME contribution was calculated post-simulation, where the contribution was calculated every 0.5 ps.

3.2.1.2 Local lipid director (trace)

Lipid traces are constructed by first accumulating a histogram (on a fine mesh) of both the density, $\rho_i(r)$, and average height, $z_i(r)$, of each atom site i of the lipid as a function of r, the distance from the peptide. Each trace averages over n lipids. Second, the value r_0 is computed:

$$\int_0^{r_0} \rho_i(r)\,\mathrm{d}r = n, \tag{3.9}$$

that is, such that n lipids are accounted for between 0 and r_0. Then, the average radius and height of the lipid site are computed as:

$$\{\langle r_i \rangle, \langle z_i \rangle\} = n^{-1} \int_0^{r_0} \rho_i(r)\{r, z_i(r)\}\,\mathrm{d}r. \tag{3.10}$$

These sets of $\{\langle r_i \rangle, \langle z_i \rangle\}$ pairs are then connected along the chemical connectivity of the lipid to yield the trace. The vector between sites $\{\langle r_i \rangle, \langle z_i \rangle\}$ and $\{\langle r_{i+1} \rangle, \langle z_{i+1} \rangle\}$ then defines a locally averaged lipid director. For the next trace, r_0 is used as the lower limit, and the upper limit of integration is computed to account for the next n lipids. Continuous connectivity is achieved by

using only the *sn-2* chain and ignoring atoms that branch off the path to the head group (**Figure 3.2**). The slope at the protein-bilayer interface taken from the trace is used to fix *s* for the 3D-CEM.

Figure 3.2: Lipid deformations from the all-atom molecular model. The purple line represents the hydrophobic surface computed from the average position of the C21 carbon. Black lines represent lipid traces, computed as described in the main text.

3.2.2   2D-CEM methods

As mentioned in **Footnote d**, other groups and collaborators developed the 2D-CEMs (30,31,34,43-45). Since this dissertation is not intended to focus on the methodology/application of these CEMs, please see Sodt et al. for a more detailed description (9).

In short, a 2D-CEM is defined by Equation 3.2 directly. The deformation of the bilayer normal to the surface is integrated implicitly, that is the local curvature and height at any point on the surface defined the local lateral deformation. However, at the protein-bilayer interface, a decision must be made on the slope of this contact (s, or $h'(r_0)$). Thus, two models were chosen as bookends to encompass the observations of the simulations. The first, denoted 2D-CEM(s free), lets the slope of the membrane at the inclusion boundary vary to minimize the deformation energy. The second, denoted 2D-CEM($s = 0$), constrains the slope of the membrane at the inclusion boundary to be zero. A more complex energetic description of the slope near the inclusion is possible, for example, by fixing the slope to match a molecular simulation. This has been difficult, however, as shown in Lee et al. (71), where, even for a thick bilayer around gA, the upper leaflet appears to slope down before sloping back up. The use of the hydrophobic surface to extract the slope leads to negative values of s, leading to a negative curvature near the channel.

3.2.3 3D-CEM methods

The 3D-CEM extends past the surface representation of the monolayer and attributes deformation energetics through the z thickness of the monolayer (143). In this representation, the monolayer was divided into three z-regions (**Figure 3.3**). Region 1 (light grey) is composed of the hydrophobic lipid tails, Region 2 (blue) is composed of the glycerols and head groups, and Region 3 (dark grey) is composed of the hydrated head groups. Assumedly similar to a real lipid leaflet, each of these regions has different local moduli, which allows a consistent treatment of the origins of bending and compression, including their coupling away from the surface. This assumption naturally introduces more parameters.

From the perspective of the bilayer as a 3D material with local elastic properties, the area

compressibility modulus of a leaflet can be viewed as the average of the local material area compressibility modulus:

$$K_{A,m} = \int_0^{t_0} \overline{K}_A(z)\,dz, \tag{3.4}$$

where t_0 is the thickness of the leaflet and the units of $\overline{K}_A(z)$ are those of pressure (that of an elastic material modulus in three dimensions).

A bending deformation induces a lateral deformation proportional to the distance from the neutral surface (i.e., where the lateral stretch is uncoupled from the bending). The local lateral deformation energy depends on the square of the deformation through $\overline{K}_A(z)$. The bending modulus is thus:

$$k_{c,m} = \int_0^{t_0} \overline{K}_A(z)\,(z - h_{ns})^2\,dz, \tag{3.5}$$

where the z position of the neutral surface is h_{ns} (144). It is defined by:

$$\int_0^{t_0} \overline{K}_A(z)\,(z - h_{ns})\,dz = 0. \tag{3.6}$$

In the 3D-CEM, each tetrahedral volume element of the leaflet is assigned a local area compressibility modulus, $\overline{K}_A(z)$, according to its distance from the bottom of the leaflet. The quantities $K_{A,m}$, $k_{c,m}$, and h_{ns} are experimentally available and thus can parameterize three regions, each with a different constant value of $\overline{K}_A(z)$ (**Table 3.1**). The thickness of the regions is defined by the varying chemical characteristics: lipid acyl chain, glycerol / head group, and hydrated head group. With varying acyl chain length, the thickness of the acyl chain region is extended without changing the value of $\overline{K}_A(z)$. These three regions are composed of tetrahedra. When the material surrounding the cylinder is constrained to match the cylinder length, the tetrahedra deform with an energy dependent on the above material properties. The slope at the protein-bilayer interface is extracted from molecular dynamics simulation (see **Section 3.2.1.2** and **Figure 3.2**). See Sodt et al. (9) for more information on the 3D-CEM energy calculation.

Figure 3.3: Representations of the 3D-CEM developed within Sodt et al. (9). There are three distinct regions with associated moduli. The energy is dependent on the deformation of the tetrahedra.

Table 3.1. Values of the local area compressibility modulus used in the 3D-CEM.

		$\overline{K}_A(z)$	
Lower Limit (Å)	Upper Limit (Å)	kcal mol^{-1} Å^{-3}	mN m^{-1} nm^{-1}
0	h_0	4.28×10^{-3}	35.6
h_0	$h_0 + 7$	20.0×10^{-3}	97.2
$h_0 + 7$	$h_0 + 11$	8.25×10^{-3}	23.0

3.3   Computing curvature stress

The curvature stress, represented as the derivative of the free energy per unit area ($\bar{F}$) with respect to curvature, is computed in all three model types, the 2D-CEM, the 3D-CEM, and the all-atom, using the methodology defined below. In each case, it is computed as $\bar{F}'(R^{-1} = c)|_{c=0} = \bar{F}'(0)$, where $R^{-1} = c$ is the global curvature of the system.

3.3.1 $\bar{F}'(0)$ computed for the all-atom models and 3D-CEM

The value of $\bar{F}'(0)$ can be computed from a 3D model (including the all-atom) using the first moment of the lateral pressure profile (37, 49, 50, 145):

$$\bar{F}'(0)_{3D} = -\int_0^\infty z[p_T(z) - p_N(z)] \ dz. \tag{3.8}$$

The all-atom model has the added complexity that an individual monomer of gA deforms the leaflets and so induces curvature stress even when it does not form a transmembrane channel. This is accounted for by performing simulations of both the monomer and dimer systems and computing $\Delta\bar{F}'(0) = \bar{F}'_d(0) - \bar{F}'_m(0)$ under the same peptide/lipid ratio (**Figure 3.4**), where the subscript indicates the molecular simulation performed. This expression attempts to account for the change in curvature stress due to only the leaflet deformation, the closest comparison possible to the continuum models. See Sodt et al. (9) for more information on the $\bar{F}'(0)$ calculation on the 3D-CEM.

The all-atom model simulations show how the deformation changes as a function of channel-bilayer hydrophobic mismatch (**Figure 3.5**). **Table 3.2** summarizes the $\bar{F}_{pure}(0)$, $\bar{F}'_m(0)$, $\bar{F}'_d(0)$, and $\Delta\bar{F}'(0)$ values from simulation. As hydrophobic mismatch increases, it becomes increasingly unfavorable for the channel to dimerize. Not only does curvature stress increase, but compression frustration also clearly increases by visual inspection (**Figure 3.5**).

Figure 3.4: The calculation of $\Delta\bar{F}'(0)$ involves $\bar{F}'(0)$ calculations for gA (A) monomers and (B) dimers.

Figure 3.5: Average bilayer deformation profiles. Lipid leaflet deformations increase as a function of hydrophobic mismatch with the channel. With increased deformation comes increased energy penalties. The hydrophobic surface of the bilayer is shown in blue.

Table 3.2. The $\bar{F}(0)$ values for lipid-only systems ($\bar{F}_{\text{pure}}(0)$), monomer-containing systems ($\bar{F}_{\text{m}}(0)$), dimer containing systems ($\bar{F}_{\text{d}}(0)$) and the leaflet bending frustration due to curvature ($\Delta\bar{F}(0)$).

	$dC_{16:1}$	$dC_{18:1}$	$dC_{20:1}$	$dC_{22:1}$	$dC_{24:1}$
$\bar{F}_{\text{pure}}(0)$	0.0322 ± 0.0030	0.0473 ± 0.0046	0.0578 ± 0.0046	0.0558 ± 0.0030	0.0485 ± 0.0051
$\bar{F}_{\text{m}}(0)$	0.0489 ± 0.0027	0.0559 ± 0.0025	0.0786 ± 0.0037	0.0685 ± 0.0017	0.0634 ± 0.0045
$\bar{F}_{\text{d}}(0)$	0.0585 ± 0.0017	0.0648 ± 0.0018	0.0923 ± 0.0031	0.1168 ± 0.0046	0.1529 ± 0.0053
$\Delta\bar{F}(0)$	$\mathbf{0.0096 \pm 0.0032}$	$\mathbf{0.0090 \pm 0.0031}$	$\mathbf{0.0137 \pm 0.0048}$	$\mathbf{0.0483 \pm 0.0049}$	$\mathbf{0.0895 \pm 0.0070}$

3.3.2 $\bar{F}'(0)$ computed for the 2D-CEM

Curvature in the 2D-CEM is accounted for locally by the Helfrich Hamiltonian:

$$\bar{F}_{\text{lipid}}(R^{-1}) = \frac{k_{c,m}}{2}(R^{-1} - R_0^{-1})^2 \tag{3.9}$$

$$\bar{F}'(0) = -k_{c,m}R_0^{-1}, \tag{3.10}$$

where $\bar{F}$ is the local energy per unit area at a point with curvature R^{-1}. The global curvature stress is approximated by computing the integral of the local curvature of the system that is embedded in a cylinder with inverse radius $R^{-1} = R_{\text{global}}^{-1} = 0$. To obtain the $\bar{F}'$ from the 2D-CEM,

$$\bar{F}'(0)_{2D} = k_{c,m}A^{-1}\int 2\pi r[c_r(r) + c_t(r)] \, dr, \tag{3.11}$$

where $c_r(r)$ and $c_t(r)$ are the radial and tangential curvatures at radius r, respectively, and A is the area.

3.4   Results and discussion

3.4.1   3D-CEM and 2D-CEM(s = free) profiles compare well to all-atom

The hydrophobic surfaces, $h(r)$, of the four models are shown in **Figure 3.6** for gA+dC$_{22:1}$ and gA+dC$_{24:1}$. For the all-atom model, the CHARMM "C21" atom is plotted (the ester atom on the *sn-2* chain). For the CEMs, $h(r)$ is plotted for the interface between the polar and apolar regions where the matching condition is chosen. The key feature of the plot is that the 2D-CEM with slope unfixed and the 3D-CEM match the all-atom simulation reasonably well, whereas the 2D-CEM with the slope fixed at zero is qualitatively incorrect.

Figure 3.6: Hydrophobic surfaces plotted for the all-atom model and CEMs. The plots are color-coded as follows: purple for the all-atom model, red for the 2D-CEM(s=0), blue for the 2D-CEM(s free), and green for the 3D-CEM.

3.4.2 Comparison of $\bar{F}'(0)$ between models

As noted by Equation 3.10, $\bar{F}'(0)$ is proportional to the total curvature of the system and so should be sensitive to the effect of the boundary condition. Results for $\bar{F}'(0)$ computed for all models are shown in **Figure 3.7**. For both 2D-CEMs, the total curvature is equal to Ring's expression, $2\pi r_0 s$ (146). Thus, the $s = 0$ model has $\bar{F}'(0) = 0$ for all lipids, inconsistent with the simulation results. With s free, the value is qualitatively similar in that all-atom simulation and 2D-CEM(s free) both indicate net positive curvature. The 3D-CEM values are consistent with the all-atom simulation. In two ways, the 3D-CEM is a more faithful model of the all-atom simulation.

Figure 3.7: Global curvature single-leaflet stress, expressed as $\bar{F}'(0)$. Points are evaluated from the all-atom simulation. For the 2D-CEM, thin lines use the Monge gauge to evaluate curvature, whereas thick lines use the true definition.

First, the 3D-CEM boundary condition has been chosen to match the all-atom simulation. As discussed, in the 2D models, curvature frustration is directly dependent on this boundary condition, and the same is true for the 3D-CEM. Second, calculation of $\bar{F}'(0)$ from the lateral stress profile in the 3D-CEM is equivalent to the calculation in the all-atom model. Even without interpretation through the continuum models, the values of $\bar{F}'(0)$ from molecular simulation suggest that the $s = 0$ boundary condition is incorrect even without being able to unambiguously characterize the slope of the all-atom model hydrophobic surface.

3.4.3 Energetics suggest a non-zero peptide-lipid boundary slope

New simulation techniques were applied to extract the peptide-lipid boundary. First, the lipid trace method revealed the form of the boundary lipids extracted from the molecular simulation. This boundary was used by the 3D-CEM to compute energetics and curvature stress. Second, the leaflet curvature stress was computed by all-atom simulation by subtracting $\bar{F}'_m(0)$ from $\bar{F}'_d(0)$. The curvature stress of the 3D-CEM matches well with the all-atom simulation $\bar{F}'(0)$. The 2D-CEM(s free) does reasonably well, considering the simplicity of the model. These results

71

indicate the importance of the boundary for $\bar{F}'(0)$, and demonstrate that the slope at the gA-lipid interface is nonzero.

3.5   Conclusions

Intuitively, the information about the boundary condition necessary for the 2D-CEM should be available from $h(r)$ computed from the all-atom simulation. However, due to the "shape" of the inclusion, in this case, a squat cylinder, lipids go over the top of the channel such that $h(r)$ is no longer the height of the leaflet in the proper sense (i.e., there is no lipid material below the hydrophobic atom). An alternative method for computing the hydrophobic surface is to compute lipid traces. These traces indicate the shape of the channel by recording the path of the lipid from tail to head group. The trace indicates both the proper boundary condition for use by the 3D-CEM and an additional way to compute the slope (from the vector normal to the trace). These traces indicate that s is positive without the ambiguity of examining $h(r)$.

The effective matching condition is confirmed by computing the curvature stresses of the leaflet around gA, quantified by $\bar{F}'(0)$. This thermodynamic quantity indicates that indeed there is net internal positive curvature of the lipids. The sign of $\bar{F}'(0)$ shows how the free energy would change if the leaflet's curvature could be individually changed. In this case, a positive value of $\bar{F}'(0)$ means the free energy would go down on a negatively curved leaflet, partially undoing the strain of the internal positive curvature created by the inclusion. Unlike what is seen in the all-atom simulation, the 2D-CEM($s = 0$) predicts that there should be minimal curvature stress. The 2D-CEM($s = $ free) and 3D-CEM values are consistent with the sign of the all-atom simulation, yet the 2D-CEM values are too high, suggesting that the first shell curvature in the model is too high and that reliance on the apparent agreement of $h(r)$ discussed above is unwarranted. In contrast, the 3D-CEM is in quantitative agreement with the all-atom simulation.

Together, these observations point to the use of the additional parameters of the 3D-CEM to consistently model the interplay between protein shape and membrane deformation energetics. The use of the all-atom simulations to compute curvature stress indicates the matching condition (i.e., a positive slope) without interpreting a molecular feature like $h(r)$ at the atomic level. Additionally, by computing local lipid directors (lipid "traces") the effective shape of the channel is revealed in a way that can be transferred directly to the 3D-CEM. By combining the 3D-CEM and these simulation observations, the energetics of the membrane surface are determined and validated.

3.6 Supplemental information

Sodt et al. (9) contains additional information on the theory and computational method involved with the CEMs. Additionally, using a 2D- and 3D-CEM, various comparisons were made to experimental and simulation data from **Chapter 2**. This information is mentioned where applicable in **Chapter 2**. Coworkers from Sodt et al. (9) generated the majority of the information within this section.

3.6.1 Monte Carlo modeling of lipid diffusion

For each selected frame of an all-atom simulation, lipids were assigned an average position and a 2D Voronoi decomposition was computed (data in **Chapter 2**). Sites that share a border were labeled as "connected," establishing a map of neighboring sites. Each site was assigned a shell based on the minimal number of neighbors to cross to reach the inclusion.

The total energy for the model runs in computed by assigning each of the short of long lipids (labeled i) an energy (e_j) by shell j, where $s_{ij} = 1$ if lipid i is in shell j, else $s_{ij} = 0$:

$$E = \sum_i e_j s_{ij}. \tag{S3.1}$$

The shell energies were extracted from either the 2D- or 3D-CEM.

For the Monte Carlo simulations, the initial lipid map was used to establish neighbors for the ensuing model run. Lipid hops for each lipid were attempted every 0.2 ns, a period much shorter than the lipid hop time ($\tau \approx 70$ ns). The probability was $k(0.2/\tau)$, where the value of k was adjusted to be 0.8 to match the simulation value of τ. For this attempt, a lipid and one of its randomly selected neighbors switched places, and the energy was recomputed and accepted/rejected with the standard Metropolis criterion. For both the Monte Carlo simulation and the values extracted from the simulation, the enrichment of the populations of each shell for each long or short lipid chain was computed.

Table S3.1: Shell occupancy of $dC_{18:1}$ lipids in a mixed $gA+dC_{18:1}+dC_{22:1}$ simulation extracted over the last 3 μs compared to occupancies from CEMs and experiment.

Model	Shell 1	Shell 2	Shell 3
2D-CEM($s = 0$)	0.66 ± 0.03	0.58 ± 0.02	0.50 ± 0.02
2D-CEM(s free)	0.62 ± 0.03	0.54 ± 0.02	0.49 ± 0.02
3D-CEM	0.59 ± 0.03	0.51 ± 0.02	0.49 ± 0.02
All-atom	0.50 ± 0.04	0.51 ± 0.02	0.52 ± 0.01
Experiment		0.58 ± 0.04	

Table S3.2: Shell occupancy of $dC_{16:1}$ lipids in a mixed $gA+dC_{16:1}+dC_{24:1}$ simulation extracted over the last 3 μs compared to occupancies from CEMs and experiment.

Model	Shell 1	Shell 2	Shell 3
2D-CEM($s = 0$)	0.72 ± 0.03	0.63 ± 0.02	0.52 ± 0.02
2D-CEM(s free)	0.68 ± 0.03	0.56 ± 0.02	0.50 ± 0.02
3D-CEM	0.65 ± 0.03	0.55 ± 0.02	0.48 ± 0.02
All-atom	0.66 ± 0.02	0.55 ± 0.01	0.49 ± 0.01
Experiment		0.66 ± 0.02	

3.6.2 Modeling the phenomenological spring constant H from CEMs

Assuming that the deformation energy is based on:

$$E = H(2h_0 - 2h_p)^2 = H(2u_0)^2 \qquad\qquad (S3.2)$$

where H is the phenomenological spring constant, h_0 is the unperturbed leaflet thickness, h_p is one half the length of the channel (for use in this single-leaflet model), and therefore, u_0 is the single-leaflet deformation. Within H, all contributions to the energy, including the effect of boundary conditions, are gathered into this phenomenological spring constant. By obtaining the total deformation energy at different initial hydrophobic mismatch (i.e., gA in $dC_{16:1}$, $dC_{18:1}$, $dC_{20:1}$, $dC_{22:1}$, and $dC_{24:1}$), a quadratic function is formed. A fit to this quadratic gives H.

As expected, the 2D-CEM($s = 0$) has the largest value of H, as curvature is unable to reduce the large energetic penalty of first shell lipid compression. The 2D-CEM(s free) and 3D-CEM have relatively similar energetics. The 3D-CEM energy is likely lower due to the more detailed treatment of first shell compression; the path of lipid compression follows the bend of the lipid rather than being project directly onto the z-axis, where it will appear to be more extreme. Comparing a breakdown of curvature and compression from the 2D-CEM(s free) model (where they are separable) indicates that ~75% of the energy is from compression and 25% is from curvature.

Table S3.3: Phenomenological spring constants, H.

Model	H (kJ mol–1 Å–2, kcal mol–1 Å–2)
2D-CEM($s = 0$)	215.6, 0.42
2D-CEM(s free)	44.7, 0.11
3D-CEM	35.7, 0.09

4 Characterizing Residue-Bilayer Interactions Using Gramicidin A as a Scaffold and Tryptophan Substitutions as Probes[d]

4.1 Introduction

Membrane proteins often have "aromatic belts," rich in Trp and Tyr residues, at the bilayer-water interface (see Pogozheva et al. (147) for a review). These regions are important in determining protein folding (61, 62, 148), interfacial anchoring (63, 81, 149–151), and function (64, 152–156). Gramicidin A (gA) channels have been important tools in establishing these principles. Previous molecular dynamics (MD) studies on gA suggested that the deformation profiles at the two ends of a pore-forming, asymmetric monomer are quite different, i.e., the Trp-rich N-terminus deforms the adjacent lipids more than the hydrophobic C-terminus (**Figure S4.1 from Kim et al. (157)**). To explore the effects of gA residues at the bilayer-water interface on the adjacent lipids and channel itself, we therefore performed extensive all-atom MD simulations of gA dimers and monomers where the four Trp residues per monomer have been substituted with other residues. Hereinafter, wild-type gA is denoted as gA^{Trp} with a monomer sequence of *formyl*-VGA<u>LA</u><u>V</u>V<u>V</u>W^9<u>L</u>W^{11}<u>L</u>W^{13}<u>L</u>W^{15}-*ethanolamide*, where D-amino acids are underlined.

gA^{Trp} channels form by transmembrane dimerization of two anti-parallel β-helical monomers (**Figure 4.1A**) (63), which are anchored to the bilayer interface by the amphipathic Trp residues (61–63):

$$2M \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} D \qquad\qquad (4.1)$$

---

where M and D denote the non-conducting monomers and conducting dimers, and k_1 and k_{-1} are the association and dissociation rate constants, respectively. The monomer↔dimer equilibrium is described by (18, 62, 64, 65):

$$\frac{[D]}{[M]^2} = \frac{k_1}{k_{-1}} = \exp\left\{-\frac{\Delta G_{protein}^{M\to D} + \Delta G_{bilayer}^{M\to D}}{k_B T}\right\} \tag{4.2}$$

where k_B is Boltzmann's constant and T the temperature in Kelvin. $\Delta G_{protein}^{M\to D}$ encompasses the energy associated with forming/breaking the six inter-monomer hydrogen bonds formed by *formyl*-V^1GA$\underline{L}$A^5 at the dimer interface. We assume that $\Delta G_{protein}^{M\to D}$ is independent of the mutations studied here because amino acid substitutions in the C-terminal half of gA have minimal effects on the energetics of subunit dimerization at the N-terminal side (152). $\Delta G_{bilayer}^{M\to D}$ denotes the bilayer deformation energy associated with the formation of dimeric channels, which is affected by the Trp mutations. For more details on the theoretical treatment of bilayer energetics around gA, see the **Experimental Discussion** in the **Supplemental Information**.

