THE UNIVERSITY OF CHICAGO

COARSE-GRAINED APPROACHES TO MEMBRANE REMODELING AND I-BAR DOMAIN SELF-ASSEMBLY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE PRITZKER SCHOOL OF MOLECULAR ENGINEERING
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
ZACHARY RYAN JARIN

CHICAGO, ILLINOIS
JUNE 2020

To my family.

Table of Contents

List of Figures

Acknowledgements

There have many people who have helped complete the work presented here. First, I would like to thank Gregory A. Voth. As an advisor, Greg gave me interesting problems and the autonomy, resources, and support to explore new solutions and provide novel answers. His expertise coarse-graining and molecular modeling as well as the impressive research group were invaluable to me. I would also like to thank Juan J. de Pablo, my pro forma advisor and Andrew L. Ferguson and Tobin R. Sosnick for serving on my thesis committee.

I am thankful for my collaborators: Patricia Bassereau, Feng-Ching Tsai, and Michael Henderson at Insitut Curie. Patricia welcomed me into her group for three months during my doctoral research and has been an extremely insightful, encouraging, and helpful collaborator. Feng and Michael taught me everything about their experiments as well as insights into the phenomenon that required modeling. Together with the rest of the Bassereau group, they gave me a great experience in Paris and with collaboration.

Research group and their culture are a large piece of the graduate school experience. I would like to thank the entire Voth group. They are consistent and hard-working group of individuals. The many postdoctoral researchers that joined with expertise and a willingness to share it with me especially Drs. Alex Pak, Anand Srivastava, Aram Davtyan, Glen Hocky, Jesper Madsen, John Grime, and Viviana Monje-Galvan. My fellow graduate students in the Voth group have been a source of camaraderie and insight especially Aleksander Durumeric, Jacob Wagner, Jaehyeok Jin, Josh Zuchniarz, Laura Watkins, Morris Sharp, Paul Calio, Thomas Dannenhoffer-Lafage, Tim Loose, and Yining Han.

I would like to thank my friends and family. The members of the inaugural class of the Institute for Molecular Engineering were indispensable friends throughout my time at the University of Chicago especially Arin Greenwood, Emre Sevgen, Elyse Watkins, Lambert Potin, Lea

Maillat, Kevin Miao, Moshe Dolejsi, and Yu Kambe. I am grateful for all of my friendships outside of the university and throughout the city of Chicago especially Max A, Paul A, Sean M, Jordan B, Luke D, Mike C, Ben S, Bill H, and Jenna Gales. I have enjoyed this city more than I ever imagined. Finally, I would like to thank my family, Elmer and Teresa, Matt, Shakira, Liam, Miles, and Arthur, and Nick and Rosin. They are an incredible source of support and I would not be here without them. Thank you all.

Abstract

I-BAR proteins are located at the plasma membrane to sense and generate local membrane curvature. Prior to inducing membrane shape changes larger than a single protein, I-BAR domains are believed to aggregate and assemble. To discern how I-BAR domains organize and collectively stabilize membrane deformation, I developed two distinct coarse-grained membrane-protein models: one is a bottom-up model parameterized from all-atom simulations to capture low bound protein density behavior of the I-BAR domain of IRSp53 and the other is a tunable, lower resolution model used to assess the effects of various characteristics of the I-BAR family (e.g., intrinsic protein curvature). The separate approaches highlight the strengths of various coarse-graining approaches while providing key insights into I-BAR domain assembly. Together, I elucidate the role of attractive membrane-mediated forces in I-BAR domain assembly and the interplay between protein and membrane curvatures.

After using both bottom-up and top-down methodologies to understand I-BAR domain self-assembly, I assess how coarse-grained representation affects the estimation of the membrane bending modulus and the temperature-dependence in the popular top-down model, MARTINI. Issues of transferability and representability have been identified in the bottom-up coarse-grained models, but a computational demonstration of these effects in top-down coarse-grained models does not exist. I use MARTINI as a test case to understand how the coarse-grained representation affects cholesterol and amphipathic helix association in bilayers. The correctness of MARTINI in capturing these various effects is of growing interest as coarse-grained simulations are applied outside of parametrization data.

Chapter 1: Introduction

1.1 Thesis Aims and Scope

Biological membranes compartmentalize the cell to maintain chemical and mechanical gradients while continuously signaling changes in their environment. Their shape is dynamic to accommodate internal and external stresses and proteins often facilitate large membrane shape changes. The Bin/Amphiphysin/Rvs (BAR) superfamily of proteins are membrane remodeling proteins, notable for the curved dimer structure. The inverse-BAR (I-BAR) family is a subset of BAR domains responsible sensing and stabilizing outward curvature from the cell. Their physiological function has been associated with metastatic cancer and autism spectrum disorders, where well-formed protrusions are necessary. Additionally, *in vitro* experiments have shown I-BAR domains are sufficient to generate membrane tubules in the absence of other cellular machinery like the cytoskeleton. Given I-BAR domains' significance in the molecular basis of disease and interesting self-assembly behavior in synthetic systems, a variety of atomistic, coarse-grained, and continuum models have been used to glean deeper insights into their mechanism of membrane remodeling. These separate approaches focusing on different scales of the I-BAR domain-mediated membrane remodeling problem without learning and incorporating the results from the separate modeling approaches.

The aim of this thesis is two-fold: understand of I-BAR domain-mediated membrane remodeling using different multiscale approaches while evaluating accuracy of the state-of-the-art coarse-grained membrane models. Through a combination of atomistic, coarse-grained, and continuum modeling, I investigate these I-BAR domains and lipid bilayers. The rest of Chapter 1: is an introduction of biological membranes, lipids, and I-BAR domains. Chapter 2: is an overview of the computational methods used throughout this thesis and their application to membrane remodeling.

In Chapter 3:, I present insights from bottom-up coarse-grained model on I-BAR domain organization inside of membrane tubules and on vesicles along with accompanying experimental results from Feng-Ching Tsai and Patricia Bassereau at Insitut Curie. In Chapter 4: I show I-BAR domains stabilize membrane protrusions through clustering negatively charged lipids using a top-down coarse-grained model. In Chapter 5:, I investigate the popular top-down coarse-grained model, MARTINI, for issues of representability and transferability using lipid bilayers and amphipathic helix association as a test case.

1.2 Biological Membranes

Membranes throughout the cell are composed of a mixture of proteins and lipids. The plasma membrane forms a dynamic barrier, constantly changing shape to separate the inside and outside of the cell in response to internal and external stimuli.(McMahon and Gallop, 2005) Chemical and mechanical stresses begin a cascade of responses that lead to changes in the shape and composition of the plasma membrane. The fluid nature of biological membrane is key to communicating stresses across the membrane while maintaining a barrier to free diffusion.

Composition and Chemistry

Lipid are a key component of biological membranes and their formation of the bilayer phase provides a semipermeable scaffold for the assembly of signaling proteins. The makeup of biological membranes can vary quite drastically throughout the same cell and between different cells but there remain a few common building blocks between membranes.(Harayama and Riezman, 2018) Most commonly there are glycerol-based lipids, sphingolipids, and sterols. Within the glycerol-based lipids, there are di- or tri-acylglycerols, where two or three adjacent hydroxyl hydrogen are replaced by a long acyl group. Oleyl (16:0) and palmitoyl (18:1) groups are among the most common lipid tail groups in the body, but lipid tail groups can have 16 to 22 carbons with up to 6 double bonds.(Harayama and Riezman, 2018) Additionally, there are glycerophospholipids, where two adjacent hydroxyl hydrogens are replaced with long fatty chains and the final is phosphate group (e.g., phosphocholine shown in Figure 1-1A). The general structure of phospholipids is amphiphilic and composed of two hydrophobic tails and a hydrophobic head group that gives rise to complex self-assembly behavior in aqueous environment.

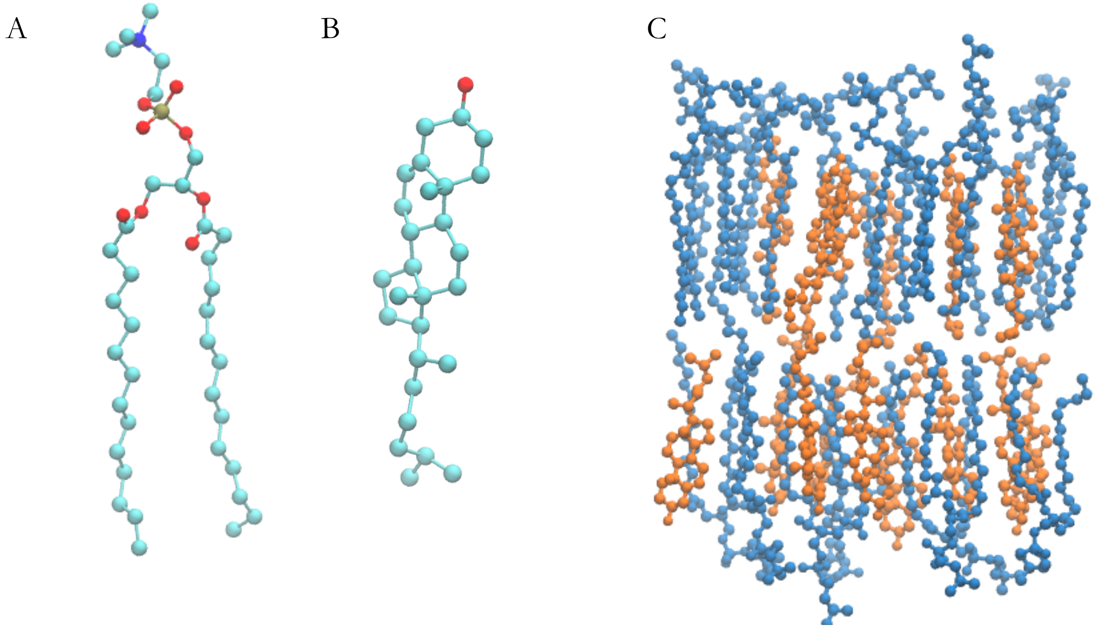

Figure 1-1: Example structure of a phospholipid (A), sterol (B), and generic bilayer of DMPC and cholesterol
Example configurations of 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) (A) and cholesterol (B) with carbon atoms in cyan, oxygen atoms in red, nitrogen atoms in blue, phosphorus atoms in gold, and hydrogen atoms omitted. A liquid order bilayer of DMPC (blue) and cholesterol (orange).

Lipids stabilize a variety of phases with hydrophilic head groups pointing toward an aqueous environment and the lipid tail groups forming a hydrophobic core. Lipid structure is a key indicator of preferred phase. Indeed, a ratio of total tail group volume to tail group surface area is correlated with the packing of the lipid, which in turn affects the stability of various phases.(Boal, 2012) For example, if the volume of the tail group is much smaller than the surface area, then the individual lipid can be approximated as a conical shape and a micellar phase is likely most stable. However, the majority of lipids in human cells are well approximated by a cylindrical shape and subsequently, stabilize a flat bilayer phase. However, within the bilayer phase, there are the liquid ordered, liquid disordered, and ripple phases and often, a coexistence between them. (Veatch and Keller, 2005) Again, lipid tails provide an empirical basis for predicting phase behavior. Generally, the lipids with longer, unsaturated (i.e., no double bonds) tend to be more ordered. Thus, lipid structure and specifically the tail groups have a significant role in phase behavior.

While the lipid tail groups make up a hydrophobic core of a bilayer, head groups face outward to the environment to form an interface and signaling platform. The most common head group is phosphatidycholine (PC) group, which is zwitterionic and relatively small, but in order to recruit membrane proteins, lipid head groups vary in charge and size to facilitate specific membrane protein interactions. Phosphatidylinositol (PI) head groups are a common messenger because the inositol sugar ring can be phosphorylated in different positions to facilitate specific binding interactions as well as nonspecific electrostatic interactions due to a localized negative charge. The PI head group is also quite large compared to PC head group and it reaches farther into the solvent. Indeed, the large head group contributes a spontaneous curvature in the membrane (i.e., preference for curved bilayer). The larger the head group the larger the deviation from a cylindrical lipid, and a more significant preference for curved membranes with a higher hydration of the large head group.(Callan-Jones et al., 2011) Thus, the head group size also contributes to its role as a messenger molecule because the lipid will localize to areas of curvature.

Inside the hydrophobic core of the bilayer, sterols localize and interact with the lipid tail groups of phospholipids to affect the phase behavior and subsequently, mechanical properties of the membrane. Sterols are nonpolar lipids with ring structures that cause packing defects in ordered lipids and conversely, cause ordering in disordered lipids.(Rog et al., 2009) For example, 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) has two 14 carbon lipid tails and forms a disorder bilayer phase at physiological temperature. When DMPC forms a bilayer with at least 10 mol% cholesterol, cholesterol will change the packing of the lipid tails, reducing the permeability of the membrane and induce a phase transition to an ordered bilayer phase, shown in Figure 1-1C. The transition from disordered to ordered phase increases the bilayer thickness and increases the membrane rigidity, which serves an important physiological function. Transmembrane proteins that span the hydrophobic core of the bilayer can sort based on the membrane thickness.(Simons and Sampaio, 2011) Thus, cholesterol has a crucial role as a signaling molecule and specifically, in the recruitment of the transmembrane proteins, which is further reinforced by the complex cellular apparatus to control cholesterol levels in the cell.

Deformations of the Lipid Membranes

Membrane deformations are generated throughout the cell and is required for many cellular functions. However, in the absence of cellular machinery, a membrane composed of cylindrical lipids forms a flat bilayer phase. Thus, a barrier to macroscopic changes in membrane shape are rooted in the microscopic behavior of individual lipids and the aqueous environment. Lipid behavior is not the only barrier to cellular membrane deformation. Indeed, membrane proteins and the cytoskeleton may prevent or facilitate membrane deformations, but significant insight into cellular membrane behavior can be gained by first considering simpler lipid bilayers. The microscopic environment of lipids in a bilayer gives rise to behavior approximated by a quasi-2D elastic sheet and provides a basis for understanding into two relevant mechanical stresses: stretching/compressing in the plane of the bilayer and bending out of the plane.

Cellular membranes are constantly under tension. Osmotic pressure, direct and indirect interactions to membrane proteins, and external substrates can all stretch the membrane.(Gauthier et al., 2012) However, lipid bilayers composed of cylindrical lipids are at equilibrium when at zero tension and resist the stretching due to the environment. In a bilayer, stretching causes lipids to pull away from each other, and subsequently, increased solvation of the head groups as well as the hydrophobic core of the bilayer.(Evans and Rawicz, 1990) Regions where the hydrophobic tails can be directly solvated by the aqueous environment are energetically unfavorable defects in bilayer and present a significant energetic barrier to tension on the membrane.

In the cellular environment, there are a variety of noncylindrical lipids. One such example is cardiolipin, which is a large lipid with four lipid tails and is highly concentrated in mitochondrial membranes. Its conical shape gives rise to curvature sorting behavior and induces a spontaneous curvature in lipid bilayers.(Beltran-Heredia et al., 2019; Tian and Baumgart, 2009) The spontaneous curvature of cardiolipin is anomalous compared to the majority cylindrical lipids that prefer zero curvature and resistant bending. Curved bilayers composed of cylindrical lipids are unfavorable due to stretching and compression of opposite leaflets. Thus, the microscopic basis for resistance to out of plane deformations is related to the stretching deformations. Bending a lipid bilayer causes one leaflet to be stretched, forcing head groups farther apart and allowing water to solvate the hydrophobic core while the opposite leaflet is compressed, forcing head groups together and also inducing unfavorable solvation.

1.3 I-BAR Domains

The Bin/Amphipphysin/Rvs (BAR) domain superfamily regulates membrane shape in a range of cellular processes ranging from endocytosis to filopodia formation.(Frost et al., 2009; Mim et al., 2012; Simunovic et al., 2019; Simunovic et al., 2015) Across the superfamily, full-length BAR proteins may have actin-binding domains, specific lipid targeting domains, or terminal amphipathic helices that are key to the proper function and localization of BAR proteins.(Salzer et al., 2017) Notably, BAR domains share a consistent dimeric structure of bundled, kinked helices, and net positive charge. Among the many factors that distinguish BAR domain function are the membrane-binding interface and protein intrinsic curvature. The three families comprising the BAR superfamily are the N-, F- and I-BAR families: N- and F-BAR domains have similar concave membrane-binding interfaces while I-BAR domains are distinct in their convex binding surface.(Mim and Unger, 2012) Subsequently, the function of I-BAR proteins is dissimilar as it is responsible for forming extracellular membrane protrusions (e.g., filopodia) instead of forming cytosolic membrane structures.(Ahmed et al., 2010; Linkner et al., 2014; Mattila and Lappalainen, 2008; Millard et al., 2005; Simunovic et al., 2015) Furthermore, I-BAR protein function is crucial to cell physiology, which is demonstrated by I-BAR proteins' central role as drug targets in metastatic cancer, and malformed or non-existent filopodia having a role in neurological disorders including Autism Spectrum Disorders. (Chen et al., 2012; Kang et al., 2016)

Experimentally, the mechanism by which I-BAR domains induce membrane curvature has been interrogated by *in vivo* and *in vitro* studies. *In vivo* experiments show that I-BAR domains are essential to protrusion formation and the over expression of I-BAR proteins results in the formations of membrane protrusions from the cell.(Mattila et al., 2007) *In vitro* experiments show I-BAR domains will sort into membrane tubules of a preferred curvature as well as cause membrane deformation and tubulate giant unilamellar vesicles (GUVs).(Barooji et al., 2016; Prevost et al., 2015) The onset of tubulation has been shown to be affected by membrane composition with phosphatidylinositol 4,5-bisphosphate (PIP_2) playing an important role in lowering the necessary surface density of I-BAR domains required to induce GUV shape change.(Chen et al., 2015) Indeed, mutagenic and fluorescence studies has been shown that I-BAR domains can cluster negatively-charged PIP_2 due to the positively-charged residues present on the ends of I-BAR domains.(Mattila et al., 2007; Saarikangas et al., 2008; Saarikangas et al., 2009)

Chapter 2: Modeling Approaches to Biological Membranes

Large membrane shape changes occur due to both local interactions between membrane proteins and individual lipids, and micron-spanning many-body effects such as fluctuation-based membrane-mediated interactions. Thus, membrane remodeling is an inherently multiscale phenomenon.(Baumgart et al., 2011; Zimmerberg and Kozlov, 2006) Unfortunately, a model that correctly captures the interactions between proteins and lipids with chemical specificity would be computationally intractable. Thus, there are methodologies aimed at atomistic, mesoscopic, and continuum scales and the following chapter delves into the various approaches taken to understand membrane remodeling. Figure 2-1 shows the time and length scales relevant to the different classes of models, which are used throughout this thesis: All-atom models that explicitly represent each atom of a system, coarse-grained (CG) models that represent groups of atoms or lipids as CG particles or beads, and continuum models that represent groups of lipids or I-BAR domains using a mean-field or density field approach.

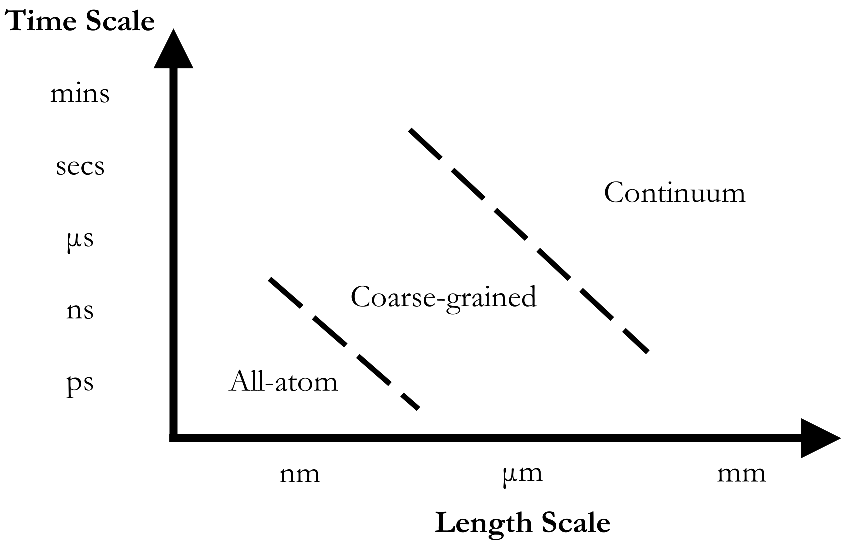

Figure 2-1: Time and length scale of models
All-atom, coarse-grained, and continuum models span different time and length scales from picoseconds to minutes and nanometer to millimeter models.

2.1 All-atom Models

All-atom molecular dynamics (MD) simulations are a conventional tool for modeling structure and dynamics with chemical specificity. Shown in the bottom left of Figure 2-1, all-atom simulations can give significant detail on the order of nanometers and nanoseconds while resolving specific interactions between chemical moieties. When used in concert with experiments, all-atom simulations are a powerful tool for molecular insights into macroscopic observables.

MD is a method for sampling molecular conformations by evaluating the dynamics of every atom in the system generating a trajectory of the particles through phase space, which can be related to experimental properties.(Allen and Tildesley, 1989; Smit and Frenkel, 2001) At each time step, the forces on every particle are evaluated and the positions are propagated using an integrator. The integration is repeated over and over to generate new positions, velocities, and forces over a time specified by the modeler. By assuming the ergodic hypothesis, a time average over the course of the trajectory can be related to thermodynamic properties and experiment providing molecular insights to experimental phenomena.

A crucial ingredient of an all-atom MD simulation is the force field, the functions and corresponding parameters that describe the interaction between atoms. There have been numerous lipid and protein force fields over the years.(Leonard et al., 2019; Lopes et al., 2015; Lyubartsev and Rabinovich, 2016) In general, all-atom force fields use simple (or computationally efficient) functional forms to describe the bonded and non-bonded interactions in the system. Their parameters determined using ab-initio calculations and sometimes tuned to experimental values. Force fields are often tested for correctness by comparing simulation results with an experimental observable of interest. For lipid systems, lipid tail order parameter, area per lipid, area compressibility, and bending modulus are common comparison. A protein model may be validated using the backbone behavior and the solvent accessible surface area may be more important.(Huang and MacKerell, 2013)

Limitations of All-atom Models

All-atom MD simulations are limited by their accuracy of the statistics generated. Thus, the force field used to model the system and the computational power required to converge the sampling of the properties of interest pose the most significant barriers to their broad application. Computationally, the current limitations lie around 1 million atoms for microseconds of the simulation time. In addition, statistical sampling of an event requires it to occur numerous times or rather, to quantify a property from an all-atom simulation, it calculated multiple times to acquire sufficient statistics. An entire subfield of enhanced sampling was born to overcome sampling barriers in the simulation and thoroughly review elsewhere.(Miao and McCammon, 2016; Mori et al., 2016; Yang et al., 2019) If statistical sampling can be acquired using enhance sampling, the remaining limitation is the accuracy of the samples generated.

Issues with all-atom force fields are being discovered as they are applied to new, increasingly complex phenomenon. In recent years, all-atom forced fields incorrectly predicted the structure intrinsically disordered proteins leading to the development of force fields geared toward unstructured or unfolded proteins.(Huang et al., 2017; Liu et al., 2018; Wu et al., 2018) Similarly, the number of polarizable lipids force fields has increased to correct issues related to the permeability of small molecules.(Chu et al., 2017; Lemkul et al., 2016; Li et al., 2017) These two examples serve to highlight issues in force field accuracy as they are applied to new systems. As is true for any model, any new application of an existing force field requires validation and should be connected to experimentally observed behavior before it is to be believed.

Applications of All-atom Models to I-BAR Domains and Membrane Remodeling

When applied to membrane remodeling, all-atom simulations are a useful tool for understanding the interactions of an individual protein or a few short peptides with lipid bilayers. Indeed, with only a single protein or few peptides, all-atom simulations can provide significant insight into mechanisms of membrane remodeling. N- and F-BAR simulations show that a single, isolated BAR domain is sufficient to drive local deformations.(Blood and Voth, 2006; Yu and Schulten, 2013) Additionally, atomistic simulations show amphipathic helices sense curvature, but may not generate curvature on their own.(Arkhipov et al., 2009; Cui et al., 2011; Cui et al., 2013) Various experimental observation agree that amphipathic helices help BAR domains sense positive curvature.(Chen et al., 2016; Fernandes et al., 2008; Jao et al., 2010) All-atom studies showcase two relevant mechanisms of curvature sensing and generation: direct scaffolding and insertion. N- and F-BAR proteins are scaffold through interacting directly the membrane and enforcing its intrinsic curvature upon the membrane while small helical structures with distinct hydrophobic and hydrophilic faces can sense membrane curvature by inserting into the membrane and finding areas of increased lipid defects.

All-atom simulations of isolated I-BAR domains show slightly different behavior. Isolated I-BAR domains in solution and bound to the membrane lose a significant amount of their intrinsic curvature seen in crystal structures.(de Groot et al., 2011; Millard et al., 2005; Takemura et al., 2017) Thus, I-BAR domains may remodel through a direct scaffolding mechanism at higher I-BAR domain surface densities, but an isolated I-BAR domain does not behave like F- and N-BAR simulations; All-atom simulations of I-BAR domain predict that an I-BAR domain flattens and loses the intrinsic curvature characteristic to the BAR domain family.

2.2 Coarse-grained Models

The goal of a CG model is to overcome the computational barriers that plague higher resolution models and to reach longer time and length scales. The computational cost of a CG system can be lower than its all-atom counterpart because a CG model will likely use fewer particles, use less complex/cheaper interactions, exhibit faster dynamics, or utilize larger integration time steps to describe the same system. (Gershenson et al., 2020; Ingólfsson et al., 2014; Noid, 2013; Pak and Voth, 2018; Saunders and Voth, 2013; Voth, 2008) The essential components of all-atom and CG models are the same: both comprise a force field and a representation. In the all-atom case, force fields are readily available, as many groups make a concerted effort to parameterize and validate all-atom force fields. At a CG resolution, the MARTINI(Bruininks et al., 2019; Marrink et al., 2004; Marrink et al., 2007; Marrink and Tieleman, 2013) and Go(Hills and Brooks, 2009; Takada, 2019; Ueda et al., 1978) models, which are parameterized at "4 heavy atoms per bead" and 1 site per amino resolutions, have had several successes in describing lipid bilayers and protein folding, respectively. However, CG models parameterized for a particular system at specific conditions are difficult to find due to issues of transferability. When considering a CG model at a lower resolution than one amino acid per bead, there is a lack of widely used and accepted models, but multiple possible methods are able to create a model at the desired resolution. Here, I provide a brief over view of these methods. There are two main approaches to coarse-graining, bottom-up and top-down, shown schematically in Figure 2-2.

