Title
Dynamics of Lipid Transport between Membranes: Insights from Simulation and Theory

Author
Rogers, Julia R

Publication Date
2021

Peer reviewed|Thesis/dissertation

Dynamics of Lipid Transport between Membranes: Insights from Simulation and Theory

by

Julia R. Rogers

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Phillip L. Geissler, Chair
Professor David T. Limmer
Professor Daniel A. Fletcher

Spring 2021

Dynamics of Lipid Transport between Membranes: Insights from Simulation and Theory

Abstract

Dynamics of Lipid Transport between Membranes: Insights from Simulation and Theory

by

Julia R. Rogers

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Phillip L. Geissler, Chair

Cell homeostasis requires spatiotemporal regulation of heterogeneous membrane compositions. One key way that proper lipid distributions are maintained is through non-vesicular transport of individual lipids between membranes. Despite its biological importance, non-vesicular transport remains poorly understood at the molecular level. While lipid transfer proteins may be largely responsible for the selective transport of individual lipids *in vivo*, lipid transfer can also occur passively. Here, we expand our biophysical knowledge of non-vesicular lipid transport mechanisms by investigating the dynamics of passive lipid transport and then using those insights to examine how lipid transfer proteins catalyze lipid transport.

A detailed understanding of passive lipid transport has remained elusive due in part to inconsistencies between experimental findings and previous molecular simulations. We resolve these discrepancies by discovering the reaction coordinate for passive lipid transport, which enables a complete biophysical characterization of the rate-limiting step of lipid transport. Through analysis of unbiased all-atom and coarse-grained molecular dynamics simulations, we find that the reaction coordinate measures the formation and breakage of hydrophobic contacts between the membrane and transferring lipid. Consistent with experiments, free energy profiles along the reaction coordinate exhibit both a rate-limiting activation barrier for lipid desorption from a membrane and a significant barrier for lipid insertion, which was entirely missed in previous computational studies. Using our newly identified reaction coordinate, we formulate an expression for the rate of passive lipid transport to enable a quantitative comparison with experiments. Most importantly, we find that the breakage of hydrophobic lipid–membrane contacts is rate limiting for passive lipid transport.

Knowledge of the reaction coordinate allows us to systematically investigate how the activation free energy of passive lipid transport depends on membrane physicochemical properties. Through all-atom molecular dynamics simulations of 11 chemically distinct glycerophospholipids, we determine how lipid acyl chain length, unsaturation, and headgroup influence the free energy barriers for lipid desorption from and insertion into liquid-crystalline and gel

phase membranes. Consistent with previous experimental measurements, we find that lipids with longer, saturated acyl chains have increased activation free energies compared to lipids with shorter, unsaturated chains. Lipids with different headgroups exhibit a range of activation free energies; however, no clear trend based solely on chemical structure can be identified, mirroring difficulties in the interpretation of previous experimental results. Compared to liquid-crystalline phase membranes, gel phase membranes exhibit substantially increased free energy barriers. Overall, we find that the activation free energy depends on a lipid's local hydrophobic environment in a membrane and that the free energy barrier for lipid insertion depends on a membrane's interfacial hydrophobicity. Both of these properties can be altered through local changes in membrane composition and phase, suggesting that variations in cell membrane hydrophobicity may be exploited to direct non-vesicular lipid traffic.

Our discovery that the rate of passive lipid exchange is limited by the disruption of a lipid's local hydrophobic environment suggests that lipid transfer proteins may catalyze lipid transport by lowering the free energy barrier for hydrophobic lipid–membrane contact breakage. To test this hypothesis, we investigate how ceramide-1-phosphate transfer protein (CPTP) catalyzes the transport of ceramide-1-phosphate (C1P), a bioactive sphingolipid, between membranes. To resolve how CPTP extracts and inserts C1P into a membrane, we utilize a multiscale simulation approach that builds upon our findings about passive lipid exchange. We find that both the apo and C1P-bound forms of CPTP bind a membrane poised to extract and insert C1P, confirming predictions based on crystal structures. Membrane binding promotes conformational changes that widen the entrance to CPTP's hydrophobic cavity, further facilitating the exchange of C1P between CPTP and a membrane. Due to its stronger electrostatic attraction to the membrane, the apo form binds deeper into the membrane, substantially disrupting a lipid's local hydrophobic environment in the membrane below. As a result, CPTP lowers the free energy barrier for the breakage of hydrophobic C1P–membrane contacts. After extracting C1P, CPTP likely unbinds a membrane through an electrostatic switching mechanism similar to that used by other lipid transfer proteins. Thus, we provide novel insights into the molecular mechanisms used by lipid transfer proteins to efficiently traffic lipids between membranes.