Table 4.1 summarizes published experimental data for gA channels where the four Trp residues of gA^{Trp} were mutated to: i) 1-methyltryptophan (gA^{mTrp}); ii) tyrosine (gA^{Tyr}); or iii) phenylalanine (gA^{Phe}); see **Figure 4.1** for more information on these mutations. The variations in channel lifetime imply that residue-specific interactions at the bilayer-water interface alter the energetics of deforming the bilayer ($\Delta G_{bilayer}^{M\to D}$). It is also apparent that lipid type affects channel lifetime (18, 48, 64). As within the rest of this article, results should be compared between structurally related residues (Trp and mTrp; Tyr and Phe).

Table 4.1: Mean lifetimes (τ) of channels (in ms) formed by gA^{Trp} and analogues in $dC_{18:1}$ (1,2-di-oleoyl-phosphatidylcholine, DOPC), 4ME $dC_{16:0}$ (1,2-di-phytanoyl-phosphatidylcholine, DPhPC), $dC_{18:2}$ (1,2-di-linoleoyl-phosphatidylcholine, DLoPC). Experimental conditions: 1.0 M CsCl, ±200 mV, and 25 ± 1 °C.

Mutant / Lipid	$dC_{18:1}$ [#]	4ME $dC_{16:0}$ [#]	$dC_{18:2}$ [&]
gA^{Trp}	600 ± 140	570 ± 57	3100 ± 150
gA^{mTrp}	–	2200 ± 220	–
gA^{Tyr}	110 ± 15	80 ± 8	520 ± 140
gA^{Phe}	670 ± 100	330 ± 30	3400 ± 760

[#] From Table 2 of Girshman et al. (18)
[&] From Table 2 of Fonseca et al. (64) (gA^{mTrp} data from Table 3 of Sun et al. (62))

Figure 4.1: Positions of the Trp residues in gramicidin A channels and the side chain structures that were examined. (A1) Side and (A2) top views of equilibrated gA^{Trp} (initial PDB: 1JNO) (94). The Trp locations are shown in yellow with the indole nitrogen in blue. Experimental data exist for (B1) tryptophan (B2) 1-methyltryptophan, (B3) tyrosine, and (B4) phenylalanine. The simulations in this article also use (B5) glutamine, (B6) leucine, and the alchemical 1-methyltryptophan without charges (same structure as B2).

MD simulations were performed on channels where experimental data exist (gA^{Trp}, gA^{mTrp}, gA^{Tyr}, gA^{Phe}), as well as: i) an alchemical 1-methyltryptophan without side-chain charges ($gA^{nc\text{-}mTrp}$); ii) glutamine (gA^{Gln}); and iii) leucine (gA^{Leu}). gA^{Trp}, gA^{mTrp}, and $gA^{nc\text{-}mTrp}$ were selected to study the roles of the indole hydrogen bond donor site (which is eliminated in mTrp) and the Trp/mTrp side chain dipole (which is nearly identical for mTrp and Trp (62), but absent in nc-mTrp). Similarly, the Tyr/Phe and Gln/Leu pairs are structurally related, but differ in their side chain dipoles and ability to form hydrogen bonds.

The influence of Trp mutations is described in terms of i) the structure and dynamics of gA bilayer-spanning dimers (the root mean squared deviations (RMSD) and fluctuations (RMSF), gA tilt, hydrogen bonding and interaction patterns, and Trp and mutant residues' side chain orientations); ii) bilayer hydrophobic thickness (and decomposition) profiles as functions of radial distance from the channel; and iii) the relationship between lipid traces, the per area free energy of bending with respect to curvature, and monomer z-positioning. These results are compared to experiment when available.

4.2   Methods

4.2.1   System setup

All systems were built, minimized, and partially equilibrated using the *Membrane Builder* module (103, 104) of CHARMM-GUI (www.charmm-gui.org) (105). The dimer gA^{Trp} structure (PDB:1JNO) was used for the gA^{Trp} simulations as well as the scaffold for Trp mutations (94). The topology and parameter information for gA^{mTrp}, not available in the standard CHARMM force field (35, 109–111), was obtained from the GAAMP gateway (http://gaamp.lcrc.anl.gov) (158). See **Table 4.2** and **Table S4.1** for system information.

Table 4.2: Lipid type and channel combinations used in this study. Phosphatidylcholine (PC) lipids are denoted by their tail type, e.g., $dC_{18:1}$ (1,2-di-oleoyl-PC, DOPC), $dC_{20:1}$ (1,2-di-eicosenoyl-PC), and $dC_{22:1}$ (1,2-di-erucoyl-PC). "Y" denotes simulations were performed for this combination; "–" denotes that no simulations were performed.

Mutant \ Lipid	$dC_{18:1}$	$dC_{20:1}$	$dC_{22:1}$
gA^{Trp}	Y	Y	Y
gA^{mTrp}	Y	Y	Y
$gA^{nc-mTrp}$	Y	Y	Y
gA^{Tyr}	Y	–	Y
gA^{Phe}	Y	–	Y
gA^{Gln}	Y	–	–
gA^{Leu}	Y	–	–

The first set of simulations contained 90 $dC_{18:1}$ lipids per leaflet and 0.15 M KCl. Three independent replicas for each mutation were simulated for 220 ns using NAMD (51) in the isothermal-isobaric (NPT) ensemble. Constant temperature was maintained at 303.15 K using Langevin dynamics with a Langevin coupling coefficient of 1 ps^{-1}. Constant pressure was maintained at 1 atm by a Nosé-Hoover Langevin piston (107, 108) with a piston period of 50 fs and a piston decay of 25 fs. The CHARMM all-atom protein force field (109) including dCMAP (110, 111) was used with the C36 lipid force field (35) and TIP3P water model (112). The SHAKE algorithm (113) was used with a 2-fs time step. Electrostatic interactions were calculated using the particle-mesh Ewald method (114) (mesh size ~1 Å, $\kappa = 0.34$ $Å^{-1}$, and sixth-order B-spline interpolation), and van der Waals interactions were switched off between 10–12 Å by a force-switching function (115).

The second set of simulations were performed with 90 $dC_{20:1}$ or $dC_{22:1}$ lipids per leaflet. The gA^{Trp} data were previously published (9), and gA^{mTrp}, $gA^{nc-mTrp}$, gA^{Tyr}, and gA^{Phe} simulations were performed with the following conditions: i) harmonic RMSD restraints on the dimer and monomers, ii) xy positional restraints to keep monomers at maximum separation, iii) 1 fs time step, and iv) constant temperature of 310.15 K. The other simulation parameters are the same as

in the above simulations with $dC_{18:1}$. Three replicas of monomer-containing and dimer-containing systems were run for 100–170 ns apiece.

4.2.2 Constructing lipid traces

To better understand the lipid conformational changes imposed by the channel, average lipid traces were calculated (9). The trace is a radial property, which assuming cylindrical symmetry, describes the average position of lipid as function of distance (r) from the center of the channel. Briefly, a histogram in r (bin width of 0.06 Å) is formed based on the coordinates of the head group through *sn-2* tail atoms of all lipids. For each radial bin, there is a density, $\rho_i(r)$, and an average height, z_i, for each atom type i. A value r_0 is calculated for the first trace (i.e., lipids closest to the channel), such that:

$$\int_0^{r_0} \rho_i(r) \, dr = n, \tag{4.3}$$

where n is a user-defined, targeted number of atom type i (e.g., $n = 3.5$, which is same value used in Sodt et al. (9)). The space between traces is based on n, so traces are spaced closer at large r compared to traces near the channel (i.e., the number of lipids per bin goes as $2\pi(r_b - r_{b-1})$, where b is some arbitrary bin). Once r_0 is known, the weighted average radial position, $\langle r_i \rangle$, and weighted average height, $\langle z_i \rangle$, of atom type i between 0 and r_0 are known:

$$n^{-1} \int_0^{r_0} \rho_i(r)\{r, z_i(r)\} \, dr = \{\langle r_i \rangle, \langle z_i \rangle\}. \tag{4.4}$$

The first radial trace is constructed with these $\{\langle r_i \rangle, \langle z_i \rangle\}$ pairs by connecting them as they are chemically from the head group through the *sn-2* lipid tail. The second trace uses r_0 as a lower limit, and the next upper limit is calculated for the next n lipids by Equation 3. This process is repeated until enough traces have been constructed to provide a clear, albeit qualitative, view of the lipid bending and tilt as a function of r. These traces should not be confused with radial lipid shells, which have been previously defined by Voronoi tessellations (48). Although the spacing

between traces (dependent on n) presented herein is arbitrary, the traces could be formed by shell. However, the radial extent of lipid shells is large enough to wash out some details associated with the finely binned traces.

Also note that traces do not reach $z = 0$ because each atom position in the trace is based on the average, and not the most likely position. **Figure S4.2**, produced from simulations of lipid-only dC$_{22:1}$ (data from Sodt et al. (9)), demonstrates why the traces do not reach $z = 0$. Although the lipid terminal carbon atoms are mostly positioned around $z = 0$ Å, there are shoulders on the probability distributions (where the terminal carbon atoms approach the head groups due to tail entropy). These shoulders raise the unsigned weighted average position of the terminal carbon atoms to $z \approx 3.5$ Å. This matches well with the z-position where the traces end when the lipids are in the effective bulk (e.g., see **Figure 4.6**; it is assumed that lipids at large r behave similarly to how they would in a lipid-only system).

4.2.3   Leaflet per area free energy change with respect to curvature, $\overline{F}'(0)$

Traces provide qualitative insight into the leaflet bending frustration; the frustration can be quantified by the leaflet per-area free energy change with respect to curvature (at zero curvature), $\overline{F}'(0)$, which is calculated by $(8, 9, 37, 54, 159)$:

$$\overline{F}'(0) = \left.\frac{d\overline{F}}{dR^{-1}}\right|_{R^{-1}=0} = -\int_0^\infty z[p_L(z) - p_N(z)] \, dz, \tag{4.5}$$

where $\overline{F}$ is the per-area Helfrich bending energy, R^{-1} is leaflet curvature at the pivotal plane, and the integrand describes the pressure within the leaflet (where $p_L(z)$ and $p_N(z)$ are the lateral and normal components of the pressure tensor, respectively). Because the leaflets are constrained to be planar by hydrophobic interactions and periodic boundary conditions, $\overline{F}'(0)$ is evaluated at $R^{-1} = 0$. As defined, $\overline{F}'(0) = 0$ means that there is no bending frustration, and the magnitude of

$\bar{F}'(0)$ provides information about the bending frustration within a leaflet. By convention, a positive $\bar{F}'(0)$ indicates a leaflet would bend toward its head groups if it was unconstrained (i.e., a negative curvature).

In bilayers that are thicker than the dimer channel, bending frustration includes three contributions. First, the lipids considered in this study have negative intrinsic curvature (i.e., positive $\bar{F}'(0)$), so forcing these lipids to be in a planar bilayer creates leaflet frustration. Second, inserting a gA monomer into a leaflet could alleviate or exacerbate the leaflet frustration (being a physical contribution to the leaflet frustration, $\bar{F}'_{\text{m}}(0)$). Last, inserting a dimer introduces a physical contribution from the two monomers, as well as a contribution from the dimerization event itself, $\bar{F}'_{\text{d}}(0)$. By simulating independent monomers (one per leaflet) and independent dimers (one per bilayer) at the same channel/lipid ratio, we can calculate leaflet $\bar{F}'_{\text{m}}(0)$ and $\bar{F}'_{\text{d}}(0)$, respectively. The difference between these values, $\Delta\bar{F}'(0) = \bar{F}'_{\text{d}}(0) - \bar{F}'_{\text{m}}(0)$, is the leaflet bending frustration due to dimerization.

4.3   Results & discussion

In this section, gA^{Trp} and mutant channels are first considered in $dC_{18:1}$ lipids bilayers (**Table 4.2**). Residue orientation (χ_1-χ_2 distributions) is described as a function of size and geometry. Then, bilayer hydrophobic thickness profiles are presented and decomposed to describe how bilayer deformations are primarily dependent on a residue's ability to form hydrogen bonds to adjacent lipids and secondarily on the residue's hydrophobicity. Finally, using simulations of a subset of channel mutants in $dC_{20:1}$ and $dC_{22:1}$ bilayers, lipid traces and the per area free energy of bending with respect to curvature are described and discussed.

4.3.1 gA channel dynamics are affected by mutation

When examined in $dC_{18:1}$ bilayers, all channels were stable for the duration of the simulations, as evident from the per-residue root-mean-square fluctuations (RMSF: **Figure S4.3**) and root-mean-square deviations (RMSD) with respect to the minimized initial structure (**Figure S4.4**). However, as expected (e.g., Ingólfsson et al. (110)), there are variations in the RMSF at the N-termini of all channels (particularly in gA^{Gln} and gA^{Leu}) as well as some transient structural variations (measured by RMSD) in the mutants. Channel tilt is also affected by mutation (see **Figure S4.5** and **Table S4.2** for the tilt distributions and average tilt angle, respectively). As expected, residues that can form hydrogen bonds with the bilayer have smaller tilt angles than their counterparts that cannot form hydrogen bonds, implying that these interactions are important for positioning/anchoring gA in the bilayer. Comparing gA^{mTrp} and $gA^{nc-mTrp}$ further suggests that the side chain dipole also helps orient/stabilize the channel.

Related residues (Trp, mTrp, and nc-mTrp; Tyr and Phe; Gln and Leu) have nearly identical major χ_1-χ_2 populations (some minor populations differ among related residues). For example, although Trp, mTrp, and nc-mTrp have different hydrophobicity and ability to form hydrogen bonds, their χ_1-χ_2 population distributions are nearly identical (**Figure 4.2**, in agreement with Sun et al. (62)). Tyr and Phe also have very similar χ_1-χ_2 distributions, as do Gln and Leu (**Figure S4.6** and **S4.7**, respectively). These distributions suggest that residue size and geometry affect preferred rotamer states more than hydrophobicity or hydrogen bond formation. To provide simple insight into the rotamer energetics, the energy differences between Trp9 rotamer populations are calculated (assuming the Trp9 per monomer sufficiently sample χ_1-χ_2 space; see **Figure S4.8** for their time series): $\Delta E = k_B T \ln(p_i/p_j)$, where and p_i and p_j are the probabilities

of being in state i or j, respectively. The ΔE values of Trp in $dC_{18:1}$ range from -1.8 kcal/mol $(E_{II} - E_{I})$ to -0.15 kcal/mol $(E_{III} - E_{IV})$.
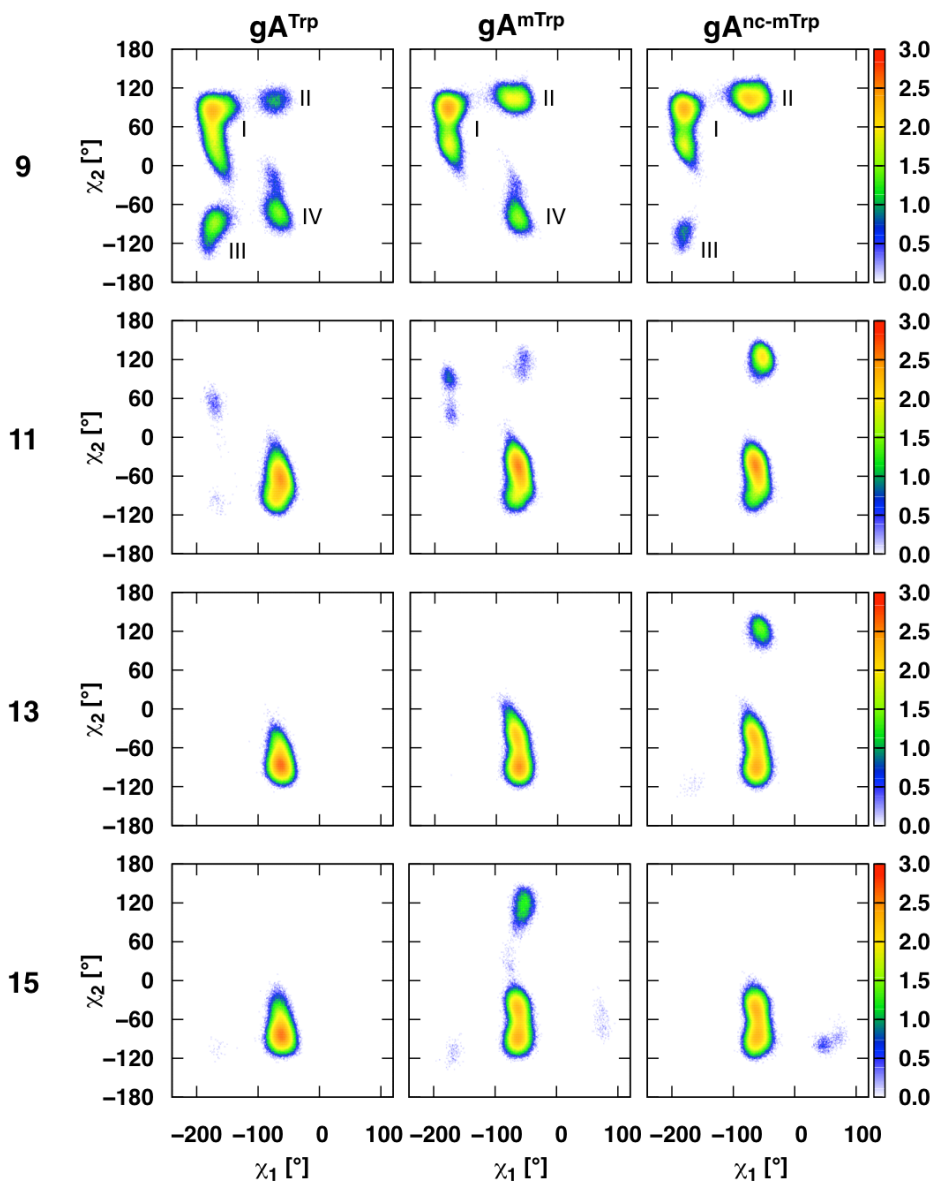
Figure 4.2: Trp, mTrp, and nc-mTrp χ_1-χ_2 dihedral angles in $dC_{18:1}$. χ_1 is the dihedral of the backbone N, Cα, Cβ, and Cγ atoms. χ_2 is the dihedral of the Cα, Cβ, Cγ, and Cδ atoms (Cδ is double bonded to Cγ and bonded to indole N). The color scheme for the heat plots is shown on the right with log{count/bin} and 1° bins in both dimensions.

4.3.2 dC$_{18:1}$ lipid slippage is affected by channel mutation

Previous MD simulations have shown that lipids tend to slip over the top of the channel, which releases some of the bilayer's frustration caused by hydrophobic matching with the channel (9, 48, 157, 160, 161). The number of lipids that slip over the top of the channel is linked to the interaction strength between the channel and adjacent lipids. This channel-bilayer coupling was explored using two complementary methods – radial distribution functions (**Figure S4.9**) and lipid contacts with the channel (**Figure S4.10**).

Radial distribution functions demonstrate that lipid acyl chains are restrained to stay outside of $r \approx 8$–10 Å (the approximate channel radius). Channels with residues capable of forming hydrogen bonds (Trp, Tyr, Gln) have low choline densities near $r = 0$ Å. Residues that cannot form hydrogen bonds (mTrp, nc-mTrp, Phe, and Leu) have high densities near $r = 0$ Å, demonstrating that the extent of lipid slippage depends on the interaction strength between the channel and the backbone of the adjacent lipids (**Table S4.3**). Next, we assessed the lipid components (acyl chain, carbonyl, phosphate, or choline) and water that were within 4 Å of any channel side chain heavy atom (**Figure S4.10**). All residues that can form hydrogen bonds (Trp, Tyr, and Gln) have more contact with the carbonyl and phosphate groups, as well as the choline (because the lipids are more restricted in z due to hydrogen bonding), as compared to residues that cannot form hydrogen bonds. Additionally, residues that cannot form hydrogen bonds have higher contact incidence with the lipid tails because of lipid slippage (allowing these more hydrophobic residues to be buried in lipid tails).

The observation that different residues prefer to be close to certain lipid chemical features, is corroborated by the frequency of hydrogen bond formation for gA^{Trp}, gA^{Tyr} and gA^{Gln} channels in dC$_{18:1}$ (**Figure 4.3**). The average fraction of time that a given residue formed hydrogen bonds

was decomposed into the contributions from the lipid carbonyl, lipid phosphate, water, and channel backbone. Trp interacts preferentially with lipid and Tyr with water. Gln stands out by its interactions with the channel backbone, as all Gln residues form occasional hydrogen bonds to backbone oxygen of adjacent residues. The Gln9 and Gln15 side chains furthermore form two relatively strong reciprocal hydrogen bonds with each other (N–H···O=C and vice versa). These distributions provide insight into how Trp and Tyr help stabilize the channel through the formation of channel-lipid hydrogen-bond interactions.
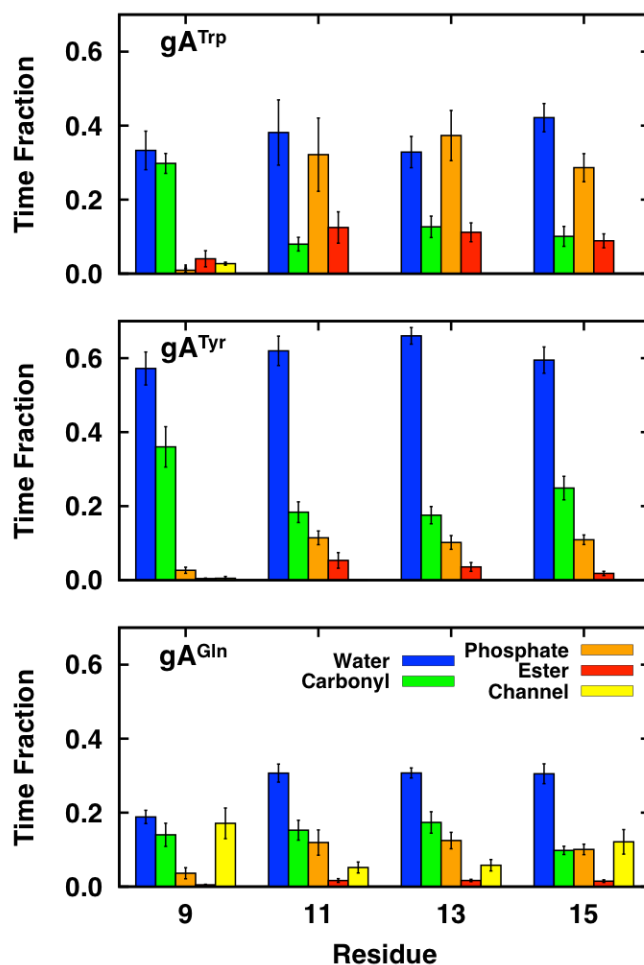

Figure 4.3: Fraction of time that gA^{Trp}, gA^{Tyr}, and gA^{Gln} form hydrogen bonds to $dC_{18:1}$. A hydrogen bond is defined as a donor and acceptor pair within 2.4 Å of each other (with no angular cutoff). The fractions do not sum to unity because the residues do not form hydrogen bonds at all times.