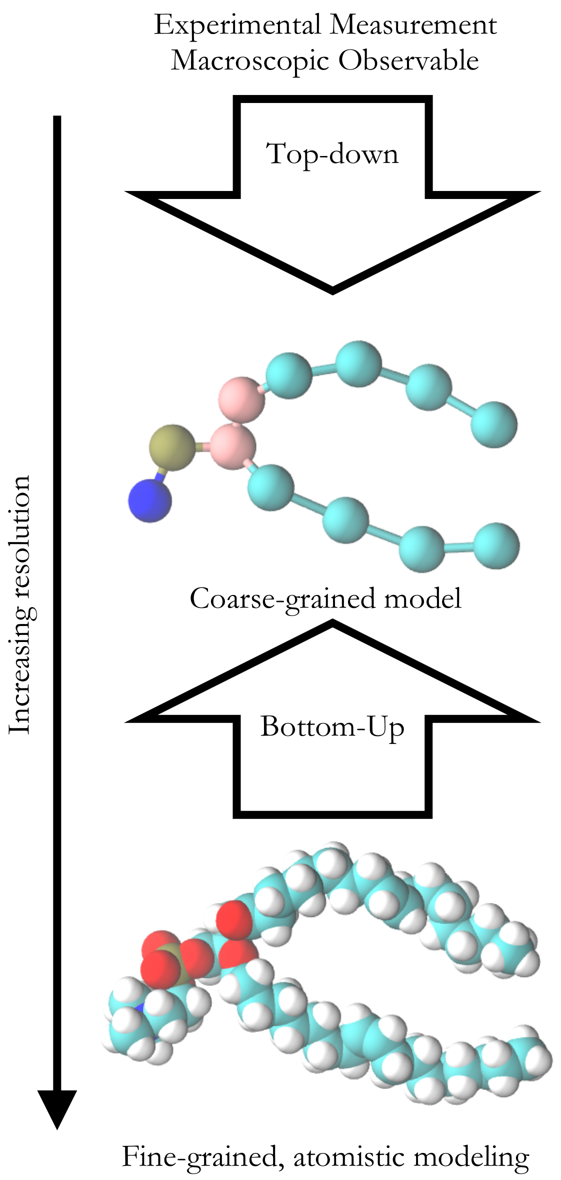

Figure 2-2: Schematic of coarse-grained methodologies
Top-down and bottom-up strategies parameterize CG models using different inputs. A bottom-up scheme uses atomistic models and a top-down scheme fits the CG model to experimental measurements and macroscopic observables. A DOPC molecule is shown to highlight difference in representations at the atomistic and MARTINI resolutions.

Bottom-up Approaches

Bottom-up approaches seek to systematically leverage higher resolution data for the parameterization and validation of a CG model. The goal of a bottom-up CG model is to retain as much information from the atomistic system as is necessary, while also extrapolating to ensembles different from the data set used during parameterization. As previously mentioned there are two components to a CG model: the representation and the force field.(Noid, 2013; Saunders and Voth, 2013; Voth, 2008) The representation should retain the important characteristics of the underlying all-atom system. For example, if the dipole of a molecule is a crucial to the CG phenomenon of interest, a potential mapping should incorporate charge.(Cao and Voth, 2015) In a bottom-up coarse-graining approach, the mapping (i.e., the coordinate transformation linking the atomistic to CG) is a key component of a coarse-graining procedure. The mapping is the mathematical representation of the assumptions inherent to the CG representation. For example, if resolving each solvent molecule is not necessary to CG phenomena, a mapping operator may not map any molecules to a CG bead (i.e., the CG model would not explicitly represent the solvent). The improvement of mapping operators is on-going as machine learning and other data driven approaches are used to quantify their effect on CG accuracy.(Chakraborty et al., 2018; Webb et al., 2019)

CG force fields are similar to their all-atom counterparts in that they are composed of bonded and nonbonded portions that describe interactions within and between molecules, respectively. There are a variety of bottom-up CG parameterization schemes to capture low-dimensional structural correlations or mapped forces. The choice of CG force field parameterization scheme contains inherent assumptions about the specific details of the underlying reference system the modeler hopes to recapitulate. Two examples of bottom-up methods to fit CG potentials are the multiscale coarse-graining method(Izvekov and Voth, 2005), which minimizes the ensemble average of the squared difference between mapped all-atom and CG forces and the relative entropy minimization method(Shell, 2008), which minimizes the relative entropy between mapped all-atom and CG distributions. Although these methods ideally produce the same CG force field with infinite sampling of all-atom statistics and a complete basis set (i.e., sufficiently complex force field), when applied to complex self-assembly the two methods have shown to differ greatly.(Pak et al., 2019; Rudzinski and

Noid, 2011) Thus, the parameterization scheme for the CG potential remains a crucial component of CG model creation and together, with the choice of mapping operator and basis set for the CG force field, encodes the assumptions of the model while incorporating information from higher resolution simulations.

Top-down Approaches

In contrast, top-down approaches seek to match the properties of a coarse-grained system to known macroscopic or experimental observables. For example, many protein models like the Go model static data from crystal structures as their primary target.(Takada, 2019) Similarly, a top-down lipid model might target mechanical properties like the bending modulus(Brannigan et al., 2005; Cooke et al., 2005), surface tension(Shinoda et al., 2010), or the octanol-water partition coefficient(Marrink et al., 2004). Similar to the bottom-up approach, top-down methodologies require a representation and a force field. However, these are no longer parameterized using finer-grained simulation data. The representation can be chosen to recapitulate a property or experimental phenomenology instead of being tailored to capture behavior from all-atom simulations. For example, if chemical specificity is required, a 3-site lipid model may not give sufficient detail because a head group is represented by a single bead. Instead a higher-resolution CG model (e.g., MARTINI) would be more appropriate. By tuning resolution to the problem, hypotheses about the essential physics of a phenomenon of interest can be tested and the relevant information captured.

Top-down force fields are similarly tuned or fit to reproduce behavior of interest. For example, MARTINI 1.0 lipids were represented by a limited set of 5 different Lennard-Jones (LJ) parameters (σ and ε) and were expanded to 10 different Lennard-Jones parameters to capture a broader chemical variety of lipids and corresponding lipid-lipid interactions.(Marrink et al., 2004; Marrink et al., 2007) Similarly, Shinoda-DeVane-Klein (SDK) lipids similarly use a 12-4 LJ potential as opposed to the standard 12-6 LJ potential to correctly capture the interfacial tension.(Shinoda et al., 2007; Shinoda et al., 2012) The force field and representation are tuned to create a minimal model that captures the necessary physics to reproduce a macroscopic or experimental observable.

Limitations of Coarse-grained Models

The application of coarse-grained models is typically hampered by limitations in the data needed for parameterization. Obviously, top-down models lacking experimental data cannot be adequately tuned and bottom-up models without sufficient atomistic simulations cannot be fit. For example, SDK lipids are fit to capture the interfacial tension and density of lipid systems; however, they were not parameterized with protein data and as a result, proteins cannot be simulated. Thus, the applicability of this model to an arbitrary system is significantly limited.(Marrink et al., 2019) Modular force fields partially ameliorate this weakness by providing mixing rules to adapt parametrizations to unseen chemical setting; however, these models suffer as said mixing rules are approximate or *ad* hoc. For example, LJ parameters in atomistic force fields are often mixed between two interacting atoms types to get an arbitrary LJ potential (using, as an example, the geometric mean of ε between two interacting atoms to parameterize the interaction potential). However, the appropriateness of additive behavior between two CG force fields is not apparent and should not be expected.(Alessandri et al., 2019) At their core, these are issues of transferability, an issue that plagues all CG models. When using a CG force field at conditions outside of its original parameterization, the model should be fully validated and may require reparameterization.

Rigorous studies of bottom-up methods have identified two major issues in CG parametrization: Transferability and representability; which encompass the issues described above. Generally, the transferability problem relates to the application of a model in conditions far from the original parameterization and limits the extensibility to different thermodynamic state points.(Jin et al., 2019) Separately, representability issues are rooted in the thermodynamic inconsistencies in CG models.(Wagner et al., 2016) Perhaps the most common example is that structure-based CG models incorrectly capture the pressure and the isothermal compressibility due to conflicting definitions of observables and the state dependence of CG force fields.(Johnson et al., 2007) These issues have been investigated in bottom-up models. Chapter 5: is dedicated to investigating these issues in top-down models.

Applications of Coarse-grained Models to I-BAR Domains and Membrane Remodeling

Other than the results shown in Chapter 3:, bottom-up modeling has not been applied to I-BAR domains. However, a bottom-up N-BAR domain model does exist, and significant insight can be gleaned from its several applications.(Simunovic et al., 2016; Simunovic et al., 2017; Simunovic et al., 2013; Simunovic and Voth, 2015) CG predictions, followed up by atomic force microscopy experiments, show N-BAR domains form linear aggregates at low surface densities. In addition, CG simulations of N-BAR domains show that surface tension plays a significant role in membrane-mediated forces, in agreement with phenomenological models of curved membrane protein aggregation.(Reynwar et al., 2007) Finally, CG simulations of N-BAR domains predict the formation of spirals of N-BAR domains around the outside of membrane tubules.

Top-down particle-based CG models have been used to understand the membrane-bound behavior of BAR superfamily. CG models of F-BAR domains show the importance of negatively charged lipids like phosphatidylinositol 4,5-bisphosphate, which can drastically affect membrane binding behavior of an F-BAR domain.(Stanishneva-Konovalova and Sokolova, 2019) Top-down CG models of N- and F-BAR domains have been parameterized using shape-based coarse-graining, and in combination with atomistic studies have elucidated a direct scaffolding mechanism deformation that is sufficient to stabilize membrane tubules at higher BAR domain densities.(Arkhipov et al., 2008, 2009; Yu and Schulten, 2013) Separately, a top-down CG model of I-BAR domains was used to show that the aggregation and collective membrane deformation of many I-BAR domains is directly dependent on the local curvature induced by a single protein.(Noguchi, 2014; Olinger et al., 2016)

When the I-BAR domain model was combined with a top-down CG model of N-BAR domains, phase segregation occurs and positive and negative curving proteins to colocalize.(Noguchi and Fournier, 2017) However, this model lacked an explicit protein representation, which limits the insights from this phenomenological model, as the crowding and individual flexibility of I-BAR domains are important factors in I-BAR domain-mediated membrane remodeling.

2.3 Continuum Approaches

Biological membranes can be treated as quasi-2D materials using a mean-field approach. An analytical description of the free energy of the membrane can be minimized to give the equilibrium state under various conditions. Through a field theoretic approach, continuum models can elucidate membrane behavior at much long time and length scales than the models discussed thus far (see Figure 2-1). Canham-Helfrich theory describes the membrane as thin elastic sheets, penalizing membrane curvature.(Canham, 1970; Helfrich, 1973)

$$E_{bend} = \int dA \left[\frac{1}{2}\kappa(c_1 + c_2) + \bar{\kappa}c_1c_2\right] \qquad 2.1$$

where A is a differential surface area element, $c_1$and c_2 are the principal curvatures of the surface, and κ and $\bar{\kappa}$ are the mean and gaussian bending moduli. κ and $\bar{\kappa}$ can be extracted from various experiments such as neutron scattering or flicker spectra.(Faucon et al., 1989; Marsh, 2006; Pan et al., 2009; Rawicz et al., 2000) When applied to cells and biological membrane, convention is that principal curvature are positive when curving toward the inside of the cell or toward a binding protein i.e., N- and F-BAR domains generate positive curvature.

Two common changes to the Canham-Helfrich free energy are the simplification of the gaussian curvature, which remains constant over closed surface (i.e., Gauss-Bonnet Theorem) and a term to describe energetic penalties to changes in surface area (shown below).

$$E_{stretch} = \sigma \int dA \tag{2.2}$$

where A is a differential surface area element and σ is the area compressibility modulus. A penalty to changes in surface area is a necessary addition curvature inducing process also induce membrane tension. Here, the continuum approach assumes elastic penalties to membrane perturbations and deformed surfaces will return to the zero-energy state (i.e., flat, tensionless surface).

As described in Chapter 1.2, resistance to bending and stretching comes from molecular interactions between lipids in the solvent. Indeed, the area compressibility and the bending moduli can be quantified from molecular simulations. Commonly, the area compressibility modulus is related to fluctuations membrane area or more specifically, is inversely proportional to the variance in the area per lipid at zero tension.(Hofsäß et al., 2003) The bending modulus can be fit to the low wavenumber behavior of the height fluctuation spectrum, which is a discrete Fourier transform of the bilayer midplane.(Brandt et al., 2011) As a result, continuum modeling is useful in relating molecular scale behavior to experimental quantities.

There are various methods to determine the minimum energy configuration of a surface described by Canham-Helfrich theory. Triangular bead-spring networks are a common description of the membrane because meshes are quickly evaluated. (Bahrami et al., 2012; Bonazzi et al., 2020; Šarić and Cacciuto, 2012; Šarić and Cacciuto, 2013) The configurations of triangular mesh are efficiently sampled using Monte Carlo schemes, which allow for significant changes in curvature through dynamic reconfiguration of the mesh. Protein density fields that induce curvature can be added to determine low energy membrane configurations using an efficient, implicit protein model.(Ramakrishnan et al., 2013; Tourdot et al., 2015; Tourdot et al., 2014) Thus, regions of the membrane with high curvature are not due to an explicit protein representation like that of the CG models discussed above, instead the highly curved regions are due to protein fields. Evaluating the protein field can reduce the overall computational cost allowing simulations to reach micron length scales.

Another family of methods are based on dynamical sampling of the membrane and do not rely upon meshes and Monte Carlo sampling. The original elastic membrane (EM) model was based on smooth particle applied mechanics, similar to the smoothed particle hydrodynamics and subsequently improved to include protein and lipid concentration fields. (Ayton et al., 2007; Ayton et al., 2009; Davtyan et al., 2017; McWhirter et al., 2004) The EM model and its extensions, EM2 and Mesoscopic Membrane with Proteins (MesM-P), stand apart from other continuum membrane models by using a grid-free approach and accounting for solvent rearrangement through hydrodynamic effects.

Limitations in Continuum Modeling

The continuum models discuss thus far are based on Canham-Helfrich theory, which assumes the membrane can be described as an elastic 2D sheet. This description is useful in many cases; however, it does have drawbacks when biological membrane undergoes inelastic deformation or when the length scales of membrane modeling are on the order of the thickness of the membrane. These three assumptions can commonly be violated in biological conditions. For example, inelastic membrane deformation can occur as lipid species reconfigure via flip-flop or curvature sorting to accommodate membrane remodeling. Another example is curvature induced by a BAR domain. This phenomenon inherently occurs on the length scale of bilayer thickness. Subsequently, any description of membrane-mediated interactions on order of a few nanometers is a violation of a key assumption of Canham-Helfrich theory.

Applications of Continuum Models to I-BAR Domains and Membrane Remodeling

Biological membranes mediate effective interactions between membrane-bound proteins. Peripheral membrane proteins cluster to induce minimal curvature and to minimally perturb thermal fluctuations.(Johannes et al., 2018) Casimir-like forces are always attractive, but only at close ranges. At longer distances, fluctuation-based forces are not strong enough to induce clustering. Forces due to curvature-coupling are typically only attractive for anisotropic particles and the length scale of these interactions depends on the mechanical properties of the membrane (e.g., area compressibility and bending modulus).(Dan, 2017; Schweitzer and Kozlov, 2015; Schweitzer et al., 2015) Membrane-mediated interactions require long length scales to sampled and accounted for. Thus, continuum models are able to give significant insights into many-body effects that cannot be fully sampled at higher resolutions.

Continuum models provide insights into the organization and assembly of membrane-bound BAR domains. In the presence of tubulating forces, continuum model show that curved I-BAR domains will sort to regions of curvature and form circumferential rings inside of membrane tubules. (Mesarec et al., 2016) Their geometric argument for I-BAR domain organization agrees with higher resolution CG models resolution CG models without tubulating forces.(Noguchi, 2016) Similarly, continuum models of I-BAR domains show that side-by-side aggregation to be the most stable configurations.(Bonazzi et al., 2020; Davtyan et al., 2017) These continuum models are exceedingly useful for predicting the minimum energy assembly of I-BAR domains when the protein curvature is constant and the same as crystallographic structures.

Chapter 3: Bottom-up Modeling of I-BAR Domain Organization

This chapter is adapted from Reference (Jarin et al., 2019).

3.1 Summary

Protein-mediated membrane remodeling is a ubiquitous and critical process for proper cellular function. Inverse Bin/Amphiphysin/Rvs (I-BAR) domains drive local membrane deformation as a precursor to large-scale membrane remodeling. We employ a multiscale approach to provide the molecular mechanism of unusual I-BAR domain driven membrane remodeling at low protein surface concentration with near-atomistic detail. We generate a bottom-up coarse-grained (CG) model that demonstrates similar membrane-bound I-BAR domain aggregation behavior as our recent Mesoscopic Membrane with Explicit Proteins (MesM-P) model. Together, these models bridge several length scales and reveal an aggregation behavior of I-BAR domains. We find that at low surface coverage, i.e., low bound protein density, I-BAR domains form transient, tip-to-tip strings on periodic flat membrane sheets. Inside of lipid bilayer tubules, we find linear aggregates parallel to the axis of the tubule. Finally, we find that I-BAR domains form tip-to-tip aggregates around the edges of membrane domes. These results are supported by *in vitro* experiments showing low curvature bulges surrounded by I-BAR domains on giant unilamellar vesicles (GUVs). Overall, our models reveal new I-BAR domain aggregation behavior in membrane tubules and on the surface of vesicles at low surface concentration that add insight into how I-BAR domain proteins may contribute to certain aspects of membrane remodeling in cells.

3.2 Introduction

Here, we use a multiscale approach combining CG and mesoscopic models (specifically, Mesoscopic Membrane with Proteins (Davtyan et al., 2017), or MesM-P) to understand the organization of I-BAR domains on membranes of various geometries.(Saunders and Voth, 2013) This multiscale approach is uniquely suited to capture protein-mediated membrane remodeling because it incorporates the interplay between near nanometer to several micron length scales. Broadly speaking (Saunders and Voth, 2013), CG models can be derived from finer-resolution, atomistic simulations (bottom-up) or by specifically reproducing experimental observables (top-down). Here, the former approach is adopted such that our protein CG model reproduces structural fluctuations observed in atomistic simulations while our lipid CG model recapitulates the properties of a representative lipid bilayer membrane. Additionally, the MesM-P model, which is lower resolution than the CG model, can be used to simulate significantly larger systems while remaining consistent with the membrane-bound protein aggregation behavior of our CG model. Together, these models span near atomistic to mesoscopic length scales, and the agreement of the two models (parameterized using separate methodologies) indicates the robustness of the CG phenomena we observe. We apply these models to understand the aggregation behavior of I-BAR domains on flat sheets that mimic the surface of giant unilamellar vesicles, membrane tubules, and spherical membrane vesicles. We focus primarily on lower density surface coverage.

Our results show that both I-BAR domain models form local membrane troughs on flat sheets and, when I-BAR domains are on membranes with global curvature, the I-BAR domains orient to minimally perturb the membrane. We predict tip-to-tip string aggregates in tubule-shaped membranes and rings at the base of bulges on the surface of vesicles. While the CG lipid and MesM-P models bear some differences, the behavior observed using the two models (each consistent with biologically relevant properties) is qualitatively similar. Both models show qualitatively similar behavior on surfaces with positive Gaussian curvature, providing a consistent mechanism for the formation of novel, low curvature deformations surrounded by I-BAR domains on giant unilamellar vesicles.

3.3 Methods

CG Model Details

The model consists of two components: a highly coarse-grained lipid bilayer and a coarse-grained IRSp53 I-BAR domain as shown in Figure 3-1. The highly coarse-grained lipid bilayer was parameterized using a hybrid multiscale coarse-graining (MS-CG) approach. The hybrid parameterization supplemented MS-CG forces from atomistic simulations with analytical CG potentials to describe the short-range interactions, which has been described previously (Srivastava and Voth, 2013). The CG model was simulated in the LAMMPS simulation engine.(Plimpton, 1995)

The coarse-grained IRSp53 I-BAR domain model was parameterized from atomistic simulation to reproduce the atomistic structure fluctuations (e.g., intrinsic curvature fluctuations), which we expect to be critical to the membrane remodeling process. The atomistic simulations composed of a lipid bilayer, IRSp53 I-BAR domain (PDB: 2YKT(de Groot et al., 2011)), and water using the CHARMM36 force field.(Best et al., 2012; Klauda et al., 2010; MacKerell et al., 1998) The initial configuration of the membrane was first generated from CHARMM-GUI(Jo et al., 2007; Jo et al., 2008; Jo et al., 2009; Lee et al., 2016; Wu et al., 2014) and equilibrated using the corresponding scheme and the GROMACS simulation engine.(Abraham et al., 2015) The membrane was composed of 75% 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 20% 1,2-dioleoyl-sn-glycero-3-phosphoserine (DOPS), and 5% Phosphatidylinositol-4,5-diphosphate (PI(4,5)P_2). Then, the I-BAR domain was added to the simulation cell and the hydration layer was increased with 150 mM NaCl using GROMACS tools.(Abraham et al., 2015) The I-BAR domain was simulated with the membrane for 300 ns and the final 100 ns were used to determine the mapping and parameterization of intraprotein forces. The map from atomistic to CG was found using essential dynamics coarse-graining (ED-CG)(Zhang et al., 2008), which divides the protein along its primary sequence. The ED-CG protocol map is determined by finding divisions in the primary sequence that preserve the dynamics of the protein, i.e., the large amplitude motions within a CG bead are minimized and the motion between beads are maximized. The divisions in the primary sequence are given in the SI. Effective harmonic potentials were used for intraprotein interactions with the parameters determined using a heterogeneous elastic network model.(Lyman et al., 2008) The spring constants and equilibrium distances are fit to reproduce the mapped structure and fluctuations from an atomistic MD trajectory and are provided in the SI.

The direct protein-protein interactions were purely repulsive to capture the excluded volume of each I-BAR domain. Additional screened electrostatic interactions between I-BAR domains were investigated in Section 3.7, and reinforce the aggregation behavior due to purely membrane-mediated attraction. The effective potential between the CG beads of the protein and the head group beads of the CG lipids was modeled as a 10-6 shifted force Lennard-Jones potential.(Allen and Tildesley, 1989) At the CG resolution, the effective potential between the protein and lipid membrane is, by its nature, a simplification of the complex electrostatic interactions at the atomistic protein-membrane interface and is meant to capture the local membrane deformation driven by an isolated, atomistically-resolved I-BAR domain.(Takemura et al., 2017) As further mechanistic studies of local membrane deformation by I-BAR domains are performed, the CG interactions could be refined to better reproduce the complex nature of membrane remodeling. Without an exhaustive study of I-BAR domain membrane remodeling at a finer resolution, the effective attraction strength was taken as a parameter and its effect on local deformation was quantified by a comparison to previous atomistic simulations.(Takemura et al., 2017)

Mesoscopic Membrane Simulations with Explicit Proteins

Here, we used a reduced and scaled down version of the recent MesM-P model (Davtyan et al., 2017) that relies on a discretized formulation of membrane elastic theory (Helfrich, 1973). In the original model, the membrane is represented as a collection of quasi-particles approximately 7 nm in diameter. Each quasi-particle describes a patch of lipid bilayer given its position and momentum with additional scalar fields to represent local protein concentration and lipid composition. As a result, MesM-P allows for efficient modeling of large-scale membrane shape changes, protein binding and unbinding, and their interplay on nearly experimental length and time scales (Davtyan et al., 2017).

In this chapter, we use only the elastic component of the MesM-P model that describes 3-dimensional membrane undulations and bending, i.e., without using mesoscopic solvent or implicit variables describing local protein concentration and lipid composition. Instead, we use an approach similar to the CG model and include explicit very highly CG representations of I-BAR proteins, which are modeled as five-bead linear chains with varied sizes shown in Figure 3-1. The size variation of the beads is included to, in part, reproduce the shape of an I-BAR domain. For the membrane, we also use smaller sized beads of approximately 3 nm, which is consistent with the width of I-BAR. We have used Lennard-Jones-like shifted force 4-2 potentials for membrane-protein interactions with various strengths (Allen and Tildesley, 1989). The MesM-P model was simulated in LAMMPS simulation engine (Plimpton, 1995).

The final set of simulations utilizes a guiding potential to replicate membrane configurations seen in the *in vitro* experiments. The flower petal structure in the latter (see Results) is likely caused by I-BAR domain driven membrane remodeling. This full phenomenon is not observable with the MesM-P model due to the use of periodic boundaries to describe the surface of a giant unilamellar vesicle. Instead the flat sheet is deformed using a spherical guiding potential. The membrane is initially flat, and the guiding potential is moved toward the surface to create a membrane deformation similar to the deformations made by CG I-BAR domains on small vesicles and the accompanying *in vitro* experiments.

More details and the parameters of each computational model are given in Section 3.7

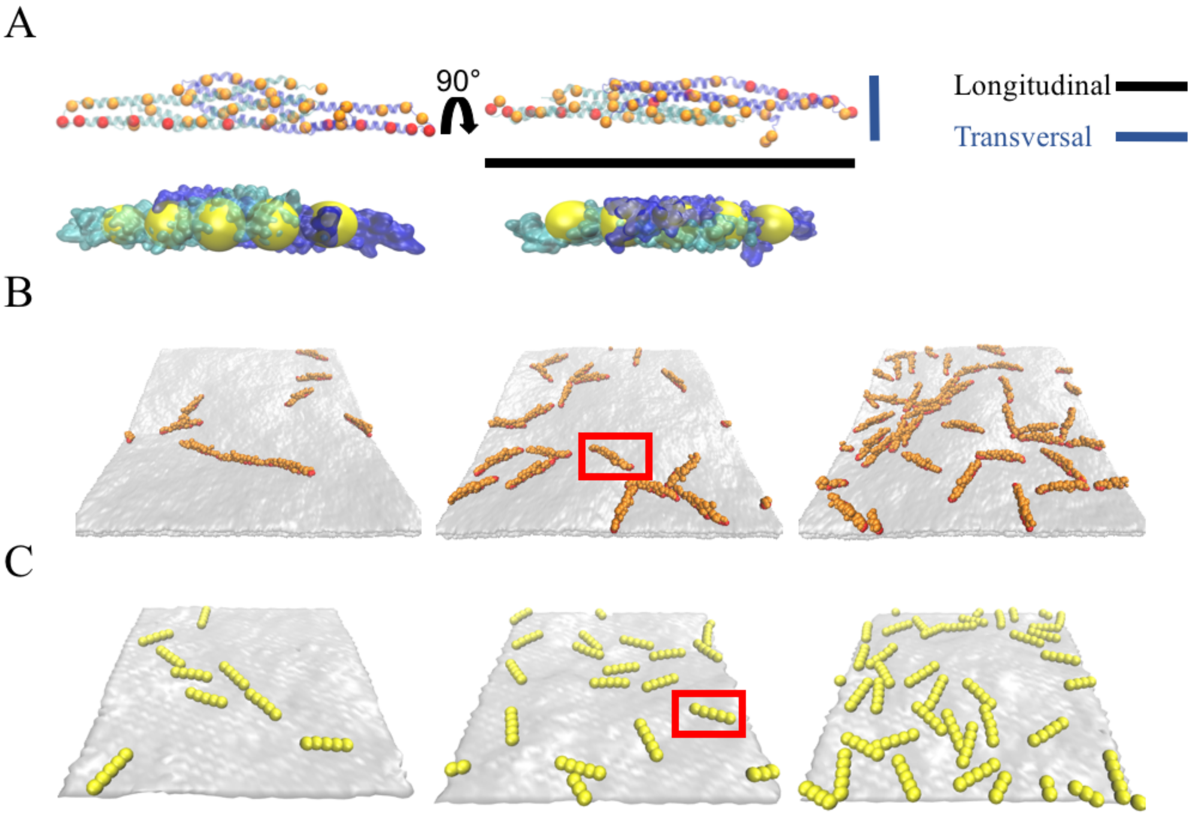

Figure 3-1: CG and MesM-P I-BAR domain representations and assembly on flat sheets. CG Side by side image x5of CG and MesM-P models, snapshots of flat sheet configurations. A) Side (left) and top (right) view of an overlay of CG I-BAR domain (orange, red) and the secondary structure of each monomer (cyan and blue) and an overlay.

Reagents

Total brain lipid extract (TBX, 131101P), brain L-α-phosphatidylinositol-4,5-bisphosphate (PI(4,5)P2, 840046P) and 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[biotinyl(polyethyleneglycol)-2000] (DSPE-PEG(2000)-biotin, 880129P) were purchased from Avanti Polar Lipids. BODIPY-TR-C5-ceramide, (BODIPY TR ceramide, D7540) and BODIPYFL C5-hexadecanoyl phosphatidylcholine (HPC*, D3803), Alexa Fluor 488 C5-Maleimide (Alexa 488) were purchased from Invitrogen. GloPIPs BODIPY® TMR-PtdIns(4,5)P2, C16 (TMR PIP2, C45M16a) was purchased from Echelon.β-casein from bovine milk (>98% pure, C6905) and other reagents were purchased from Sigma-Aldrich.