Many lipid transfer proteins, including CPTP, function in local thermodynamic equilibrium. Yet, others require an expenditure of energy to drive lipid transport, thus functioning out of equilibrium. Efforts to characterize such lipid transfer proteins are currently limited by deficiencies in nonequilibrium simulation methods. To address some of the challenges in using simulations to calculate generalized free energies, or large deviation functions, of nonequilibrium systems, we develop a new transition path sampling method. Specifically, we devise a novel set of path sampling moves based on Brownian bridges, which are stochastic trajectories constrained to start and end at specified configurations. We use our method to efficiently calculate large deviation functions of asymmetric simple exclusion processes, a paradigmatic nonequilibrium transport model that could foreseeably be used to model lipid transfer proteins with tubular structures.

The studies presented in this thesis together significantly advance our biophysical knowledge of lipid transport. Furthermore, the approaches used and methods developed provide a means to investigate the function of numerous lipid transfer proteins. By understanding how individual lipid transfer proteins function at the molecular level, we can begin to understand how they collectively function to spatiotemporally regulate cell membrane compositions.

For my grandparents.

"So in this sense, all places are the same place. And the only time you ever notice any difference, is in the moment of transition."

— Alan Watts

Contents

Acknowledgments

Over the course of my graduate studies, I've had the pleasure of working alongside many talented researchers who've, in turn, shaped my own scientific career. Below I want to acknowledge members of the Berkeley community that I've interacted closely with, but by no means is this a complete list.

Firstly, I want to thank my graduate advisor, Prof. Phill Geissler, for providing me the freedom to pursue my own ideas and grow into an independent scientist. His support and guidance along the way have been invaluable.

Secondly, I'm grateful to the many Berkeley professors whom I've learned from. In particular, I want to thank Prof. David Limmer for many valuable discussions about both my research and teaching, and for providing me opportunities to present my work during the Berkeley seminars that he organized. Additionally, I want to thank Prof. Dan Fletcher for inspiring new research directions and fostering collaborations. I also want to thank Prof. Eran Rabani for serving on my qualifying committee and continuously offering insights into academia. I want to thank Prof. Matthew Francis for being a wonderful collaborator and advocate for graduate students in the department. Finally, I want to thank Prof. Naomi Ginsberg for also being a wonderful collaborator and a strong supporter of my scientific success.

Thirdly, I want to acknowledge my fellow Geissler Group members, who've shaped my daily research experience. I want to thank Andreana Rosnik, Pratima Satish, Katie Klymko, and Dr. Steve Cox for showing me the ropes initially. I want to thank Layne Frechette, Dr. Georg Menzl, Dr. Jaffar Hasnain, and Dr. Ahmad Omar for many productive scientific discussions, especially those over lunch. I want to thank Paul Wrona, Nathan Odendahl, and Lucie Liu for being wonderful colleagues. I am especially grateful for having had the opportunity to work with many talented Berkeley undergraduate researchers. In particular, working with Rian Kormos helped initially develop many of the ideas used in Chapter 5, and Gustavo Espinoza Garcia significantly contributed to the work presented in Chapter 3. I learned a lot from informal discussions with Easun Arunachalam and Ty Perez, and I hope they learned some useful things from me too. Finally, but certainly not least of the Geissler Group members, I want to thank Leslie Dietterick, our group assistant, for managing all logistics so that I could focus solely on research and for sharing delicious homemade baked goods, which always made my day a little brighter.