4.3.3 Adjacent lipid compression is a function of interfacial residues

Radial bilayer hydrophobic thickness profiles, $d_H(r)$, determined from the average z locations of the C22 and C32 lipid tail carbon atoms (i.e., the carbon atoms bonded to the lipid carbonyl group), provide insight into the bilayer deformation caused by the channel. The profiles for gA^{Trp}, gA^{Tyr} and gA^{Gln} are similar, and distinct from the profiles around channels that cannot form hydrogen bonds with lipids. Indicating that lipids are constrained to match residues that can form hydrogen bonds, as compared to residues that cannot form hydrogen bonds. This is evidenced by deeper minima at the edge of the channel ($r \approx 10$–12 Å) for gA^{Trp}, gA^{Tyr}, and gA^{Gln} compared to gA^{mTrp}, $gA^{nc\text{-}mTrp}$, gA^{Phe}, and gA^{Leu} (i.e., the steep increase in $d_H(r)$ at $r < 10$ Å arises because lipids slip over the top of the channel; **Figure 4.4**).

The difference in minima depths between related residues (Trp, mTrp, and nc-mTrp; Tyr and Phe; Gln and Leu) qualitatively describes the leaflet frustration due to these residues. As discussed previously (9, 48), the main energetic factor directly adjacent to the channel is lipid compression (curvature frustration to be discussed later in this article). Here, we demonstrate that hydrogen bond formation appears to be the major factor in determining channel-lipid coupling (e.g., compare Trp and mTrp minima), but hydrophobicity also plays a role in the deformation (e.g., compare mTrp and nc-mTrp minima).
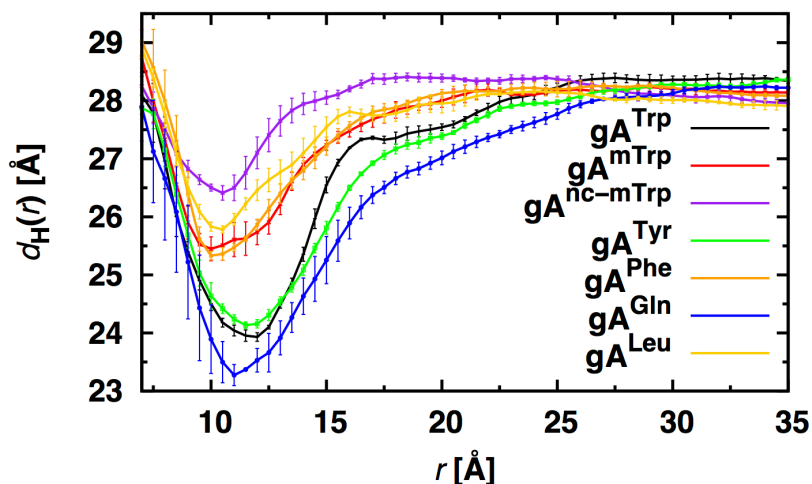
Figure 4.4: Radial hydrophobic thickness profiles of $dC_{18:1}$ bilayers with different embedded gA analogues.

Because the ability to form hydrogen bonds appears to be the key factor in deforming the bilayer adjacent to the channel, we decomposed $d_H(r)$ into contributions from lipids that do and do not form hydrogen bonds with Trp, Tyr or Gln. If a lipid forms hydrogen bonds with a target residue, it is placed in the appropriate radial bin for the "H-bond" group; otherwise, it is placed in the "Free" group (**Figure 4.5**).

The profiles for the H-bond lipids (blue curves) are similar in that $d(r) \approx 23$ Å, close to the canonical hydrophobic length of gA^{Trp} channels (66, 99, 134) at $r \approx 10$–12 Å. Because the channel and lipids are able to tilt, channel-lipid hydrogen bonds can be formed at large radial distances, as shown in the frequency distribution for hydrogen bond formation (in grey). If the channel or lipids do tilt to form hydrogen bonds, the lipids appear to be thinner, which is the case at larger r (blue curves). The "Free" lipids have a biphasic thickness profile, with a minimum at the edge of the channel, which most likely due to the hydrophobic matching requirement (to the channel, to the H-bond lipids, or to both).
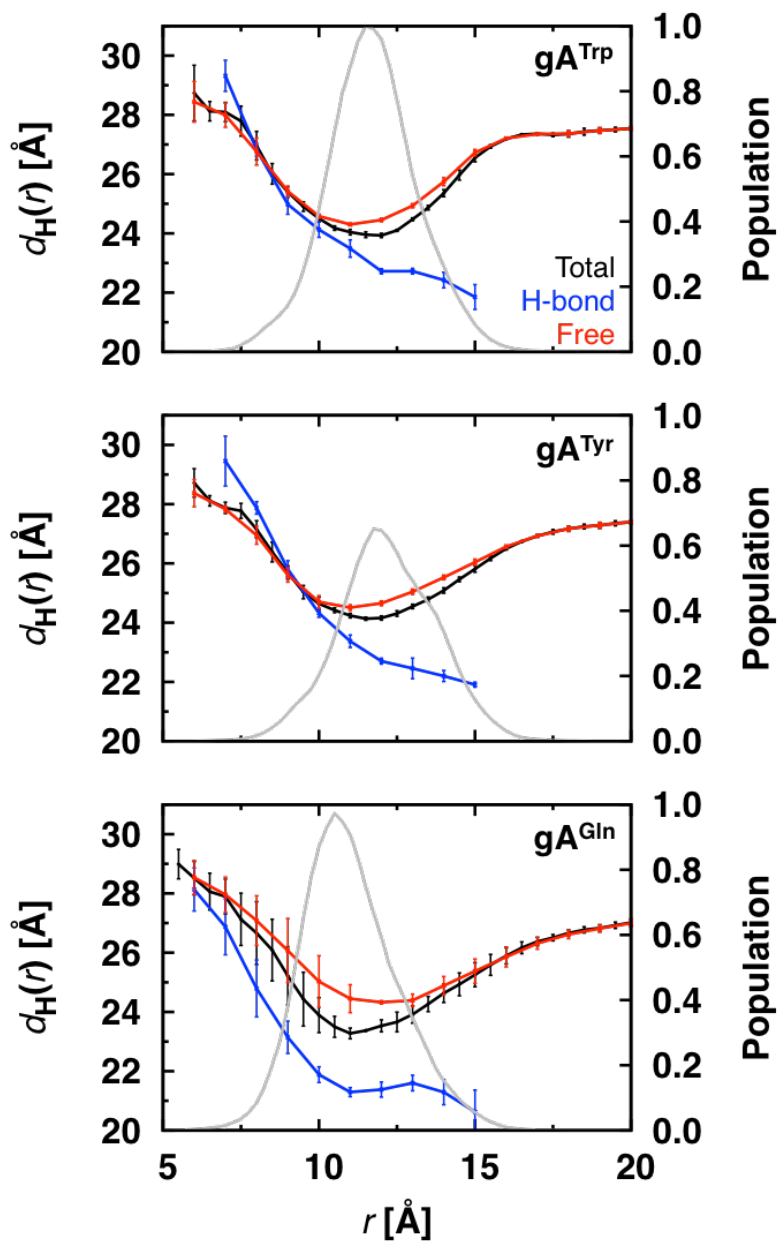
Figure 4.5: Decomposition of the radial hydrophobic thickness profiles for gA^{Trp}, gA^{Tyr}, and gA^{Gln} channels in $dC_{18:1}$. The total hydrophobic thickness profiles from **Figure 4** are decomposed into profiles from lipids that form hydrogen bonds to the channel (H-bond population) and lipids that do not form hydrogen bonds with the channel (Free population). In grey, the radial distribution of lipids that form hydrogen bonds to the channel (plots are normalized to the peak of the gA^{Trp} distribution).

With access to atomic detail, the differences in the hydrophobic thickness profile minima can clearly be attributed to hydrogen bonding between gA channels and surrounding bilayers. As worked out explicitly in the **Experimental Discussion** of the **Supplemental Information**, effective channel hydrophobic lengths can be used to explain differences in mean channel lifetimes. That is, an effectively "longer" channel would have a longer lifetime in thick bilayers. Ideally, the differences in hydrophobic thickness minima near the channel could be directly related to the channel's effective hydrophobic length, and therefore, the channel lifetime. Based on the lifetimes of gA^{Tyr} and gA^{Phe} in dC$_{18:1}$ (**Table 4.1**), one would predict ~3 Å difference in effective channel length. However, **Figure 4.4** shows a smaller difference in apparent hydrophobic thickness between gA^{Tyr} and gA^{Phe} (~1.25 Å), suggesting that more energetic factors are at play than simple compression stress. So, more energetic insight is necessary to explain $\Delta G_{bilayer}^{M \rightarrow D}$.

4.3.4 dC$_{20:1}$ and dC$_{22:1}$ bilayer energetics are affected by channel mutation

To gain energetic insight into the curvature frustration due to hydrophobic mismatch and interfacial residue influence, we calculated the per area free energy change with respect to curvature, $\bar{F}'(0)$ (see **Methods** for more details). These $\bar{F}'(0)$ calculations were performed in dC$_{20:1}$ and dC$_{22:1}$, to increase the signal-to-noise associated with these calculations; results for dC$_{18:1}$ are expected to follow the same trend, but are unfeasible due to the long simulations necessary to have acceptable uncertainties for comparison across channels (9). Note that, as mentioned in the **Methods**, an RMSD restraint was placed on all channel backbones to eliminate structural changes that are possible at large channel-bilayer hydrophobic mismatch (18, 61, 162), but the RMSD restraint applied in this study does not drastically alter Trp χ_1-χ_2 dynamics in dC$_{18:1}$ (**Figure S4.11**). **Table 4.2** summarizes $\bar{F}'(0)$ estimates for gA^{mTrp} and gA$^{nc-mTrp}$ in dC$_{20:1}$

and $dC_{22:1}$ together with estimates for gA^{Trp} from previous work (9), as well as results for gA^{Tyr} and gA^{Phe} in $dC_{22:1}$.

Lipid traces were calculated for gA dimers and monomers in $dC_{22:1}$ (lipid traces were qualitatively similar in $dC_{20:1}$, data not shown). All traces near the dimers have a similar shape, but there appears to be some lipid tilt at large r (i.e., the traces around the dimers gA^{Trp} are tilted relative to those around gA^{mTrp}). We did not pursue this issue of long-range lipid tilt produced by channel formation. There are more apparent changes in the traces directly adjacent to the monomers, but these differences are not propagated as far in r. Also, lipids form basket-like conformations as the tails go underneath the monomers, which has been reported for other peptides as well (159).
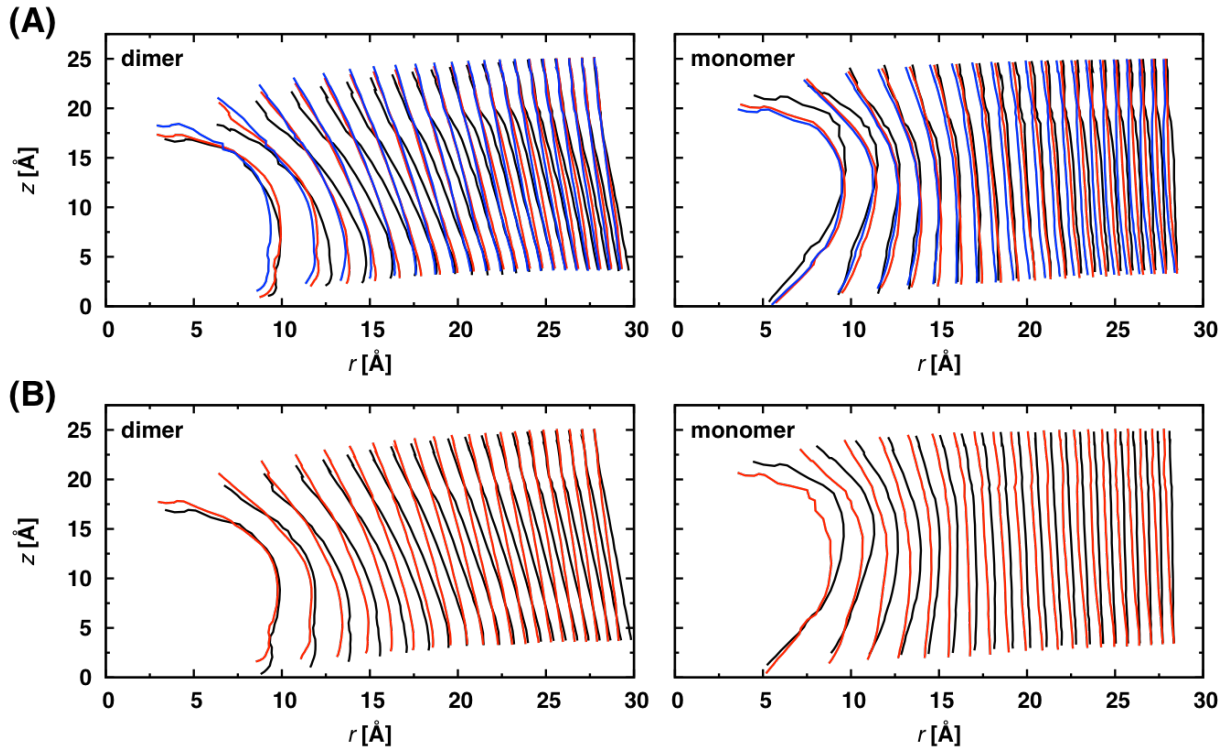

Figure 4.6: $dC_{22:1}$ lipid traces around dimers and monomers: (A) gA^{Trp} (black), gA^{mTrp} (red), and $gA^{nc\text{-}mTrp}$ (blue); (B) gA^{Tyr} (black) and gA^{Phe} (red). The x-axis is extended relative to the y-axis to make the differences between the traces clearer.

While traces provide information on the lipid conformation, the information is qualitative. Quantitative information is obtained by the per area free energy with respect to curvature, $\bar{F}'(0)$. Using the procedure detailed in the **Methods** section, $\bar{F}'_m(0)$, $\bar{F}'_d(0)$, and $\Delta\bar{F}'(0)$ were calculated and presented in **Table 4.3**. The value of $\Delta\bar{F}'(0)$ describes the leaflet curvature frustration due to a dimerization event, and such, directly tied to $\Delta G_{\text{bilayer}}^{\text{M}\rightarrow\text{D}}$ and the channel lifetime. $\Delta\bar{F}'(0)$, varies as $gA^{\text{Trp}} \approx gA^{\text{Tyr}} > gA^{\text{mTrp}} \approx gA^{\text{Phe}} > gA^{\text{nc-mTrp}}$ in $dC_{22:1}$ (and $gA^{\text{Trp}} > gA^{\text{mTrp}} \approx gA^{\text{nc-mTrp}}$ in $dC_{20:1}$). As expected, the $\Delta\bar{F}'(0)$ is larger in $dC_{22:1}$ than in $dC_{20:1}$ for all channels due to hydrophobic matching considerations (9). In either membrane environment, removing an interfacial residue's ability to form hydrogen bonds reduces the leaflet curvature frustration. In $dC_{22:1}$, the curvature frustration $gA^{\text{Trp}} > gA^{\text{mTrp}} > gA^{\text{nc-mTrp}}$ suggests that hydrogen bond formation and hydrophobic interactions are important. Comparing with a residue family (Trp, mTrp; Tyr and Phe), we see that monolayer curvature frustration is linked to channel lifetime. It is energetically more costly for gA^{Trp} to remain as a dimer than gA^{mTrp}, and the same is true between gA^{Tyr} and gA^{Phe}.

Table 4.3: Leaflet curvature frustration $\bar{F}'(0)$ (in kcal/mol/Å) for gA^{Trp}, gA^{mTrp}, $gA^{\text{nc-mTrp}}$, gA^{Tyr}, and gA^{Phe} monomers and dimers in $dC_{20:1}$ and $dC_{22:1}$. The $\bar{F}'(0)$ are 0.0578 ± 0.0046 kcal/mol/Å and 0.0558 ± 0.0030 kcal/mol/Å for pure $dC_{20:1}$ and pure $dC_{22:1}$, respectively (9).

		gA^{Trp}	gA^{mTrp}	$gA^{\text{nc-mTrp}}$	gA^{Tyr}	gA^{Phe}
	$\bar{F}'_m(0)$	0.0786 ± 0.0037	0.0838 ± 0.0018	0.0865 ± 0.0025	–	–
$dC_{20:1}$	$\bar{F}'_d(0)$	0.0927 ± 0.0031	0.0919 ± 0.0017	0.0927 ± 0.0018	–	–
	$\Delta\bar{F}'(0)$	$\mathbf{0.0141 \pm 0.0048}$	$\mathbf{0.0081 \pm 0.0024}$	$\mathbf{0.0062 \pm 0.0031}$	–	–
	$\bar{F}'_m(0)$	0.0685 ± 0.0017	0.0765 ± 0.0025	0.0830 ± 0.0026	0.0622 ± 0.0024	0.0716 ± 0.0024
$dC_{22:1}$	$\bar{F}'_d(0)$	0.1169 ± 0.0046	0.1104 ± 0.0021	0.1099 ± 0.0020	0.1081 ± 0.0023	0.1035 ± 0.0025
	$\Delta\bar{F}'(0)$	$\mathbf{0.0484 \pm 0.0049}$	$\mathbf{0.0340 \pm 0.0032}$	$\mathbf{0.0269 \pm 0.0033}$	$\mathbf{0.0459 \pm 0.0033}$	$\mathbf{0.0319 \pm 0.0035}$

The major changes in $\Delta\bar{F}'(0)$ arise from the changes in $\bar{F}'_m(0)$. Because the location of peptides along the z-axis influences the value of $\bar{F}'(0)$ (8, 159), we constructed symmetrized z-density plots to determine gA's preferred positioning in the bilayer (**Figure S4.12** for $dC_{20:1}$ and **Figure 4.7** for $dC_{22:1}$). Here, gA monomer can be thought of as a wedge, whose position relative to the lipid pivotal plane (approximately the lipid C22 atom) affects the preferred bending of the lipid leaflet (see **Figure 4.8**). The monomers studied here have high density below the pivotal plane, meaning that they all induce negative curvature frustration (i.e., a more positive $\bar{F}'_m(0)$ compared to the lipid-only value), and the insertion depth correlates to the value of $\bar{F}'_m(0)$. For example, gA^{Trp} monomers frustrate the leaflets less than gA^{mTrp} monomers (**Table 4.3** and **Figures 4.7** & **4.8**).
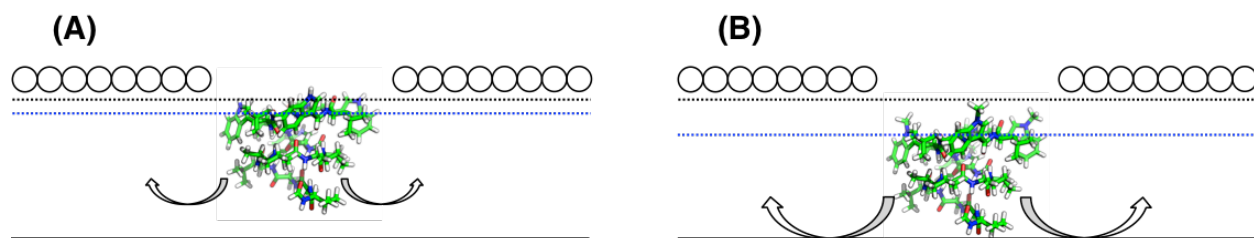