Protein Purification and Labeling

Recombinant mouse IRSp53 I-BAR domain was purified and labeled with AX488 dyes as previously described (Saarikangas et al., 2009).

GUVs Preparation and Observation

For all experiments, GUVs composed of brain total lipid extract (TBX) (Yu et al., 2006) supplemented with 5 mole% brain PI(4,5)P_2, 0.2 mole% DSPE-PEG(2000)-biotin and 0.5 mole% BODIPY TR ceramide were prepared by electroformation on platinum electrodes overnight at 4 °C in a physiologically relevant salt buffer. The salt buffer outside GUVs was 20 mM Tris pH 7.5, 60 mM NaCl and 100 mM sucrose. The salt buffer inside GUVs was 20 mM Tris pH 7.5, 60 mM NaCl and 100 mM glucose.

GUVs were incubated with IRSp53 I-BAR domain at a bulk concentration of 0.02-0.1 μM for at least 30 min at room temperature before observation. For all experiments, microscope slides and coverslips were washed with water and ethanol followed by passivation with a β-casein solution at a concentration of 5 mg/mL for at least 5 min at RT. GUVs were observed by Nikon Eclipse Ti microscope (Nikon) equipped with Yokogawa CSU-X1 confocal head, 100X CFI Plan Apo VC objective (Nikon) and QuantEM:512SC camera (Photometrics).

3.4 Results

Planar Membranes

First, we compared the two modeling approaches by separate simulations of both I-BAR domain models on tension-free, periodic flat lipid bilayers at various I-BAR protein surface densities. Infinite flat sheets are a close approximation to the surface of giant unilamellar vesicles (GUVs) that have quasi-null local curvature. The CG and MesM-P simulations demonstrated a preference for forming relatively linear aggregates, as shown in Figure 3-1B and C respectively.

At low surface coverage, transient short linear strings of I-BAR form in both the CG and MesM-P models (Figure 3-1). As surface coverage increases from 5% to 20%, the transient linear strings change into crowded strings where each I-BAR domain of the linear aggregate can switch orientation from one neighbor to another. We observe comparable behavior in both models, indicating our results are robust with respect to model resolution.

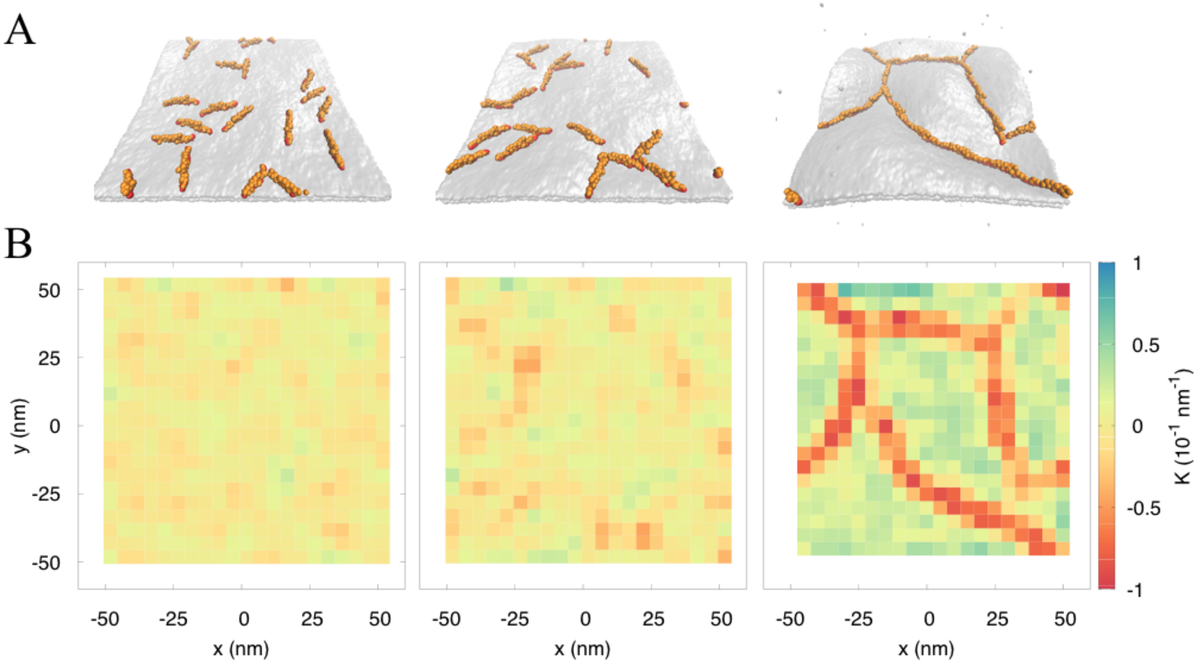

Figure 3-2: Effect of I-BAR-lipid interaction strength on local membrane curvature generate at 10% coverage.
A) Snapshots of CG model with increasing interaction strength between attractive I-BAR CG beads and lipid head bead. From left to right: the strengths increase from 0.5, 1.0 and 1.5 kcal/mol. B) Mean curvature calculated as a function of position on the membrane for the snapshots in panel A.

Next, we investigated the effects of protein-lipid interaction on protein organization and curvature generated. The effective protein-lipid interaction be changed in an in vitro experiment by changing the local concentration of negatively charged lipids or phosphoinositides. Figure 3-2 shows final protein configurations using a constant coverage of 10% as a function of protein-lipid interaction strengths. In Figure 3-2, we observe that at increased protein-lipid interactions strength, significantly more curvature is generated, and linear aggregates are much more likely to form. It is reasonable to expect that, as the local curvature increases, the subsequent aggregation also increases given the nature of membrane-mediated protein interactions (van der Wel et al., 2016). In the CG models parameterized here, membrane-mediated interactions between proteins can broadly defined as effective protein interactions caused by perturbations to the membrane after protein binding events. The response of the membrane to protein binding (e.g., induced curvature or dampened fluctuations) results in an effective protein interaction that drives membrane aggregation (Aranda-Espinoza et al., 1996; M. Goullian et al., 1993). In fact, we demonstrate here that membrane-mediated interactions are sufficient for protein aggregation (Simunovic et al., 2017; Simunovic et al., 2013; Simunovic and Voth, 2015).

We calculated the mean principal curvature generated on the x- and y- axes for each model to further quantify the effects of interaction strength on I-BAR domain organization. Both models exhibit curvature parallel and perpendicular to the longitudinal axis of the I-BAR domain (see Figure 3-1). In the presence of thermal fluctuations, deformation along either axis results in transient Gaussian curvature and variations in the principal curvature. Plots of mean principal curvature are shown in Figure 3-2B, corresponding to the snapshots of Figure 3-2A. With weak protein-lipid interactions, we see little to no aggregation and the I-BAR domains are disordered. With intermediate interactions, we find that there are linear aggregates and the mean curvature driven by isolated I-BAR domains to be around 0-0.5 10-1/nm, similar to the curvature observed near a single I-BAR domain in atomistic simulations.(Takemura et al., 2017) With stronger protein-lipid interactions, we observed complete linear aggregate formation and significantly more membrane deformation than atomistic simulation of isolated I-BAR domains, which is to be expected. The local membrane deformation is crucial to understanding I-BAR domain curvature sensing, and so in the following sections we address how globally curved surfaces affect the aggregation of I-BAR domains.

Tubular Membranes

We mimicked the experimental conditions of Prévost et al. (Prevost et al., 2015) by simulating each I-BAR domain model inside of a membrane tubule. In the experiments, tubules are formed by pulling on a micropipette-aspirated GUV using optical tweezers, with the curvature of the tubule controlled using the applied pressure in the pipette. In the tubule simulations, the surface tension is initially zero as the radius and length of the tubule are allowed to equilibrate before I-BAR domains are introduced to the system and, subsequently, bind to the lipid bilayer. After this initial equilibration, the length of the tubule is then held constant with the radius allowed to fluctuate as we seek to understand I-BAR organization in tubules pulled from GUVs, which are not tension-free.

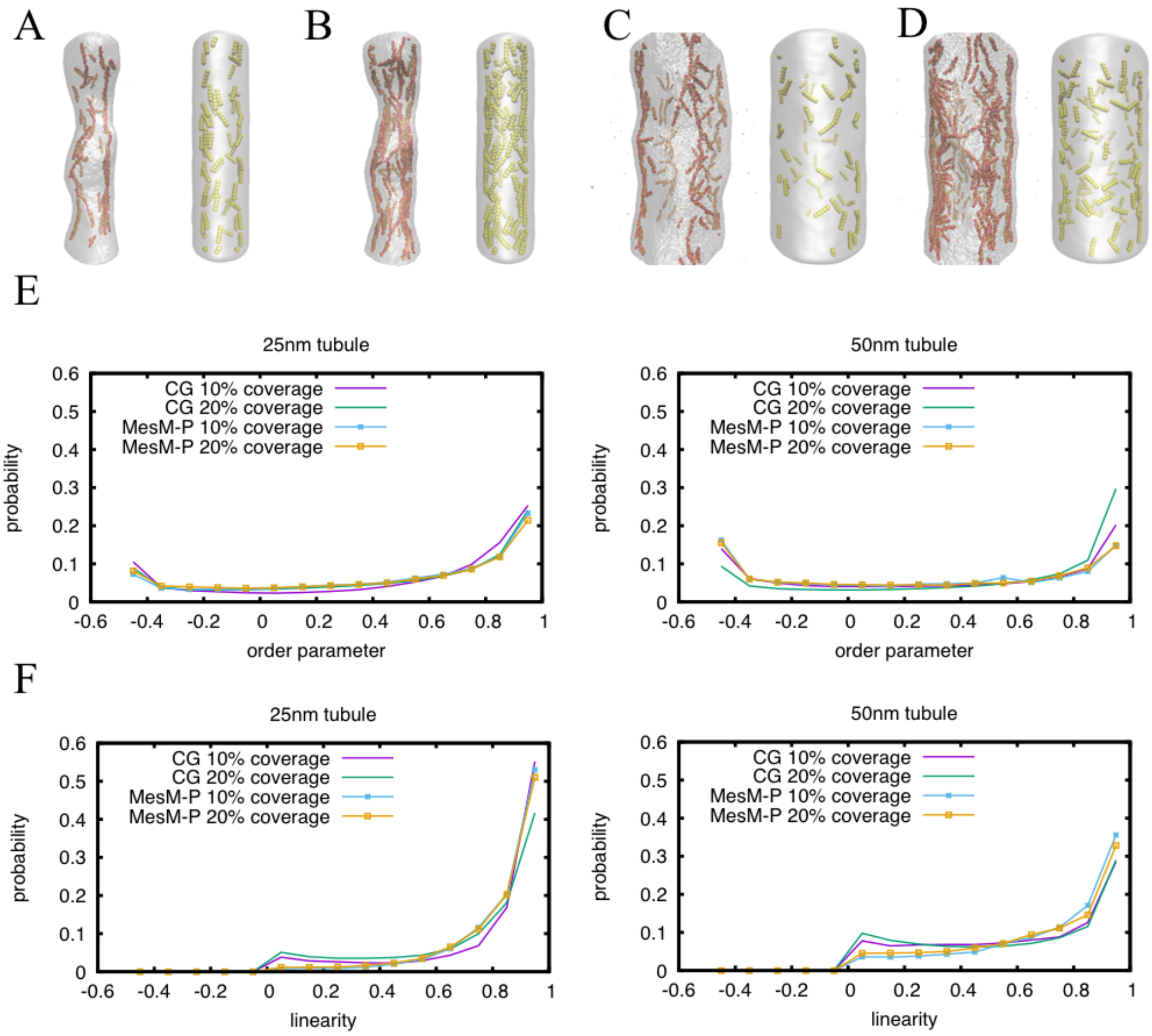

Figure 3-3: CG and MesM-P tubule snapshots and probability of ordering with the z-axis. A) 10% and B) 20% coverage in a periodic ~25nm radius tubules, C) 10% and D) 20% surface coverage in a ~50nm radius tubule (left CG snapshots, right MesM-P snapshots). E) Normalized histogram of order parameter for various bound densities and tubule radii. F) Normalized histogram of cosine of the angle formed by a single protein and z-axis for various bound densities and tubule radii.

To understand the nature of the curvature sorting property of the I-BAR domain, we probed the I-BAR domain organization inside tubules with two different radii. Interestingly, CG simulations at these lower surface coverages produced rather linear aggregates of I-BAR along the major axis of the tubule (Figure 3-3). We quantified the ordering between proteins by plotting probability density of the order parameter, $S = (3\cos^2\theta - 1)/2$, where θ is the angle formed between the long dimension of two I-BAR domains. The bimodal probability densities indicate perpendicular ($S = -0.5$) and parallel ($S = 1$) aggregates are extremely prevalent as shown in Figure 3-3E. As the tubule radius is increased from 25nm to 50nm, however, we noticed an increased stability of the perpendicular aggregates. The increased stability of parallel aggregates in the narrower tubule is likely due to the coupling between the curvature generated by each I-BAR domain and the inherent curvature of the tubule itself. We quantified ordering due to tubule radius by plotting the probability of the linearity shown in Figure 3-3F. Linearity is defined as the cosine of the angle formed by each I-BAR domain and the axis of the tubule (i.e., the z-axis). Linearity is 0 when the protein is perpendicular to the axis of the tubule and 1 when parallel to the axis. Figure 3-3F shows that the protein has a significant preference for the axial direction in the narrower tubules, i.e., perpendicular or spiral aggregates are less likely as I-BAR domains prefer axial aggregation. Axial aggregates are preferential because CG I-BAR domains form local membrane troughs and the axial orientation requires less membrane deformation than the perpendicular direction. In other words, the trough formed by a single I-BAR domain is more stable when formed in the axial direction, and as I-BAR domains aggregate, they are already axially aligned.

Spherical Membranes

We investigated surfaces with global Gaussian curvature in order to understand how I-BAR protein bind and induce large-scale deformation. Considering the limits and boundary conditions of the two simulation models, we used two separate approaches to approximate the surface of a GUV; we simulated a small CG vesicle of 200 nm diameter and a large MesM-P sheet with curvature driven by a guiding potential. The approximations made here are necessary to construct simulations that are more computationally tractable than a complete GUV. These simulations test the stability of both linear I-BAR aggregates on surfaces that display positive Gaussian curvature.

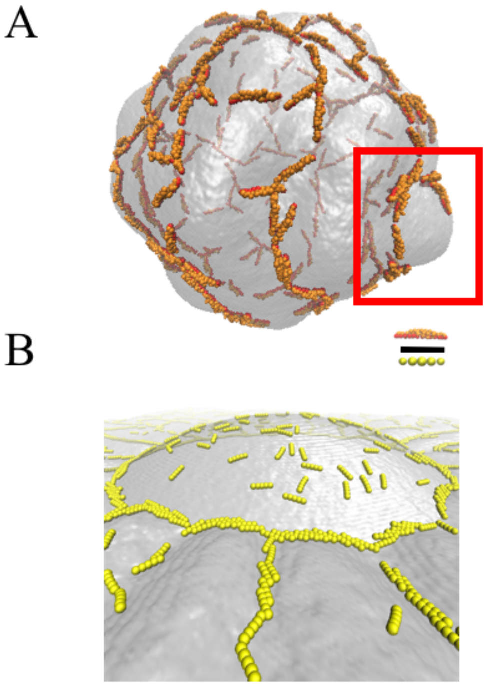

Figure 3-4: CG and MesM-P Snapshots on Curved Membranes.
A) CG organization on a 200nm diameter vesicle at ~10% coverage. In the red box: Liner aggregates of I-BAR domain organized around the base of a membrane bulge. B) MesM-P organization around a ~325nm diameter dome at ~10% coverage. Black scale bar is 15 nm.

The aggregation behavior of CG I-BAR shows significant deformation of the surface of small vesicles surrounded by I-BAR domains. The linear aggregates form strings on the membrane and form bulges out of the membrane with I-BAR on the periphery. Next, we employed the MesM-P based approach which used a guiding potential (see Section 3.7 for more details) to drive Gaussian curvature and the organization of I-BAR domains is investigated. These simulations probe the aggregation behavior on the surface of a deformed vesicle with quasi-null local curvature. When the membrane is perturbed into a ~ 325nm diameter dome structure with a spherical guiding potential, I-BAR domains preferentially sort to the edge of the surface forming a tip-to-tip ring as shown in Figure 3-4B. The preference for a tip-to-tip ring can be understood again as a way for the I-BAR domains to lie in a membrane trough; the edge of the dome structure is the region of the membrane that requires minimal perturbation to form a trough.

Experimental Results

Finally, we experimentally studied the I-BAR-driven protein membrane deformations on a GUV using fluorescence microscopy. We found that upon binding to PIP2-containing GUVs, IRSp53 I-BAR domain deforms the GUV membranes into tubular invaginations towards the interior of the GUVs, where the I-BAR domain decorates the inner surface of the tubules (Figure 3-5A and C), as previously reported (Barooji et al., 2016; Saarikangas et al., 2009). The bulk of the I-BAR domain fluorescence, as seen in the maximal projections in Figure 3-5C, is inside of the invaginated tubules. This observation shows that the I-BAR domains are enriched in the tubules compared to rather flat GUV membranes, consistent with the previous study (Prevost et al., 2015). Moreover, we observed that the invaginated tubules are localized at the intersections of the membrane indentation, i.e., the inward deformation of the circular cross-section of the GUV (Figure 3-5A, arrows). Besides tubulation, to our surprise, we observed that the I-BAR domain deforms GUV membranes into bulges where the I-BAR domain accumulates around their bases, which appear as local indentations. The bulges vary in size from a few microns in Figure 3-5B to a few hundred nanometers in Figure 3-5D. Many bulges with accumulated I-BAR domains produce a "flower-like" structure as shown in Figure 3-5B and D. This behavior appears to be quite similar to that predicted by the simulations shown in Figure 3-4B.

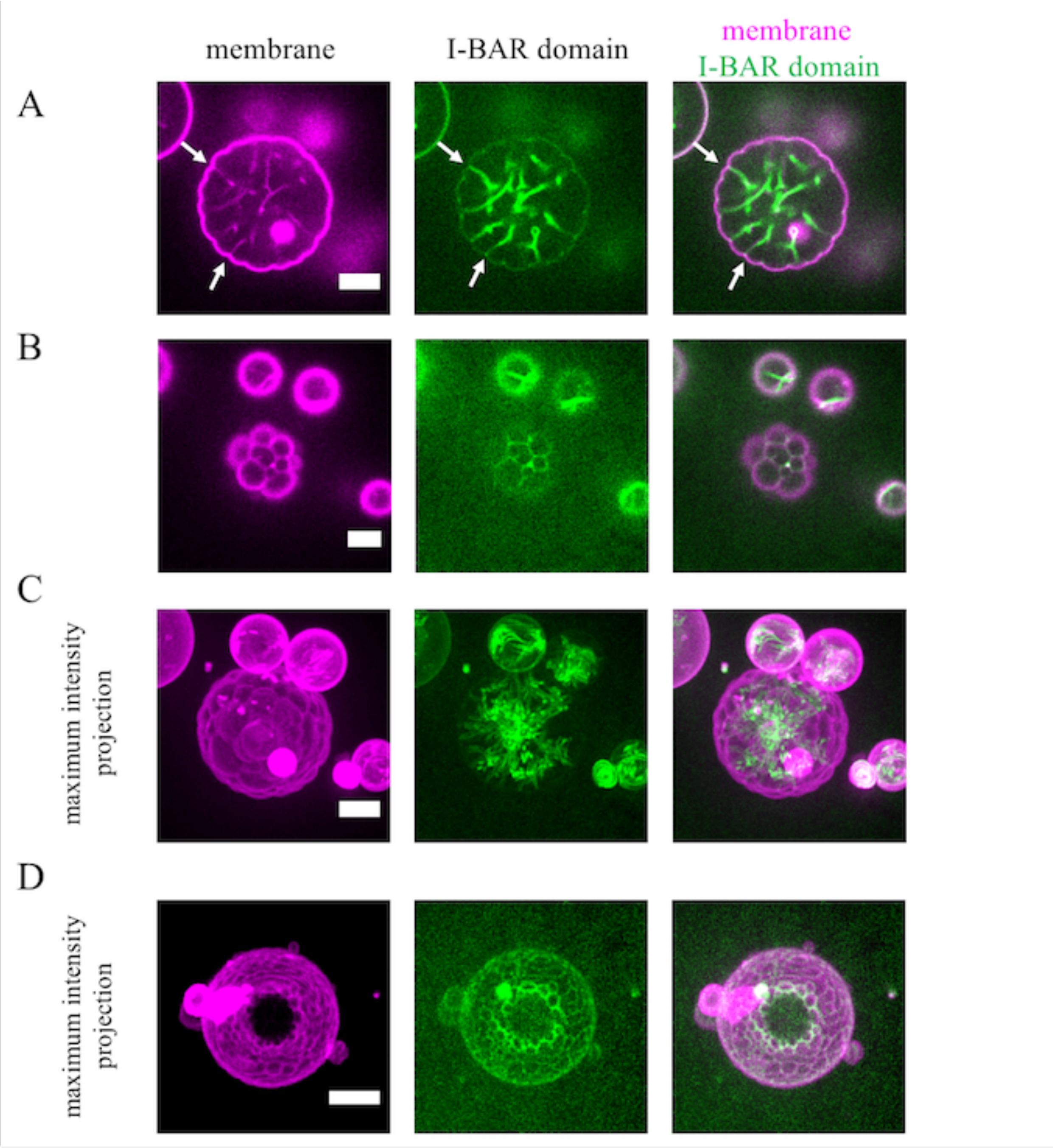

Figure 3-5: Microscopy Images of IRSp53 I-BAR Domain-Induced Flower-Like GUV Membranes

Representative confocal images of a GUV in the presence of IRSp53 I-BAR domain (A and B). Confocal images were taken at the equator of the GUV (A) and at the top of the GUV (B). Arrows in (A) indicate some membrane indentations. (C) Maximum intensity projection of the same GUV shown in (A) and (B). (D) Maximum intensity projection of a flower-like GUV in the presence of IRSp53 I-BAR domain. Protein concentrations: (A-C) 0.02 μM (70% unlabeled and 30% AX488 labeled I-BAR domain), and (D) 0.1 μM AX488 labelled I-BAR domain. Scale bars, 5 μm.

3.5 Discussion

In this chapter, we utilized a combined multiscale simulation and experimental approach to understand I-BAR domain organization at low surface coverage on lipid bilayers in vitro. We showed that I-BAR domains at low surface coverage prefer to organize into axial aggregates inside membrane tubules and organize around the periphery of membrane bulges on simulated vesicles and experimentally imaged GUVs. Using separate and complementary CG and mesoscopic approaches to capture protein shape and membrane curvature, we find evidence for robust levels of I-BAR organization. We find transient linear strings at low protein density on the membrane surface through a purely membrane-mediated attraction. As each I-BAR domain deforms the membrane locally, multiple I-BAR domains are attracted to generate long troughs on the surface of GUVs, inside of tubules, and on 200nm diameter vesicles. We modeled a variety of geometries to approximate the fundamental conditions and curvatures found in experiments, especially in GUVs as shown in Figure 3-5. When the membrane has inherent curvature, i.e., is not locally flat, both the CG and MesM-P I-BAR domain models couple to the curvature of membrane and preferentially orient to minimally deform the membrane. The minimal perturbation results in axial aggregates in tubules and rings around membrane bulges.

I-BAR Domains Form Axial Aggregates at Low Coverage

The aggregation behavior shown in the vesicle and tubule simulations suggest a mechanism by which tubules are initiated. Areas where multiple I-BAR aggregates contact are the regions of highest curvature in the flat membrane sheet simulations shown in Figure 3-2B and in the 200nm vesicle simulations shown in Figure 3-4A. In the simulation models presented here, the effective attraction between the ends of I-BAR domains is membrane-mediated such that the coupling of curvature minimizes the system free energy. In a more realistic bilayer, such as the in vitro assays presented here, the positively-charged ends of I-BAR domains could also be electrostatically attracted by the clustering of negatively-charged lipids, e.g., PIP2. After the initial nucleation of tubule, it is unclear how the growth is driven. While the models here show that the axial aggregate is preferred orientation, further studies (e.g., simulations of nascent tubules or experiments investigating orientation of the I-BAR domains at moderate density inside tubules) will be required in the future to fully elucidate the tubulation mechanism.

The axial aggregates of I-BAR domains inside tubular membranes seen in our simulations are unexpected. We find that I-BAR domains organize such that the grooves in the membrane formed by I-BAR domains produce minimal deformations of the membrane, resulting in long axial troughs that are parallel to the tubule axis. At the surface densities simulated here, the differences between tubule radii most notably affects the orientation of the individual I-BAR domains and the propensity for axial aggregates to form. Our results suggest that the preferred curvature of I-BAR domains may arise from the balance between energetic preference of axial aggregates, which minimally perturb the membrane, and entropic penalty due to reduced rotational freedom inside of the tubule.

As protein surface density increases, we would expect that the axial aggregates would instead form the proposed perpendicular aggregates (Mesarec et al., 2016; Noguchi, 2016) similar to those formed by other members of the BAR family (Cui et al., 2013; Daum et al., 2016; Frost et al., 2008; Mim et al., 2012). At high surface densities, the properties of a single I-BAR domain would also be different due to the presence of the neighboring I-BAR domains, and the resultant local deformation caused by each I-BAR domain could change from the deformations generated by the current model. Given the computational cost of resolving many I-BAR domains at a finer resolution, the current model parameterization also does not take into account direct interactions between neighboring I-BAR domains (except excluded volume) and, therefore, high density conditions are outside of the scope of the present model. We also note that axial aggregates are in contrast to previous computational studies of I-BAR domains (Mesarec et al., 2016; Noguchi, 2016) that show perpendicular aggregates inside of membrane tubules, but these previous studies used models unlike the models presented here. Among the several differences including membrane representation and protein-membrane interactions, the CG I-BAR model used in this work reproduces the properties of a single, isolated, and atomically-resolved I-BAR domain, which flattens when bound to the membrane and is outside the curvature ranges previously studied. Our model shows a new phenomenology because our multiscale approach considers I-BAR domain properties outside the scope of the previous models to date.

I-BAR Domains Aggregate Around Bulges on Vesicular Membrane at Low Coverage

We experimentally revealed low curvature bulges that are surrounded by I-BAR domains at their bases shown in Figure 3-5B and D. Low surface coverage membrane remodeling of this kind has not been shown before, to our knowledge. In addition, in the simulations we observed consistent aggregation behavior of I-BAR domains on the surface of small vesicles and around the base of preformed membrane bulges in planar membranes. The bulges due to I-BAR domain aggregation here bear a striking resemblance to deformation modeled by endophilin N-BAR domains (Simunovic et al., 2017) and spherical nanoparticles (Šarić and Cacciuto, 2012; Šarić and Cacciuto, 2013). This common phenomenology suggests the formation of membrane bulges are not specific to I-BAR domains but are the result emergent phenomenon due to linear aggregation. Furthermore, Figure 3-4A shows significant deformation of the surface of small vesicles similar to the experimental images shown in Figure 3-5B, although on a smaller length scale. As these simulated vesicles are 100 times smaller than the experimentally-relevant vesicles shown in Figure 3-5, we also employed the MesM-P based approach using a spherical guiding potential to from a “dome” structure with Gaussian curvature on a planar membrane. When the membrane is perturbed into a ~ 325nm diameter dome, which is much closer to the size of membrane bulges shown in Figure 3-5D, I-BAR domains aggregate around the edge of the surface forming a tip-to-tip ring as shown in Figure 3-4B. This result is striking because it resembles the CG result at shorter length scale, where Gaussian curvature was generated on a small vesicle, as well as the experimental result that shows significant aggregation around the large indentations created by the membrane-bound I-BAR domains.