Fourthly, I want to thank members of the Rabani and Limmer Groups whom I've had the pleasure of getting to know over lunches, at conferences, and outside of the office. From the Rabani Group, I especially want to acknowledge Dipti Jasrasaria, Lyran Kidon, (and Paul Wrona, a joint Rabani-Geissler Group member) for their continued friendship; John Philbin, Dr. Ming Chen, Dr. Ty Takeshita for many lively discussions and scientific support; and Nathan Ng and Dr. Amikam Levy for being wonderful colleagues. From the Limmer group, I especially want to thank Avishek Das for his introduction to the nonequilibrium model explored in Chapter 5; Dr. Sam Niblett, Dr. Mirza Galib, Trevor Grand Pre, Addison Schile, and Chloe Gao for being wonderful colleagues and great company at conferences.

Fifthly, I want to thank my many fantastic collaborators whom I've had the pleasure of conducting research with. I enjoyed investigating a biomimetic artificial photosynthetic system with members of the Ginsberg Group, Dr. Milan Delor, Trevor Roberts, and Leo Hamerlynck; members of the Francis Group, Jing Dai and Amanda Bischoff; and Prof. Jeffrey Neaton, and members of his group, Samia Hamed and Caroline McKeon. I have greatly benefited from Prof. Juan Garrahan's insights and ideas about using path sampling methods to simulate nonequilibrium systems; he significantly contributed to the work presented in Chapter 5. I have enjoyed discussing all things about lipid membranes and continuing to collaborate with members of the Fletcher Group, especially Dr. Sungmin Son, Carmen Chan, and Dr. Eva Schmid.

Finally, I want to acknowledge my fellow graduate students who've served on departmental committees and conducted scientific outreach with me. I'm grateful to have served on the 2019-2020 Junior Chemistry Faculty Student Hiring Committee alongside dedicated physical chemistry students Dipti Jasrasaria, Scott Stonemeyer, Julia Carlstad, Dip Hait, and Valerie McGraw. I enjoyed organizing the 2021 Student-Hosted Physical Chemistry Seminar with Jacki Shea, Emma Berger, and Daria Balatsky. I valued the time I spent teaching BASIS lessons to local elementary students with Sheena Louisia, Kiera Wilhelm, and many other fantastic volunteer science teachers. I have been inspired by my fellow graduate students' exceptional commitment to improving the culture of academic research and their continued efforts to involve more people in scientific discovery.

Chapter 1

An Introduction to Membrane Biophysics and Lipid Transport

Lipids: Chemically Diverse Amphiphiles

Lipids are one of the essential building blocks of all cells. As a class of biomolecules, all lipids, regardless of their specific chemical structure, share common physical features. Lipids have elongated hydrophobic tails attached to a hydrophilic headgroup. The hydrophilic headgroup favorably interacts with water, for example by incorporating itself into water's hydrogen bond network, whereas the hydrophobic tails prefer to be shielded from water. Lipids are amphiphiles since they are composed of groups that have conflicting relationships with water. Due to their amphiphilicity, lipids readily assemble into bilayers so that the hydrophobic tails are shielded from solvent by the polar headgroups (Figure 1.1A).

Over a thousand chemically distinct lipid species make up cellular membranes.[1,2] The main classes of eukaryotic lipids are glycerophospholipids (GPLs),[3] sphingolipids,[4] and sterols[5] (Figure 1.1B). Such chemical diversity arises from mixing and matching different headgroups and tails. Indeed, lipids have modular structures: GPLs are composed of a phosphate functionalized headgroup, glycerol backbone, and two acyl chains derived from fatty acids; sphingolipids are composed of a headgroup, sphingoid base (which is both the backbone and one acyl chain), and one N-acyl chain derived from a fatty acid; and sterols are composed of a hydroxyl group, fused ring backbone, and hydrocarbon chain. Combinations of chemically distinct building blocks create a large diversity of lipids. Chemical variations in the acyl chain(s) and headgroup contribute to the diversity of both GPLs and sphingolipids. Fatty acid chains, such as those shown in Figure 1.1C, differ in length, double bond number, double bond position, and hydroxylation. Headgroups, such as those shown in Figure 1.1D, vary in net charge, polarity, and size. Among GPLs, the glycerol backbone and *sn1* acyl chain are linked through either an ester, ether, or vinyl-ether linkage. Among sphingolipids, the sphingosine base can be hydroxylated to form 4-hydroxy-sphingosine or unsaturated to form sphinganine. Among sterols, the fused rings vary in the number of methyl substituents