Figure 4.7: Illustration of (A) gA^{Trp} and (B) gA^{mTrp} monomers in a leaflet in which there is a large channel-bilayer hydrophobic mismatch (e.g., $dC_{22:1}$). The monolayer pivotal plane is represented by the dotted line, the bilayer center (i.e., $z = 0$) is the solid line, and the lipid head groups are circles. gA^{Trp} resides closer to the lipid head groups and has more density above the pivotal plane, while gA^{mTrp} has more density below the pivotal plane. When there is more gA density below the pivotal plane, the lipids will more strongly prefer to bend toward the head groups. The extent of the bending will be constrained by the opposing leaflet because the two leaflets are coupled by hydrophobic constraints, but a curvature frustration will be produced. The magnitude of frustration will depend on the position of the monomer in its leaflet. Note that this, too, is an approximation because if the intrinsic curvature in the upper leaflet were high enough it would also affect the lower leaflet; also see Phillips et al. (163) Although not observed in this article, theoretically a monomer could position itself so that it introduced no frustration to the monolayer, i.e., $\bar{F}'_m(0) = \bar{F}'_{pure\ lipid}(0)$. Similarly, a monomer could position itself with high density above the pivotal plane and relax some frustration within the leaflet.
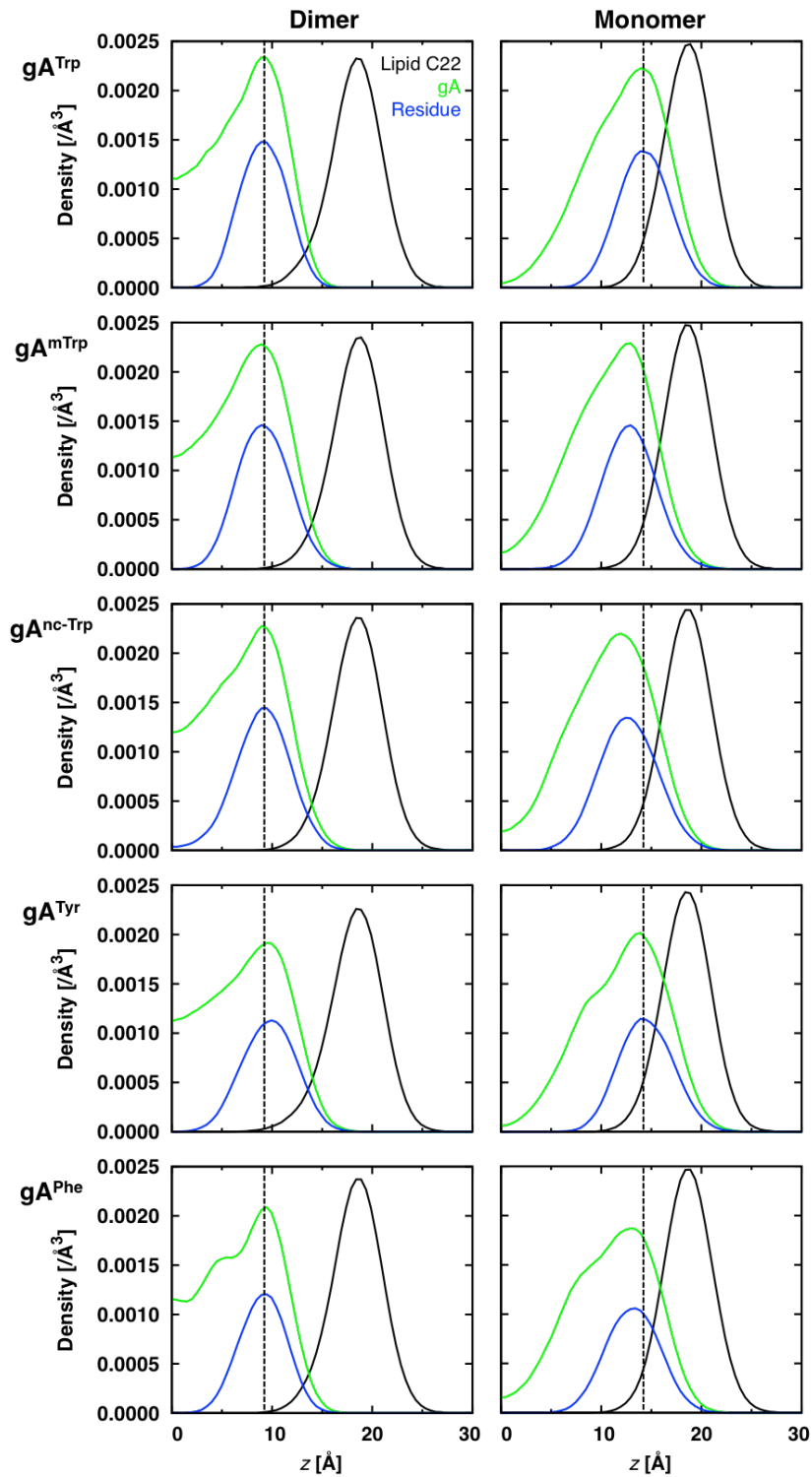
Figure 4.8: Heavy-atom *z*-density plots for the lipid C22 atom (approximate pivotal plane position), entire channel, and the specific residues in $dC_{22:1}$. Dotted black lines are shown to accentuate the peak shifts relative to gA^{Trp}. Data is plotted in 0.5 Å bins. Systems were centered by shifting the bilayer's center of mass to $z = 0$ Å.

The monomer z-positions vary with side chain preference for the interface relative to the bilayer core, which can be evaluated using a suitable free energy scale (164). For example, Tyr prefers the interface over the bilayer core (165), as evident by the large amount of hydrogen bonding to water (**Figure 4.3**). Additionally, Trp similarly prefers to be at the interface, whereas mTrp and nc-mTrp prefer to be in the core. Therefore, gA^{mTrp}, $gA^{nc\text{-}mTrp}$, and gA^{Phe} monomers are shifted closer to $z = 0$ compared to gA^{Trp} and gA^{Tyr} monomers. Hydrophobicity affects monomer position in the bilayer, which affects the leaflet bending frustration.

4.4 Conclusions

Experimentally, mutating the four Trp residues of gA^{Trp} changes the mean channel lifetime, i.e., the channel function. Using MD simulations we have shown dimer backbone structure (RMSD) and dynamics (RMSF) are mostly unchanged by mutation in $dC_{18:1}$ bilayers on a typical MD timescale (~200 ns). Each of a residue's key characteristics (size/geometry, ability to form hydrogen bonds, and hydrophobicity) directly affects characteristics of the dimer and monomer.

On a dimer, a residue at a particular location (9, 11, 13, 15) has its orientation determined by its size and geometry, as evidenced in χ_1-χ_2 plots. Whereas, a residue's ability to form hydrogen bonds has direct correlation to the bilayer deformation caused by the channel. Trp, Tyr, and Gln perturb $dC_{18:1}$ bilayers similarly, while mTrp, Phe, and Leu perturb the bilayer less than their counterparts that can form hydrogen bonds with lipids. Hydrophobicity impacts bilayer deformations as well, but does not seem to have as large of effect as the ability to hydrogen bond. For example, the hydrophobic nc-mTrp perturbs the bilayer less than mTrp, but the perturbation difference is less than comparing Trp to mTrp. Indeed, decomposition of radial bilayer hydrophobic thickness profiles demonstrates that residue-lipid hydrogen bonds play a large role in the overall deformation. Since a channel's lifetime is determined in part by $\Delta G_{bilayer}^{M \rightarrow D}$,

which involves the compression energy, we attribute residue-lipid hydrogen bond formation as part of the differences between gA^{Trp}/gA^{mTrp} and gA^{Tyr}/gA^{Phe} lifetimes. This is further discussed by a theoretical treatment of the effective channel lengths.

Overall, there was a distinct trend in the leaflet curvature frustration due to dimerization, $\Delta \bar{F}'(0)$. Residues that can form hydrogen bonds frustration leaflets more than those that cannot. Furthermore, residue hydrophobicity plays a role in determining curvature frustration (gA^{mTrp} compared with $gA^{nc\text{-}mTrp}$). As was the case with compression contributions, curvature frustration also contributes to $\Delta G_{bilayer}^{M \to D}$. Indeed, within a residue family (Trp and mTrp; Tyr and Phe), the $\Delta \bar{F}'(0)$ correlate well with experimental channel lifetimes.

Although $\bar{F}'_d(0)$ was equal within error for all mutant dimers, there are differences in $\bar{F}'_m(0)$ between the mutant monomers due to lipid conformational differences near the channel (including basket-formations). Monomer insertion was understood by considering residue hydrophobicity. gA monomers with hydrophobic residues (nc-mTrp > Phe ≈ mTrp) embed deeper in their leaflet compared to Trp and Tyr. The more monomer density there was below the lipid pivotal plane, the more curvature frustration was induced. These results were supported by z-density plots.

The results here were obtained on a simple channel, but the overall effects of mutations at the bilayer-water interface should apply generally to other membrane proteins. A mutation involving a change in residue size/geometry, ability to form hydrogen bonds with lipids, and/or hydrophobicity could change the protein's orientation, adjacent bilayer deformation, and/or preferred positioning relative to the bilayer core.

4.5   Supplemental information

4.5.1   Experimental discussion

At equilibrium, gA monomers associate to form dimers with rate constants k_1 and k_{-1}:

$$2M \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} D \tag{S4.1}$$

The monomer$\leftrightarrow$dimer equilibrium is described by:

$$\frac{[D]}{[M]^2} = \frac{k_1}{k_{-1}} = \exp\left\{-\frac{\Delta G_{protein}^{M \to D} + \Delta G_{bilayer}^{M \to D}}{k_B T}\right\} \tag{S4.2}$$

where k_B is Boltzmann's constant, T the temperature in Kelvin, and k_1 and k_{-1} the association and dissociation rate constants, respectively. $\Delta G_{protein}^{M \to D}$ encompasses the energy associated with forming/breaking the six inter-monomer hydrogen bonds, involving residues 1–5 at the dimer interface, and is assumed to be independent of the mutations studied here because amino acid substitutions in the C-terminal half of gA have minimal effects on the energetics of subunit dimerization (152). $\Delta G_{bilayer}^{M \to D}$ denotes the bilayer deformation energy associated with the formation of dimeric channels.

Because $\Delta G_{protein}^{M \to D}$ is considered to be constant, changes in the rate constants are determined by $\Delta G_{bilayer}^{M \to D}$. The bilayer free energy change due to a channel-caused deformation can be approximated as (48, 124):

$$\Delta G_{bilayer}^{M \to D}(u_0) = H(d_0 - l)^2 = H(2u_0)^2 \tag{S4.3}$$

where u_0 is the bilayer deformation equal to the difference in thickness between the unperturbed (d_0) lipid bulk and effective channel length (l). H is the phenomenological spring coefficient that describes the bilayer "stiffness."

Working from transition state theory, the dissociation rate constant (and therefore, the channel lifetime, τ) is related to the transition free energy $\Delta G^{\ddagger}$ (specifically, the activation energy due to a dimer reaching the dissociation transition state):

$$\ln(k_{-1}) = -\ln(\tau) = -\frac{\Delta G^{\ddagger}}{RT} - \ln(\tau_0) \tag{S4.4}$$

where R is the ideal gas constant and τ_0 is the pre-exponential coefficient. When the channel reaches the transition state during dissociation, the inter-monomer separation has increased by δ (generally accepted to be ~1.6 Å), which is associated with the initial steps of breaking the inter-monomer hydrogen bonds:

$$\Delta G^{\ddagger} \propto \Delta\Delta G^0_{\text{bilayer}} = H((2u_0 - \delta)^2 - (2u_0)^2) = H(4u_0 - \delta)\delta \tag{S4.5}$$

where $\Delta\Delta G^0_{\text{bilayer}}$ describes the bilayer energy difference when the channel is fully associated and when it is at δ separation. The $\Delta G^{\ddagger}$ might include energy from the channel, but those are ignored here.

Substituting Equation S4.5 into S4.4, and taking the derivative with respect to u_0 (124):

$$\frac{d(\ln(k_{-1}))}{du_0} = \frac{d(-\ln(\tau))}{du_0} = \frac{4H\delta}{RT} \tag{S4.6}$$

Equation A4.6 relates the channel lifetime to the bilayer deformation, bilayer stiffness and inter-monomer separation at the transition state. Assuming $H\delta$ is constant, Equation A4.6 can be integrated, leading to an expression that relates the effective channel lengths of two channel types to their mean lifetimes:

$$\frac{-[\ln(\tau_2) - \ln(\tau_1)]}{\frac{1}{2}[(d_0 - l_2) - (d_0 - l_1)]} = \frac{4H\delta}{RT} \tag{S4.7}$$

which can be rearranged to:

$$l_2 - l_1 = \frac{RT}{2H\delta}[\ln(\tau_2) - \ln(\tau_1)] \tag{S4.8}$$

99

Equation A4.8 states that the difference between the effectives lengths of two channels is related to the difference in lifetimes of the two channels. Therefore, using the previously published lifetimes in **Table 1**, the difference in effective channel lengths can be estimated. In $dC_{18:1}$:

$$l_{\text{Phe}} - l_{\text{Tyr}} = \frac{RT}{2H\delta}[\ln(0.670) - \ln(0.110)] \approx 3 \text{ Å} \tag{S4.9}$$

where $H\delta$ is defined to be 784 J/(mol·Å) (48). That is, gA^{Phe} has a longer effective hydrophobic length than gA^{Tyr}.

4.5.2 Supplemental tables

Table S4.1: System information. All systems have 90 lipids / leaflet. A "d" denotes a dimer simulation and an "m" denotes a monomer simulation. The gA^{Trp} simulations in dC$_{20:1}$ and dC$_{22:1}$ are taken from Sodt et al. (9).

Mutant	Lipid Type	ID	System Size$^\$$	Water	Total Atoms	Sim. Time [ns]
gA^{Trp} d	dC$_{18:1}$	1	79.7×79.7×70.0	6,802	45,835	220
		2	79.7×79.7×70.0	6,756	45,697	220
		3	79.7×79.7×70.0	6,743	45,658	220
	dC$_{20:1}$	1	80.0×80.0×86.0	8,148	52,050	110
	dC$_{22:1}$	1	80.0×80.0×87.0	8,063	53,957	110
gA^{Trp} m	dC$_{20:1}$	1	79.3×79.3×88.0	8,119	51,965	110
	dC$_{22:1}$	1	79.2×79.2×90.0	8,194	54,352	110
gA^{mTrp} d	dC$_{18:1}$	1	79.8×79.8×80.0	8,928	52,249	220
		2	79.8×79.8×80.0	9,000	52,465	220
		3	79.8×79.8×80.0	8,870	52,075	220
	dC$_{20:1}$	1	80.0×80.0×86.0	7,789	50,997	140
		2	80.0×80.0×86.0	7,774	50,952	170
		3	80.0×80.0×86.0	7,796	51,018	130
	dC$_{22:1}$	1	80.0×80.0×87.0	7,775	53,117	130
		2	80.0×80.0×87.0	7,789	53,159	130
		3	80.0×80.0×87.0	7,803	53,201	130
gA^{mTrp} m	dC$_{20:1}$	1	79.4×79.4×88.0	8,164	52,124	150
		2	79.4×79.4×88.0	8,170	52,142	130
		3	79.4×79.4×88.0	8,160	52,112	80
	dC$_{22:1}$	1	79.3×79.3×90.0	8,241	54,517	130
		2	79.3×79.3×90.0	8,217	54,445	130
		3	79.3×79.3×90.0	8,196	54,382	130
gA$^{nc-mTrp}$ d	dC$_{18:1}$	1	79.8×79.8×80.0	8,917	52,216	220
		2	79.8×79.8×80.0	8,887	52,126	220
		3	79.8×79.8×80.0	8,870	52,075	220
	dC$_{20:1}$	1	79.4×79.4×86.0	7,708	50,754	100
		2	79.4×79.4×86.0	7,689	50,697	100
		3	79.4×79.4×86.0	7,683	50,679	100
	dC$_{22:1}$	1	79.3×79.3×87.0	7,634	52,692	100
		2	79.3×79.3×87.0	7,651	52,743	100
		3	79.3×79.3×87.0	7,656	52,758	100
gA$^{nc-mTrp}$ m	dC$_{20:1}$	1	79.4×79.4×88.0	8,172	52,148	100
		2	79.4×79.4×88.0	8,170	52,142	100
		3	79.4×79.4×88.0	8,181	52,175	100
	dC$_{22:1}$	1	79.3×79.3×90.0	8,205	54,409	100
		2	79.3×79.3×90.0	8,216	54,442	100
		3	79.3×79.3×90.0	8,241	54,517	100
gA^{Tyr} d	dC$_{18:1}$	1	79.5×79.5×80.0	8,893	52,094	220
		2	79.5×79.5×80.0	8,835	51,920	220
		3	79.5×79.5×80.0	8,927	52,196	220
	dC$_{22:1}$	1	79.0x79.0x87.0	7,554	52,404	100

		2	79.0x79.0x87.0	7,568	52,446	100
		3	79.0x79.0x87.0	7,544	52,374	100
gA^{Tyr} m	dC$_{22:1}$	1	79.0x79.0x90.0	8,208	54,370	100
		2	79.0x79.0x90.0	8,208	54,370	100
		3	79.0x79.0x90.0	8,208	54,370	100
gA^{Phe} d	dC$_{18:1}$	1	79.5×79.5×80.0	8,889	52,074	220
		2	79.5×79.5×80.0	8,852	51,963	220
		3	79.5×79.5×80.0	8,852	51,963	220
	dC$_{22:1}$	1	79.0x79.0x87.0	7,541	52,357	100
		2	79.0x79.0x87.0	7,528	52,318	100
		3	79.0x79.0x87.0	7,545	52,370	100
gA^{Phe} m	dC$_{22:1}$	1	79.0x79.0x90.0	8,161	54,221	100
		2	79.0x79.0x90.0	8,149	54,185	100
		3	79.0x79.0x90.0	8,112	54,075	100
gA^{Gln} d	dC$_{18:1}$	1	79.4×79.4×80.0	8,856	51,951	220
		2	79.4×79.4×80.0	8,827	51,864	220
		3	79.4×79.4×80.0	8,845	51,918	220
gA^{Leu} d	dC$_{18:1}$	1	79.4×79.4×80.0	8,818	51,853	220
		2	79.4×79.4×80.0	8,832	51,895	220
		3	79.4×79.4×80.0	8,834	51,901	220

Table S4.2: Average channel tilt angles (mean ± standard error). Channels with residues (at positions 9, 11, 13 an 15) that can hydrogen bond have less tilt.

Channel	Tilt Angle [°]
gA^{Trp}	9.3 ± 0.7
gA^{mTrp}	12.4 ± 0.3
gA$^{nc-mTrp}$	15.2 ± 0.9
gA^{Tyr}	12.2 ± 0.8
gA^{Phe}	13.8 ± 0.4
gA^{Gln}	10.9 ± 1.3
gA^{Leu}	15.0 ± 1.7

Table S4.3: Average count of phospholipid choline nitrogen atoms within r = 3 Å of the center of the channel pore (where they would block ion permeation). The channel's pore is aligned with the z-axis ($xy = 0$) and the number of choline N is counted as a function of time. The presented value provides the time-averaged fraction of choline N in the vicinity of the pore. Residues that cannot hydrogen bond, and therefore impose fewer constraints on the lipids adjacent to the channel) increase the probability of finding a choline N near the pore.

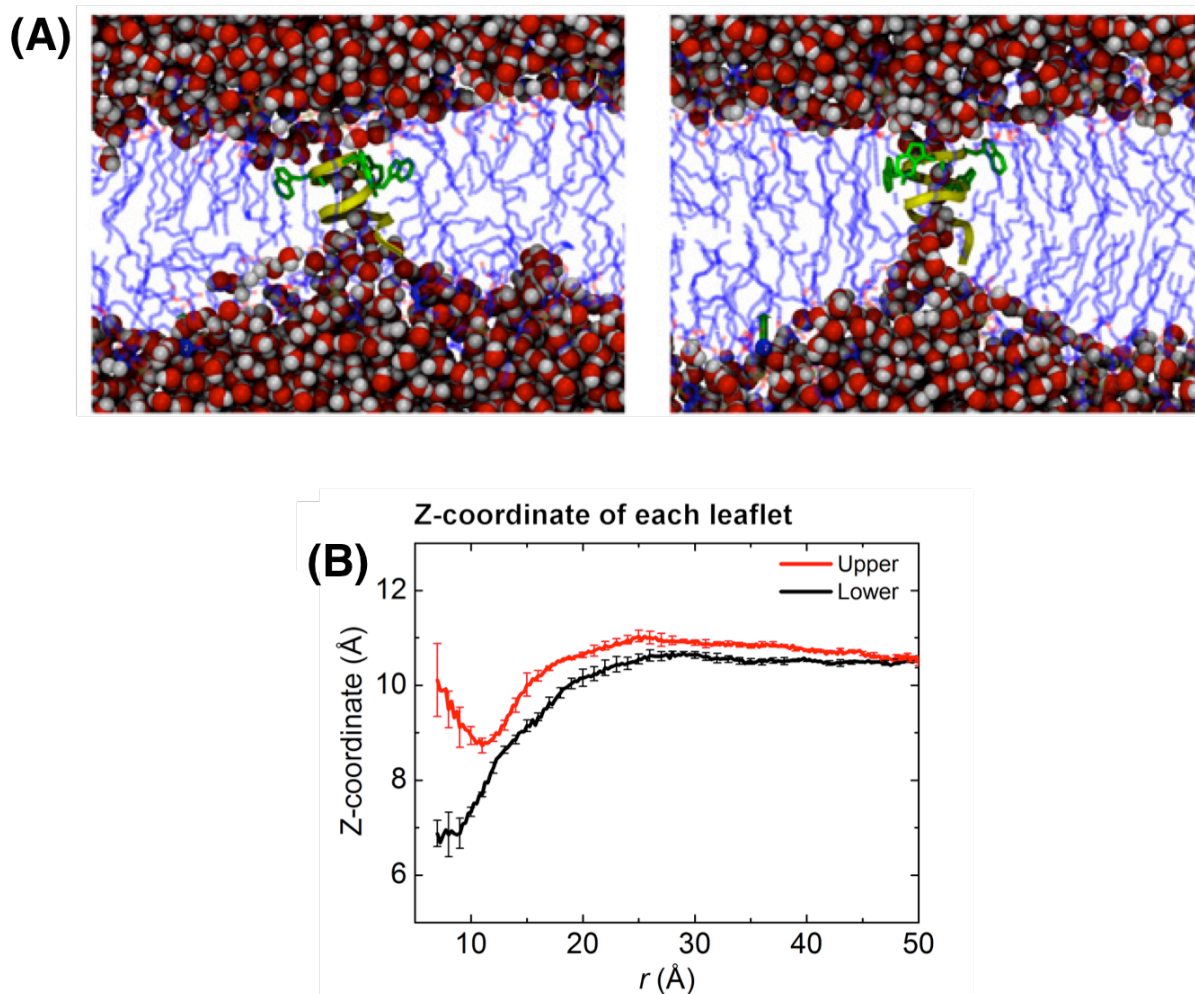| Channel	Avg ± St Err
gA^{Trp}	0.12 ± 0.02
gA^{mTrp}	0.26 ± 0.03
$gA^{nc\text{-}mTrp}$	0.30 ± 0.02
gA^{Tyr}	0.14 ± 0.02
gA^{Phe}	0.27 ± 0.02
gA^{Gln}	0.14 ± 0.02
gA^{Leu}	0.29 ± 0.01