The phenomenological similarities between I-BAR domain aggregation around the small bulge on the CG vesicle, tip-to-tip ring around the preformed dome on the MesM-P sheet, and the aggregation at the periphery of flower-like membrane structures in experimental images suggest a common driving force acting at different length scales. In the case of the simulations, I-BAR domains bind to the membrane and generate local deformation that leads to aggregation. Given the similar phenomenology, we suggest that the bulges seen in experiment are a minimally perturbative conformation due to I-BAR domain aggregation.

3.6 Conclusions

Our multiscale modeling approach captures the local behavior of an isolated, membrane-bound I-BAR domain of IRSp53 and is used to model I-BAR domains on lipid bilayers of various geometries, and scales to micron-size systems. We demonstrate that lipid bilayer geometry is an important factor in I-BAR domain aggregation. At low surface coverage, we show I-BAR domains generate local troughs leading to transient tip-to-tip aggregates on flat surfaces and the preference for trough formation leads to string aggregates along the long axis of tubules and a ring of I-BAR domains at the base of membrane bulges on the surface of vesicles. The aggregation behavior demonstrated in the CG and MesM-P simulations provides a mechanism for the intriguing low curvature membrane bulges that we observe experimentally.

3.7 Supplemental Information

The CG models used in this chapter were systematically parameterized from reference all-atom simulations. We first describe the atomistic simulations used, then the procedure to train and simulate the coarse-grained models.

All-atom Simulation Details

All-atom simulations of a single I-BAR domain of IRSp53 interacting with a fully solvated, periodic lipid bilayer were run using GROMACS (version 5.0) (Abraham et al., 2015). Initially, a lipid bilayer composed of 75% 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 20% 1,2-dioleoyl-sn-glycero-3-phosphoserine (DOPS), and 5% Phosphatidylinositol-4,5-diphosphate (PI(4,5)P_2) was generated using the CHARMM-GUI and equilibrated using the CHARMM-GUI scheme of sequentially restrained simulations (Jo et al., 2007; Jo et al., 2008; Jo et al., 2009; Lee et al., 2016; Wu et al., 2014). Next, the I-BAR domain of IRSp53 (PDB: 2YKT(de Groot et al., 2011)) was added to the membrane, and solvated with 150 mM NaCl using GROMACS tools (Abraham et al., 2015). Production simulations were run at 298 K at 1 atm using a Nose-Hoover thermostat and the Parrinello-Rahman barostat implemented with corresponding 1 ps and 5.0 ps coupling constants in GROMACS (Abraham et al., 2015; Parrinello and Rahman, 1981b; Smit and Frenkel, 2001). The CHARMM36 force-field was used in all simulations as well (Best et al., 2012; Klauda et al., 2010; MacKerell et al., 1998).

CG Model Details

A coarse-grained map of I-BAR domain of IRSp53 was created using Essential Dynamics Coarse-graining (ED-CG) using 24 beads per monomer (Zhang et al., 2008). This number of beads was chosen to reproduced to maintain a similar resolution between the protein and lipid model, which has been thoroughly described ref (Srivastava and Voth, 2013). The ED-CG protocol generates a mapping dividing the protein along its primary sequence by minimizing the residual (shown in Equation 3.1) describing the fluctuations between beads.

$$\chi^2 = \frac{1}{3N}\sum_{I=1}^{N}\frac{1}{n_t}\sum_{t=1}^{n_t}\left(\sum_{i\in I}\sum_{j\geq i\in I}\left|\Delta r_i^{ED}(t) - \Delta r_j^{ED}(t)\right|^2\right) \qquad 3.1$$

where N is the number of CG sites, $n_t$is the number of configurations, and $\Delta r_i^{ED}$is the displacement from equilibrium of the ith CG site at configuration t. The protocol used here first maps each residue to the carbon-alpha atom of the residue backbone, then maps multiple carbon-alpha atoms to a single bead resulting in approximately 10 residues per bead. For example, the boundaries for the first bead are the first residue and the fourth residue i.e. the first three residues' carbon alphas map to the first bead. As a result of a steepest descent and simulated annealing minimization scheme, the residue boundaries between beads are as follows, 1st, 4th, 8th, 20th 35th, 48th, 63rd, 77th, 90th, 105th, 114th, 127th, 138th, 147th, 152nd, 156th, 160th, 167th, 177th, 190th, 202nd, 216th, 226th, and 236th. The spacing is not regular as ED-CG optimizes these boundaries to maximally capture the atomistic fluctuations in the protein. The intraprotein interactions are parameterized using a hetero-elastic network model parameterized using the all-atom simulations (Lyman et al., 2008). Parameters, k_{ij} and b_{ij} shown in Equation 3.2, between each pair of CG protein beads are determined to reproduce the fluctuations of the CG beads in the mapped all-atom trajectory. **Error! Reference source not found.** gives the intraprotein elastic network parameters.

$$U_b(r_{ij}) = k_{ij}(r_{ij} - b_{ij})^2 \tag{3.2}$$

The interprotein and protein-lipid (nonbonded) interactions use a shifted-force 10-6 Lennard Jones potential defined in Equations 3.3 and 3.4 (Allen and Tildesley, 1989). Both interprotein and protein-lipid interactions used a sigma value of 1.5 nm, which was the most probable distance between mapped protein beads to the mapped lipid head groups. The nonbonded interaction cutoff was assumed to be 2σ. The epsilon values of the protein-lipid interactions were investigated as described earlier in Chapter 3:. The interprotein interactions were the repulsive portion of the protein-lipid interaction.

$$U_{NB} = U_{LJ}(r_{ij}) - U_{LJ}(r_C) - (r - r_c)\, F_{LJ}(r_c), \qquad r < r_C \tag{3.3}$$

$$U_{LJ}(r_{ij}) = 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{10} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \qquad \text{and} \qquad F_{LJ}(r_{ij}) = -U'_{LJ}(r_{ij}) \tag{3.4}$$

The initial configurations for each system were created by placing lipids on equally spaced points in the desired geometry (e.g., cylinder) and equilibrating under zero membrane tension when applicable. Proteins were subsequently added to the equilibrated lipid bilayer system. CG simulations were run using the LAMMPS MD engine and the Langevin thermostat with temperature dampening parameter of 5000fs and Parrinello-Rahman barostat with pressure dampening parameter of 50,000fs (Dunweg and Paul, 1991; Parrinello and Rahman, 1981b; Plimpton, 1995). The timescale of the production simulations was at least 30 million steps with a CG timestep of 5 fs but varied based on the geometry and the time required to converge order parameters statistics described earlier in this chapter.

Mesoscopic Membrane with Explicit Proteins Model Details

The extension of Mesoscopic Membrane with Explicit Proteins (MesM-P) model presented here leverages successes of the original MesM-P model with three major changes: no solvent, softer 10-2 Lennard-Jones potential, and shorter discretization length (Davtyan et al., 2017). We set the discretization of the membrane mesh to approximately 3 nm, based on the dimensions of the I-BAR domain of IRSp53 (de Groot et al., 2011). Next, we parametrize the I-BAR domain model as a linear chain of MesM-P particles with varied radii. The protein-membrane interactions use a shifted-force 4-2 Lennard Jones potential similar to that shown above. In the protein-membrane interactions, we use values of sigma equal to 2.67, 2.90, and 3.12 nm and corresponding values of epsilon equal to 1.0, 1.25, and 1.5 kcal/mol, which capture the shape and aggregation behavior of the I-BAR domain as shown earlier in this chapter. The protein-protein interactions were the repulsive portion of the protein membrane interaction with the same varied sigma values to maintain the shape of the I-BAR domains.

The initial configurations for each system were made by placing membrane beads on a hexagonal lattice in the desired geometry and equilibrating under zero tension using the Langevin thermostat with temperature dampening parameter of 5000fs and Berendsen barostat with pressure dampening parameter of 50000fs. All systems ran for 5 million timesteps to converge order parameter statistics described in main text. The guiding potential used to generate the membrane bulge as described in the main text was a 325 nm spherical guiding potential slowly moved toward the membrane under zero tension. The spherical surface had a potential form of a 9-3 Lennard Jones similar to the potential shown in Equation 3.4 with a value of 1 kcal/mol.

Coarse-grained Protein-Protein Interactions

Thus far, we have assumed that the protein-protein interactions are purely repulsive, i.e., direct attraction between protein beads is not included. In reality, there would be some degree of protein-protein interaction. I-BAR domains are significantly charged proteins, which directly affects the membrane binding behavior. If each CG bead is assigned the mapped net charge, i.e., a simple charge mapping, we find a positively charge protein surface near the membrane and a positive high charge density near the ends of the I-BAR domains, as shown in Figure 3-6. The protein shown similarly to Figure 3-1, and the membrane would lie directly below the I-BAR domain in this view. We note a positively charge protein surface near the membrane and the high charge density near the ends of the I-BAR domain. Additionally, there is a net +3 charge on a CG single bead, which in an atomistic resolution would be delocalized over several residues. In the absence of explicit solvent, we model screened electrostatic interactions between two I-BAR domains with a Yukawa potential with the unknown screening length, κ, shown in Equation 3.5.

$$V_{Yukawa}(r) = -\frac{q_1 q_2}{4\pi\epsilon_0}\frac{e^{-\kappa r}}{r} \tag{3.5}$$

In the bulk, we could approximate the experimental screening length corresponding to 20 mM Tris pH 7.5, 60 mM NaCl and 100 mM sucrose buffer, but published studies(Böckmann et al., 2003; Pandit et al., 2003) have shown the effective ion concentration near the membrane is higher than in the bulk. We ran simulations using the possible screening lengths from near bulk ion concentration (κ = 0.15 Å^{-1}) to ~10x increase in effective ion concentration (κ = 0.45 Å^{-1}). We characterize these aggregates using the same order parameter analysis described earlier in this chapter and the order parameter probability distribution is shown in Figure 3-6B.

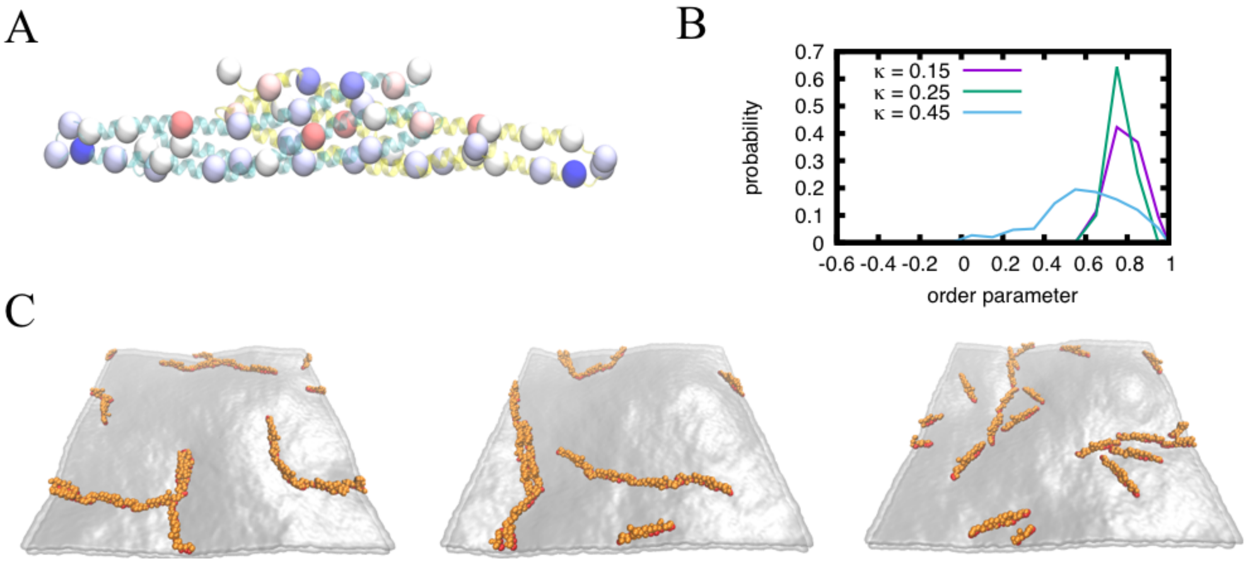

Figure 3-6: CG I-BAR domain charge distribution, order parameter distribution, and snapshots of I-BAR domain aggregation on flat sheet.
Charge distribution on CG model and snapshots of flat sheet configurations with screened electrostatic interaction protein-protein interactions. A) Side view of an overlay of the secondary structure of each monomer (cyan and yellow) and CG I-BAR domain colored by charge (+3 charge [blue], -2 [red]) B) probability distribution of inter-protein order parameter for varied screening lengths. C) snapshot of 10% coverage of a 100nm by 100nm flat at varied screening length: 0.15 Å-1, 0.25 Å-1, and 0.45 Å-1 from left to right.

As seen Figure 3-6B and C, at low effective ion concentration (κ = 0.15 Å^{-1}), the I-BAR domains begin to form bundled linear aggregates. This behavior is due to an overestimation of the protein-protein interactions because of the low effective salt concentration and the localization of charge on a single CG bead. At 3x bulk ion concentration (κ = 0.25 Å^{-1}), I-BAR domains again form more unbundled linear aggregates with some additional side-to-side character, as additional I-BAR domains move to align opposite charges on the I-BAR domains. This behavior is expected at more reasonable bulk ion concentration. Finally, at 10x bulk ion concentration (κ = 0.45 Å^{-1}), we find mostly linear aggregates again. Thus, when we model the electrostatic interactions between I-BAR domains in this way, we find that aggregation behavior qualitatively agrees with that from purely repulsive excluded volume protein interactions, which is to say that linear tip-to-tip aggregation is dominant and that this behavior is the result of an indirect membrane-mediated force.

Chapter 4: Phenomenological I-BAR Domain Model

4.1 Summary

Collective action by Inverse-BAR (I-BAR) domains drive micron-scale membrane remodeling. The macroscopic curvature sensing and generation behavior of I-BAR domains is well characterized, and computational models have suggested various mechanisms on simplified membrane systems, but there remain missing connections between the complex environment of the cell and the models proposed thus far. Here, we show a connection between the role of protein curvature and lipid clustering in the stabilization of large membrane deformations. We find lipid clustering provides a directional membrane-mediated interaction between membrane-bound I-BAR domains. Lipid clusters stabilize I-BAR domain aggregates that would not arise through membrane fluctuation-based or curvature-based interactions. Inside of membrane protrusions, lipid cluster-mediated interaction draws long side-by-side aggregates together resulting in more cylindrical protrusions as opposed to bulbous, irregularly shaped protrusions.

4.2 Introduction

In this study, we aim to understand membrane-mediated forces involved in I-BAR domain aggregation and membrane remodeling. We introduce highly coarse-grained membrane and protein models. We characterize the phenomenological aggregation behavior and membrane protrusion stabilization of I-BAR domains by investigating the effects of I-BAR domain curvature, protein-membrane interaction strength, and membrane composition. Together, we gain further insight into the membrane-mediated forces that drive protein aggregation and stabilization of membrane protrusions.

4.3 Methods

The membrane model is a solvent-free coarse-grained model recently developed by Grime and Madsen that is represented by a monolayer.(Grime and Madsen, 2019) The model represents the membrane patches by three beads with tunable fluidity, rigidity, and intrinsic curvature as modulated by the inter- and intra-membrane patch interactions. The membrane-membrane and membrane-protein interactions follow the form shown in Equation 4.1, which is a soft, sine-based pair potential.

$$U_{pair}(r) = \begin{cases} -\left(\frac{A}{a}\right)\sin(r \cdot a), & if\ r \leq r_0 \\ -\left(\frac{B}{b}\right)\sin\left(\frac{\pi}{2} + (r - r_c) \cdot b\right), & if\ r_0 < r \leq r_c \\ 0, & r_c \end{cases} \quad 4.1$$

The membrane-membrane interactions of the upper and lower beads (purple beads in Figure 4-1A) are purely repulsive (i.e., B value of 0 k_BT) and, in this study, both upper and lower beads have an r_0 value of 1.125 nm resulting in no intrinsic curvature. The center bead (grey beads in Figure 4-1A) has an interaction strength or B value of 9.0 k_BT to other center beads and B value of 0 k_BT to other membrane beads and a r_0 value of 1.5 nm to all beads. The membrane parameter set results in a fluid monolayer with a bending modulus of ~10 k_BT. The bonded interactions are harmonic with zero energy distance of 1.5 nm and spring constant of 25 k_BT or ~15 kcal/mol. The angular potential is harmonic with zero energy angle of 180 degrees and spring constant of 10 k_BT or ~6 kcal/mol. At this coarse-grained resolution, the membrane beads do not correspond to individual lipids, but represent a patch of the lipid bilayer. From the perspective of simpler field theoretic models, the membrane acts as an undulating sheet that mediates protein-protein interactions via curvature coupling and Casimir-like forces. The highly coarse-grained nature of the membrane allows for large system sizes necessary for the simulation of the collective behavior of proteins at near micron length scales.

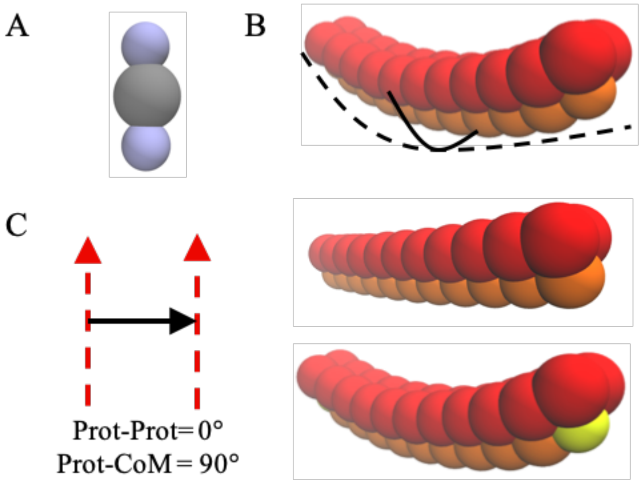

Figure 4-1: Phenomenological membrane and I-BAR domain representation with schematic of characteristic aggregation angles.
Membrane presentation (A) and three protein representations with first, curved I-BAR domain with uniform binding surface (B upper), flattened I-BAR domain with uniform binding surface (B center), and curved I-BAR domain with specific lipid binding surface (B lower) highlighting the transverse (solid) and longitudinal (dashed) curvatures. Schematic of two angles used to characterize protein aggregates: the angle formed by vectors describing the long dimension of two neighboring proteins (Prot-Prot) and the angle formed by a vector describing the long dimension of a protein and the center of mass between two neighboring proteins (Prot-CoM).

The protein model is represented by three linear strings with tunable longitudinal and transverse curvatures. The intra-protein interactions are an elastic network holding the protein rigid. The width and length of each I-BAR domain is consistent with the dimensions from the crystal structure of the I-BAR of IRSp53, while the curvature ranges from 1/15 nm^{-1} to 1/50 nm^{-1} (i.e. nearly flat).(de Groot et al., 2011; Millard et al., 2005) The protein-protein interactions are purely repulsive and are of the same form as the membrane-membrane interactions with bead sizes that are 1.7 nm.

The protein-membrane interactions are purely repulsive except in the case of the center string of I-BAR domain beads (see orange beads in Figure 4-1B) and the middle monolayer bead (see grey beads in Figure 4-1A). These interactions are varied from having an interaction strength of ~0.6 to ~3.2 kcal/mol, which effectively modulates the binding energy of the I-BAR domain and the local curvature induced by an individual I-BAR domain. By spanning a range of membrane-protein attractions, we assess the onset of protein clustering and characterize the resultant I-BAR domain clusters, which, for example, implicitly include membrane-protein electrostatic interactions. Additionally, in select cases, the end beads of the I-BAR domain (see yellow beads in Figure 4-1B) have specific attraction to a subset of the membrane beads, e.g., to represent selective binding to PIP_2-containing lipid patches. We modulate the membrane binding surface of the protein to probe how attractive membrane bead clusters affect aggregation behavior of I-BAR domains and the protrusion stabilization.

We run these simulations using the LAMMPS molecular dynamics simulator (Plimpton, 1995) with a timestep of 200 fs using the Langevin thermostat (Schneider and Stoll, 1978) with temperature dampening of 20 ps and an anisotropic Parrinello-Rahman barostat (Parrinello and Rahman, 1981a) with pressure dampening of 200 ps. The periodic flat sheet simulations fluctuate around 160 nm by 160 nm with 44,409 membrane beads. The larger protrusion simulations contained 188,031 membrane beads and vary in lateral dimension based on the radius of the protrusion made. The protrusions were formed using an external time-dependent repulsive cylindrical and spherical potentials implemented in LAMMPS to enforce a minimum protrusion radius. The I-BAR domains coated the protrusion and were equilibrated by ramping the protein-protein repulsion down to 5 or 10% and back up, which facilitates equilibration of high-density coats of proteins. We ramped the repulsion down, ran with weak repulsion, and ramped back three times over the course of 16.2 μs. Then, we ran with full repulsive interactions for 5 μs. We analyzed the last 5 μs for I-BAR domain statistics in protrusions formed by external potentials. During this time, the protein aggregate can deform the shape of the tubule outward, away from a cylindrical protrusion because the external potential only enforces a minimum protrusion radius. Finally, the external bias is removed, and both the membrane and proteins are relaxed for another 25 μs. We use the last 5 μs for statistics of relaxed protrusions.

We quantify I-BAR domain aggregation using two characteristic angles shown in Figure 1C: The angle between two adjacent proteins (Prot-Prot) and the angle formed by the protein and the center-of-mass vector between two adjacent proteins (Prot-CoM). We characterize the stability of protrusions by calculating the local curvature over the entire protrusion. We calculate local curvature by clustering membrane beads, performing a local dimensionality reduction with the isomap algorithm, performing a Delaunay triangulation in the reduced dimensions, and finally, using the triangles in the full three dimensions, to calculate the curvature at each cluster center.(Rusinkiewicz, 2004) All characterization was done in Python using numpy,(van der Walt et al., 2011) scipy,(Millman and Aivazis, 2011; Oliphant, 2007; Virtanen et al., 2019) and scikit-learn(Pedregosa et al., 2011).

4.4 Results

Flat Sheet Organization as a Function of Binding Strength and Protein Curvature

Initially, we run flat sheet simulations of the I-BAR domains to explore and characterize the aggregation behavior of the current I-BAR domain model in comparison to previous particle-based BAR domain models. We find that I-BAR domain curvature and binding strength, as modulated by the membrane-protein interaction strength, control the aggregation behavior. At I-BAR domain radii greater than 30 nm and interaction strengths below 1.9 kcal/mol, the proteins do not aggregate (see black lines in Figure 4-2A). As membrane-protein interaction strength increases, the I-BAR domains form end-to-end linear aggregates (see yellow squares in Figure 4-2A). At low interaction strength and increasing curvature, the I-BAR domains form partial side-by-side aggregates. At high curvature and increasing interaction strength, the preferred aggregate morphology changes from end-to-end, through an intermediate aggregate of side-by-side/end-to-end aggregate and eventually, side-by-side aggregates (see purple circles in Figure 4-2A). By tuning the protein-membrane interactions and protein curvature, we show a rich phenomenology of aggregation behavior similar to previous models of I-BAR domains.(Jarin et al., 2019; Noguchi and Fournier, 2017) We provide further evidence that membrane-mediated forces (i.e. curvature-mediated and Casimir-like attraction) are sufficient to drive membrane aggregation in the absence of direct protein-protein attraction. Next, we introduce an additional membrane-mediated force due to attractive membrane bead clustering.

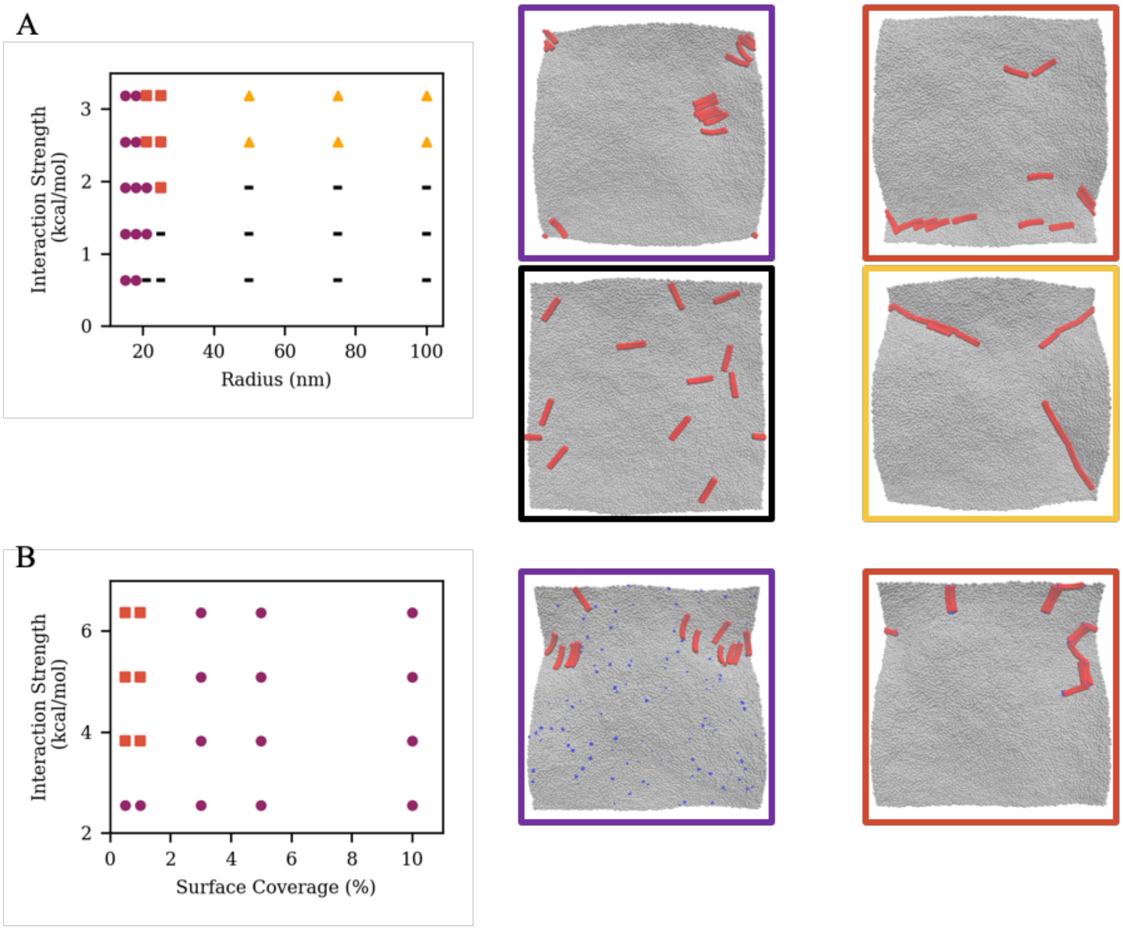

Figure 4-2: Phase Diagrams of I-BAR Domain Aggregation Behavior and Characteristic Snapshots of Each Phase.

Characterization of aggregation behavior as a function of membrane curvature and protein-membrane interaction strength (A) and PIP_2-like membrane bead interaction strength and PIP_2-like surface coverage at an I-BAR domain radius of 18nm (B). Representative snapshots of no aggregate (black lines), end-to-end aggregate (yellow triangle), intermediate aggregates (red square) and side-by-side aggregate (purple circles) without PIP_2-like membrane beads without PIP_2-like membrane beads (A) and with PIP_2-like membrane beads (B).

Flat Sheet Organization with Attractive Membrane Beads

We seek to understand the role of lipid clustering and how it gives rise to a new local membrane-mediated force. To capture this effect, e.g., for PIP_2 clustering, we make two changes to our low-resolution model. We introduce a PIP_2-like membrane bead that has stronger attraction to the ends of I-BAR domains (see Figure 4-1B lower), thereby making the membrane binding surface of the I-BAR domain non-uniform. We choose a protein curvature of 1/18 nm^{-1} and protein-membrane interaction strength of 0.6 kcal/mol for the rest of the membrane-binding surface. Then, we characterize the aggregation behavior as a function of the I-BAR domain-attractive membrane bead interaction strength and the effective surface coverage of attractive membrane beads.