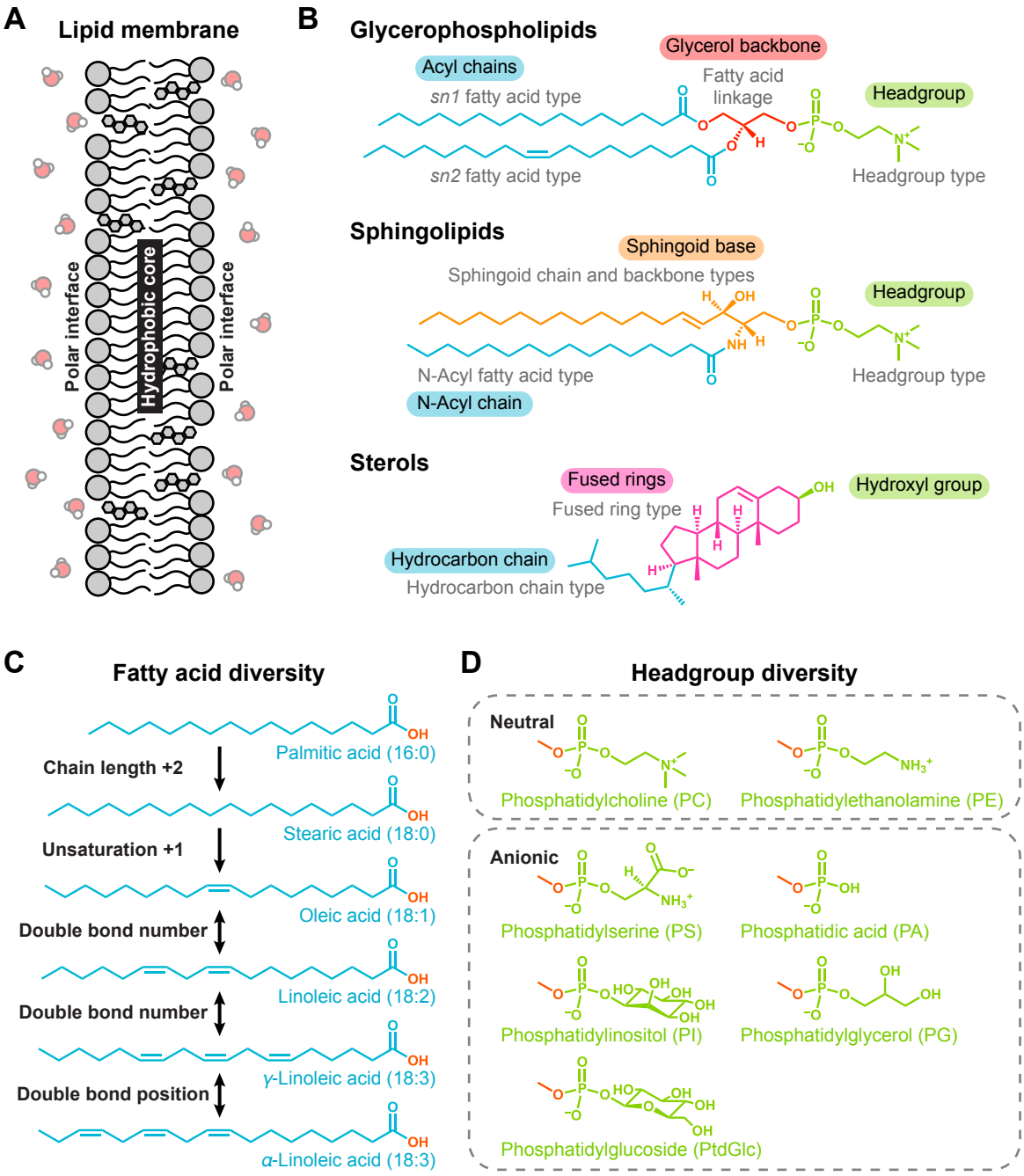


Figure 1.1: Cell membranes assemble from a diversity of lipids. (A) Due to their amphiphilic structure, lipids form bilayers. The hydrophobic tails create a hydrophobic core sandwiched between polar surfaces of headgroups. (B) Glycerophospholipids (GPLs), sphingolipids, and sterols are the main classes of cellular lipids. Each lipid has a modular structure composed of building blocks, which are labeled in boxes of different colors. Chemical variations in those building

(Continued on next page)

Figure 1.1: *(Continued from previous page)*
blocks create a large diversity of lipids. GPLs have two acyl chains attached to the *sn1* and *sn2* positions of a glycerol backbone. A headgroup, which is composed of a phosphate and variable functional group, is attached to the *sn3* position of the backbone. GPLs are named according to both their tails and headgroup. Sphingolipids have an N-acyl chain attached to a sphingoid base, which is simultaneously the backbone and a second chain. A headgroup, which may not include a phosphate, is attached to the sphingoid base. Sphingolipids are named according to their headgroup. Sterols have a hydrocarbon chain attached to fused rings, which are functionalized with a hydroxyl group at the other end. The chemical structure of cholesterol, the major mammalian sterol, is shown. Variations in acyl chain, backbone (or sphingoid base), and headgroup chemistry result in a diversity of GPLs and sphingolipids. For example, (C) acyl chains, which are derived from fatty acids, vary in length, degree of unsaturation, and position of any double bonds, and (D) headgroups vary in chemical structure.

and number of double bonds, and the hydrocarbon chain varies in length, number of double bonds, and location of methyl branching.