4.5.3 Supplemental figures

Figure S4.1: Previously published (157) (A) gA^{Trp} monomer pore formation in dC$_{12:0}$ (1,2-di-lauroyl-PC, DLPC) in the left panel and dC$_{14:0}$ (1,2-di-myristoyl-PC, DMPC) in the right panel. (B) Monolayer hydrophobic thickness profiles, which show an asymmetry in monolayer deformations due to an asymmetric channel. Reprinted from Biophysical Journal, 102, T. Kim, K.I. Lee, P. Morris, R.W. Pastor, O.S. Andersen, and W. Im, Influence of Hydrophobic Mismatch on Structures and Dynamics of Gramicidin A and Lipid Bilayers, 1551–1560, 2012, with permission from Elsevier.
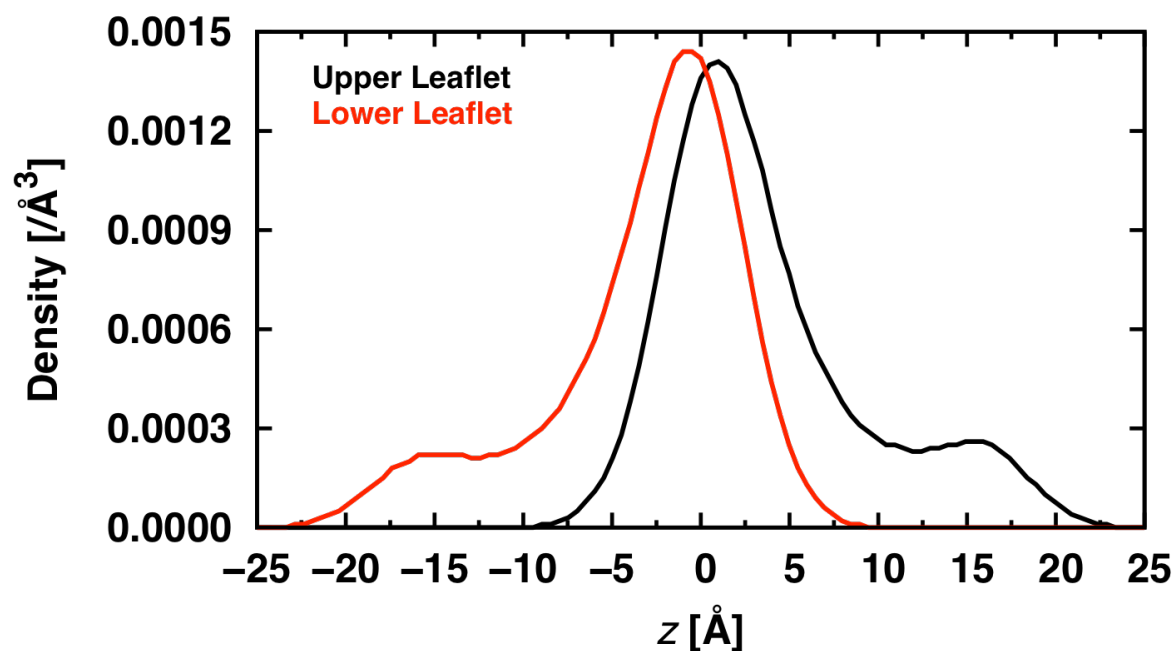
Figure S4.2: $dC_{22:1}$ lipid *sn-2* terminal carbon atom densities along the *z*-axis for both leaflets. The most populated *z* (for the bilayer) is at ~0 Å, but the unsigned weighted averages are at ~3.7 Å and ~3.3 Å for the upper and lower leaflets, respectively.
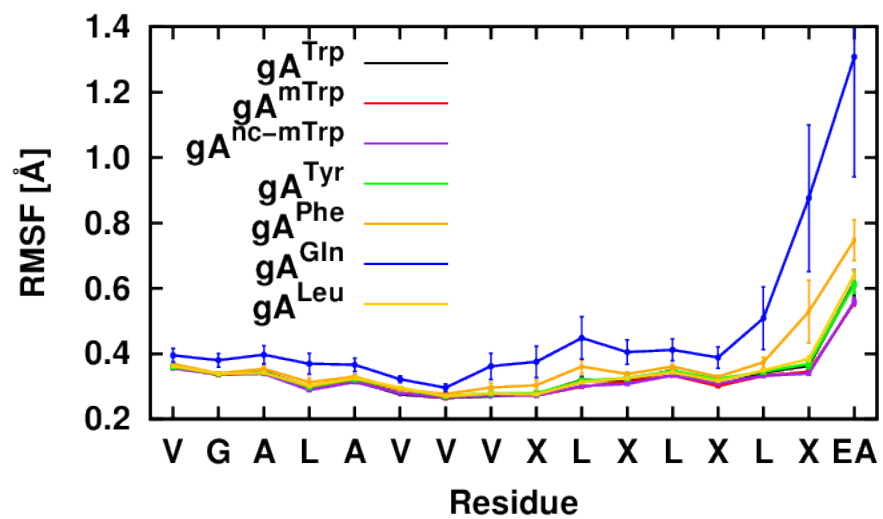
Figure S4.3: Heavy atom backbone root mean squared fluctuations (RMSFs) for all channel types in dC$_{18:1}$. The "X" residues are those that were mutated, "EA" is ethanolamide.
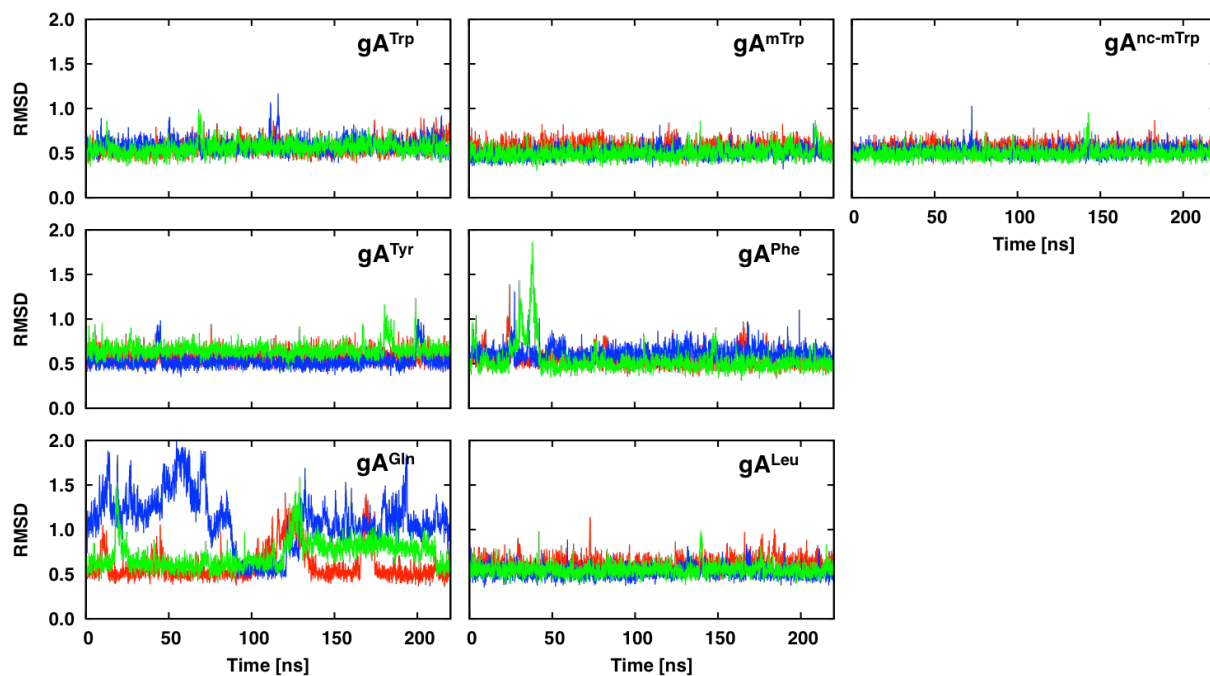
Figure S4.4: Time series of heavy atom backbone root mean square deviations (RMSDs) for each channel type in dC$_{18:1}$ calculated at each point relative to the minimized PDB:1JNO structure. Red, blue, and green are replicas 1, 2, and 3, respectively.
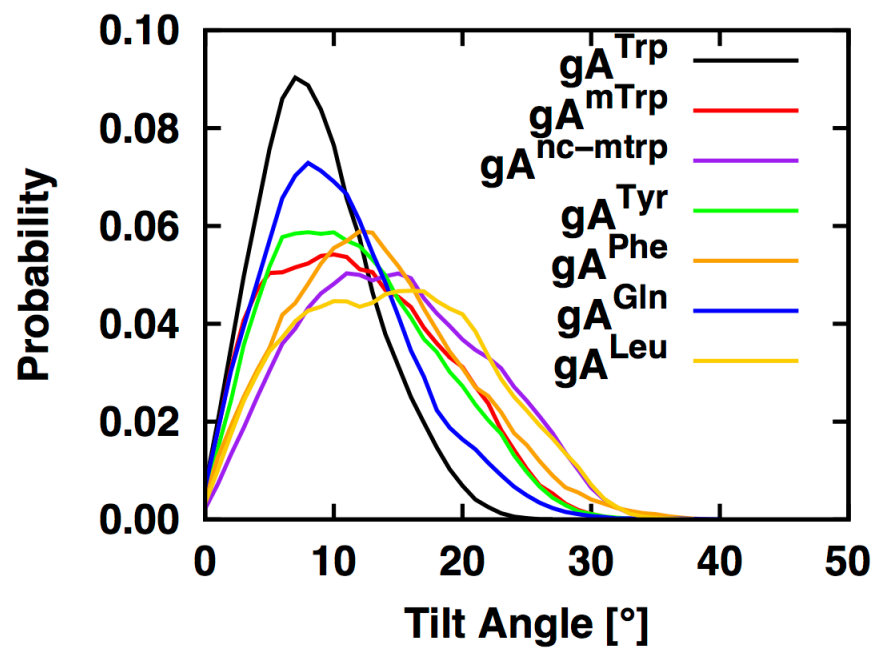
Figure S4.5: Channel tilt angle distributions for each channel type in dC$_{18:1}$. 1° bin width. Area under each curve equals unity.

Figure S4.6: Tyr and Phe χ_1-χ_2 dihedral angles in dC$_{18:1}$. χ_1 is the dihedral of the backbone N, Cα, Cβ, and Cγ atoms. χ_2 is the dihedral of the Cα, Cβ, Cγ, and Cδ atoms (Cδ is double bonded to Cγ and bonded to indole N). The color scheme for the heat plots is shown on the right with log{count/bin} and 1° bins in both dimensions.

110

Figure S4.7: Gln and Leu χ_1-χ_2 dihedral angles in dC$_{18:1}$. See the **Figure S4.6** caption for the figure notation.

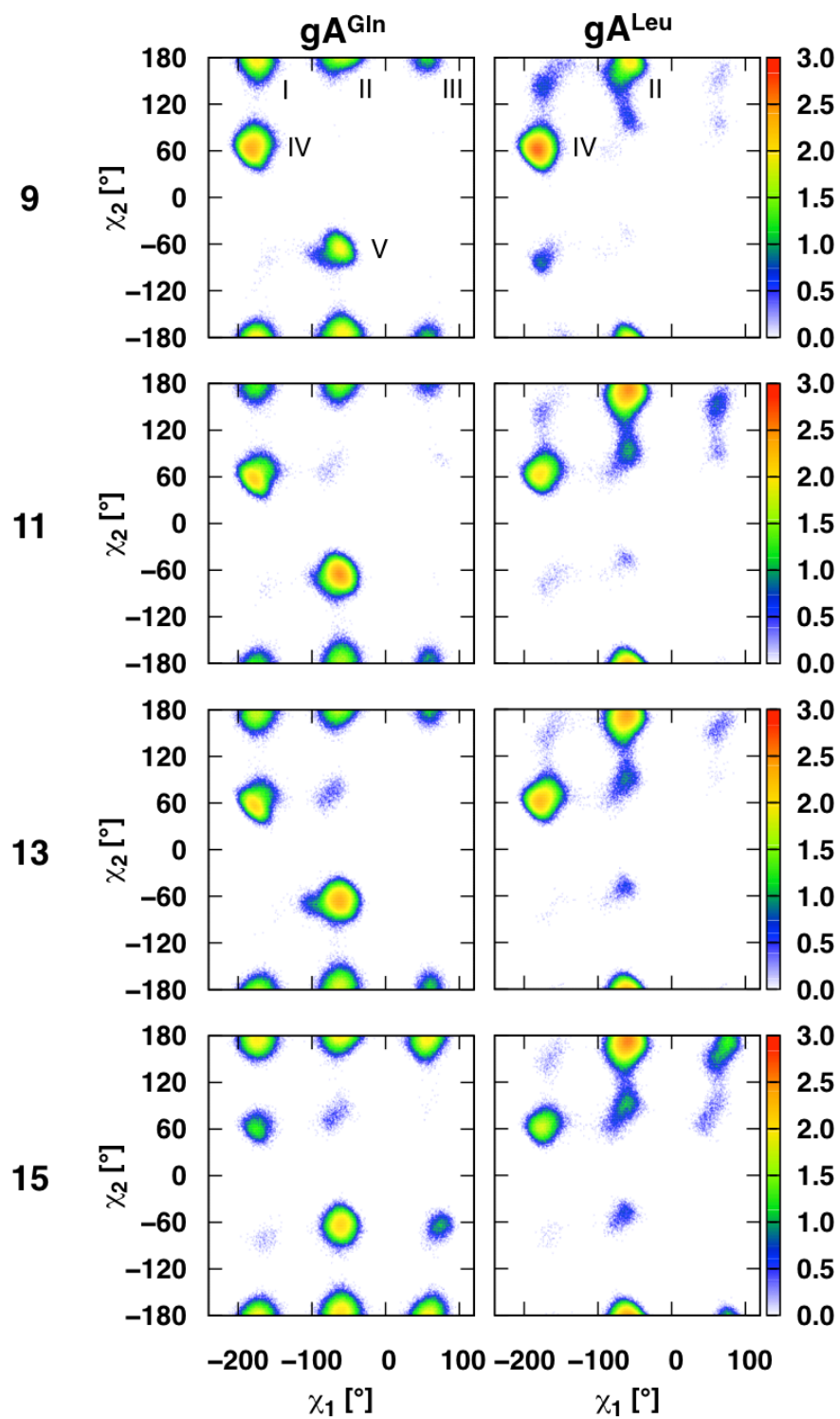**Figure S4.8:** Time series of the Trp9 χ_1-χ_2 dihedral angles of gA$^{\text{Trp}}$ in dC$_{18:1}$. Red circles indicate the χ_1 angle and green circles indicate the χ_2 angle. (A–C) are the three independent simulation replicas. The left and right panels are for each monomer.

Figure S4.9: Radial RDFs for dC$_{18:1}$ lipids around each channel type with 0.5 Å bins. The center of mass of the total lipid (COM) is in green, the location of the nitrogen (N) is in blue, and the center of mass of the tails only (TAIL) is in black. RDFs are normalized by the expected bulk concentration per radial bin: $g(r) = \rho(r)/\rho_{\text{bulk}}$.

Figure S4.10: Contact plots for each channel type in $dC_{18:1}$. The sum of the frequencies of the considered interactions is normalized to 1.0 for each residue.

Figure S4.11: Trp χ_1-χ_2 dihedral angles in $dC_{18:1}$. Since it has been shown that backbone restraints (like dCMAP) affect residue dynamics, we compare (A) a previously published simulation(9) of gA^{Trp} in $dC_{18:1}$ using an RMSD backbone restraint (100 ns of sampling) to (B) a simulation from this study of gA^{Trp} in $dC_{18:1}$ using only dCMAP (600 ns of sampling). The RMSD restraint possibly reduces the dynamics of the Trp9 residue (it fluctua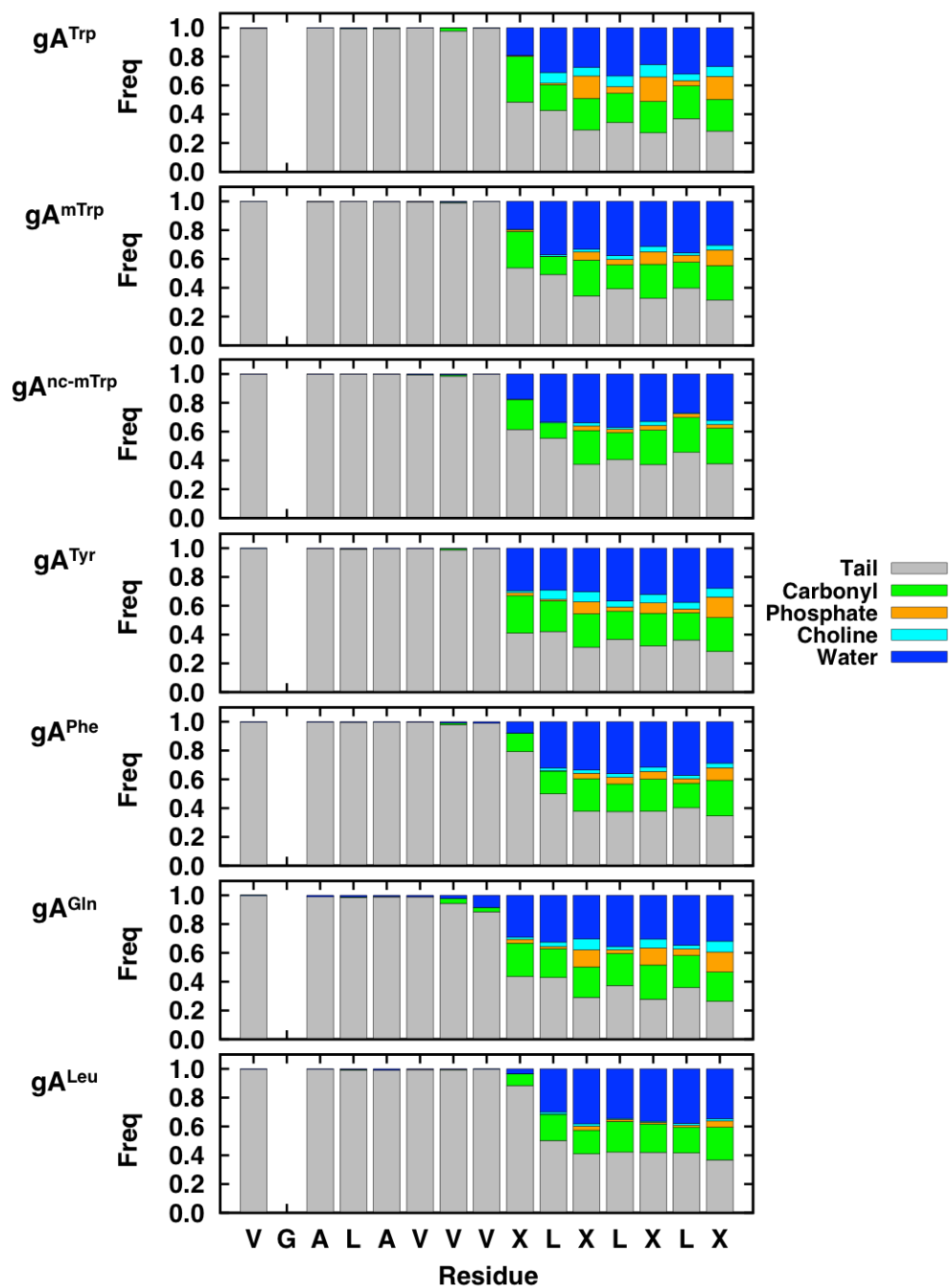tes around the initial χ_1-χ_2, which is the preferred orientation of this residue), whereas the dynamics of the other residues appear unchanged. The relative lack of sampling in the system with the RMSD restraint means that it is possible that simulation simply did not have enough time to sample other rotamer states. The true effect of a backbone restraint would have to be studied by a long timescale simulation, which is beyond the scope of this paper. Based on the χ_1-χ_2 plots, we conclude that while the backbone RMSD restraint might affect residue dynamics, it is likely minimal. See the **Figure S6** caption for the figure notation.

Figure S4.12: Heavy-atom *z*-density plots for the lipid C22 atom (approximate pivotal p plane position), entire channel, and the specific residues in dC$_{20:1}$. Dotted black lines are shown to accentuate the peak shifts relative to gA$^{\mathrm{Trp}}$. Data is plotted in 0.5 Å bins. Systems were centered by shifting the bilayer's center of mass to *z* = 0 Å. The summed peak density of all Trp residues is ~9 Å. Presumably, the hydrophobic extent of the channel would be larger than the peak density. Therefore, these peaks are in agreement with the canonical hydrophobic length of gA$^{\mathrm{Trp}}$ (~22 Å) (66, 99, 134).

5 Protocol of CHARMM-GUI *Hex Phase Builder*[e]

5.1 Introduction

Biological processes, such as fusion/fission events, and the general shapes of cells/organelles are dependent on curved lipid structures, but obtaining information on the energetics of membrane curvature is nontrivial (see reviews by Kozlov and coworkers (166–168)). A convenient method for studying highly curved lipid structures is using the inverse hexagonal (H_{II}) phase (31, 32, 37, 72). The H_{II} phase is a lipid polymorphism where the polar head groups face inward toward hexagonally packed water cylinders. Gaps between lipid tails (the so-called interstitial space that arises from hexagonally packing cylinders) can be filled with alkane to help make these structures more favorable. Due to the geometry of the polymorphism, these structures exist with some amount of lipid curvature that is a balance between the pressure of the water inside the cylinder and the intrinsic lipid curvature.

Lipid monolayers have an intrinsic curvature (R_0^{-1}) that is the curvature of minimum free energy (31, 32, 37, 72). To determine the amount of energy needed to deform a monolayer from R_0^{-1}, it is useful to assign a bulk material property, the bending modulus k_C as an effective "force constant" against changes in curvature. Cylindrically bending the monolayer away from the intrinsic curvature to a new curvature R^{-1} adds energy per unit area $\bar{F}_H$ by way of the Helfrich Hamiltonian (55, 56):

$$\bar{F}_H = \frac{k_{c,m}}{2}(R^{-1} - R_0^{-1})^2. \tag{5.1}$$

Therefore, energies can be calculated for lipid monolayers at some R^{-1}, but before the energies can be estimated, $k_{c,m}$ and R_0^{-1} themselves must be known.

Although getting $k_{c,m}$ and R_0^{-1} from experiment via the H_{II} phase has been richly documented (21, 24, 30–32, 72, 123), it has not been possible to computationally calculate these bending parameters until recently. The oldest computational methods to calculate $k_{c,m}$ rely on measuring the fluctuating z-displacement of the bilayer from a reference plane (49, 145, 169). In a related method (46), lipid angle fluctuations are used to obtain $k_{c,m}$. Both procedures have yielded promising results, but both are subject to drawbacks: currently calculations may only be done with single-component bilayers, and these simulations only provide information on $k_{c,m}$.

Alternatively, both $k_{c,m}$ and R_0^{-1} can be calculated by molecular dynamics (MD) simulation of the inverse hexagonal phase (37). Similar to experiment, the bending parameters are extracted by varying the lipid hydration (i.e., changing the measured lipid curvature) and measuring the change in the osmotic pressure. This methodology allows for both $k_{c,m}$ and R_0 to be calculated for any lipid mixture or lipid-protein mixture.