As expected, we find that attractive membrane beads cluster to the ends of I-BAR domains. We find that the highly attractive membrane patch significantly affects the I-BAR domain organization by introducing an area where I-BAR domains can co-localize. In the case of curved I-BAR domains, the side-by-side aggregates can be destabilized, and end-to-end aggregates preferred because the I-BAR domains organize around a central attractive membrane bead (see red squares of Figure 4-2B).

In Figure 4-2, we show snapshots of the 18 nm radius I-BAR domain with a single site to bind attractive lipids while the rest of the membrane binding surface is unchanged from previous simulations. These choices are used to closely model the membrane binding surface of I-BAR domain of IRSp53. The flat sheet aggregation behavior changes drastically with the addition of attractive membrane beads. The attraction strength between the ends of I-BAR domains to the PIP_2-like lipid patches is an open parameter, which we investigated from 2.5 kcal/mol to 6.4 kcal/mol. When the interaction strength between the end of the I-BAR domain and the attractive membrane beads is five or six times stronger than other membrane beads (i.e., interaction strength greater than 10), there is an onset of clustering and the aggregation behavior is affected (see red squares in Figure 4-2B). When there is a uniform binding surface, we find the 18nm radius I-BAR domain with weak membrane-protein interactions would form weak side-by-side aggregates. As we add PIP_2-like membrane beads (i.e., the surface coverage of the PIP_2-like lipids increases), we find that both the side-by-side and end-to-end aggregates are stabilized. Furthermore, as we increase the surface coverage of attractive membrane beads, we find that end-to-end aggregates are less probable side-by-side aggregates. Thus, the clustering effect is less pronounced as the amount of attractive membrane beads increases.

Protrusion Stability

Next, we probe the stability of the protrusions coated in I-BAR domains. We form a protrusion using an external potential, remove the external potential, and allow the protrusion to relax under the influence of only I-BAR domains. We sweep over a range of protein and protrusion curvatures, Figure 4-3 shows representative snapshots for 15, 21, and 50 nm radius I-BAR domains inside of 15 and 21 nm radius protrusions as well as the characterization of membrane Gaussian curvature and protein number density. Over this range of radii, we find a variety of relaxed protrusion configurations. Generally, we find that the curved I-BAR domains change the radius of a protrusions towards the radius of the I-BAR domains and flat I-BAR domains did not stabilize the protrusions and sort toward the flat area surround the protrusions.

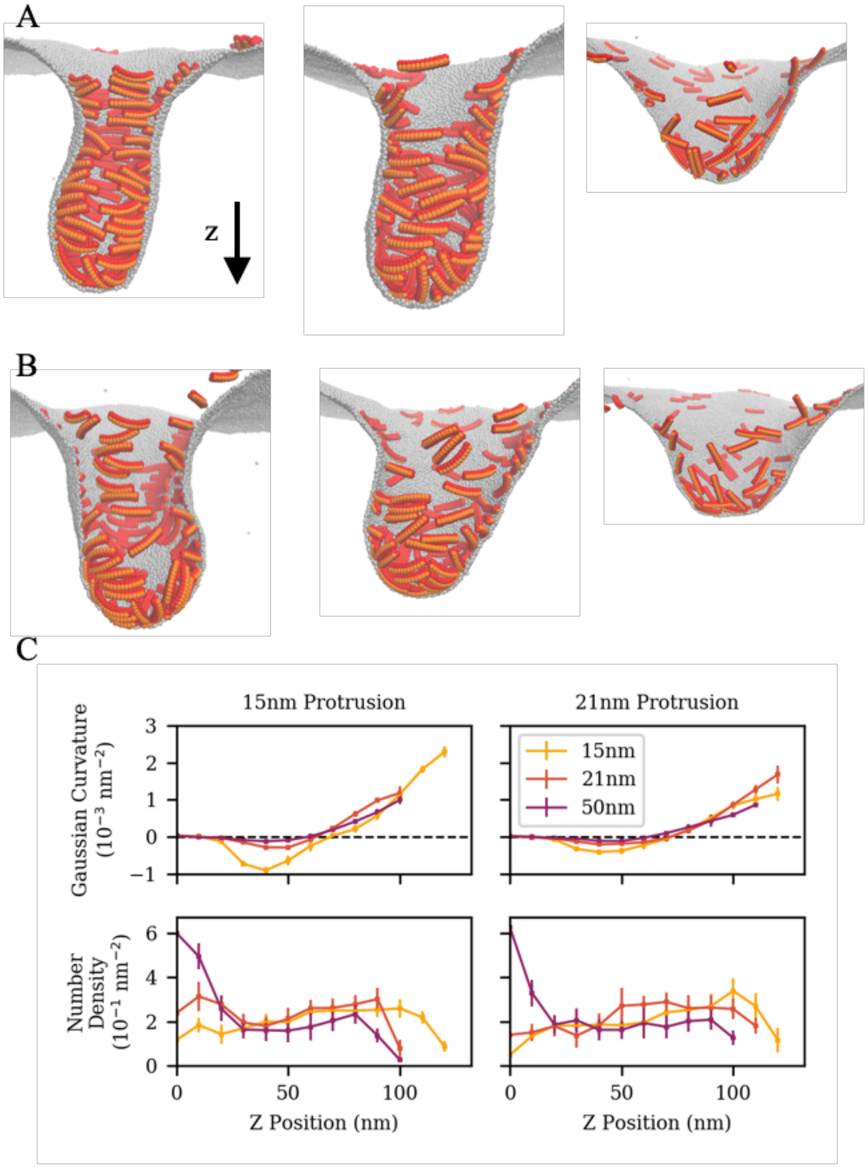

Figure 4-3: Snapshots of I-BAR Domain Assemblies Stabilizing Protrusions and Quantification of Gaussian Curvature and Number Density.
Snapshots of protrusions of 15nm (A) and 21nm (B) radii stabilized by I-BAR domains of 15nm, 21nm, and 50nm (left to right) radii. Gaussian curvature and number density as position along the length of the protrusion (C) for the corresponding protrusion and I-BAR domain radii.

In the case of 15 nm I-BAR domains inside of 15 nm radius protrusions, aggregates stabilize a bulbous protrusion (see Figure 4-3A left). Indeed, when we calculate the Gaussian curvature profile along the length of the protrusion, the protrusion does not have a region of 0 curvature, which is a characteristic of cylinders. We see constantly varying Gaussian curvature along the length of the protrusion. Additionally, we see a relatively flat number density of I-BAR domains meaning there is little sorting in or out of the protrusion. Inside of a wider, 21 nm protrusion (see Figure 4-3B left), the I-BAR domain aggregate stabilizes the bulbous protrusion with constantly varying Gaussian curvature of a lower magnitude as compared to the 15 nm protrusion. In the intermediate case of 21 nm I-BAR domains (see Figure 4-3A center), we find that the stabilized protrusion is similarly bulbous as well as widened by the I-BAR domain aggregate towards the radius of the I-BAR domain. When the I-BAR domain radius and protrusion radius are both 21 nm, the stabilized protrusion is similar to the 15 nm protrusion case. Although, we do find that the number density on the flat sheet to be slightly lower when the initial protrusion is 21 nm as compared to 15 nm. The difference is small, but suggests that sorting effect between the 15 nm and 21 nm protrusions. In the case of low curvature, I-BAR domains (i.e. flat I-BAR domains) shown in right snapshots of Figure 4-3A and B, the protein aggregate is not sufficient to stabilize the protrusion. Indeed, when considering the Gaussian curvature and number density profiles, we find that there is little to no negative Gaussian curvature and the I-BAR domains have a significantly higher density on the flat membrane surrounding the protrusion.

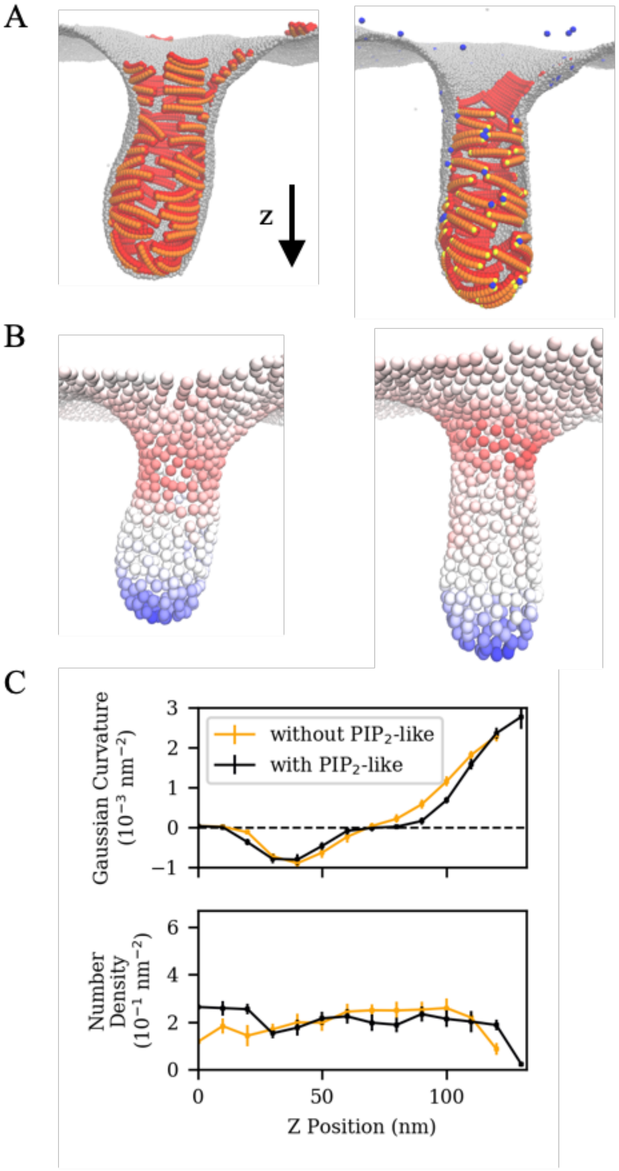

Figure 4-4: Snapshots of I-BAR Domain Assemblies Stabilizing Protrusions With and Without PIP$_2$-like membrane beads and Quantification of Gaussian Curvature and Number Density. Snapshots of protrusions stabilized by 15nm I-BAR domains with and without PIP2-like membrane bead (A: left and right, respectively), the reduced representation of membrane (B: red (-) and blue (+)), and Gaussian curvature and number density as position along the length of the protrusion (C).

We compare protrusion stabilization in the presence of attractive membrane patches. We find that the I-BAR domain aggregates continue to cluster the attractive membrane beads in between side-by-side aggregates (See Figure 4-4A). When comparing Gaussian curvature profiles shown in Figure 4-4C, protrusions with PIP$_2$-like membrane beads have a shoulder in the Gaussian curvature profile while protrusions without PIP_2-like membrane beads do not. In other words, protrusions with the attractive membrane beads are more cylindrical than those without. The attractive membrane beads mediate an attraction between side-by-side aggregates, constricting the protrusion towards the curvature of the I-BAR domains. Thus, we show that PIP_2-like membrane beads serve a functional role in stabilizing cylindrical membrane protrusions.

4.5 Discussion

PIP_2-like Membrane Beads Alter Flat Sheet Organization Through Cluster-mediated Interactions

Our goal was to determine the essential physics of I-BAR domain-mediated membrane remodeling. The generic nature of our model captured only the key characteristics of membrane and I-BAR domain family. We begin by accounting for shape, binding energy of I-BAR domains, and the deformation energy of the membrane. We show that our model can express a variety of organization and aggregation behavior that is consistent with previous BAR domain models using strictly pair-wise interactions. As the interactions between I-BAR domains are completely repulsive, membrane-mediated attraction is attributed to curvature-inducing and thermal Casimir-like forces. Indeed, when membrane-protein interactions are weak, the I-BAR domain induces little to no local membrane deformation and does not aggregate. When protein curvatures and membrane-protein interaction strengths are varied, we observe a variety of side-by-side and end-to-end I-BAR domain aggregates. The end-to-end aggregates are curvature-driven primarily due to the transverse curvature of the I-BAR domain. Each flattened protein generates local deformations in the form of a membrane trough and the proteins align to create a single long membrane trough. As the curvature along the long-dimension of the protein increases, the curvature coupling in the long dimension dominates and end-to-end aggregates give way to more stable side-by-side aggregates. When side-by-side aggregates form, we find that there is an increase in membrane curvature, but in our flat sheet simulations, we do not find any tubulation.

The absence of tubulation suggests that our model does not include details that are required for membrane tubulation. For example, there may be direct protein-protein interactions, or a more complex membrane binding surface (e.g., specific PIP_2 binding rather than) not being taken into considered in our generic model. We hypothesize that the competition between side-by-side and end-to-end aggregates to be important for the onset of tubulation as we found solely side-by-side or end-to-end aggregates to be insufficient to induce tubulation.

Lipid Clustering Helps Stabilize Membrane Protrusions

We introduced an attractive membrane bead meant to capture the effect of a PIP_2-like lipid that is specifically attracted to the ends of the I-BAR domain, thereby representing an additional membrane-mediated force. After characterization of I-BAR domain aggregates, we find that the PIP_2-like membrane beads drastically change the aggregation behavior on flat sheets. When comparing representative snapshots in Figure 4-2B, we find clustered PIP_2-like membrane beads are located at the ends of the I-BAR domain causing a competition between end-to-end and side-by-side aggregates at low surface coverages of PIP_2-like membrane beads.

When modeling protrusion stability, we notice that the protein aggregation behavior inside of tubules is directly affected by the addition of PIP_2-like, attractive membrane beads. When the protein has a uniform membrane binding surface, curved proteins will form side-by-side aggregates inside of the protrusions. However, gaps emerge between side-by-side aggregates that give rise to irregularly shaped membrane protrusions with constantly varying Gaussian curvature. When we introduce the attractive membrane beads and relatedly, the membrane binding surface, the PIP_2-like membrane beads cluster and serve to co-localize side-by-side aggregates, which collectively induce membrane tubule constriction. Ultimately, the presence of PIP_2 appears to yield protrusions with regular cylindrical shapes and reduced membrane curvatures. We find that PIP_2-like membrane beads provide a new membrane-mediated driving force that stabilizes protrusions. We hypothesize that the membrane-mediated interaction gives rise to the experimental observation that PIP_2 reduces the amount of bound I-BAR domains necessary to induce significant shape change.(Chen et al., 2015)

Connection to Experiments

We directly investigate the sensitivity of protein aggregation and tubule stabilization across a wide range of physically-relevant parameters. We consider this an advantage of phenomenological models. We now aim to connect various parameters of the model to experimentally measurable quantities. For example, the membrane parameters are adjusted such that the membrane bending rigidity is consistent with that of *in vitro* membranes.(Prevost et al., 2015) Another key parameter was the protein curvature, which we varied to reflect the diversity throughout the I-BAR domain family and the flexibility seen in atomistic simulations.(Jarin et al., 2019; Saarikangas et al., 2009; Takemura et al., 2017) As we investigated protein curvature, we found that proteins with insufficiently small curvature fail to stabilize protrusions, which would be an interesting hypothesis to test experimentally. Finally, we note that the external potentials used to initiate tubule formation is conceptually similar to the force applied in tube pulling assays. In these experiments, the membrane tubules are not generated spontaneously by I-BAR domains, but are formed by beads pulled by pipettes or optical traps, allowing the I-BAR domains to sort into the tubules.(Heinrich et al., 2010; Prevost et al., 2017) Analogously, the protrusions in our simulations are initially generated by an external potential and the I-BAR domains sort into/out of the tubules depending the protein and protrusion curvature. Our results show protein sorting as curved and flattened proteins have higher number densities outside of the membrane protrusions instead of inside.

The strength of membrane-protein interactions is another key parameter that our simulations identified as a modulator for protein aggregation behavior. We probe the nonspecific attraction to the membrane as well as the specific attraction to a PIP_2-like membrane bead. Experimentally, both specific and nonspecific interactions may be tuned by adjusting electrostatic forces between I-BAR domains and the membrane. The specific interactions could be tuned by changing the location positively-charged residues in the I-BAR domain sequence to change the charge density of the membrane binding surface while the nonspecific interactions could be tuned by varying the membrane composition with fewer anionic lipids, varying salt content, or mutating the I-BAR domain to change the total number of positively-charged residues. Thus, our findings provide intuition for future experiments investigating the role of lipid clustering and the membrane binding surface.

4.6 Conclusion

Our phenomenological model shows the stabilization of membrane protrusions by an I-BAR domain interaction mediated by clustering lipids to the ends of I-BAR domains. We demonstrate I-BAR domain aggregation and organization of flat membranes is significantly different when PIP_2-like membrane patches are present. When I-BAR domains have a uniform membrane binding surface, curved I-BAR domains form side-by-side aggregates. When I-BAR domains have preferred lipid binding domains, curved I-BAR domains form a variety of aggregates due to competition between membrane-mediated forces (i.e., forces due to curvature induction, thermal fluctuations and lipid clustering). Thus, we show that membrane composition directly affects membrane aggregation resulting in differing membrane protrusion stabilization.

Chapter 5: Evaluation of top-down model from the statistical mechanical perspective

5.1 Summary

We use the popular MARTINI coarse-grained model as a test case to analyze issues of transferability and representability in top-down models. Despite the theoretical and numerical investigation in bottom-up modeling, these issues have not been investigated in the top-down models (i.e. models parameterized to match experimental observables). Specifically, we compare the height fluctuation spectra and the temperature dependence of various potentials of mean force between MARTINI 2.0, Dry MARTINI, and mapped all-atom simulations in lipid bilayers. We find that while the amplitudes of fluctuation modes are different between models, we do not find that these issues arise from the representability of the height fluctuation spectrum. Although, we find that MARTINI models do not capture the lipid structure seen in atomistic simulations, and issues of temperature transferability may arise due to incorrect enthalpy-entropy decomposition of pair-wise potentials of mean force.

5.2 Introduction

As discussed in 0, there are two general methodologies to parameterizing a CG model. MARTINI models are commonly used top-down CG models most notably for their computational efficiency. The significant speedup compared to all-atom models arises, in part, from the reduced resolution, i.e. mapping approximately 4 heavy atoms-to-1 CG bead. Specifically, the reduced representation significantly decreases the computational effort needed for a given simulation by reducing the calculations per time step, while also increasing the diffusion coefficients of the system by a factor of 3-6, further increasing the temporal sampling of a system.(Marrink et al., 2004; Marrink et al., 2007) Additionally, employing an integration time step much longer than typical atomistic systems balances the efficiency and accuracy when sampling the underlying energy landscape. The overall computational saving has proven attractive when modeling complex systems, which require longer length and time scales than available with atomistic models.

Additionally, the parameter set necessary to run the diverse set of molecules is reduced to only the polar (P), nonpolar (N), apolar (C), and charged (Q) bead types, which leads to a simpler parameterization and optimization strategy. The original parameters were fit using the octanol-water partition coefficient to capture the free energy of transfer from the hydrophobic region of a lipid bilayer to the fully hydrated exterior.(Marrink et al., 2007) The simplicity of bead types and potentials gives MARTINI models a modular and additive representation that provides an initial guess for force field parameters for new molecules, which is another factor for its wide spread use for complex systems.

Although MARTINI is widely accepted, there are obvious drawbacks in accuracy due to lack of a rigorous mapping, and the top-down parameterization. The approximation of 4 heavy atoms-to-1 CG bead is a poor description for chemically similar lipids because lipids with different tails are represented by the same MARTINI models e.g. a single model represents C12:0 dilauroyl (DLPC) and C14:0 dimyristoyl (DMPC) tails. There is no rigorous correspondence between a MARTINI model and the atomistic model. The MARTINI model is a general model with loose phenomenological connections to the underlying atomistic models. This is further reinforced in the parameterization scheme, which is based on matching experimental data and not the atomistic observables.

The problem in mapping 4 heavy atoms-to-1 CG bead is most apparent when describing water i.e. 4 waters-to-1 CG bead. As has been discussed elsewhere(Hadley and McCabe, 2010; Marrink et al., 2010; Winger et al., 2009; Yesylevskyy et al., 2010), MARTINI water has several problems, including freezing point, diffusion, and hydration behavior. This is apparent in our analyses of lipid bilayer structure, but is not the focus of the current study. Instead, we focus on issues of transferability and representability as they pertain to properties of lipid bilayers. We analyze lipid structure and bilayer undulations of the MARTINI models and compare to analyses from mapped atomistic trajectories.

In this study, we use frequently studied C18:1 dioleoyl (DOPC) and cholesterol bilayers as test cases because lipid bilayers provide adequate configurational sampling at MARTINI and atomistic resolutions, and clear comparisons to fundamental statistical mechanical quantities e.g. radial distribution function. Additionally, several properties of lipid bilayers calculable from molecular simulation have a direct experimental analogue i.e. bending modulus from the fluctuation spectra. These quantities and their connection to the CG theory issues of transferability and representability are thoroughly described in the Theory section. Finally, we present a more applied example of differences between MARTINI and atomistic resolutions, the lateral association of two amphipathic helices embedded in a lipid bilayer. We characterize the lateral association energy by quantifying the potential of mean force (PMF) with respect to the center of mass (CoM) distance between two H0 helices of amphiphysin.

5.3 Theory

Generally speaking, CG potentials are parameterized to capture the behavior of a system using a reduced representation compared to the underlying finer-grained (e.g., atomistic) system. Since the reduced number of degrees of freedom in the CG system cannot fully capture the entropy of the system, a state dependence arises, and the CG potential is conditionally parameterized for a specific ensemble. The state dependence becomes an issue when a CG potential is used outside of the specific ensemble for which the CG potential is necessarily parametrized (e.g., at a higher temperature, with different composition, or at an increased surface tension).

In the case of the MARTINI CG force field, one might expect the same transferability problem to arise considering the approximate 4 to 1 mapping scheme and the interaction parameterization scheme detailed elsewhere.(Marrink et al., 2007) Here, we look first at temperature transferability. More specifically, we calculate the entropic contribution to a CG PMF. We will calculate a midplane-projection of the pairwise PMF in a single leaflet and its decomposition into the enthalpic and entropic contributions (see Methods for further computational details).(Lu and Voth, 2011) This decomposition of the PMF gives direct insight into the nature of the temperature dependence and provides deeper understanding of the missing contributions from the underlying atomistic scale. This statistical mechanical insight into the nature of MARTINI CG interactions will provide a better understanding of how issues of transferability may occur in top-down models.

Representability issues in CG models, often more opaque, are related to the roles of resolution and the parameterization scheme in defining observables and calculating them.(Dunn et al., 2016; Johnson et al., 2007; Wagner et al., 2016) We seek here to understand the nature of representability as it applies to a common lipid observable, the height fluctuation spectrum.(Brandt et al., 2011; Watson et al., 2012) This technique is widely used to relate molecular simulations to an analytical, continuum representation of an elastic sheet. The height fluctuation spectrum provides a direct connection to Helfrich theory and a means to estimate the bending modulus.(Helfrich, 1973) This method has been applied to all-atom and CG model alike, with the general expectation that the spectrum converges to the continuum result at low wavenumbers. However, the relationship between the resolution of the molecular model and the height fluctuation spectrum is poorly understood.

We expect mapped all-atom (i.e. all-atom models mapped to the resolution of MARTINI models) and MARTINI fluctuation spectra to converge to Helfrich-like behavior in the low wave number regime, where the corresponding length scale is much larger than a single phosphorus atom (or phosphate bead) so we do not expect an issue. Instead we are interested in the intermediate regimes, which can be related to molecular level detail as the corresponding length scale is of the order of these simulations. Unless one makes a specific effort to use a large simulation box, the spectrum does not converge and presents difficulties in fitting the bending modulus. The height fluctuations in an intermediate regime, 0.3 nm^{-1} to 3 nm^{-1}, are of specific interest because they are on the length scale of membrane-mediated interactions and have implications for using MARTINI to understand association of lipids or proteins.

5.4 Methods

Simulation Details

All simulations were performed using GROMACS simulation suite (versions 5.1.4 and 2016.3)(Abraham et al., 2015) using the CHARMM36 atomistic force field,(Best et al., 2012) MARTINI 2.0 and Dry MARTINI CG force fields.(Arnarez et al., 2015; Marrink et al., 2007) The simulations for structural analysis of DOPC and cholesterol were 270 DOPC molecules and 68 cholesterol molecules with ~36 waters per lipid (12,180 water molecules). The initial structures were generated using the CHARMM-GUI.(Jo et al., 2007; Jo et al., 2008; Lee et al., 2016) Each simulation ran for a microsecond at 1 atm using the Nose-Hoover thermostat with coupling of 1ps and the semi-isotropic Parrinello-Rahman barostat(Parrinello and Rahman, 1981a) at 1.0 atm with coupling of 5ps. The simulations to study the undulation spectrum had 1152 DOPC molecules for the smaller 20nm by 20nm box size and 28,800 DOPC molecules for the larger 100nm by 100nm box size. The amphipathic helix simulations contain two H0 helices of endophilin in a bilayer of 140 DOPC lipids hydrated by 4,200 water molecules and 0.15 M of KCl. After embedding, umbrella sampling is used to retrain the CoM distances every 0.1 nm from 1.5nm to 3nm using a harmonic umbrella potential with force constant 15 kJ/mol implemented in PlumeD v2.3.(Torrie and Valleau, 1977; Tribello et al., 2014) Finally, weighted histogram analysis method was used to reproduce the PMF from the final 100ns of a 150ns run at each window.(Grossfield, 2013; Kumar et al., 1992)

XY-projection of Per Leaflet Radial Distribution Function

We analyze the xy-projection of the radial distribution function (RDF) for lipids of each leaflet as a way to understand the lateral association of lipids and cholesterol in each leaflet. First, the xy-projection is used because lipid bilayers are not spherically symmetric and the normalization of a radial distribution function assuming spherical symmetry results in an RDF that decays to 0 rather than the typical 1. Second, distinguishing between leaflets provides deeper insights to the near 0 distance behavior. The xy-projection for both leaflets shows significant nonzero behavior at distances near 0 corresponding to the lipid in the opposite leaflet. The contribution from lipid in the opposite leaflet is not significant but complicates distinguishing overlap between a choline-group of a lipid and a hydroxyl-group of cholesterol in the same leaflet, which is significant. Thus, the per leaflet xy-projection provides deep insights into the lateral behavior of the bilayer and correctly attributes xy-overlap to headgroup interaction instead of effects due to the opposite leaflet. Given the significant amount of cholesterol flip-flop, we assign each lipid to a leaflet by determining if the tail bead (C4A bead of DOPC or C2 bead of cholesterol) is at least 0.5 nm above or below the head group bead (NC3 bead of DOPC or ROH bead of cholesterol). Comparison between leaflets also provides a means of determining convergent behavior as both leaflets of a symmetric bilayer should show similar behavior. Notably, beads that are not interacting in the same leaflet and can share xy-position. Thus, their corresponding RDF will be relatively flat. Similarly, little information can be gleaned from the RDF of a head group bead (e.g., choline bead) and corresponding tail bead (e.g., terminal hydrophobic bead). The RDF remains near 1, as the position of the head group bead has a small effect on the corresponding position of the tail bead in disordered lipid bilayers.