To facilitate the building process of lipid hexagonal phase systems and to reduce their equilibration time, a build protocol for *Hex Phase Builder* in CHARMM-GUI (http://charmm-gui.org) (105) is developed and described. In the future, the builder will function with the 182 available lipid species, 5 interstitial alkane species (tetradecane, squalene, 3-ethylpentane, n-decane, and n-hexadecane), and 4 ion pairs (KCl, NaCl, $CaCl_2$, and $MgCl_2$). Based on *Bilayer Builder* (103, 104), *Hex Phase Builder* builds systems using a six-step protocol that includes placing lipids, water, ions, and alkane in as close to equilibrium positioning as possible and running a short minimization and equilibration. In this chapter, the *Hex Phase Builder* methodology is detailed along with some preliminary results for DOPC and DOPE lipids.

5.2 Methods

5.2.1 Lipid-only system build

The methodology of *Hex Phase Builder* closely follows that of the *Bilayer Builder* within CHARMM-GUI, and the general workflow is summarized in **Figure 5.1**. For a lipid-only system, a user begins by selecting: i) a target radius (r, in xy) and height (h, along the z-axis) with a specific lipid ratio (e.g., 1:1 DOPC and DOPE), or ii) a target radius and specific number of lipids (h is calculated from these values and lipid area). Thus, for the first listed option:

$$n_{\text{lipids}} = \frac{A_{\text{cyl}}}{A_{\text{lipid}}} = \frac{2\pi r h}{w_1 A_1 + w_2 A_2 + \cdots} \tag{5.2}$$

where n_{lipids} is the number of lipids, A_{cyl} is the area of the lipid cylinder to be used, A_{lipid} is the area per lipid, and w_1 is the weight of the area A_1, and so on (assuming the sum of the weights is unity). That is, for a 1:1 molar ratio of DOPC and DOPE, $A_{\text{lipid}} = (0.5)(69.7\,\text{Å}^2) + (0.5)(63.4\,\text{Å}^2) = 66.6\,\text{Å}^2$. In this test case, the number of lipids is calculated (and adjusted if necessary) and evenly split between the DOPC and DOPE components.

$$h = \frac{n_1 A_1 + n_2 A_2 + \cdots}{2\pi r} \tag{5.3}$$

where n_1 and A_1 are the number and area of lipid type 1.

Once r and h are determined, the size of the unit cell is determined. The r and the estimated length of the lipid tails (l_{lipid}) determine the side-to-side size in xy ($2R$) of the unit cell (**Figure 5.1**):

$$R = r + l_{\text{lipid}}. \tag{5.4}$$

The l_{lipid} is used within *Bilayer Builder*, so the values are internally contained within CHARMM-GUI.

Figure 5.1: Schematic of a lipid hexagonal phase system. The radius of the water cylinder is r, the length of the lipid is l_{lipid}, and the summation of these parts is R. R is also one half the hexagonal unit cell size along the y-axis.

Figure 5.2: *Hex Phase Builder* protocol. (A) Dummy spheres are placed on a cylinder of radius R and height h. (B) Dummy spheres are replaced by lipids. (C) Pore water is added. (D) Alkanes are added to interstitial spaces. (E) After minimization and brief dynamics are performed. (F) The simulation unit cell is shown (emphasized by the black hexagon) surrounded by periodic images.

After the system size is determined, dummy spheres are randomly placed on a cylinder with radius r. The dummy spheres have a certain van der Waals radius, so with a radial restraint placed on the dummy spheres, they are first minimized and briefly equilibrated using Langevin dynamics. Then, the dummy spheres are replaced one-by-one with lipids from a structure library containing 2,000 lipids of different conformations. The replacement scheme works by: i) rotating the lipid so that its major moment of inertia lies pointing alo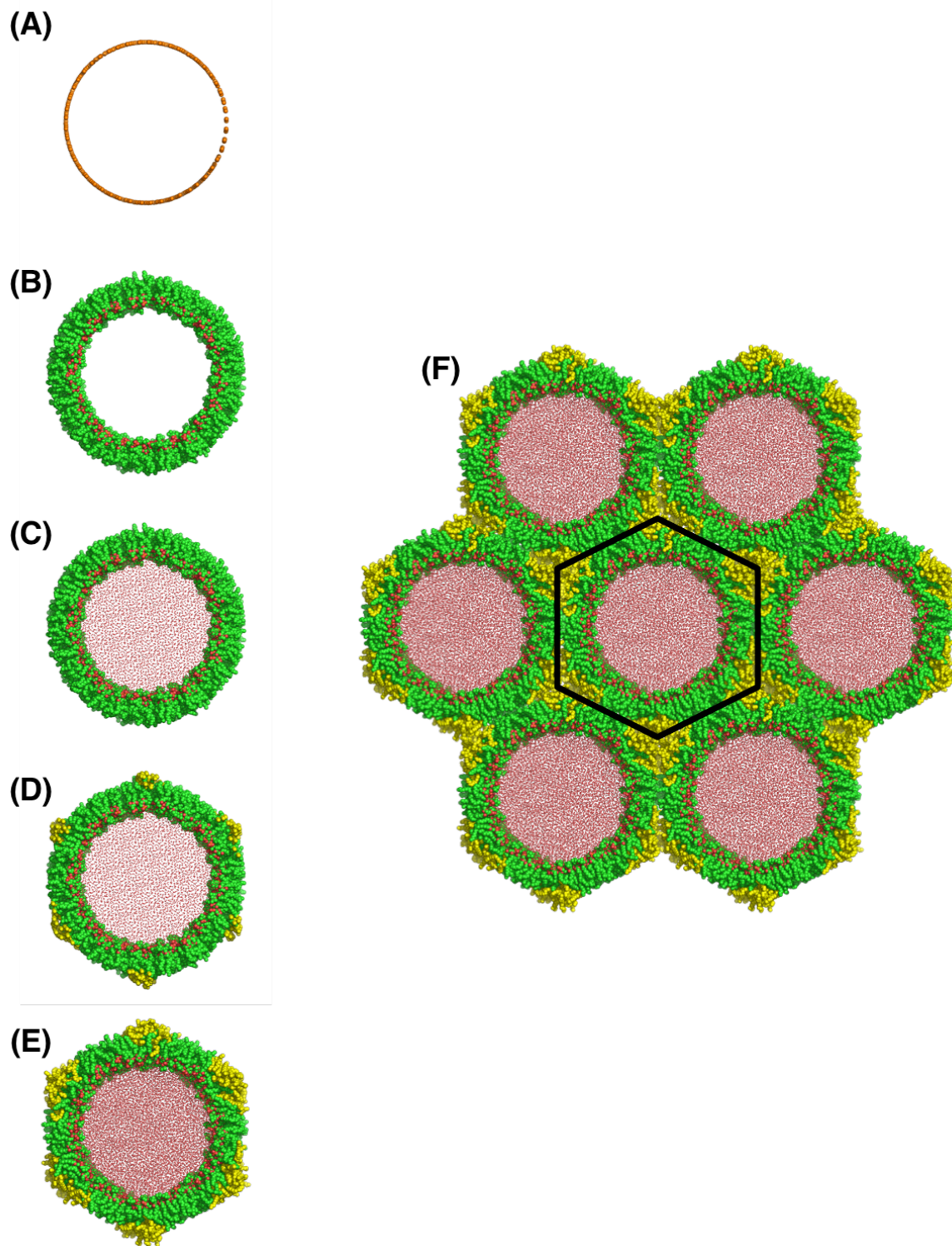ng the x-axis; ii) moving the approximate pivotal plane atom to the origin; iii) obtaining the xy vector from the origin to the dummy sphere that will be replaced; iv) rotating the lipid around the z-axis so that it is pointing toward the dummy sphere that will be replaced; v) translating the lipid in xyz so that the pivotal plane atom lies where the dummy sphere is located; and vi) deleting the dummy sphere. This process is repeated until lipids have replaced all dummy spheres.

The exact placement of the lipid on the cylinder must be considered. Since there is a specific region of lipids at which they bend with constant area (the so-called pivotal plane) (31, 170), lipids are placed so that their pivotal plane location has the same area it would in a bilayer. Previous work has suggested that the pivotal plane for DOPC and DOPE is near the C22 atom (i.e., the carbon below the carbonyl on the *sn-2* chain) (37). Therefore, when a user requests a radius, the C22 atom is placed at this radius. If necessary, further refinements will be made in the future for better system building.

Next, water and ions (if desired) are added to the system. The amount of water is estimated by using: $V_{\text{water}} = \pi(r-2)^2 h$. That is, the amount of water used is slightly less than the initial radius of the lipids. This is an empirical fix, and should be better calibrated in the future. Briefly, minimized water slabs are stacked and then cut so that they form a cylinder. Ions can be randomly added or semi-equilibrated using a Monte Carlo simulation.

Due to the geometry of the H_{II} phase, there are interstitial spaces between the tails of neighboring unit cells (**Figure 5.2**). Often, interstitial alkane is used to fill these voids. With lipid types that energetically prefer the H_{II} phase, these alkanes are unnecessary, however, alkanes are often used to promote easier H_{II} phase formation. In experiment, the amount of interstitial alkane is usually added by weight without consideration of the system's geometry. With the atomistic resolution available from MD simulations, the amount of necessary alkane can be better estimated:

$$n_{alkane} = \rho_{alkane} \cdot V_{alkane} = \rho_{alkane} \cdot \left[V_{hex} - \left(V_{water} + V_{lipid} \right) \right] \tag{5.5}$$

where n_{alkane} is the number of alkane molecules to be placed in the system and ρ_{alkane} is the density of the target alkane. The $V_{water} + V_{lipid}$ is estimated from R (Equation 5.4).

Although the alkane addition seems straightforward, it is complicated by periodic boundary conditions necessary for MD simulation. If an alkane is placed improperly, there can be major alkane-alkane and/or alkane-lipid atomic clashes (**Figure 5.3**). Therefore, an alkane's placement is determined by geometric considerations. A vertex is randomly selected, and a random xy position (within the unit cell, and outside R) is calculated. A dummy sphere is placed at this location. This random addition is done for all alkane molecules to be added to the unit cell.

An energy minimization and short Langevin dynamics equilibration are performed to let the dummy spheres adjust to each other (including across the periodic boundaries), and a restraint is applied to keep the spheres in the approximated interstitial space. Finally, dummy spheres are replaced by alkanes that are in all *trans* conformations and oriented parallel to the z-axis to ensure there will not be clashes across the periodic boundaries. Using this method, the number and severity of inter- and intra-cell atomic clashes are significantly reduced and the remaining clashes can be easily resolved using energy minimization using CHARMM.
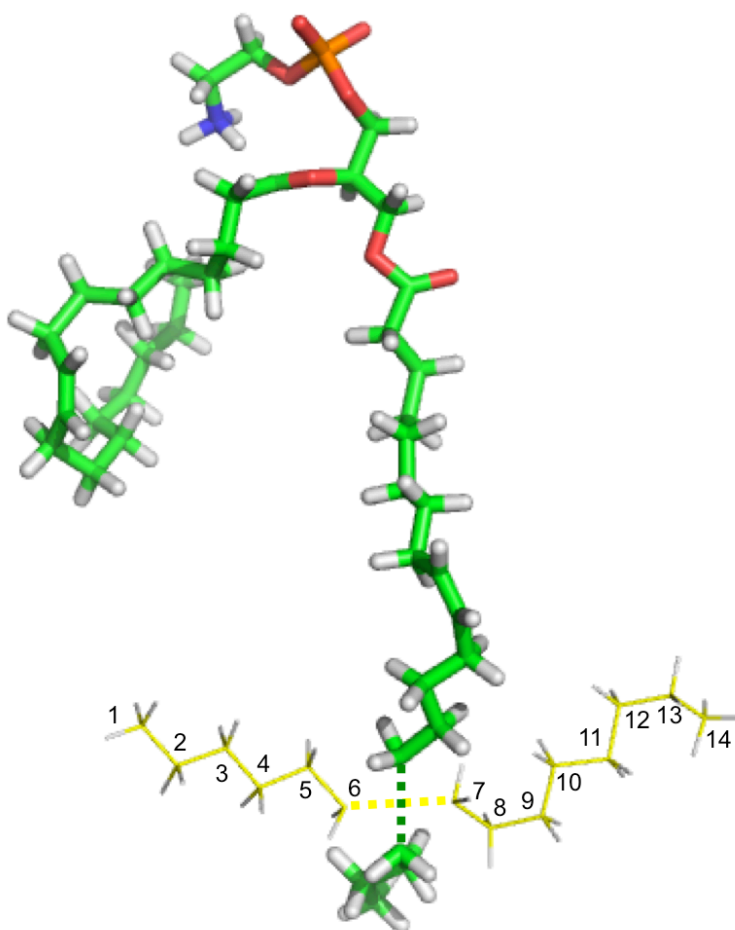
Figure 5.3: A major atomic clash between a DOPE lipid and tetradecane molecule. This type of clash cannot be resolved by CHARMM minimization techniques or unconstrained simulation. The massive energy associated with this clash eventually causes the simulation to be killed.

At this point, the coordinates of all components (lipids, water, ions, and alkanes) have been determined. The final step is the component assembly, generation of input files, and generation of appropriate restraints. Restraints are used to keep all lipid tail double bonds *cis* and all lipid chiral centers fixed during the minimization and equilibration (i.e., bad atomic clashes can force *cis* bonds into *trans* and affect chirality) and to keep the lipids at the desired radius. The restraints on the lipid chirality and lipid tail conformation are initially strong and relaxed through the six-step equilibration procedure. The lipid position restraint is much weaker and only used for the first two steps of the procedure.

5.2.2   Lipid-protein system build

If a user desires to have a system including protein, the first step is to read the protein coordinates from a PDB file. The protein is oriented with its principle axis along the y-axis and its center of mass at the origin. It is then translated by R along the y-axis so that the protein's center of mass lies on the periodic boundary between the main cell and neighboring cell. This location (symmetric to the other five faces of the hexagon) has a bilayer-like quality, in that it is furthest away from the interstitial areas (**Figure 5.1**). Clearly, this method is only good for transmembrane proteins. Therefore, CHARMM-GUI provides a user option in which the protein can be translated back along the y-axis. For example, if the user wants to simulate a membrane surface-bound protein, the user could translate back along y some distance until it has reached an optimal starting location. This requires some guess-and-check work, but does not require a substantial workload for the user.

The addition of dummy spheres is slightly modified to account for the protein. Instead of being randomly assigned across the surface area of the cylinder, dummy spheres are instead randomly assigned across two arcs of the cylinder. This ensures that no dummy spheres will be

placed inside or initially too close to a protein. After a minimization and brief Langevin dynamics (where the dummy spheres have a cylindrical position restraint), the dummy spheres are well-packed around the protein. This method is optimal for any type of protein. Wherever the protein touches the cylinder, dummy spheres will be excluded. The rest of the procedure (water and alkane addition, and the generation of input files) is the same as for a lipid-only system (**Section 5.2.1.1**). A resulting H_{II} system with a protein is shown in **Figure 5.4**.

Figure 5.4: An H_{II} phase system with embedded protein (PDB: 2HAC). (A) Top view. (B) Side showing protein embedded in lipid and in contact with water. (C) Opposite side of (B), showing the lipid void where the protein will be embedded by periodic boundary conditions.

5.3   Analysis

As written elsewhere in this dissertation (**Introduction**, **Chapter 3**, and **Chapter 4**), the Helrich Hamiltonian describes the energy associated with bending lipids to some curvature (R^{-1}) away from their intrinsic curvature (R_0^{-1}):

$$\bar{F}_H = \frac{k_{c,m}}{2}(R^{-1} - R_0^{-1})^2 \tag{5.6}$$

where the bending modulus is more strictly written for a lipid monolayer, $k_{c,m}$. Also note that due to the cylindrical geometry of the lipids in the hexagonal phase, there is only one curvature axis, i.e., no Gaussian curvature.

Experimentalists with the lipid hexagonal phase connected the Helfrich Hamiltonian to changes in osmotic pressure (Π) within the water cylinder (72):

$$R^2\Pi = -2k_{c,m}(R^{-1} - R_0^{-1}) \tag{5.7}$$

Equation 5.7 can be similarly tied to simulation, as was shown by Sodt and Pastor by simulations of the H_{II} phase:

$$R^2\Delta p = -2k_{c,m}(R^{-1} - R_0^{-1}) = -2\frac{d\bar{F}}{dR^{-1}} \tag{5.8}$$
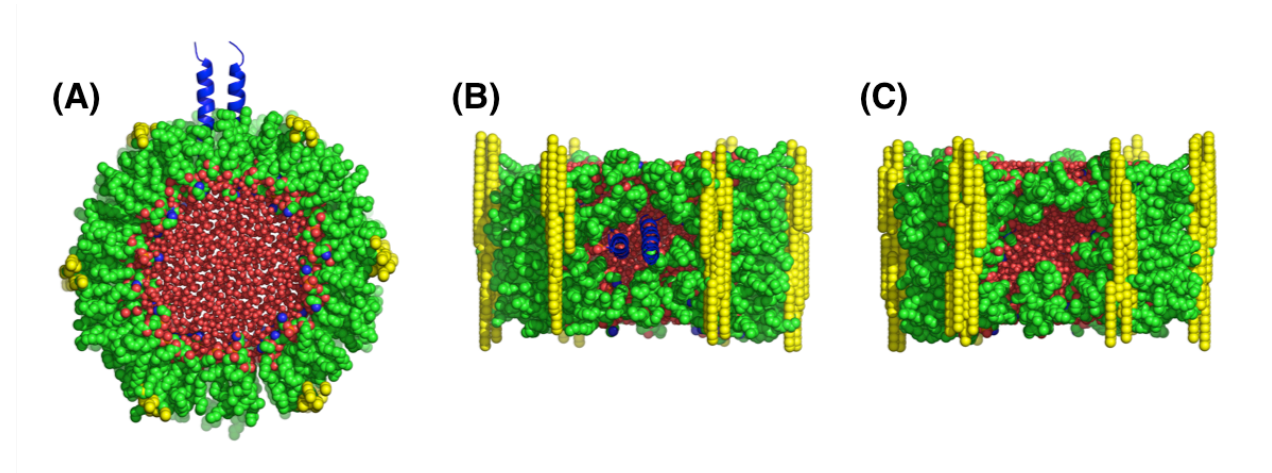
where Δp is equivalent to Π, and measures the pressure difference between the acyl chain region and center of the water cylinder. From Equation 5.8, plotting $R^2\Delta p$ against R^{-1} provides a linear relationship with the slope $-2k_{c,m}$ and y-intercept of $2k_{c,m}R_0^{-1}$. Therefore, both $k_{c,m}$ and R_0^{-1} can be obtained from this method. Additionally, the pivotal plane (where a lipid bends with constant area) can be determined by comparing to a planar bilayer area.

5.4   Preliminary results

Figures 5.5 and **5.6** show time series for the average radius of the C22 atoms of all lipids. Since we are interested in building these systems as close to equilibrium as possible, three independent replicas were used to monitor convergence. All lipids of a given type were placed with the same initial area per lipid, therefore, we would expect radius instabilities in all systems if there was a problem with this value. Since the radii only change in some systems, it suggests that the empirical addition of $V_{water} = \pi(r - 2)^2h$ is reasonable for larger systems, but is not reasonable for smaller systems. More testing will be done to determine how the amount of water affects the system evolution as a function of time.
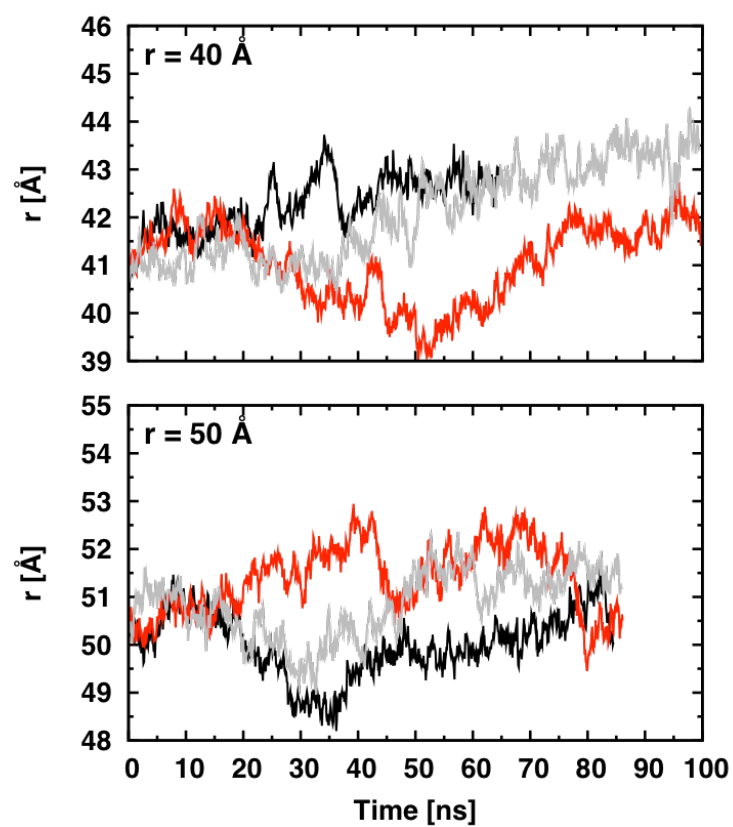
Figure 5.5: Lipid cylinder radius time series for DOPC. The three independent replicas are indicated by black, red, and grey.
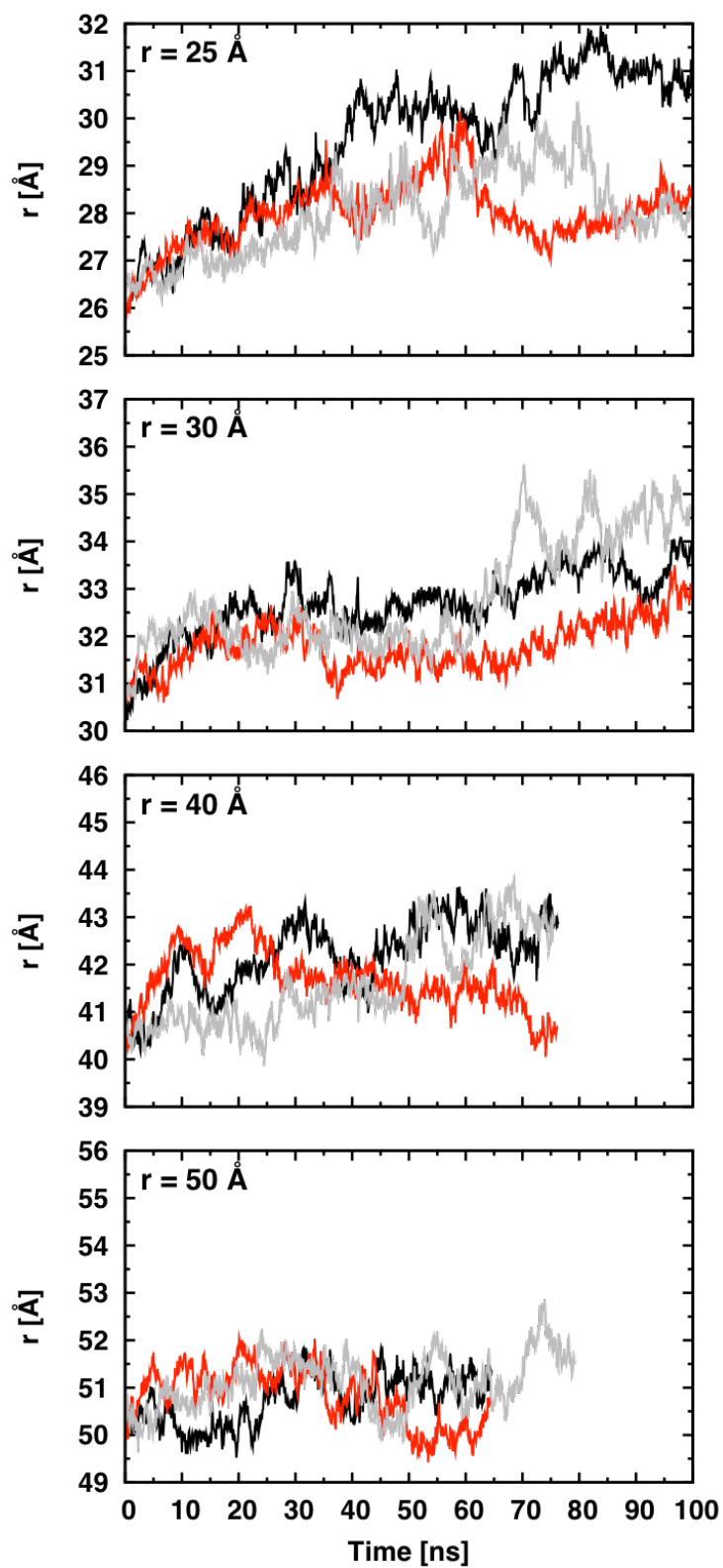
Figure 5.6: Lipid cylinder radius time series for DOPE. The three independent replicas are indicated by black, red, and grey.

129

6   Conclusion

6.1   Summary of dissertation

This dissertation has focused on the energetics of the bilayer with and without embedded gA channel. In **Chapter 1**, the origins of lipid packing and the penalties associated with deforming lipids were discussed. **Chapter 2** detailed lipid redistribution in a two-component bilayer to minimize compression penalties. **Chapter 3** introduced the concept of a curvature frustration due to dimerization and contact slope at the gA-bilayer interface. **Chapter 4** discussed how channel and bilayer properties are modulated as a function of channel mutation of residues at the bilayer-water interface. Finally, **Chapter 5** introduced CHARMM-GUI *Hex Phase Builder*, which will provide users a fast and replicable method to build lipid systems in the inverse hexagonal phase. Each of these chapters together have hopefully provided a solid introduction and reasoning for studying protein-bilayer interactions and bilayer energetics.

6.2   Remaining problems and research

The gA channel has been extensively studied for biophysical applications due to its simplicity and similarity to larger channels. Although gA's golden age seems to be in the past, biophysicists continue to use it as a simple model to extract information about how the bilayer influences channel function (i.e., lifetime). Hopefully, the work presented here has provided evidence that there is still plenty to learn from gA. The rest of this section discusses future problems and the research that could be built from this dissertation. Primarily, the following topics are of interest: i) further study of the biological meaning of protein-bilayer coupling strength; ii) the problem of obtaining direct experimental evidence of bilayer deformations; and iii) the effects of small molecules/peptides on gA channel function.

One of the largest questions generated from the work in this dissertation is how channel-

bilayer coupling affects bilayer energetics and channel lifetime. **Chapter 3** discusses the importance of the coupling (i.e., slope) at the channel-monolayer interface and how it theoretically describes the deformation when considered alongside the material properties of the monolayer. Recent work on this topic (not included in this dissertation) has demonstrated that the interfacial monolayer slope is in fact linked to the interfacial coupling strength (**Chapter 4**). In other words, the interfacial residues have a profound effect on determining the interfacial slope, and therefore, the entirety of the monolayer deformation.

In the near future, it would be of interest to study other semi-cylindrical proteins (e.g., rhodopsin and MscL) to determine whether predictions from gA are applicable to proteins of biological importance. Ultimately, it could be of biophysical and possibly therapeutic importance to study how profoundly interfacial residues affect bilayer deformations, and therefore, the protein function. These concepts are intuitively related to protein "anchoring," which has been previously discussed in the literature. However, the results from my doctoral studies suggest that an "anchoring" mechanism is simply the result of interfacial interaction strength, which in turn determines how the bilayer is deformed, which in turn helps determine the protein's conformation.

As we continue to improve computational techniques, experimentalists must also gather direct experimental descriptions of the channel-bilayer interface and the resulting deformation. Obviously, experimental results should be held as the standard for simulation to be compared against, however, some methodology (e.g., relating gA mean channel lifetimes to bilayer deformations) depends on inference instead of direct evidence. That is, mean channel lifetimes are observed in different lipid environments, but it is currently experimentally unknown how lipids really respond to gA. Key problems for experimentalists are that bilayer structural

information is difficult to obtain due to the dynamic nature of the lipids, and additionally, gA dimers themselves are transient. A possible method to experimentally study protein-/peptide-induced deformations is by neutron scattering, which has been recently used to study other protein-/peptide-bilayer interactions. Since these experiments provide insight into the bilayer perturbing properties of embedded proteins/peptides, and the results could be intuitively linked to simulation.

Finally, although briefly mentioned in **Chapter 2**, small molecules in and/or associated with the bilayer affect bilayer properties. These small molecules could be alkanes (necessary for experimental reasons), drugs, different ions, or peptides. If a molecule is a bilayer modulator, its bilayer association could in turn affect protein function. Indeed, laboratories around the world are currently studying these ideas. Possibly one of the more interesting and biologically relevant examples would be studying how drugs become associated with and modulate the bilayer. For example, a drug could potentially produce broad side effects due to bilayer modulation, resulting in modulation of off-target transmembrane protein function. Computational studies regarding these interactions will be necessary to explain the molecular origins of the interactions and possibly help guide better drug design.

Overall, the projects within this dissertation have provided quantitative information on a simple channel in model lipid membranes. Although there is still plenty to learn from gA, the usefulness of these model systems is that the results should be applicable to a range of biologically relevant protein-bilayer systems.

References

1. Finkelstein, A. 1976. Water and nonelectrolyte permeability of lipid bilayer membranes. J. Gen. Physiol. 68: 127–135.

2. Walter, A., and J. Gutknecht. 1986. Permeability of small nonelectrolytes through lipid bilayer membranes. J. Membr. Biol. 90: 207–217.

3. Zhuang, X., E.M. Dávila-Contreras, A.H. Beaven, W. Im, and J.B. Klauda. 2016. An extensive simulation study of lipid bilayer properties with different head groups, acyl chain lengths, and chain saturations. Biochim. Biophys. Acta - Biomembr. 1858: 3093–3104.

4. Huang, H.W. 1986. Deformation free energy of bilayer membrane and its effect on gramicidin channel lifetime. Biophys. J. 50: 1061–1070.

5. Helfrich, P., and E. Jakobsson. 1990. Calculation of deformation energies and conformations in lipid membranes containing gramicidin channels. Biophys. J. 57: 1075–1084.

6. Lundbaek, J.A., A.M. Maer, and O.S. Andersen. 1997. Lipid bilayer electrostatic energy, curvature stress, and assembly of gramicidin channels. Biochemistry. 36: 5695–5701.

7. Nielsen, C., M. Goulian, and O.S. Andersen. 1998. Energetics of inclusion-induced bilayer deformations. Biophys. J. 74: 1966–1983.

8. Sodt, A.J., and R.W. Pastor. 2014. Molecular Modeling of Lipid Membrane Curvature Induction by a Peptide: More than Simply Shape. Biophys. J. 106: 1958–1969.

9. Sodt, A.J., A.H. Beaven, O.S. Andersen, W. Im, and R.W. Pastor. 2017. Gramicidin A Channel Formation Induces Local Lipid Redistribution II: A 3D Continuum Elastic Model. Biophys. J. 112: 1198–1213.

10. Cooper, G. 2000. The Chemistry of Cells: Cell Membranes. In: The Cell: A Molecular Approach. Sunderland (MA): Sinauer Associates.

11. Andersen, O.S., and R.E. Koeppe. 2007. Bilayer thickness and membrane protein function: an energetic perspective. Annu. Rev. Biophys. Biomol. Struct. 36: 107–130.

12. Wiener, M.C., and S.H. White. 1992. Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. III. Complete structure. Biophys. J. 61: 434–447.

13. Wenk, M.R. 2005. The emerging field of lipidomics. Nat. Rev. Drug Discov. 4: 594–610.

14. Sud, M., E. Fahy, D. Cotter, A. Brown, E.A. Dennis, C.K. Glass, A.H.J. Merrill, R.C. Murphy, C.R.H. Raetz, D.W. Russell, and S. Subramaniam. 2007. LMSD: LIPID MAPS structure database. Nucleic Acids Res. 35: D527-32.

15. Caffrey, M., and G.W. Feigenson. 1981. Fluorescence quenching of model membranes. 3. Relationship between calcium adenosinetriphosphatase enzyme activity and the affinity of the protein for phosphatidylcholines with different acyl chain characteristics. Biochemistry. 20: 1949–1961.

16. Baldwin, P.A., and W.L. Hubbell. 1985. Effects of lipid environment on the light-induced conformational changes of rhodopsin. 2. Roles of lipid chain length, unsaturation, and phase state. Biochemistry. 24: 2633–2639.

17. Brown, M.F. 1994. Modulation of rhodopsin function by properties of the membrane

bilayer. Chem. Phys. Lipids. 73: 159–180.

18. Girshman, J., D. V Greathouse, R.E. Koeppe, and O.S. Andersen. 1997. Gramicidin channels in phospholipid bilayers with unsaturated acyl chains. Biophys. J. 73: 1310–1319.

19. Jensen, M.O., and O.G. Mouritsen. 2004. Lipids do influence protein function-the hydrophobic matching hypothesis revisited. Biochim. Biophys. Acta. 1666: 205–226.

20. Kim, T., K. Il Lee, P. Morris, R.W. Pastor, O.S. Andersen, and W. Im. 2012. Influence of Hydrophobic Mismatch on Structures and Dynamics of Gramicidn A and Lipid Bilayers. Biophys. J. 102: 1551–1560.

21. Teague, W.E., O. Soubias, H. Petrache, N. Fuller, K.G. Hines, R.P. Rand, and K. Gawrisch. 2013. Elastic properties of polyunsaturated phosphatidylethanolamines influence rhodopsin function. Faraday Discuss. 161: 383–395.

22. Cullis, P.R., and B. De Kruijff. 1979. Lipid polymorphism and the functional roles of lipids in biological membranes. Biochim. Biophys. Acta - Rev. Biomembr. 559: 399–420.

23. Lindblom, G., and L. Rilfors. 1989. Cubic phases and isotropic structures formed by membrane lipids — possible biological relevance. Biochim. Biophys. Acta - Rev. Biomembr. 988: 221–256.

24. Gruner, S.M. 1989. Stability of lyotropic phases with curved interfaces. J. Phys. Chem. 93: 7562–7570.

25. Seddon, J.M. 1990. Structure of the inverted hexagonal (HII) phase, and non-lamellar phase transitions of lipids. Biochim. Biophys. Acta. 1031: 1–69.

26. Brown, M.F. 2012. Curvature forces in membrane lipid-protein interactions. Biochemistry. 51: 9782–9795.

27. van Meer, G., D.R. Voelker, and G.W. Feigenson. 2008. Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol. 9: 112–124.

28. Landau, E.M., and J.P. Rosenbusch. 1996. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. Proc. Natl. Acad. Sci. . 93: 14532–14535.

29. Caffrey, M. 2015. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. Acta Crystallogr. Sect. F, Struct. Biol. Commun. 71: 3–18.

30. Kirk, G.L., and S.M. Gruner. 1985. Lyotropic effects of alkanes and headgroup composition on the lα -Hii lipid liquid crystal phase transition: hydrocarbon packing versus intrinsic curvature. J. Phys. Fr. 46: 761–769.

31. Rand, R.P., N.L. Fuller, S.M. Gruner, and V.A. Parsegian. 1990. Membrane curvature, lipid segregation, and structural transitions for phospholipids under dual-solvent stress. Biochemistry. 29: 76–87.

32. Fuller, N., and R.P. Rand. 2001. The influence of lysolipids on the spontaneous curvature and bending elasticity of phospholipid membranes. Biophys. J. 81: 243–254.

33. Hui, S.W., T.P. Stewart, P.L. Yeagle, and A.D. Albert. 1981. Bilayer to non-bilayer transition in mixtures of phosphatidylethanolamine and phosphatidylcholine: Implications for membrane properties. Arch. Biochem. Biophys. 207: 227–240.

34. Separovic, F., and K. Gawrisch. 1996. Effect of unsaturation on the chain order of phosphatidylcholines in a dioleoylphosphatidylethanolamine matrix. Biophys. J. 71: 274–

282.

35. Klauda, J.B., R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell, and R.W. Pastor. 2010. Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. J. Phys. Chem. B. 114: 7830–7843.

36. Jämbeck, J.P.M., and A.P. Lyubartsev. 2012. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. J. Chem. Theory Comput. 8: 2938–2948.

37. Sodt, A.J., and R.W. Pastor. 2013. Bending Free Energy from Simulation: Correspondence of Planar and Inverse Hexagonal Lipid Phases. Biophys. J. 104: 2202–2211.

38. Stubbs, C.D., and A.D. Smith. 1984. The modification of mammalian membrane polyunsaturated fatty acid composition in relation to membrane fluidity and function. Biochim. Biophys. Acta - Rev. Biomembr. 779: 89–137.

39. Alberts, B., A. Johnson, and J. Lewis. 2002. The Lipid Bilayer. In: Molecular Biology of the Cell. Garland Science.

40. Evans, E., W. Rawicz, and B.A. Smith. 2013. Concluding remarks Back to the future: mechanics and thermodynamics of lipid biomembranes. Faraday Discuss. 161: 591–611.

41. Feller, S.E., and R.W. Pastor. 1999. Constant surface tension simulations of lipid bilayers: The sensitivity of surface areas and compressibilities. J. Chem. Phys. 111: 1281–1287.

42. May, S., Y. Kozlovsky, A. Ben-Shaul, and M.M. Kozlov. 2004. Tilt modulus of a lipid monolayer. Eur. Phys. J. E. Soft Matter. 14: 299–308.

43. Kessel, A., N. Ben-Tal, and S. May. 2001. Interactions of Cholesterol with Lipid Bilayers: The Preferred Configuration and Fluctuations. Biophys. J. 81: 643–658.

44. Hamm, M., and M.M. Kozlov. 2000. Elastic energy of tilt and bending of fluid membranes. Eur. Phys. J. E. 3: 323–335.

45. Hamm, M., and M.M. Kozlov. 1998. Tilt model of inverted amphiphilic mesophases. Eur. Phys. J. B - Condens. Matter Complex Syst. 6: 519–528.

46. Levine, Z.A., R.M. Venable, M.C. Watson, M.G. Lerner, J.-E. Shea, R.W. Pastor, and F.L.H. Brown. 2014. Determination of Biomembrane Bending Moduli in Fully Atomistic Simulations. J. Am. Chem. Soc. 136: 13582–13585.

47. Feller, S.E., and R.W. Pastor. 1996. On simulating lipid bilayers with an applied surface tension: periodic boundary conditions and undulations. Biophys. J. 71: 1350–1355.

48. Beaven, A.H., A.M. Maer, A.J. Sodt, H. Rui, R.W. Pastor, O.S. Andersen, and W. Im. 2017. Gramicidin A Channel Formation Induces Local Lipid Redistribution I: Experiment and Simulation. Biophys. J. 112: 1185–1197.

49. Lindahl, E., and O. Edholm. 2000. Spatial and energetic-entropic decomposition of surface tension in lipid bilayers from molecular dynamics simulations. J. Chem. Phys. 113: 3882–3893.

50. Sonne, J., F.Y. Hansen, and G.H. Peters. 2005. Methodological problems in pressure profile calculations for lipid bilayers. J. Chem. Phys. 122: 124903.

51. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten. 2005. Scalable Molecular Dynamics with NAMD. J.

Comput. Chem. 26: 1781–1802.

52. Szleifer, I., D. Kramer, and A. Ben-Shaul. 1990. Molecular theory of curvature elasticity in surfactant films. J. Chem. Phys. 92: 6800–6817.

53. Ollila, S., T. Róg, M. Karttunen, and I. Vattulainen. 2007. Role of sterol type on lateral pressure profiles of lipid membranes affecting membrane protein functionality: Comparison between cholesterol, desmosterol, 7-dehydrocholesterol and ketosterol. J. Struct. Biol. 159: 311–323.

54. Park, S., A.H. Beaven, J.B. Klauda, and W. Im. 2015. How Tolerant are Membrane Simulations with Mismatch in Area per Lipid between Leaflets? J. Chem. Theory Comput. 11: 3466–3477.

55. Canham, P.B. 1970. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. J. Theor. Biol. 26: 61–81.

56. Helfrich, W. 1973. Elastic properties of lipid bilayers: theory and possible experiments. Z. Naturforsch. C. 28: 693–703.

57. Deserno, M. 2015. Fluid lipid membranes: From differential geometry to curvature stresses. Chem. Phys. Lipids. 185: 11–45.

58. Niphakis, M.J., K.M. Lum, A.B. Cognetta III, B.E. Correia, T.-A. Ichu, J. Olucha, S.J. Brown, S. Kundu, F. Piscitelli, H. Rosen, and B.F. Cravatt. 2015. A Global Map of Lipid-Binding Proteins and Their Ligandability in Cells. Cell. 161: 1668–1680.

59. Mouritsen, O.G., and M. Bloom. 1984. Mattress model of lipid-protein interactions in membranes. Biophys. J. 46: 141–153.

60. Killian, J.A. 1998. Hydrophobic mismatch between proteins and lipids in membranes. Biochim. Biophys. Acta. 1376: 401–415.

61. Salom, D., E. Pérez-Payá, J. Pascal, and C. Abad. 1998. Environment- and Sequence-Dependent Modulation of the Double-Stranded to Single-Stranded Conformational Transition of Gramicidin A in Membranes. Biochemistry. 37: 14279–14291.

62. Sun, H., D. V Greathouse, O.S. Andersen, and R.E. Koeppe II. 2008. The Preference of Tryptophan for Membrane Interfaces: Insights From N-Methylation of Tryptophans in Gramicidin Channels. J. Biol. Chem. 283: 22233–22243.

63. O'Connell, A.M., R.E. Koeppe II, and O.S. Andersen. 1990. Kinetics of Gramicidin Channel Formation in Lipid Bilayers: Transmembrane Monomer Association. Science (80-.). 250: 1256–1259.

64. Fonseca, V., P. Daumas, L. Ranjalahy-Rasoloarijao, F. Heitz, R. Lazaro, Y. Trudelle, and O.S. Andersen. 1992. Gramicidin channels that have no tryptophan residues. Biochemistry. 31: 5340–5350.

65. Providence, L.L., O.S. Andersen, D. V. Greathouse, R.E. Koeppe II, and R. Bittman. 1995. Gramicidin channel function does not depend on phospholipid chirality. Biochemistry. 34: 16404–16411.

66. Huang, H.W. 1986. Deformation free energy of bilayer membrane and its effect on gramicidin channel lifetime. Biophys. J. 50: 1061–1070.

67. Nielsen, C., and O.S. Andersen. 2000. Inclusion-induced bilayer deformations: effects of monolayer equilibrium curvature. Biophys. J. 79: 2583–2604.

68. Dowhan, W. 1997. Molecular basis for membrane phospholipid diversity: why are there

so many lipids? Annu. Rev. Biochem. 66: 199–232.

69. Myher, J.J., A. Kuksis, and S. Pind. 1989. Molecular species of glycerophospholipids and sphingomyelins of human erythrocytes: improved method of analysis. Lipids. 24: 396–407.

70. Dill, K.A., T.M. Truskett, V. Vlachy, and B. Hribar-Lee. 2005. Modeling water, the hydrophobic effect, and ion solvation. Annu. Rev. Biophys. Biomol. Struct. 34: 173–199.

71. Lee, K.I., R.W. Pastor, O.S. Andersen, and W. Im. 2013. Assessing smectic liquid-crystal continuum models for elastic bilayer deformations. Chem. Phys. Lipids. 169: 19–26.

72. Gruner, S.M. 1985. Intrinsic curvature hypothesis for biomembrane lipid composition: a role for nonbilayer lipids. Proc. Natl. Acad. Sci. . 82: 3665–3669.

73. London, E., and G.W. Feigenson. 1981. Fluorescence quenching in model membranes. 2. Determination of local lipid environment of the calcium adenosinetriphosphatase from sarcoplasmic reticulum. Biochemistry. 20: 1939–1948.

74. East, J.M., and A.G. Lee. 1982. Lipid selectivity of the calcium and magnesium ion dependent adenosine triphosphatase, studied with fluorescence quenching by a brominated phospholipid. Biochemistry. 21: 4144–4151.

75. Dumas, F., M.M. Sperotto, M.C. Lebrun, J.F. Tocanne, and O.G. Mouritsen. 1997. Molecular sorting of lipids by bacteriorhodopsin in dilauroylphosphatidylcholine / distearoylphosphatidylcholine lipid bilayers. Biophys. J. 73: 1940–1953.

76. Lewis, B.A., and D.M. Engelman. 1983. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. J. Mol. Biol. 166: 211–217.

77. Ryba, N.J., and D. Marsh. 1992. Protein rotational diffusion and lipid/protein interactions in recombinants of bovine rhodopsin with saturated diacylphosphatidylcholines of different chain lengths studied by conventional and saturation-transfer electron spin resonance. Biochemistry. 31: 7511–7518.

78. Williamson, I.M., S.J. Alvis, J.M. East, and A.G. Lee. 2002. Interactions of phospholipids with the potassium channel KcsA. Biophys. J. 83: 2026–2038.

79. Ridder, A.N.J.A., R.E.J. Spelbrink, J.A.A. Demmers, D.T.S. Rijkers, R.M.J. Liskamp, J. Brunner, A.J.R. Heck, B. de Kruijff, and J.A. Killian. 2004. Photo-crosslinking analysis of preferential interactions between a transmembrane peptide and matching lipids. Biochemistry. 43: 4482–4489.