Next, we use the reversible work theorem to determine the corresponding potential of mean force (PMF) from the RDF. This PMF is a measure of the stability of two lipid beads sharing the same xy-position in a leaflet of the bilayer. For lipid beads with the same preferred z-position in the membrane (i.e., overlapping projected z-density), the PMF is a proxy for the lateral association energy of the two beads. After assuming negligible heat capacity change, the PMFs can be decomposed across different temperatures according to equation 5.1(Lu and Voth, 2011).

$$\Delta W(r,T) = \Delta H(r) - T\Delta S(r) \tag{5.1}$$

Numerically, we evaluate the decomposition by performing a linear fit of PMF and temperature at a given distance value to determine the slope (i.e., the entropy term) and y-intercept (i.e., the enthalpic term). The enthalpic and entropic terms can subsequently be plotted at each distance value and compared across models to gain a deeper understanding of the lipid association behavior and bead-wise interactions in the bilayer.

5.5 Results

Lateral Association

When considering the lateral organization of cholesterol in a bilayer and the main driving forces, we first analyze the previously described per leaflet xy-projection of the RDF. This is a subset of the RDFs shown in Figure 5-1 using the MARTINI bead naming convention (i.e., NC3 corresponds to the choline group of the headgroup).(Marrink et al., 2007)

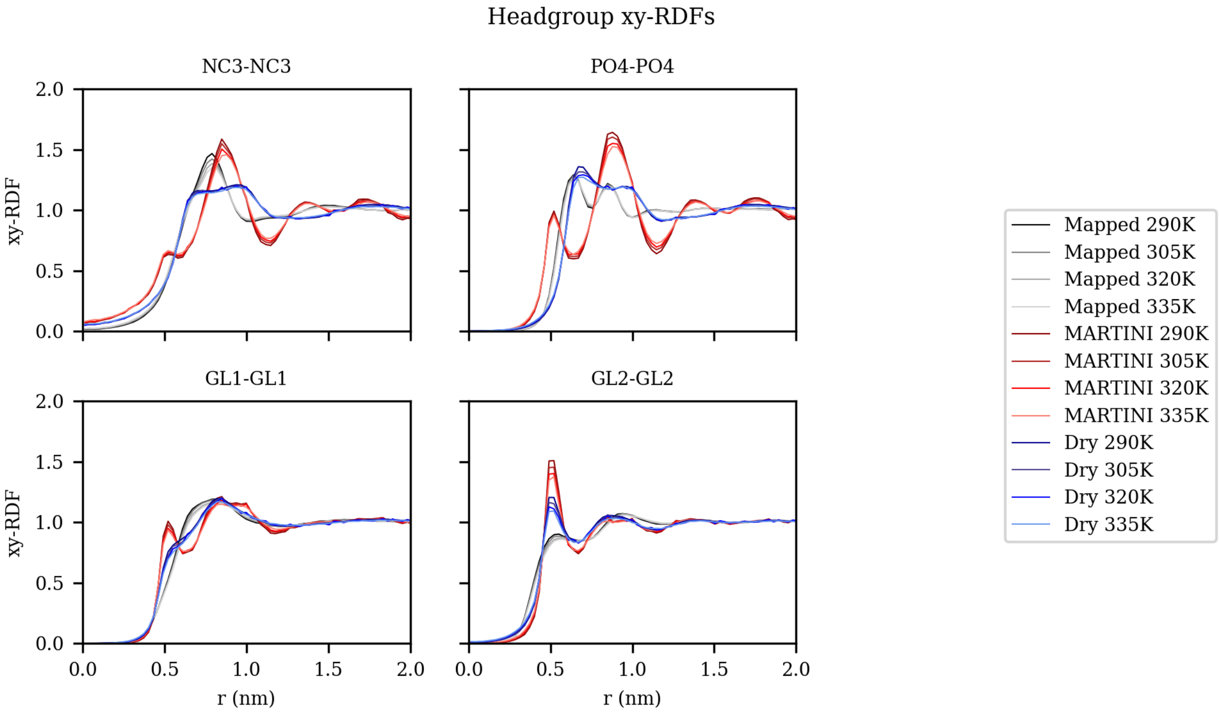

Figure 5-1: DOPC Headgroup xy-RDFs

XY-projection of the Radial Distribution Function for DOPC head group beads comparing the Mapped All-Atom, MARTINI, and Dry MARTINI models at various temperature averaged per leaflet.

Generally, when comparing the MARTINI RDFs to atomistic reference for head group beads, we find that MARTINI 2.0 over-structures and Dry MARTINI under-structures. More specifically, when comparing choline-choline (NC3-NC3) RDFs to the mapped atomistic reference (shown in Figure 5-1), MARTINI 2.0 model produces an initial shoulder and secondary features after the first peak and Dry MARTINI models produce a broader, shorter first peak. Additionally, we see little qualitative variation across the 45K temperature range analyzed, but significant quantitative differences in the first peak heights. The phosphate-phosphate (PO4-PO4) RDF shows similar behavior: the MARTINI 2.0 model has a significant initial feature before the first peak and the Dry MARTINI model produces broader peaks. There are strong differences between the DOPC structures.

When considering the DOPC-cholesterol structures, we find significant over-structuring in both the MARTINI 2.0 and Dry MARTINI. The glycerol-hydroxyl per leaflet xy-projection of the RDF shown in Figure 5-2 shows differences between the models and drastic temperature variations in the CG models that are not present to the same extent in the mapped atomistic RDFs.

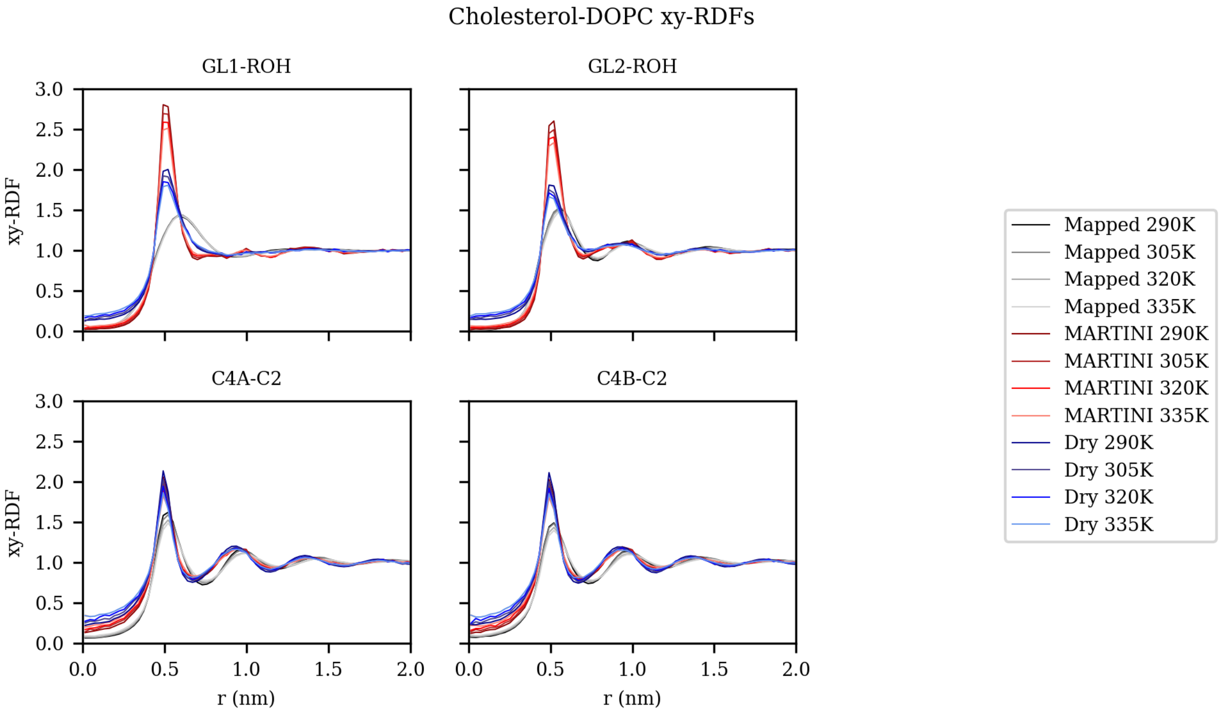

Figure 5-2: Cholesterol-DOPC xy-RDFs
XY-projection of the radial distribution function of cholesterol hydroxyl bead (ROH) and DOPC glycerol beads (GL1, GL2) and cholesterol tail (C2) and DOPC terminal carbons (C4A, C4B) comparing the mapped all-atom, MARTINI, and Dry MARTINI models at various temperature averaged per leaflet.

In the remaining RDFs shown in Section 5.8, we clearly see very good agreement between the mapped atomistic reference and both MARTINI models. This is expected considering other analyses of the tail group behavior of MARTINI models.(Baron et al., 2006) Furthermore, this suggests that the tail group behavior is a packing problem and one that is well captured by the Lennard-Jones-like interactions of the MARTINI models.

Enthalpy -Entropy Decomposition of Lateral Potential of Mean Force

By simulating the DOPC/Cholesterol system at various temperatures, we can use a finite difference calculation (see section 5.4) to decompose the PMF into the enthalpic and entropic contributions. In Figure 5-3, we first look at the head group interactions and generally find there is enthalpy-driven structuring. In MARTINI and Dry MARTINI, we see over-structuring and under-structuring, respectively. We note the entropic contribution mirrors the enthalpic contributions and provides a driving force to reduce structuring as temperature increases i.e. the entropic contributions are most negative where the enthalpic contributions are most positive.

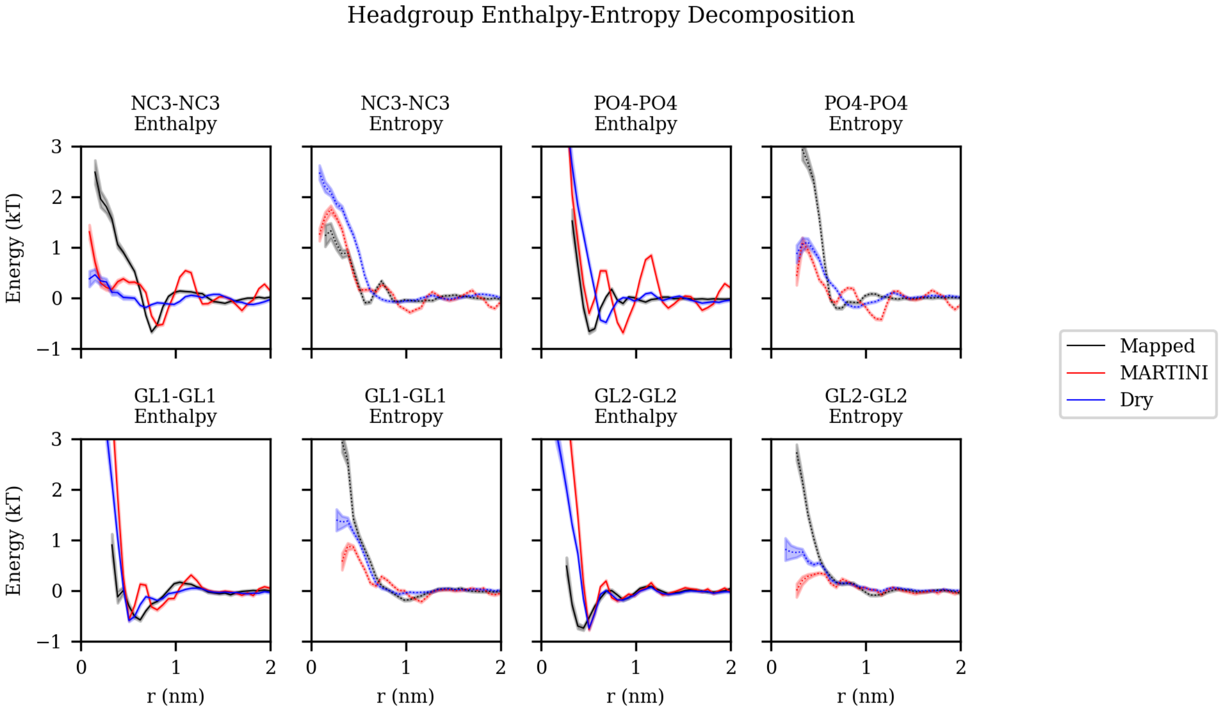

Figure 5-3: DOPC Headgroup Enthalpy-Entropy Decomposition
Entropy-enthalpy decomposition of potential of mean force between DOPC head group and glycerol beads comparing mapped all-atom, MARTINI, and Dry MARTINI models.

Next, we investigate the DOPC-cholesterol interactions (Figure 5-4). Based on the previous decompositions, we expect there to be strong enthalpic contributions that lead to the over-structuring and commensurate entropic contributions to reduce structure as temperature increases. Indeed, this is what we see in the MARTINI models. Specifically, we note that the MARTINI 2.0 decomposition has larger magnitude enthalpy and entropy than the other models at close distances. However, this is the opposite of what we see for the mapped atomistic model. There are small enthalpic contributions to the PMF and significant enthalpic contributions at the less than 0.5 nm distances. This suggests that the preferred distance between glycerol beads and hydroxyl bead of cholesterol is entropically driven in the case of the mapped atomistic system and the opposite is true for the MARTINI systems.

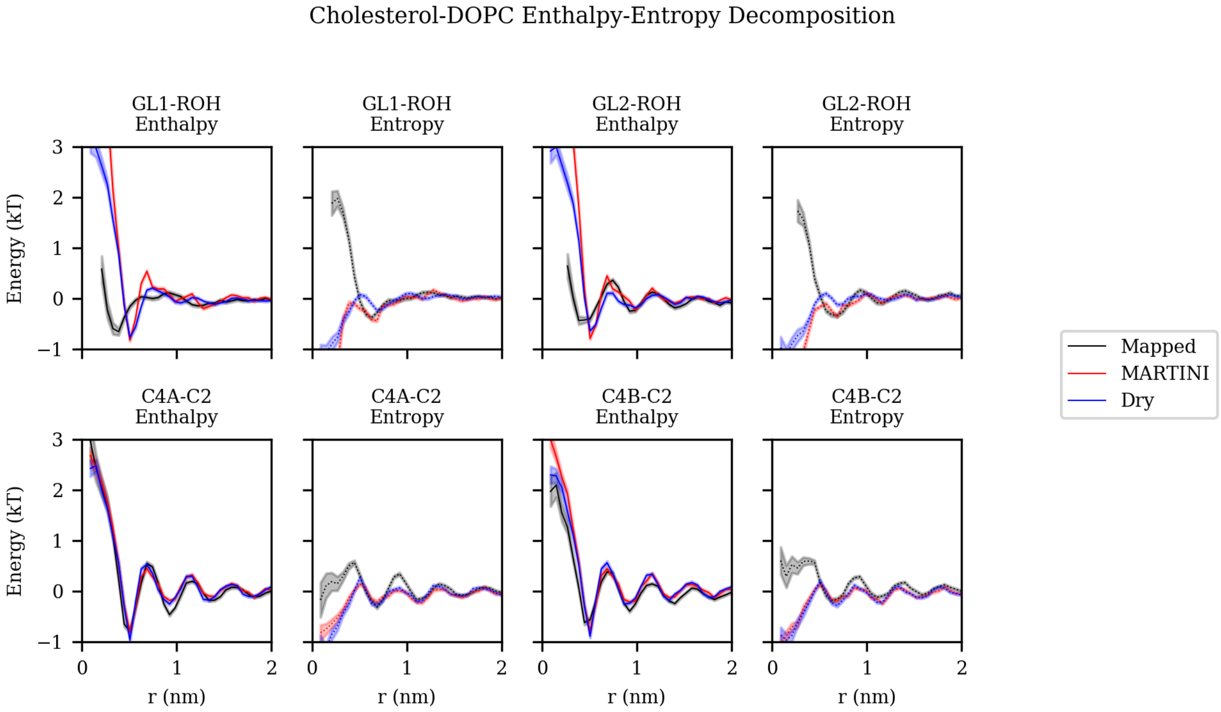

Figure 5-4: Cholesterol-DOPC Enthalpy-Entropy Decomposition
Entropy-enthalpy decomposition of potential of mean force between cholesterol hydroxyl group and DOPC glycerol beads and cholesterol tail group bead and DOPC terminal tail beads comparing mapped all-atom, MARTINI, and Dry MARTINI models.

Additionally, in Section 5.8 we again see clear agreement between the mapped atomistic reference and both MARTINI models. We expect good agreement between enthalpic contributions to the PMF between the various models because we have generally seen that enthalpy drives the structuring and there is good agreement between the mapped atomistic, MARTINI, and Dry MARTINI tail structure. Given other studies of the entropy contributed by the tail of MARTINI models, we also expect entropic contributions to be very similar.(Baron et al., 2006)

Large System Undulation Spectrum

Using a 20nm by 20nm lipid bilayer patch of DOPC, we investigate the effects of mapping and model resolution on the height fluctuation spectrum. The height fluctuation spectrum is most commonly used to approximate the bending modulus as described by Helfrich theory. Additionally, bilayer height fluctuations are also an important driving force of protein and raft aggregation because the undulations in the bilayer have highest number of accessible states when the proteins that dampen fluctuations are closer together. From our calculation of the height fluctuation spectrum in Figure 5-5, the mapping itself does not affect height fluctuation spectrum and subsequently, does not affect the estimation of the bending modulus.

Next, we see how the model resolution can affect the height fluctuation spectrum. As seen in Figure 5-5, we find that the spectra from the mapped all-atom, MARTINI 2.0, and Dry MARTINI models do not qualitatively differ in the lower q region, which is used to estimate the bending modulus, but differ quantitatively in their estimate of bending modulus. The mapped atomistic model had the lowest bending modulus of 22kT, then MARTINI 2.0 at 27kT, and Dry MARTINI at 38kT. MARTINI systems permit a much larger system (e.g. 100nm x 100nm) to be simulated and show that the height fluctuation spectrum converges to Helfrich theory. There is some difference between MARTINI 2.0 and Dry MARTINI; The shape and the magnitude of both the CG spectra differ from the mapped atomistic spectrum in the intermediate q regime i.e. 3nm^{-1} to 0.3nm^{-1}. This regime lacks a direct correspondence between the MARTINI and mapped atomistic models but is on the order of membrane-mediate protein interactions.

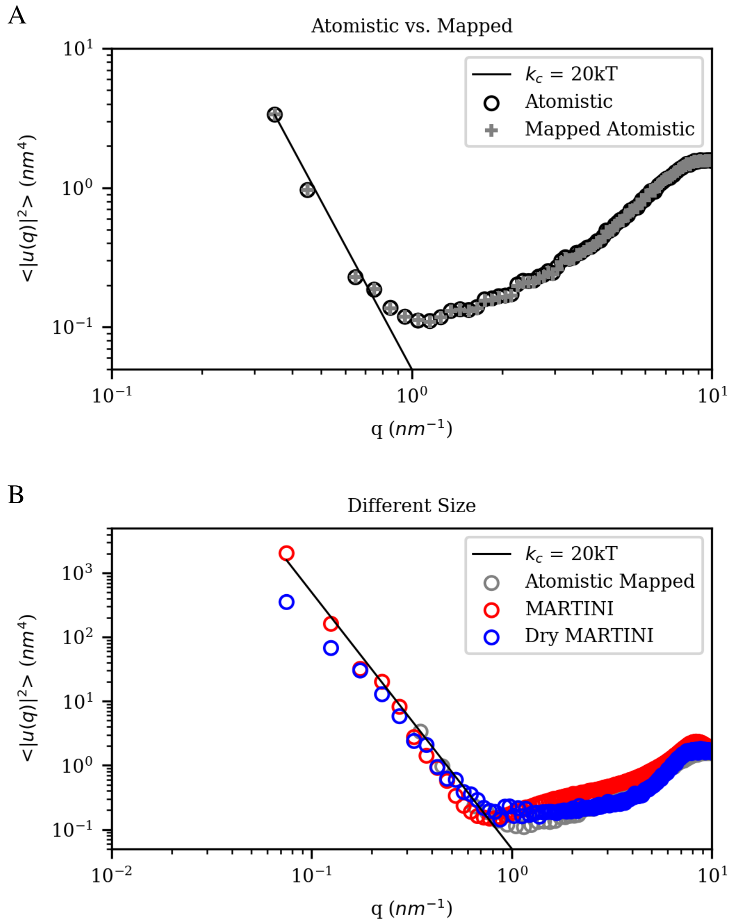

Figure 5-5: Height Fluctuation Spectra Comparisons
Comparison of undulation spectra. Spectra calculated using the mapped phosphate bead compared to spectrum calculated using phosphorus atom (A). Comparison of spectra calculated from 20nm by 20nm mapped atomistic, 100nm by 100nm MARITNI, and 100nm by 100nm Dry MARTINI models (B).

H0 Lateral Association

As a final test case to assess the differences between atomistic and MARTINI coarse-grained models, we analyze the lateral association energy of two embedded helical peptides by computing the PMF with respect to the CoM distance. The embedded peptides have nonpolar residues pointing toward the hydrophobic core of the bilayer while the polar or charged residues interact with surround solvent or lipid head groups. The system is large enough such that the PMF between them has decayed to 0 kJ/mol. In this configuration, the dipoles of the amphipathic helices are parallel resulting in the purely repulsive behavior, shown in Figure 6B. The decay of the association energy is similar in both atomistic and MARTINI models. There is a significant magnitude difference in the repulsion energy between the two models with atomistic model reaching a repulsive energy of 10 kJ/mol at a CoM distance of 2.0 nm while the MARTINI model never reaches 10 kJ/mol within the region sampled.

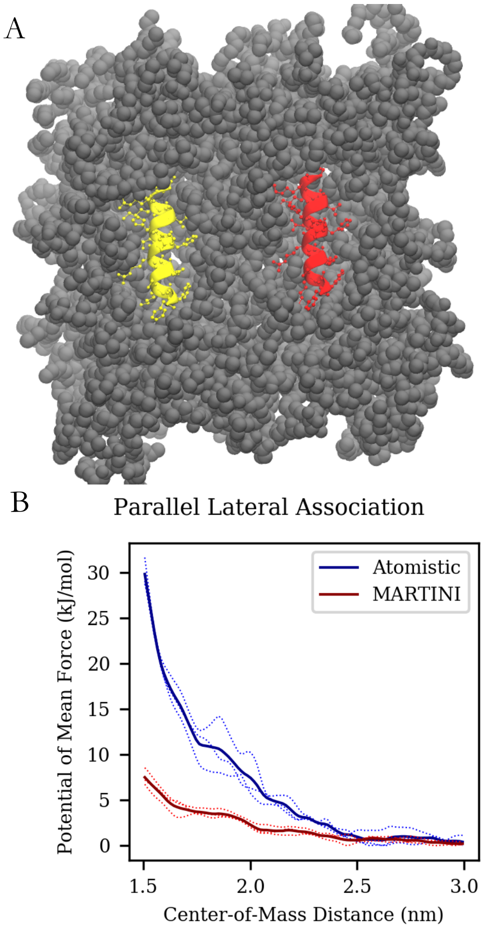

Figure 5-6: Snapshot of embedded amphipathic helices and a center of mass distance PMF between the embedded helices.
Snapshot of H0 helices (red and yellow) embedded in DOPC bilayer (gray) with a center-of-mass distance of 3.0nm (A) and potential of mean force as a function of center of mass distance between two embedded helices (B).

5.6 Discussion

Driving Forces of Lateral Association

In Figure 5-1 and Figure 5-2, we seek to describe the lateral structure of a DOPC-cholesterol bilayer and in Figure 5-3 and Figure 5-4, the decomposition of the enthalpic and entropic contributions to the lateral PMF that give rise to this association to evaluate the temperature transferability. We see that there is good agreement in the hydrophobic region of the bilayer and significant disagreement in the head groups in both structure and driving force. This observation is best highlighted by considering structure and driving force between glycerol beads of DOPC and the hydroxyl-group bead of cholesterol. We show in Figure 5-4 that not only are both MARTINI systems drastically over-structuring, but that interaction is enthalpically driven in the MARTINI systems and entropically driven in the mapped atomistic system. Thus, the temperature behavior of the MARTINI systems will be inherently wrong due to incorrect attribution of driving force. Additionally, the rather simple mixing rules for MARTINI beads may be missing key behavior.

Effects of Mapping and Model Resolution on Estimation of Bending Modulus

We show that the mapping has little or no effect on the height fluctuation spectrum and the subsequent estimate of bending modulus. We attribute this to our original definition of the bilayer mid-plane, which is the average position of the Phosphorus atoms in each leaflet. Thus, when the Phosphorus atom is mapped with the bonded Oxygen atoms to the Phosphate bead, there is relatively little change in the estimation of the midplane and thus, little to no problems arise due to representability. This would not be true if our definition of the midplane was more sensitive to mapping. For example, if the tail carbon was chosen, there would be a significant effect because the position of the last tail carbon is significantly more dynamic than the last tail bead. As a result, the mapped spectrum would have an effective filter upon the short wavelength (high q) fluctuations, which should not affect the estimate of the bending modulus. However, this should only act as a short wavelength filter and the undulation spectrum should converge to q^{-4} as expected from Helfrich theory.

The role of model resolution on the height fluctuation spectrum is of utmost importance especially when considering the growing number of coarse-grained lipid models. We show in Figure 5-5that the height fluctuation spectrum at long wavelengths (low q) does not have significant effects due to model resolution and does converge to q^{-4} as expected from Helfrich theory. When considering the role of mappings and the definition of the bilayer midplane, this result is not surprising. However, a surprising finding here is the behavior in the intermediate q regime of the height fluctuation spectra. The intermediate regime corresponds to distances on the order of membrane-mediated protein-protein interactions and the role of lipid bilayer fluctuations is exceedingly important. We note that there are much higher fluctuations in the case of the MARTINI 2.0 model and Dry MARTINI model spectra have a qualitative shape difference as compared to the mapped atomistic spectra. This suggests that the nature of these fluctuations may not be the same. Indeed, we expect the solvent (or lack thereof) and the interactions between adjacent lipids to play a significant role in the nature of lipid height fluctuations. When considering the case of MARTINI 2.0, these fluctuations are dampened by the solvent and interactions between adjacent lipids while the fluctuations in the Dry MARTINI model are dampened solely by interactions of adjacent lipids. While the reduced representation of water in the MARTINI 2.0 model was not a primary interest of this work, it cannot be ignored when considering the fluctuations and potentially over-stabilizing lipids moving out of the membrane. We suggest that an overly attractive force between water and lipids pulls lipids out of the membrane and stabilizes local fluctuations. In the Dry MARTINI case, there is no solvent preventing lipids moving out of the bilayer, and these local fluctuations are not hampered at all. In both cases, it is concerning that these intermediate fluctuations differ qualitatively and quantitatively from the mapped atomistic.

Lateral Association of Embedded Amphipathic Helices

Amphipathic helices are an important motif for membrane protein targeting and assembly.(Cui et al., 2013; Fernandes et al., 2008) The association energy between two amphipathic helices is an essential test case because it is affected by the over-structuring behavior seen in Figures 1 and 2 and the height fluctuation difference seen in Figure 5. Additionally, it may also be affected by the solvent behavior of the 4 waters to 1 bead mapping as solvating the amphipathic helices is different in the two models as well. Finally, it is a multicomponent system and suffer from the lateral association issues seen in Figure 4 (i.e., the model may be erroneously attributing the repulsion to either enthalpic or entropic effects). Therefore, we do not expect the lateral association energy to be quantitatively correct and Figure 6 shows the repulsive energy between the two amphipathic helices is underestimated. Although, this relatively simple membrane-protein system compared to the systems that have been are expected in coming years.(Marrink et al., 2019)

5.7 Conclusion

The MARTINI models have become widely used due to ease of use and applicability to a wide variety of systems. We investigate the lateral ordering and driving forces of a relatively simple case of DOPC and cholesterol by calculating the per leaflet xy-projection of the RDF and the enthalpy-entropy decomposition of the corresponding PMF. We find that the hydrophobic region of the bilayer is described as well by the MARTINI models as a mapped atomistic system. Perhaps unsurprisingly, we find significant disagreement in the head group region, both in structuring and enthalpy-entropy decomposition. We find the cross interactions between the components to be erroneously driven by enthalpy as opposed to entropy as seen in the mapped atomistic system. This should give pause to a growing community of modelers investigating complex coarse-grained systems across wide temperature ranges. The enthalpy-entropy balance is commonly delicate in biological systems and the misattribution of enthalpy-entropy in interactions will result in erroneous conclusions when the model is used to simulate systems at different state points than initially parameterized.

Additionally, we investigate the role of mapping and model resolution on the frequently calculated height fluctuation spectrum. We find at low wavenumbers (or long wavelengths) that the spectrum is only quantitatively affected by mapping or model resolution, while at intermediate wavenumbers corresponding to wavelengths of 0.3nm to 3nm, there are deviations between the models. These differences are not due to the representability of the height fluctuation spectrum, but we suggest that these differences are more likely due to model differences in the inherent stability of undulating membranes. We suggest that the solvent plays a significant role. Indeed, the 4 waters to 1 bead mapping in the case of the MARTINI 2.0 or no solvent in the case of Dry MARTINI do not provide the correct energetic barriers to bilayer undulations, resulting in higher magnitude fluctuations in the MARTINI models. We expect this to impact directly membrane-mediated interactions between anything that dampens membrane fluctuations. Indeed, when we investigate the lateral association of two embedded helices we find an underestimation of their repulsion energy.

5.8 Supplemental Information

It has been discussed elsewhere that the entropy of the lipid tail groups is well captured by the MARTINI model.(Baron et al., 2006) As such, we expect the radial distribution functions (RDFs) and the atomistic temperature behavior to be recapitulated. The following figures presenting the tail bead RDFs and corresponding enthalpy-entropy decompositions, which are calculated as described in the Section 5.4.

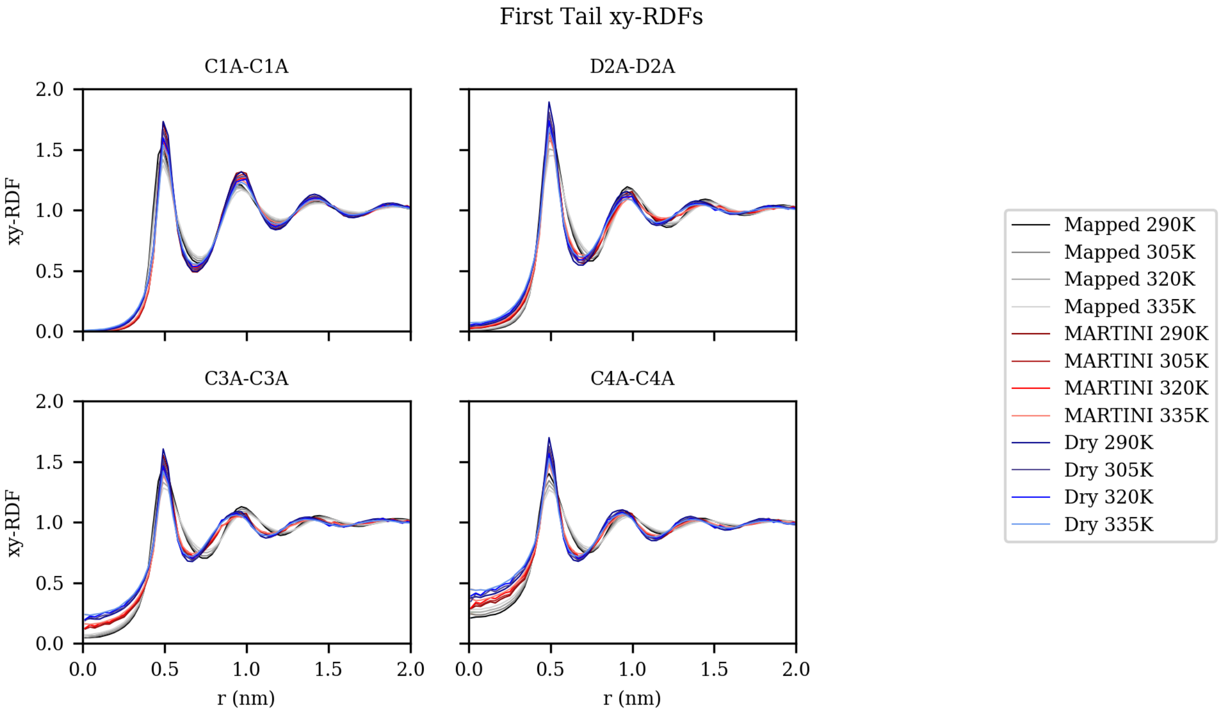

Figure 5-7: First Tail of DOPC xy-RDFs

XY-projection of the radial distribution function for DOPC first tail beads comparing the mapped all-atom, MARTINI, and Dry MARTINI models at various temperature averaged per leaflet.

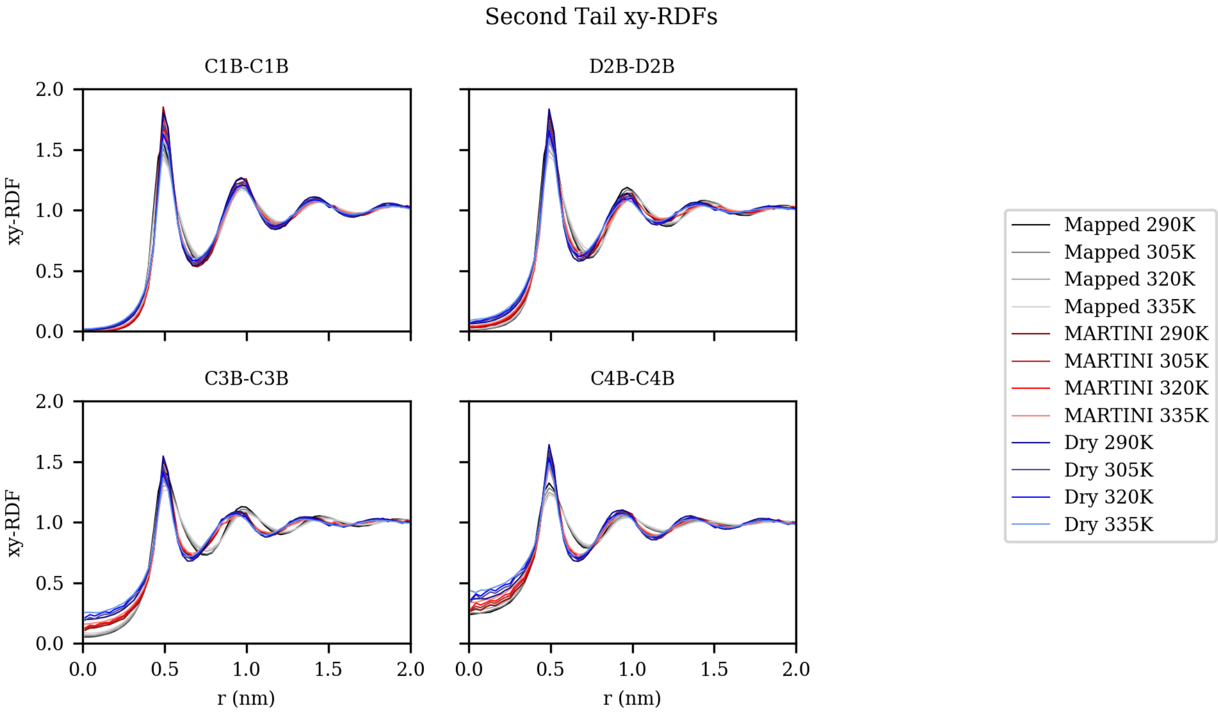

Figure 5-8: Second Tail of DOPC xy-RDFS
XY-projection of the radial distribution function for DOPC second tail beads comparing the mapped all-atom, MARTINI, and Dry MARTINI models at various temperature averaged per leaflet.

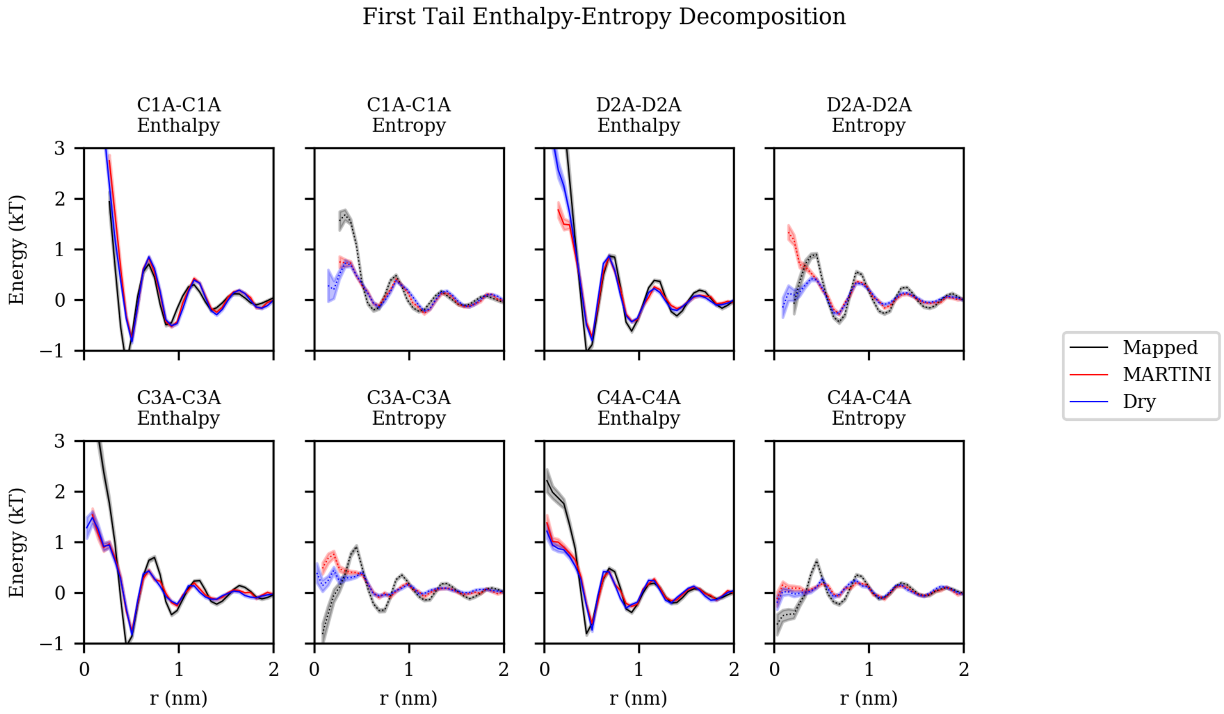

Figure 5-9: First Tail of DOPC Enthalpy-Entropy Decompositions
Entropy-enthalpy decomposition of potential of mean force between first tail beads of DOPC comparing mapped all-atom, MARTINI, and Dry MARTINI models.

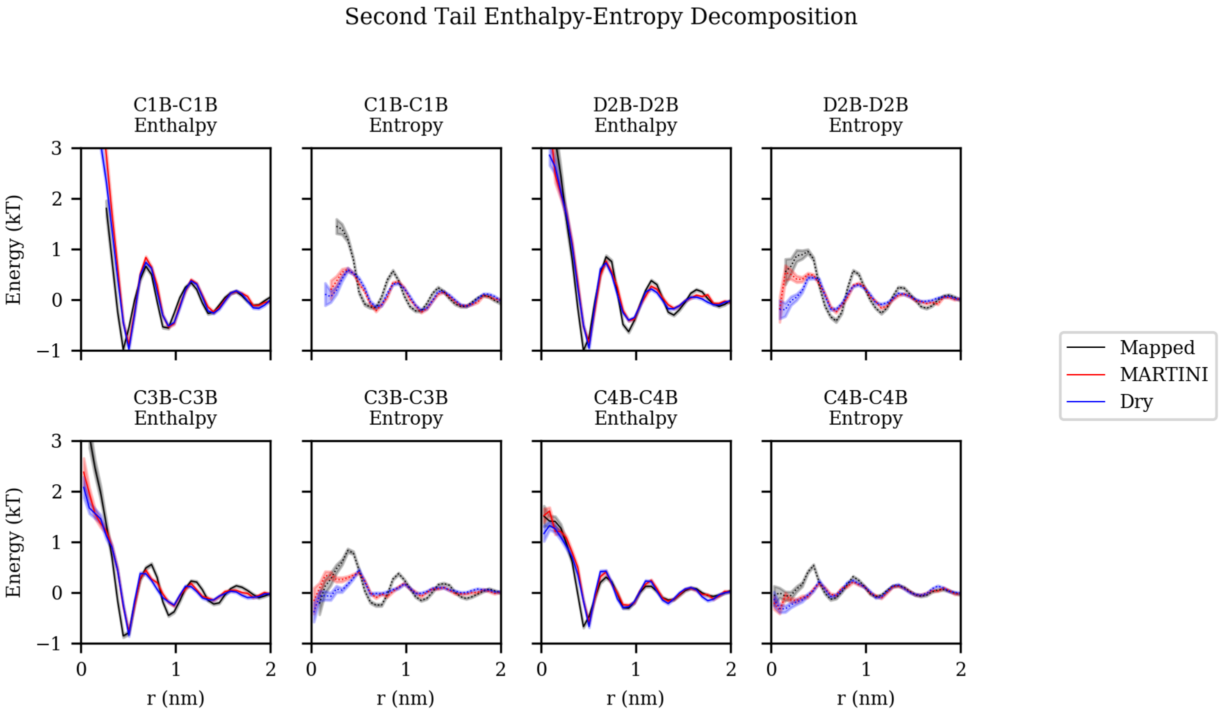

Figure 5-10: Second Tail of DOPC Enthalpy-Entropy Decompositions

Entropy-enthalpy decomposition of potential of mean force between second tail beads of DOPC comparing mapped all-atom, MARTINI, and Dry MARTINI models.

Chapter 6: Conclusions and Future Directions

6.1 Conclusions

Proteins facilitate reshaping membranes for many cellular processes and the BAR superfamily is a key group of membrane remodeling proteins. In particular, I-BAR proteins help form protrusions from the cell. The goal of this thesis is to elucidate the mechanism of membrane remodeling by I-BAR domains and the self-assembly that precedes deformations on the order of microns and probe systemic issues in coarse-grained modeling. To this end, all-atom, coarse-grained, and field theoretic models are used to elucidate the membrane-bound behavior of I-BAR domains. Coarse-grained models are given particular emphasis because coarse-grained models resolve assembly of individual membrane-bound proteins on near-micron length scales.

Chapter 3: is dedicated to a multiscale modeling approach to understanding I-BAR domain aggregation and the influence of membrane geometry. The coarse-grained model of I-BAR domains was parameterized using all-atom simulation data and combined with previously parameterized lipids. Inside of membrane tubules, I-BAR domains form axial aggregates and on vesicles, I-BAR domains create and surround membrane domes, in contrast to other phenomenological models that show circumferential aggregation. The observed coarse-grained behavior is similar to images from confocal microscopy, which show I-BAR domains creating flower pedal shapes out of giant unilamellar vesicles rather than membrane tubules.

Chapter 4: presents a phenomenological approach to modeling I-BAR domains assembly and the influence of negatively charged lipids in stabilization of protrusions. The coarse-grained model developed varies the curvature and membrane binding surface to produce a broad emergent phenomenology. Additionally, the model shows I-BAR domains aggregate through an effective interaction via clustered lipids in addition to curvature-coupling and fluctuation-mediated effective interactions mediated by the membrane. Clustered lipids mediate a directional effective interaction between the ends of the I-BAR domains. Inside of a membrane protrusion, curved I-BAR domains can stabilize irregular protrusions while the addition of negatively charged lipids induces an effective interaction between I-BAR domains that constricts the membrane protrusion toward a cylindrical shape.

Chapter 5: represents from a departure from I-BAR domain modeling and instead investigates issues of representability and transferability in the MARTINI family of top-down coarse-grained models. MARTINI models are popular due to computational efficiency and inherent modular application. However, the application of a force field to a system or state point outside the scope of the parameterization suffer from accuracy issues. Although, the nature of the inaccuracies has not been fully explained in top-down coarse-grained models. As an example, the lateral association behavior of cholesterol and DOPC shows that the entropic and enthalpic driving force in MARTINI and atomistic systems are distinctly different. In the MARITNI system, the DOPC-cholesterol repulsion is incorrectly attributed to enthalpic driving forces rather than entropic driving forces seen in atomistic systems. Thus, the application of MARTINI models from a single component to multicomponent systems is dubious based on issues of transferability.

6.2 Future Directions

Improving Bottom-up Models of I-BAR Domains

The bottom-up model of I-BAR domains in Chapter 3: can be readily improved using more all-atom simulation data. The current I-BAR domain model is parameterized from a single I-BAR domain on a membrane of a single composition. The coarse-grained and experimental environments of an I-BAR domain is drastically more complicated than the all-atom system. Thus, increasing the amount of all-atom data toward the target systems would improve the applicability of the coarse-grained models. Two facets of the model that can be improved through further all-atom simulations are the flatness and rigidity of the model, and the membrane-protein interactions.

Current simulations of I-BAR domains are on flat membrane. However, tube-pulling assays are the exact opposite of this condition.(Prevost et al., 2017; Prevost et al., 2015) In these experiments, the curvature of the membrane tubules can be tuned and adjusted and the I-BAR domains sense and sort into the membrane tubules. Thus, all-atom simulations that can approach these conditions and model I-BAR domains on curved surfaces are of interest. First, there is a prevailing hypothesis that I-BAR domains bound to curved surfaces will have the same curvature seen in crystal structures, which one would expect to affect the assembly of I-BAR domains. Indeed, the model presented in Chapter 4: shows that protein curvature directly affects the protein assembly. All-atom simulations of I-BAR domains on curved surfaces will help understand the curvature of I-BAR domain and corresponding assembly behavior should be expected. Secondly, all-atom simulations can be used to calculate energetic barriers of membrane remodeling with and without an I-BAR domain. All-atom simulations of I-BAR domain on flat membrane remain flat, but membrane could be deformed using enhanced sampling (e.g. umbrella sampling using midplane shape as a collective variable). By comparing the potentials of mean force of a membrane with and without an I-BAR domain, the free energy of membrane remodeling could be directly evaluated and provide mechanistic insights into remodeling by different I-BAR domains. Further elucidating the cellular requirements of I-BAR domains and the design rules membrane remodelers.

The parameterization scheme used in Chapter 3: assumes membrane-binding surface of the I-BAR domain and functional form. Although, the sensitivity of the aggregation behavior to the energetics of the membrane-protein interactions are investigated, further all-atom simulations can elucidate the nature of membrane-protein interactions. Experiments show that the I-BAR domain of IRSp53 is a peripheral membrane protein and does not embed deeply into the membrane.(Saarikangas et al., 2009) Thus, the model assumed the CG model only interacted with the head group bead of the lipids, but that is not broadly true of the I-BAR family. Further all-atom simulations of I-BAR domains interacting with lipids can further refine which residues of the I-BAR domain interact with head group or glycerol and how strongly they interact. From the coarse-grained models of Chapter 3:, it is clear that the interaction strength affects the curvature generated and the resulting aggregation behavior. However, the diffusivity of lipids near I-BAR domains is unknown and could have significant implications for lipid mediated interactions like those seen in Chapter 4:. Thus, the bottom-up model presented in Chapter 3: benefit from a more thorough investigation into the underlying assumptions of nature of the membrane-protein interactions.

Extending Top-down Models of I-BAR Proteins

The phenomenological model presented in Chapter 4: captures the structural diversity of I-BAR domains and the variety of membrane binding surfaces, but experimentally, it is known that the behavior of I-BAR domains changes when in the presence of other proteins (e.g., actin) or when grafted to other protein domains (e.g, intrinsically disordered proteins [IDPs]). In the cellular environment, I-BAR proteins sense and generate membrane curvature and interact with cytoskeletal proteins. The additional domains of I-BAR proteins and interactions with the cytoskeleton likely change the membrane remodeling behavior. Additionally, *in vitro* experiments of I-BAR domain of Missing-in-Metastasis (MIM) show that I-BAR domain with amphipathic helices form wider membrane tubules than the I-BAR domain without amphipathic helices.(Saarikangas et al., 2009) Similarly, I-BAR domains with grafted IDPs vesiculate rather than tubule giant unilamellar vesicles.(Snead et al., 2019; Zeno et al., 2018) The model from Chapter 4: represents the stabilization of membrane protrusions by I-BAR domains and could be improved by adding a representation of the cytoskeleton or increasing the complexity of the I-BAR protein.

An example of cytoskeleton representation could be implicit like an external potential to approximate actin forces on the membrane. Continuum models have approximated actin's role in protrusion formation this way and is similar to the way the protrusions are formed in Chapter

4:.(Mesarec et al., 2016) However, a particle-based representation of the cytoskeleton could better test hypotheses related to the mechanisms of protrusion formation and the necessary actin nucleators or proteins are minimally required to mediate the interactions between the cytoskeleton and membrane remodeling proteins. The goal of my phenomenological protein was to add complexity as to determine the minimal representation required for protrusion stabilization. Thus, by sequentially adding the force due to actin, an explicit representation of actin, actin nucleators, and proteins that mediate membrane-cytoskeleton interactions, minimal proteins and their explicit functions can be determined. This would an incredible boon to *in vitro* reconstitution experiments attempting the same goal through experimental means.

Through entropic driving forces, IDPs generate positive curvature, opposing the curvature generated by I-BAR domains.(Zeno et al., 2018) Thus, when attached to I-BAR domains, there will be an equilibrium between direct scaffolding and entropically driving membrane remodeling. The balance between these opposing curvature generating forces is unclear and coarse-grained modeling can shed light on the competing driving forces. While the systemic issues of atomistic IDP force fields continue to be ameliorated and would plague a bottom-up coarse-grained model, a top-down model could approximate the entropic forces through explicit IDP representations. Polymer models (e.g. self-avoiding freely jointed chains) can capture the flexibility IDPs and recapitulate the entropic driving forces that generate membrane curvature. Through an explicit IDP representation grafted to the top-down I-BAR domain model, one could gain insights into the onset of vesiculation and tubulation.

Comparing Top-down Coarse-grained Models to Corresponding Atomistic Systems

Bottom-up models are by definition compared to the underlying all-atom counterpart via parameterization scheme. Top-down models, on the other hand, do not have the same requirement. Therefore, their representation and force field could be improved from insights into their corresponding atomistic systems. As an example, the MARTINI model investigated in Chapter 5: uses

4 waters to 1 coarse-grained bead representation and an interaction potential of the Lennard-Jones form.(Marrink et al., 2007) Subsequently, the MARTINI model obviously struggles to accurately represent systems that require singular waters (e.g., water pores in lipid membranes). However, the planar lipid bilayer systems also stuffer from the coarse representation of water, which is demonstrated by the difference in headgroup ordering shown in Chapter 5:. The MARTINI water model is infamous for its many problems, but there have been various attempts to improve coarse-grained water models. (Hadley and McCabe, 2010; Hadley and McCabe, 2012; Yesylevskyy et al., 2010) However, coarse-grained water models can be further improved from insights into water behavior around solutes. Specifically, water near lipid head groups diffuses slower and slowly exchanges with bulk water. The residence time of water near lipid headgroups and the ordering around lipid defects (i.e., regions where the hydrophobic core is directly solvated) necessitates a water representation and force field that exhibits drastically different behavior near the solute than in the bulk. Thus, coarse-grained water could be improved by incorporating insights from all-atom solutions or more directly, through hybrid parameterization schemes.

References

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX *1-2*, 19-25.

Ahmed, S., Goh, W.I., and Bu, W. (2010). I-BAR domains, IRSp53 and filopodium formation. Semin Cell Dev Biol *21*, 350-356.

Alessandri, R., Souza, P.C.T., Thallmair, S., Melo, M.N., de Vries, A.H., and Marrink, S.J. (2019). Pitfalls of the Martini Model. J Chem Theory Comput *15*, 5448-5460.

Allen, M.P., and Tildesley, D.J. (1989). Computer simulation of liquids (Oxford: Clarendon Press).

Aranda-Espinoza, H., Berman, A., Dan, N., Pincus, P., and Safran, S. (1996). Interaction between inclusions embedded in membranes. Biophysical Journal *71*, 648-656.

Arkhipov, A., Yin, Y., and Schulten, K. (2008). Four-scale description of membrane sculpting by BAR domains. Biophys J *95*, 2806-2821.

Arkhipov, A., Yin, Y., and Schulten, K. (2009). Membrane-bending mechanism of amphiphysin N-BAR domains. Biophys J *97*, 2727-2735.

Arnarez, C., Uusitalo, J.J., Masman, M.F., Ingolfsson, H.I., de Jong, D.H., Melo, M.N., Periole, X., de Vries, A.H., and Marrink, S.J. (2015). Dry Martini, a coarse-grained force field for lipid membrane simulations with implicit solvent. J Chem Theory Comput *11*, 260-275.

Ayton, G.S., Blood, P.D., and Voth, G.A. (2007). Membrane remodeling from N-BAR domain interactions: insights from multi-scale simulation. Biophys J *92*, 3595-3602.

Ayton, G.S., Lyman, E., Krishna, V., Swenson, R.D., Mim, C., Unger, V.M., and Voth, G.A. (2009). New insights into BAR domain-induced membrane remodeling. Biophys J *97*, 1616-1625.

Bahrami, A.H., Lipowsky, R., and Weikl, T.R. (2012). Tubulation and Aggregation of Spherical Nanoparticles Adsorbed on Vesicles. Physical Review Letters *109*.

Baron, R., de Vries, A.H., Hunenberger, P.H., and van Gunsteren, W.F. (2006). Configurational entropies of lipids in pure and mixed bilayers from atomic-level and coarse-grained molecular dynamics simulations. J Phys Chem B *110*, 15602-15614.

Barooji, Y.F., Rorvig-Lund, A., Semsey, S., Reihani, S.N., and Bendix, P.M. (2016). Dynamics of membrane nanotubes coated with I-BAR. Sci Rep *6*, 30054.

Baumgart, T., Capraro, B.R., Zhu, C., and Das, S.L. (2011). Thermodynamics and mechanics of membrane curvature generation and sensing by proteins and lipids. Annu Rev Phys Chem *62*, 483-506.

Beltran-Heredia, E., Tsai, F.C., Salinas-Almaguer, S., Cao, F.J., Bassereau, P., and Monroy, F. (2019). Membrane curvature induces cardiolipin sorting. Commun Biol *2*, 225.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E., Mittal, J., Feig, M., and Mackerell, A.D., Jr. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput *8*, 3257-3273.

Blood, P.D., and Voth, G.A. (2006). Direct observation of Bin/amphiphysin/Rvs (BAR) domain-induced membrane curvature by means of molecular dynamics simulations. Proceedings of the National Academy of Sciences *103*, 15068-15072.

Boal, D.H. (2012). Mechanics of the Cell (Cambridge University Press).

Böckmann, R.A., Hac, A., Heimburg, T., and Grubmüller, H. (2003). Effect of Sodium Chloride on a Lipid Bilayer. Biophysical Journal *85*, 1647-1655.

Bonazzi, F., Hall, C., and Weikl, T.R. (2020). Membrane morphologies induced by mixtures of arc-shaped particles with opposite curvature. Soft Matter.

Brandt, E.G., Braun, A.R., Sachs, J.N., Nagle, J.F., and Edholm, O. (2011). Interpretation of fluctuation spectra in lipid bilayer simulations. Biophys J *100*, 2104-2111.

Brannigan, G., Philips, P.F., and Brown, F.L. (2005). Flexible lipid bilayers in implicit solvent. Phys Rev E Stat Nonlin Soft Matter Phys *72*, 011915.

Bruininks, B.M.H., Souza, P.C.T., and Marrink, S.J. (2019). A Practical View of the Martini Force Field. Methods Mol Biol *2022*, 105-127.

Callan-Jones, A., Sorre, B., and Bassereau, P. (2011). Curvature-driven lipid sorting in biomembranes. Cold Spring Harb Perspect Biol *3*.