80. Weiss, T.M., P.C.A. van der Wel, J.A. Killian, R.E. Koeppe II, and H.W. Huang. 2003. Hydrophobic Mismatch between Helices and Lipid Bilayers. Biophys. J. 84: 379–385.

81. Kim, T., and W. Im. 2010. Revisiting Hydrophobic Mismatch with Free Energy Simulation Studies of Transmembrane Helix Tilt and Rotation. Biophys. J. 99: 175–183.

82. Fahsel, S., E.-M. Pospiech, M. Zein, T.L. Hazlet, E. Gratton, and R. Winter. 2002. Modulation of concentration fluctuations in phase-separated lipid membranes by polypeptide insertion. Biophys. J. 83: 334–344.

83. Sperotto, M.M., and O.G. Mouritsen. 1993. Lipid enrichment and selectivity of integral membrane proteins in two-component lipid bilayers. Eur. Biophys. J. 22: 323–328.

84. Yin, F., and J.T. Kindt. 2010. Atomistic Simulation of Hydrophobic Matching Effects on Lipid Composition near a Helical Peptide Embedded in Mixed-Lipid Bilayers. J. Phys. Chem. B. 114: 8076–8080.

85. Yin, F., and J.T. Kindt. 2012. Hydrophobic Mismatch and Lipid Sorting Near OmpA in Mixed Bilayers: Atomistic and Coarse-Grained Simulations. Biophys. J. 102: 2279–2287.

86. Nielsen, S.O., C.F. Lopez, I. Ivanov, P.B. Moore, J.C. Shelley, and M.L. Klein. 2004. Transmembrane Peptide-Induced Lipid Sorting and Mechanism of L(α)-to-Inverted Phase Transition Using Coarse-Grain Molecular Dynamics. Biophys. J. 87: 2107–2115.

87. Klingelhoefer, J.W., T. Carpenter, and M.S.P. Sansom. 2009. Peptide nanopores and lipid bilayers: interactions by coarse-grained molecular-dynamics simulations. Biophys. J. 96: 3519–3528.

88. Andersen, O.S. 1983. Ion movement through gramicidin A channels. Single-channel measurements at very high potentials. Biophys. J. 41: 119–133.

89. Busath, D.D., O.S. Andersen, and R.E. Koeppe. 1987. On the conductance heterogeneity in membrane channels formed by gramicidin A. A cooperative study. Biophys. J. 51: 79–88.

90. Sawyer, D.B., R.E. 2nd Koeppe, and O.S. Andersen. 1989. Induction of conductance heterogeneity in gramicidin channels. Biochemistry. 28: 6571–6583.

91. Arseniev, A.S., I.L. Barsukov, V.F. Bystrov, A.L. Lomize, and O. YuA. 1985. 1H-NMR study of gramicidin A transmembrane ion channel. Head-to-head right-handed, single-stranded helices. FEBS Lett. 186: 168–174.

92. Killian, J.A., M.J. Taylor, and R.E. 2nd Koeppe. 1992. Orientation of the valine-1 side chain of the gramicidin transmembrane channel and implications for channel functioning. A 2H NMR study. Biochemistry. 31: 11283–11290.

93. Rijs, A.M., M. Kabelac, A. Abo-Riziq, P. Hobza, and M.S. de Vries. 2011. Isolated gramicidin peptides probed by IR spectroscopy. Chemphyschem. 12: 1816–1821.

94. Townsley, L.E.. T.W.A.. S.S.. H.J.F. 2001. Structures of Gramicidins A, B, and C Incorporated Into Sodium Dodecyl Sulfate Micelles. Biochemistry. 40: 11676.

95. Ulrich, W.-P., and H. Vogel. 1999. Polarization-Modulated FTIR Spectroscopy of Lipid/Gramicidin Monolayers at the Air/Water Interface. Biophys. J. 76: 1639–1647.

96. Allen, T.W., O.S. Andersen, and B. Roux. 2004. Energetics of ion conduction through the gramicidin channel. Proc. Natl. Acad. Sci. . 101: 117–122.

97. Roux, B. 1996. Valence selectivity of the gramicidin channel: a molecular dynamics free energy perturbation study. Biophys. J. 71: 3177–3185.

98. Roux, B., B. Prod'hom, and M. Karplus. 1995. Ion transport in the gramicidin channel: molecular dynamics study of single and double occupancy. Biophys. J. 68: 876–892.

99. Harroun, T.A., W.T. Heller, T.M. Weiss, L. Yang, and H.W. Huang. 1999. Experimental evidence for hydrophobic matching and membrane-mediated interactions in lipid bilayers containing gramicidin. Biophys. J. 76: 937–945.

100. Wallace, B.A., W.R. Veatch, and E.R. Blout. 1981. Conformation of gramicidin A in phospholipid vesicles: circular dichroism studies of effects of ion binding, chemical modification, and lipid structure. Biochemistry. 20: 5754–5760.

101. Bruno, M.J., R.E. Koeppe, and O.S. Andersen. 2007. Docosahexaenoic acid alters bilayer elastic properties. Proc. Natl. Acad. Sci. . 104: 9638–9643.

102. Marsh, D. 2008. Protein modulation of lipids, and vice-versa, in membranes. Biochim. Biophys. Acta - Biomembr. 1778: 1545–1575.

103. Jo, S., T. Kim, and W. Im. 2007. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. PLoS One. 2: e880.

104. Jo, S., J.B. Lim, J.B. Klauda, and W. Im. 2009. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. Biophys. J. 97: 50–58.

105. Jo, S., T. Kim, V.G. Iyer, and W. Im. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. J. Comput. Chem. 29: 1859–1865.

106. Brooks, B.R., C.L. Brooks, A.D. MacKerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus. 2009. CHARMM: The Biomolecular Simulation Program. J. Comput. Chem. 30: 1545–1614.

107. Feller, S.E., Y. Zhang, R.W. Pastor, and B.R. Brooks. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. J. Chem. Phys. 103.

108. Martyna, G.J., D.J. Tobias, and M.L. Klein. 1994. Constant pressure molecular dynamics algorithms. J. Chem. Phys. 101.

109. MacKerell A. D., D. Bashford, M. Bellott, Dunbrack R. L., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J. Phys. Chem. B. 102: 3586–3616.

110. Ingólfsson, H.I., Y. Li, V. V Vostrikov, H. Gu, J.F. Hinton, R.E. Koeppe, B. Roux, and O.S. Andersen. 2011. Gramicidin A Backbone and Side Chain Dynamics Evaluated by Molecular Dynamics Simulations and Nuclear Magnetic Resonance Experiments. I: Molecular Dynamics Simulations. J. Phys. Chem. B. 115: 7417–7426.

111. Mackerell, A.D., M. Feig, and C.L. Brooks. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J. Comput. Chem. 25: 1400–1415.

112. Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79: 926.

113. Ryckaert, J.-P., G. Ciccotti, and H.J.. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23: 327–341.

114. Essmann, U., L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen. 1995. A smooth particle mesh Ewald method. J. Chem. Phys. 103: 8577.

115. Steinbach, P.J., and B.R. Brooks. 1994. New spherical-cutoff methods for long-range forces in macromolecular simulation. J. Comput. Chem. 15: 667–683.

116. Shaw, D.E., R.O. Dror, J.K. Salmon, J.P. Grossman, K.M. Mackenzie, J.A. Bank, C. Young, M.M. Deneroff, B. Batson, K.J. Bowers, E. Chow, M.P. Eastwood, D.J. Ierardi, J.L. Klepeis, J.S. Kuskin, R.H. Larson, K. Lindorff-Larsen, P. Maragakis, M.A. Moraes, S. Piana, Y. Shan, and B. Towles. 2009. Millisecond-scale Molecular Dynamics

Simulations on Anton. In: Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis. New York, NY, USA: ACM. p. 39:1--39:11.

117. Berendsen, H.J.C., J.P.M. Postma, W.F. van Gunsteren, A. DiNola, and J.R. Haak. 1984. Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81.

118. Kräutler, V., W.F. van Gunsteren, and P.H. Hünenberger. 2001. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. J. Comput. Chem. 22: 501–508.

119. Shan, Y., J.L. Klepeis, M.P. Eastwood, R.O. Dror, and D.E. Shaw. 2005. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. J. Chem. Phys. 122: 54101.

120. Tuckerman, M., B.J. Berne, and G.J. Martyna. 1992. Reversible multiple time scale molecular dynamics. J. Chem. Phys. 97.

121. Andoh, Y., S. Okazaki, and R. Ueoka. 2013. Molecular dynamics study of lipid bilayers modeling the plasma membranes of normal murine thymocytes and leukemic GRSL cells. Biochim. Biophys. Acta. 1828: 1259–1270.

122. Martinez-Seara, H., T. Rog, M. Karttunen, I. Vattulainen, and R. Reigada. 2010. Cholesterol induces specific spatial and orientational order in cholesterol/phospholipid membranes. PLoS One. 5: e11162.

123. Rand, R.P., and V.A. Parsegian. 1997. Chapter 4 Hydration, Curvature, and Bending Elasticity of Phospholipid Monolayers. In: Membranes RMEBT-CT in, editor. Lipid Polymorphism and Membrane Properties. Academic Press. pp. 167–189.

124. Lundbaek, J.A., and O.S. Andersen. 1999. Spring constants for channel-induced lipid bilayer deformations. Estimates using gramicidin channels. Biophys. J. 76: 889–895.

125. Lundbaek, J.A., R.E. 2nd Koeppe, and O.S. Andersen. 2010. Amphiphile regulation of ion channel function by changes in the bilayer spring constant. Proc. Natl. Acad. Sci. U. S. A. 107: 15427–15430.

126. Kučerka, N., J. Gallová, D. Uhríková, P. Balgavý, M. Bulacu, S.-J. Marrink, and J. Katsaras. 2009. Areas of Monounsaturated Diacylphosphatidylcholines. Biophys. J. 97: 1926–1932.

127. Benz, R., and K. Janko. 1976. Voltage-induced capacitance relaxation of lipid bilayer membranes Effects of membrane composition. Biochim. Biophys. Acta - Biomembr. 455: 721–738.

128. Durkin, J.T., L.L. Providence, R.E. Koeppe, and O.S. Andersen. 1993. Energetics of Heterodimer Formation among Gramicidin Analogues with an NH2-terminal Addition or Deletion. J. Mol. Biol. 231: 1102–1121.

129. Wanasundara, S.N., V. Krishnamurthy, and S.-H. Chung. 2011. Free energy calculations of gramicidin dimer dissociation. J. Phys. Chem. B. 115: 13765–13770.

130. Miloshevsky, G. V, and P.C. Jordan. 2004. Gating Gramicidin Channels in Lipid Bilayers: Reaction Coordinates and the Mechanism of Dissociation. Biophys. J. 86: 92–104.

131. Rabinowitch, E. 1937. Collision, co-ordination, diffusion and reaction velocity in condensed systems. Trans. Faraday Soc. 33: 1225–1233.

132. Lewis, R.N., B.D. Sykes, and R.N. McElhaney. 1988. Thermotropic phase behavior of model membranes composed of phosphatidylcholines containing cis-monounsaturated

acyl chain homologues of oleic acid: differential scanning calorimetric and 31P NMR spectroscopic studies. Biochemistry. 27: 880–887.

133. Mitra, K., I. Ubarretxena-Belandia, T. Taguchi, G. Warren, and D.M. Engelman. 2004. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. Proc. Natl. Acad. Sci. U. S. A. 101: 4083–4088.

134. Elliott, J.R., D. Needham, J.P. Dilger, and D.A. Haydon. 1983. The effects of bilayer thickness and tension on gramicidin single-channel lifetime. Biochim. Biophys. Acta - Biomembr. 735: 95–103.

135. Yoo, J., and Q. Cui. 2013. Three-Dimensional Stress Field around a Membrane Protein: Atomistic and Coarse-Grained Simulation Analysis of Gramicidin A. Biophys. J. 104: 117–127.

136. Argudo, D., N.P. Bethel, F. V Marcoline, and M. Grabe. 2016. Continuum descriptions of membranes and their interaction with proteins: Towards chemically accurate models. Biochim. Biophys. Acta - Biomembr. 1858: 1619–1634.

137. Pan, J., S. Tristram-Nagle, N. Kučerka, and J.F. Nagle. 2008. Temperature Dependence of Structure, Bending Rigidity, and Bilayer Interactions of Dioleoylphosphatidylcholine Bilayers. Biophys. J. 94: 117–124.

138. Evans, E., and D. Needham. 1987. Physical properties of surfactant bilayer membranes: thermal transitions, elasticity, rigidity, cohesion and colloidal interactions. J. Phys. Chem. 91: 4219–4228.

139. Rawicz, W., K.C. Olbrich, T. McIntosh, D. Needham, and E. Evans. 2000. Effect of Chain Length and Unsaturation on Elasticity of Lipid Bilayers. Biophys. J. 79: 328–339.

140. Henriksen, J.R., and J.H. Ipsen. 2004. Measurement of membrane elasticity by micro-pipette aspiration. Eur. Phys. J. E. 14: 149–167.

141. Dimova, R. 2014. Recent developments in the field of bending rigidity measurements on membranes. Adv. Colloid Interface Sci. 208: 225–234.

142. Venable, R.M., F.L.H. Brown, and R.W. Pastor. 2015. Mechanical properties of lipid bilayers from molecular dynamics simulation. Chem. Phys. Lipids. 192: 60–74.

143. Campelo, F., H.T. McMahon, and M.M. Kozlov. 2008. The Hydrophobic Insertion Mechanism of Membrane Curvature Generation by Proteins. Biophys. J. 95: 2325–2339.

144. M. M. Kozlov, and M. Winterhalter. 1991. Elastic moduli for strongly curved monoplayers. Position of the neutral surface. J. Phys. II Fr. 1: 1077–1084.

145. Goetz, R., G. Gompper, and R. Lipowsky. 1999. Mobility and Elasticity of Self-Assembled Membranes. Phys. Rev. Lett. 82: 221–224.

146. Ring, A. 1996. Gramicidin channel-induced lipid membrane deformation energy: influence of chain length and boundary conditions. Biochim. Biophys. Acta - Biomembr. 1278: 147–159.

147. Pogozheva, I.D., H.I. Mosberg, and A.L. Lomize. 2014. Life at the border: adaptation of proteins to anisotropic membrane environment. Protein Sci. 23: 1165–1196.

148. Fiedler, S., J. Broecker, and S. Keller. 2010. Protein folding in membranes. Cell. Mol. Life Sci. 67: 1779–1798.

149. Killian, J.A., and G. von Heijne. 2000. How proteins adapt to a membrane-water interface. Trends Biochem. Sci. 25: 429–434.

150. Ridder, A.N.J.A., S. Morein, J.G. Stam, A. Kuhn, B. de Kruijff, and J.A. Killian. 2000. Analysis of the Role of Interfacial Tryptophan Residues in Controlling the Topology of Membrane Proteins. Biochemistry. 39: 6521–6528.

151. de Planque, M.R.R., B.B. Bonev, J.A.A. Demmers, D. V Greathouse, R.E. Koeppe II, F. Separovic, A. Watts, and J.A. Killian. 2003. Interfacial anchor properties of tryptophan residues in transmembrane peptides can dominate over hydrophobic matching effects in peptide-lipid interactions. Biochemistry. 42: 5341–5348.

152. Becker, M.D., D. V Greathouse, R.E. Koeppe II, and O.S. Andersen. 1991. Amino acid sequence modulation of gramicidin channel function: effects of tryptophan-to-phenylalanine substitutions on the single-channel conductance and duration. Biochemistry. 30: 8830–8839.

153. Garcia, J.C., M. Strube, K. Leingang, K. Keller, and M.M. Mueckler. 1992. Amino acid substitutions at tryptophan 388 and tryptophan 412 of the HepG2 (Glut1) glucose transporter inhibit transport activity and targeting to the plasma membrane in Xenopus oocytes. J. Biol. Chem. 267: 7770–7776.

154. Navedo, M., M. Nieves, L. Rojas, and J.A. Lasalde-Dominicci. 2004. Tryptophan substitutions reveal the role of nicotinic acetylcholine receptor alpha-TM3 domain in channel gating: differences between Torpedo and muscle-type AChR. Biochemistry. 43: 78–84.

155. Hong, H., S. Park, R.H.F. Jimenez, D. Rinehart, and L.K. Tamm. 2007. Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. J. Am. Chem. Soc. 129: 8320–8327.

156. Chattopadhyay, A., S.S. Rawat, D. V Greathouse, D.A. Kelkar, and R.E. Koeppe. 2008. Role of Tryptophan Residues in Gramicidin Channel Organization and Function. Biophys. J. 95: 166–175.

157. Kim, T., K. Il Lee, P. Morris, R.W. Pastor, O.S. Andersen, and W. Im. 2012. Influence of hydrophobic mismatch on structures and dynamics of gramicidin A and lipid bilayers. Biophys. J. 102: 1551–1560.

158. Huang, L., and B. Roux. 2013. Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. J. Chem. Theory Comput. 9: 3543–3556.

159. Perrin, B.S.J., A.J. Sodt, M.L. Cotten, and R.W. Pastor. 2015. The Curvature Induction of Surface-Bound Antimicrobial Peptides Piscidin 1 and Piscidin 3 Varies with Lipid Chain Length. J. Membr. Biol. 248: 455–467.

160. Chiu, S.W., S. Subramaniam, and E. Jakobsson. 1999. Simulation study of a gramicidin/lipid bilayer system in excess water and lipid. II. Rates and mechanisms of water transport. Biophys. J. 76: 1939–1950.

161. Allen, T.W., O.S. Andersen, and B. Roux. 2003. Structure of Gramicidin A in a Lipid Bilayer Environment Determined Using Molecular Dynamics Simulations and Solid-State NMR Data. J. Am. Chem. Soc. 125: 9868–9877.

162. Sychev, S. V, L.I. Barsukov, and V.T. Ivanov. 1993. The double $\pi\pi5.6$ helix of gramicidin A predominates in unsaturated lipid membranes. Eur. Biophys. J. 22: 279–288.

163. Phillips, R., T. Ursell, P. Wiggins, and P. Sens. 2009. Emerging roles for lipids in shaping membrane-protein function. Nature. 459: 379–385.

164. MacCallum, J.L., and D.P. Tieleman. 2011. Hydrophobicity scales: a thermodynamic looking glass into lipid-protein interactions. Trends Biochem. Sci. 36: 653–662.

165. MacCallum, J.L., W.F.D. Bennett, and D.P. Tieleman. 2008. Distribution of Amino Acids in a Lipid Bilayer from Computer Simulations. Biophys. J. 94: 3393–3404.

166. Chernomordik, L. V, and M.M. Kozlov. 2003. Protein-lipid interplay in fusion and fission of biological membranes. Annu. Rev. Biochem. 72: 175–207.

167. Zimmerberg, J., and M.M. Kozlov. 2006. How proteins produce cellular membrane curvature. Nat Rev Mol Cell Biol. 7: 9–19.

168. Chernomordik, L. V, and M.M. Kozlov. 2008. Mechanics of membrane fusion. Nat Struct Mol Biol. 15: 675–683.

169. Marrink, S.J., and A.E. Mark. 2001. Effect of Undulations on Surface Tension in Simulated Bilayers. J. Phys. Chem. B. 105: 6122–6127.

170. Marsh, D. 2011. Pivotal surfaces in inverse hexagonal and cubic phases of phospholipids and glycolipids. Chem. Phys. Lipids. 164: 177–183.

Personal list of publications

1. **Beaven, A.H.**, A.J. Sodt, R.W. Pastor, R.E. Koeppe II, O.S. Andersen, and W. Im. 2017. Characterizing residue-bilayer interactions using gramicidin A as a scaffold and tryptophan substitutions as probes. *Submitted*.

2. Sodt, A.J.[*], **A.H. Beaven**[*], O.S. Andersen, W. Im, and R.W. Pastor. 2017. Gramicidin A channel formation induces local lipid redistribution II: A 3D continuum elastic model. Biophys. J. 112: 1198–1213. DOI: 10.1016/j.bpj.2017.01.035. ([*] indicates equal contribution)

3. **Beaven, A.H.**, A.M. Maer, A.J. Sodt, H. Rui, R.W. Pastor, O.S. Andersen, and W. Im. 2017. Gramicidin A channel formation induces local lipid redistribution I: experiment and simulation. Biophys. J. 112: 1185–1197. DOI: 10.1016/j.bpj.2017.01.028.

4. Zhuang, X., E.M. Dávila-Contreras, **A.H. Beaven**, W. Im, and J.B. Klauda. 2016. An extensive simulation study of lipid bilayer properties with different head groups, acyl chain lengths, and chain saturations. Biochim. Biophys. Acta - Biomembr. 1858: 3093–3104. DOI: 10.1016/j.bbamem.2016.09.016.

5. Jo, S., X. Cheng, J. Lee, S. Kim, S.-J. Park, D.S. Patel, **A.H. Beaven**, K.I. Lee, H. Rui, S. Park, H.S. Lee, B. Roux, A.D. MacKerell Jr., J.B. Klauda, Y. Qi, and W. Im. 2016. CHARMM-GUI 10 years for biomolecular modeling and simulation. J. Comput. Chem. DOI: 10.1002/jcc.24660.

6. Park, S., **A.H. Beaven**, J.B. Klauda, and W. Im. 2015. How tolerant are membrane simulations with mismatch in area per lipid between leaflets? J. Chem. Theory Comput. 11: 3466–3477. DOI: 10.1021/acs.jctc.5b00232.
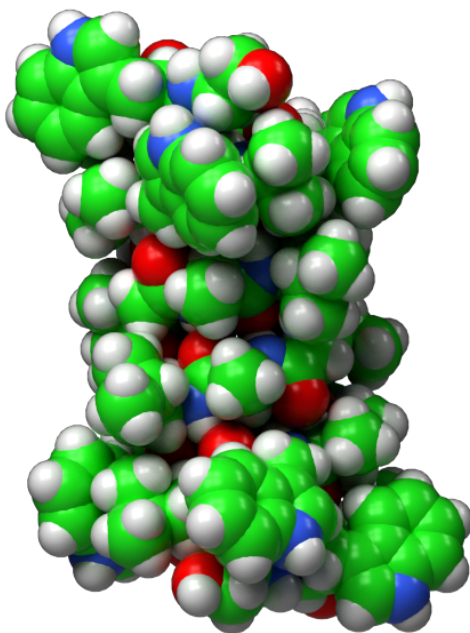
Image rendered by the Indigo Molecule Portrait Maker service.
www.indigomolecularimages.com/molecule-portrait-maker/

144