Canham, P.B. (1970). The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. Journal of Theoretical Biology *26*, 61-81.

Cao, Z., and Voth, G.A. (2015). The multiscale coarse-graining method. XI. Accurate interactions based on the centers of charge of coarse-grained sites. J Chem Phys *143*, 243116.

Chakraborty, M., Xu, C., and White, A.D. (2018). Encoding and selecting coarse-grain mapping operators with hierarchical graphs. J Chem Phys *149*, 134106.

Chen, Y., Aardema, J., Misra, A., and Corey, S.J. (2012). BAR proteins in cancer and blood disorders. International journal of biochemistry and molecular biology *3*, 198-208.

Chen, Z., Shi, Z., and Baumgart, T. (2015). Regulation of membrane-shape transitions induced by I-BAR domains. Biophys J *109*, 298-307.

Chen, Z., Zhu, C., Kuo, C.J., Robustelli, J., and Baumgart, T. (2016). The N-Terminal Amphipathic Helix of Endophilin Does Not Contribute to Its Molecular Curvature Generation Capacity. J Am Chem Soc *138*, 14616-14622.

Chu, H., Cao, L., Peng, X., and Li, G. (2017). Polarizable force field development for lipids and their efficient applications in membrane proteins. Wiley Interdisciplinary Reviews: Computational Molecular Science *7*.

Cooke, I.R., Kremer, K., and Deserno, M. (2005). Tunable generic model for fluid bilayer membranes. Phys Rev E Stat Nonlin Soft Matter Phys *72*, 011506.

Cui, H., Lyman, E., and Voth, G.A. (2011). Mechanism of membrane curvature sensing by amphipathic helix containing proteins. Biophys J *100*, 1271-1279.

Cui, H., Mim, C., Vazquez, F.X., Lyman, E., Unger, V.M., and Voth, G.A. (2013). Understanding the role of amphipathic helices in N-BAR domain driven membrane remodeling. Biophys J *104*, 404-411.

Dan, N. (2017). Membrane-induced interactions between curvature-generating protein domains: the role of area perturbation. AIMS Biophysics *4*, 107-120.

Daum, B., Auerswald, A., Gruber, T., Hause, G., Balbach, J., Kuhlbrandt, W., and Meister, A. (2016). Supramolecular organization of the human N-BAR domain in shaping the sarcolemma membrane. J Struct Biol *194*, 375-382.

Davtyan, A., Simunovic, M., and Voth, G.A. (2017). The mesoscopic membrane with proteins (MesM-P) model. The Journal of Chemical Physics *147*, 044101.

de Groot, J.C., Schluter, K., Carius, Y., Quedenau, C., Vingadassalom, D., Faix, J., Weiss, S.M., Reichelt, J., Standfuss-Gabisch, C., Lesser, C.F.*, et al.* (2011). Structural basis for complex formation between human IRSp53 and the translocated intimin receptor Tir of enterohemorrhagic E. coli. Structure *19*, 1294-1306.

Dunn, N.J., Foley, T.T., and Noid, W.G. (2016). Van der Waals Perspective on Coarse-Graining: Progress toward Solving Representability and Transferability Problems. Acc Chem Res *49*, 2832-2840.

Dunweg, B., and Paul, W. (1991). Brownian Dynamics Simulations without Gaussian Random Numbers. International Journal of Modern Physics C *2*, 817-827.

Evans, E., and Rawicz, W. (1990). Entropy-driven tension and bending elasticity in condensed-fluid membranes. Phys Rev Lett *64*, 2094-2097.

Faucon, J.F., Mitov, M.D., Méléard, P., Bivas, I., and Bothorel, P. (1989). Bending elasticity and thermal fluctuations of lipid membranes. Theoretical and experimental requirements. Journal de Physique *50*, 2389-2414.

Fernandes, F., Loura, L.M., Chichon, F.J., Carrascosa, J.L., Fedorov, A., and Prieto, M. (2008). Role of helix 0 of the N-BAR domain in membrane curvature generation. Biophys J *94*, 3065-3073.

Frost, A., Perera, R., Roux, A., Spasov, K., Destaing, O., Egelman, E.H., De Camilli, P., and Unger, V.M. (2008). Structural basis of membrane invagination by F-BAR domains. Cell *132*, 807-817.

Frost, A., Unger, V.M., and De Camilli, P. (2009). The BAR domain superfamily: membrane-molding macromolecules. Cell *137*, 191-196.

Gauthier, N.C., Masters, T.A., and Sheetz, M.P. (2012). Mechanical feedback between membrane tension and dynamics. Trends Cell Biol *22*, 527-535.

Gershenson, A., Gosavi, S., Faccioli, P., and Wintrode, P.L. (2020). Successes and challenges in simulating the folding of large proteins. Journal of Biological Chemistry *295*, 15-33.

Grime, J., and Madsen, J.J. (2019). Efficient Simulation of Tunable Lipid Assemblies Across Scales and Resolutions. arXiv preprint arXiv:191005362.

Grossfield, A. (2013). WHAM: the weighted histogram analysis method, version 2.0. 9. Available at membrane urmc rochester edu/content/wham Accessed November *15*, 2013.

Hadley, K.R., and McCabe, C. (2010). On the investigation of coarse-grained models for water: balancing computational efficiency and the retention of structural properties. J Phys Chem B *114*, 4590-4599.

Hadley, K.R., and McCabe, C. (2012). Coarse-grained molecular models of water: a review. Molecular Simulation *38*, 671-681.

Harayama, T., and Riezman, H. (2018). Understanding the diversity of membrane lipid composition. Nat Rev Mol Cell Biol *19*, 281-296.

Heinrich, M., Tian, A., Esposito, C., and Baumgart, T. (2010). Dynamic sorting of lipids and proteins in membrane tubes with a moving phase boundary. Proc Natl Acad Sci U S A *107*, 7208-7213.

Helfrich, W. (1973). Elastic properties of lipid bilayers: theory and possible experiments. Z Naturforsch C *28*, 693-703.

Hills, R.D., Jr., and Brooks, C.L., 3rd (2009). Insights from coarse-grained Go models for protein folding and dynamics. Int J Mol Sci *10*, 889-905.

Hofsäß, C., Lindahl, E., and Edholm, O. (2003). Molecular Dynamics Simulations of Phospholipid Bilayers with Cholesterol. Biophysical Journal *84*, 2192-2206.

Huang, J., and MacKerell, A.D., Jr. (2013). CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem *34*, 2135-2145.

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmuller, H., and MacKerell, A.D., Jr. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods *14*, 71-73.

Ingólfsson, H.I., Lopez, C.A., Uusitalo, J.J., de Jong, D.H., Gopal, S.M., Periole, X., and Marrink, S.J. (2014). The power of coarse graining in biomolecular simulations. Wiley Interdisciplinary Reviews: Computational Molecular Science *4*, 225-248.

Izvekov, S., and Voth, G.A. (2005). A multiscale coarse-graining method for biomolecular systems. The Journal of Physical Chemistry B *109*, 2469-2473.

Jao, C.C., Hegde, B.G., Gallop, J.L., Hegde, P.B., McMahon, H.T., Haworth, I.S., and Langen, R. (2010). Roles of amphipathic helices and the bin/amphiphysin/rvs (BAR) domain of endophilin in membrane curvature generation. J Biol Chem *285*, 20164-20170.

Jarin, Z., Tsai, F.C., Davtyan, A., Pak, A.J., Bassereau, P., and Voth, G.A. (2019). Unusual Organization of I-BAR Proteins on Tubular and Vesicular Membranes. Biophys J *117*, 553-562.

Jin, J., Pak, A.J., and Voth, G.A. (2019). Understanding Missing Entropy in Coarse-Grained Systems: Addressing Issues of Representability and Transferability. J Phys Chem Lett *10*, 4549-4557.

Jo, S., Kim, T., and Im, W. (2007). Automated builder and database of protein/membrane complexes for molecular dynamics simulations. PLoS One *2*, e880.

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

Jo, S., Lim, J.B., Klauda, J.B., and Im, W. (2009). CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. Biophys J *97*, 50-58.

Johannes, L., Pezeshkian, W., Ipsen, J.H., and Shillcock, J.C. (2018). Clustering on Membranes: Fluctuations and More. Trends Cell Biol *28*, 405-415.

Johnson, M.E., Head-Gordon, T., and Louis, A.A. (2007). Representability problems for coarse-grained water potentials. J Chem Phys *126*, 144509.

Kang, J., Park, H., and Kim, E. (2016). IRSp53/BAIAP2 in dendritic spine development, NMDA receptor regulation, and psychiatric disorders. Neuropharmacology *100*, 27-39.

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

Kumar, S., Rosenberg, J.M., Bouzida, D., Swendsen, R.H., and Kollman, P.A. (1992). THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. Journal of Computational Chemistry *13*, 1011-1021.

Lee, J., Cheng, X., Swails, J.M., Yeom, M.S., Eastman, P.K., Lemkul, J.A., Wei, S., Buckner, J., Jeong, J.C., Qi, Y.*, et al.* (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. J Chem Theory Comput *12*, 405-413.

Lemkul, J.A., Huang, J., Roux, B., and MacKerell, A.D., Jr. (2016). An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. Chem Rev *116*, 4983-5013.

Leonard, A.N., Wang, E., Monje-Galvan, V., and Klauda, J.B. (2019). Developing and Testing of Lipid Force Fields with Applications to Modeling Cellular Membranes. Chem Rev *119*, 6227-6269.

Li, H., Chowdhary, J., Huang, L., He, X., MacKerell, A.D., and Roux, B. (2017). Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. Journal of Chemical Theory and Computation *13*, 4535-4552.

Linkner, J., Witte, G., Zhao, H., Junemann, A., Nordholz, B., Runge-Wollmann, P., Lappalainen, P., and Faix, J. (2014). The inverse BAR domain protein IBARa drives membrane remodeling to control osmoregulation, phagocytosis and cytokinesis. J Cell Sci *127*, 1279-1292.

Liu, H., Song, D., Lu, H., Luo, R., and Chen, H.F. (2018). Intrinsically disordered protein-specific force field CHARMM36IDPSFF. Chem Biol Drug Des *92*, 1722-1735.

Lopes, P.E., Guvench, O., and MacKerell, A.D., Jr. (2015). Current status of protein force fields for molecular dynamics simulations. Methods Mol Biol *1215*, 47-71.

Lu, L., and Voth, G.A. (2011). The multiscale coarse-graining method. VII. Free energy decomposition of coarse-grained effective potentials. J Chem Phys *134*, 224107.

Lyman, E., Pfaendtner, J., and Voth, G.A. (2008). Systematic multiscale parameterization of heterogeneous elastic network models of proteins. Biophys J *95*, 4183-4192.

Lyubartsev, A.P., and Rabinovich, A.L. (2016). Force Field Development for Lipid Membrane Simulations. Biochim Biophys Acta *1858*, 2483-2497.

M. Goullian, R. Bruinsma, and Pincus, P. (1993). Long-Range Forces in Heterogeneous Fluid Membranes. Europhysics Letters *22*, 145-150.

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S.*, et al.* (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B *102*, 3586-3616.

Marrink, S.J., Corradi, V., Souza, P.C.T., Ingolfsson, H.I., Tieleman, D.P., and Sansom, M.S.P. (2019). Computational Modeling of Realistic Cell Membranes. Chem Rev *119*, 6184-6226.

Marrink, S.J., de Vries, A.H., and Mark, A.E. (2004). Coarse Grained Model for Semiquantitative Lipid Simulations. The Journal of Physical Chemistry B *108*, 750-760.

Marrink, S.J., Periole, X., Tieleman, D.P., and de Vries, A.H. (2010). Comment on "On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models" by M. Winger, D. Trzesniak, R. Baron and W. F. van Gunsteren, Phys. Chem. Chem. Phys., 2009, 11, 1934. Phys Chem Chem Phys *12*, 2254-2256; author reply 2257-2258.

Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H. (2007). The MARTINI force field: coarse grained model for biomolecular simulations. J Phys Chem B *111*, 7812-7824.

Marrink, S.J., and Tieleman, D.P. (2013). Perspective on the Martini model. Chem Soc Rev *42*, 6801-6822.

Marsh, D. (2006). Elastic curvature constants of lipid monolayers and bilayers. Chem Phys Lipids *144*, 146-159.

Mattila, P.K., and Lappalainen, P. (2008). Filopodia: molecular architecture and cellular functions. Nat Rev Mol Cell Biol *9*, 446-454.

Mattila, P.K., Pykalainen, A., Saarikangas, J., Paavilainen, V.O., Vihinen, H., Jokitalo, E., and Lappalainen, P. (2007). Missing-in-metastasis and IRSp53 deform PI(4,5)P2-rich membranes by an inverse BAR domain-like mechanism. J Cell Biol *176*, 953-964.

McMahon, H.T., and Gallop, J.L. (2005). Membrane curvature and mechanisms of dynamic cell membrane remodelling. Nature *438*, 590-596.

McWhirter, J.L., Ayton, G., and Voth, G.A. (2004). Coupling Field Theory with Mesoscopic Dynamical Simulations of Multicomponent Lipid Bilayers. Biophysical Journal *87*, 3242-3263.

Mesarec, L., Gozdz, W., Iglic, V.K., Kralj, S., and Iglic, A. (2016). Closed membrane shapes with attached BAR domains subject to external force of actin filaments. Colloids Surf B Biointerfaces *141*, 132-140.

Miao, Y., and McCammon, J.A. (2016). Unconstrained Enhanced Sampling for Free Energy Calculations of Biomolecules: A Review. Mol Simul *42*, 1046-1055.

Millard, T.H., Bompard, G., Heung, M.Y., Dafforn, T.R., Scott, D.J., Machesky, L.M., and Futterer, K. (2005). Structural basis of filopodia formation induced by the IRSp53/MIM homology domain of human IRSp53. EMBO J *24*, 240-250.

Millman, K.J., and Aivazis, M. (2011). Python for Scientists and Engineers. Computing in Science & Engineering *13*, 9-12.

Mim, C., Cui, H., Gawronski-Salerno, J.A., Frost, A., Lyman, E., Voth, G.A., and Unger, V.M. (2012). Structural basis of membrane bending by the N-BAR protein endophilin. Cell *149*, 137-145.

Mim, C., and Unger, V.M. (2012). Membrane curvature and its generation by BAR proteins. Trends Biochem Sci *37*, 526-533.

Mori, T., Miyashita, N., Im, W., Feig, M., and Sugita, Y. (2016). Molecular dynamics simulations of biological membranes and membrane proteins using enhanced conformational sampling algorithms. Biochim Biophys Acta *1858*, 1635-1651.

Noguchi, H. (2014). Two- or three-step assembly of banana-shaped proteins coupled with shape transformation of lipid membranes. EPL (Europhysics Letters) *108*.

Noguchi, H. (2016). Membrane tubule formation by banana-shaped proteins with or without transient network structure. Sci Rep *6*, 20935.

Noguchi, H., and Fournier, J.B. (2017). Membrane structure formation induced by two types of banana-shaped proteins. Soft Matter *13*, 4099-4111.

Noid, W.G. (2013). Perspective: Coarse-grained models for biomolecular systems. The Journal of Chemical Physics *139*, 090901.

Olinger, A.D., Spangler, E.J., Kumar, P.B., and Laradji, M. (2016). Membrane-mediated aggregation of anisotropically curved nanoparticles. Faraday Discuss *186*, 265-275.

Oliphant, T.E. (2007). Python for Scientific Computing. Computing in Science & Engineering *9*, 10-20.

Pak, A.J., Dannenhoffer-Lafage, T., Madsen, J.J., and Voth, G.A. (2019). Systematic Coarse-Grained Lipid Force Fields with Semiexplicit Solvation via Virtual Sites. Journal of Chemical Theory and Computation *15*, 2087-2100.

Pak, A.J., and Voth, G.A. (2018). Advances in coarse-grained modeling of macromolecular complexes. Curr Opin Struct Biol *52*, 119-126.

Pan, J., Tristram-Nagle, S., and Nagle, J.F. (2009). Effect of cholesterol on structural and mechanical properties of membranes depends on lipid chain saturation. Phys Rev E Stat Nonlin Soft Matter Phys *80*, 021931.

Pandit, S.A., Bostick, D., and Berkowitz, M.L. (2003). Mixed Bilayer Containing Dipalmitoylphosphatidylcholine and Dipalmitoylphosphatidylserine: Lipid Complexation, Ion Binding, and Electrostatics. Biophysical Journal *85*, 3120-3131.

Parrinello, M., and Rahman, A. (1981a). Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics *52*, 7182-7190.

Parrinello, M., and Rahman, A. (1981b). Polymorphic Transitions in Single-Crystals - a New Molecular-Dynamics Method. Journal of Applied Physics *52*, 7182-7190.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V*., et al.* (2011). Scikit-learn: Machine Learning in Python. J Mach Learn Res *12*, 2825-2830.

Plimpton, S. (1995). Fast Parallel Algorithms for Short-Range Molecular Dynamics. Journal of Computational Physics *117*, 1-19.

Prevost, C., Tsai, F.C., Bassereau, P., and Simunovic, M. (2017). Pulling Membrane Nanotubes from Giant Unilamellar Vesicles. J Vis Exp.

Prevost, C., Zhao, H., Manzi, J., Lemichez, E., Lappalainen, P., Callan-Jones, A., and Bassereau, P. (2015). IRSp53 senses negative membrane curvature and phase separates along membrane tubules. Nat Commun *6*, 8529.

Ramakrishnan, N., Sunil Kumar, P.B., and Ipsen, J.H. (2013). Membrane-mediated aggregation of curvature-inducing nematogens and membrane tubulation. Biophys J *104*, 1018-1028.

Rawicz, W., Olbrich, K.C., McIntosh, T., Needham, D., and Evans, E. (2000). Effect of Chain Length and Unsaturation on Elasticity of Lipid Bilayers. Biophysical Journal *79*, 328-339.

Reynwar, B.J., Illya, G., Harmandaris, V.A., Muller, M.M., Kremer, K., and Deserno, M. (2007). Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. Nature *447*, 461-464.

Rog, T., Pasenkiewicz-Gierula, M., Vattulainen, I., and Karttunen, M. (2009). Ordering effects of cholesterol and its analogues. Biochim Biophys Acta *1788*, 97-121.

Rudzinski, J.F., and Noid, W.G. (2011). Coarse-graining entropy, forces, and structures. J Chem Phys *135*, 214101.

Rusinkiewicz, S. (2004). Estimating curvatures and their derivatives on triangle meshes. Paper presented at: Proceedings 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004 3DPVT 2004 (IEEE).

Saarikangas, J., Hakanen, J., Mattila, P.K., Grumet, M., Salminen, M., and Lappalainen, P. (2008). ABBA regulates plasma-membrane and actin dynamics to promote radial glia extension. J Cell Sci *121*, 1444-1454.

Saarikangas, J., Zhao, H., Pykalainen, A., Laurinmaki, P., Mattila, P.K., Kinnunen, P.K., Butcher, S.J., and Lappalainen, P. (2009). Molecular mechanisms of membrane deformation by I-BAR domain proteins. Curr Biol *19*, 95-107.

Salzer, U., Kostan, J., and Djinovic-Carugo, K. (2017). Deciphering the BAR code of membrane modulators. Cell Mol Life Sci *74*, 2413-2438.

Šarić, A., and Cacciuto, A. (2012). Fluid membranes can drive linear aggregation of adsorbed spherical nanoparticles. Phys Rev Lett *108*, 118101.

Šarić, A., and Cacciuto, A. (2013). Self-assembly of nanoparticles adsorbed on fluid and elastic membranes. Soft Matter *9*, 6677.

Saunders, M.G., and Voth, G.A. (2013). Coarse-graining methods for computational biology. Annu Rev Biophys *42*, 73-93.

Schneider, T., and Stoll, E. (1978). Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. Physical Review B *17*, 1302-1322.

Schweitzer, Y., and Kozlov, M.M. (2015). Membrane-mediated interaction between strongly anisotropic protein scaffolds. PLoS Comput Biol *11*, e1004054.

Schweitzer, Y., Shemesh, T., and Kozlov, M.M. (2015). A Model for Shaping Membrane Sheets by Protein Scaffolds. Biophys J *109*, 564-573.

Shell, M.S. (2008). The relative entropy is fundamental to multiscale and inverse thermodynamic problems. The Journal of chemical physics *129*, 144108.

Shinoda, W., DeVane, R., and Klein, M.L. (2007). Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. Molecular Simulation *33*, 27-36.

Shinoda, W., DeVane, R., and Klein, M.L. (2010). Zwitterionic lipid assemblies: molecular dynamics studies of monolayers, bilayers, and vesicles using a new coarse grain force field. J Phys Chem B *114*, 6836-6849.

Shinoda, W., DeVane, R., and Klein, M.L. (2012). Computer simulation studies of self-assembling macromolecules. Curr Opin Struct Biol *22*, 175-186.

Simons, K., and Sampaio, J.L. (2011). Membrane organization and lipid rafts. Cold Spring Harb Perspect Biol *3*, a004697.

Simunovic, M., Evergren, E., Callan-Jones, A., and Bassereau, P. (2019). Curving Cells Inside and Out: Roles of BAR Domain Proteins in Membrane Shaping and Its Cellular Implications. Annu Rev Cell Dev Biol *35*, 111-129.

Simunovic, M., Evergren, E., Golushko, I., Prevost, C., Renard, H.F., Johannes, L., McMahon, H.T., Lorman, V., Voth, G.A., and Bassereau, P. (2016). How curvature-generating proteins build scaffolds on membrane nanotubes. Proc Natl Acad Sci U S A *113*, 11226-11231.

Simunovic, M., Saric, A., Henderson, J.M., Lee, K.Y.C., and Voth, G.A. (2017). Long-Range Organization of Membrane-Curving Proteins. ACS Cent Sci *3*, 1246-1253.

Simunovic, M., Srivastava, A., and Voth, G.A. (2013). Linear aggregation of proteins on the membrane as a prelude to membrane remodeling. Proc Natl Acad Sci U S A *110*, 20396-20401.

Simunovic, M., and Voth, G.A. (2015). Membrane tension controls the assembly of curvature-generating proteins. Nat Commun *6*, 7219.

Simunovic, M., Voth, G.A., Callan-Jones, A., and Bassereau, P. (2015). When Physics Takes Over: BAR Proteins and Membrane Curvature. Trends Cell Biol *25*, 780-792.

Smit, B., and Frenkel, D. (2001). Understanding Molecular Simulation: From Algorithms to Applications, 2nd edn (Elsevier).

Snead, W.T., Zeno, W.F., Kago, G., Perkins, R.W., Richter, J.B., Zhao, C., Lafer, E.M., and Stachowiak, J.C. (2019). BAR scaffolds drive membrane fission by crowding disordered domains. J Cell Biol *218*, 664-682.

Srivastava, A., and Voth, G.A. (2013). A Hybrid Approach for Highly Coarse-grained Lipid Bilayer Models. J Chem Theory Comput *9*, 750-765.

Stanishneva-Konovalova, T.B., and Sokolova, O.S. (2019). Effects of PI(4,5)P2 concentration on the F-BAR domain membrane binding as revealed by coarse-grained simulations. Proteins *87*, 561-568.

Takada, S. (2019). Go model revisited. Biophys Physicobiol *16*, 248-255.

Takemura, K., Hanawa-Suetsugu, K., Suetsugu, S., and Kitao, A. (2017). Salt Bridge Formation between the I-BAR Domain and Lipids Increases Lipid Density and Membrane Curvature. Sci Rep *7*, 6808.

Tian, A., and Baumgart, T. (2009). Sorting of lipids and proteins in membrane curvature gradients. Biophys J *96*, 2676-2688.

Torrie, G.M., and Valleau, J.P. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics *23*, 187-199.

Tourdot, R.W., Ramakrishnan, N., Baumgart, T., and Radhakrishnan, R. (2015). Application of a free-energy-landscape approach to study tension-dependent bilayer tubulation mediated by curvature-inducing proteins. Phys Rev E Stat Nonlin Soft Matter Phys *92*, 042715.

Tourdot, R.W., Ramakrishnan, N., and Radhakrishnan, R. (2014). Defining the free-energy landscape of curvature-inducing proteins on membrane bilayers. Phys Rev E Stat Nonlin Soft Matter Phys *90*, 022717.

Tribello, G.A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014). PLUMED 2: New feathers for an old bird. Computer Physics Communications *185*, 604-613.

Ueda, Y., Taketomi, H., and G, N. (1978). Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A. Three-dimensional lattice model of lysozyme. Biopolymers *17*, 1531-1548.

van der Walt, S., Colbert, S.C., and Varoquaux, G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation. Computing in Science & Engineering *13*, 22-30.

van der Wel, C., Vahid, A., Saric, A., Idema, T., Heinrich, D., and Kraft, D.J. (2016). Lipid membrane-mediated attraction between curvature inducing objects. Sci Rep *6*, 32825.

Veatch, S.L., and Keller, S.L. (2005). Seeing spots: complex phase behavior in simple membranes. Biochim Biophys Acta *1746*, 172-185.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., and Bright, J. (2019). SciPy 1.0--fundamental algorithms for scientific computing in Python. arXiv preprint arXiv:190710121.

Voth, G.A. (2008). Coarse-graining of condensed phase and biomolecular systems (CRC press).

Wagner, J.W., Dama, J.F., Durumeric, A.E., and Voth, G.A. (2016). On the representability problem and the physical meaning of coarse-grained models. J Chem Phys *145*, 044108.

Watson, M.C., Brandt, E.G., Welch, P.M., and Brown, F.L. (2012). Determining biomembrane bending rigidities from simulations of modest size. Phys Rev Lett *109*, 028102.

Webb, M.A., Delannoy, J.Y., and de Pablo, J.J. (2019). Graph-Based Approach to Systematic Molecular Coarse-Graining. J Chem Theory Comput *15*, 1199-1208.

Winger, M., Trzesniak, D., Baron, R., and van Gunsteren, W.F. (2009). On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. Phys Chem Chem Phys *11*, 1934-1941.

Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Davila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V., Venable, R.M.*, et al.* (2014). CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. J Comput Chem *35*, 1997-2004.

Wu, H., Wolynes, P.G., and Papoian, G.A. (2018). AWSEM-IDP: A Coarse-Grained Force Field for Intrinsically Disordered Proteins. J Phys Chem B *122*, 11115-11125.

Yang, Y.I., Shao, Q., Zhang, J., Yang, L., and Gao, Y.Q. (2019). Enhanced sampling in molecular dynamics. J Chem Phys *151*, 070902.

Yesylevskyy, S.O., Schafer, L.V., Sengupta, D., and Marrink, S.J. (2010). Polarizable water model for the coarse-grained MARTINI force field. PLoS Comput Biol *6*, e1000810.

Yu, H., and Schulten, K. (2013). Membrane sculpting by F-BAR domains studied by molecular dynamics simulations. PLoS Comput Biol *9*, e1002892.

Yu, S., Cho, K., Kim, Y.H., Park, S., Kim, J., and Oh, H.B. (2006). Identification of phospholipid molecular species in porcine brain extracts using high mass accuracy of 4.7 Tesla Fourier transform ion cyclotron resonance mass spectrometry. Bulletin of the Korean Chemical Society *27*, 793-796.

Zeno, W.F., Baul, U., Snead, W.T., DeGroot, A.C.M., Wang, L., Lafer, E.M., Thirumalai, D., and Stachowiak, J.C. (2018). Synergy between intrinsically disordered domains and structured proteins amplifies membrane curvature sensing. Nat Commun *9*, 4152.

Zhang, Z., Lu, L., Noid, W.G., Krishna, V., Pfaendtner, J., and Voth, G.A. (2008). A systematic methodology for defining coarse-grained sites in large biomolecules. Biophys J *95*, 5073-5083.

Zimmerberg, J., and Kozlov, M.M. (2006). How proteins produce cellular membrane curvature. Nat Rev Mol Cell Biol *7*, 9-19.