Such a large chemical diversity allows lipids to serve multiple cellular functions. First, lipid membranes act as fluid boundaries that delineate a cell from its surrounding environment and that organize and compartmentalize intracellular functions (Figure 1.2A).[2] Such compartmentalization enhances biochemical efficiency by colocalizing interrelated chemical reactions within the same cellular volume and by creating microenvironments tailored to each organelle's function. Because lipids help establish particular microenvironments, organelles' membrane compositions vary depending on their function. For example, the membrane of the endoplasmic reticulum (ER) contains a larger fraction of phosphatidylcholine (PC) lipids (Figure 1.2A, red box) compared to the plasma membrane (PM), which contains a significant fraction of anionic phosphatidylinositol (PI) and phosphatidylserine (PS) lipids (Figure 1.2A, blue box).[2] Throughout the cell, membranes also serve as 2D planar environments for proteins, and, thus, they can directly modulate the activities of integral and peripheral membrane proteins[6–8] (Figure 1.2A, gray box). Due to their inherently 2D structures, membranes can also enhance the assembly of membrane-bound protein complexes that otherwise occurs slowly in 3D.[9] Due to their fluidity, membranes, and hence cellular boundaries, are malleable and dynamic. This enables many essential cellular processes, such as cell division, biological reproduction, and intracellular membrane trafficking, to occur through membrane budding, tubulation, fission, and fusion.

Second, lipids serve as chemical energy sources. Due to their reduced and anhydrous from, triacylglycerols (TAGs) along with cholesterol esters (CEs) are predominantly used for energy storage within lipid droplets (Figure 1.2B). Lipid droplets function as both caloric reserves and stocks of precursors for membrane biogenesis. Thus, lipid droplets and their lipid contents are essential for both energy and membrane homeostasis.[10]

Third, lipids act as signaling molecules.[11] Phosphoinositide GPLs (PIPs)[12–14] and sph-

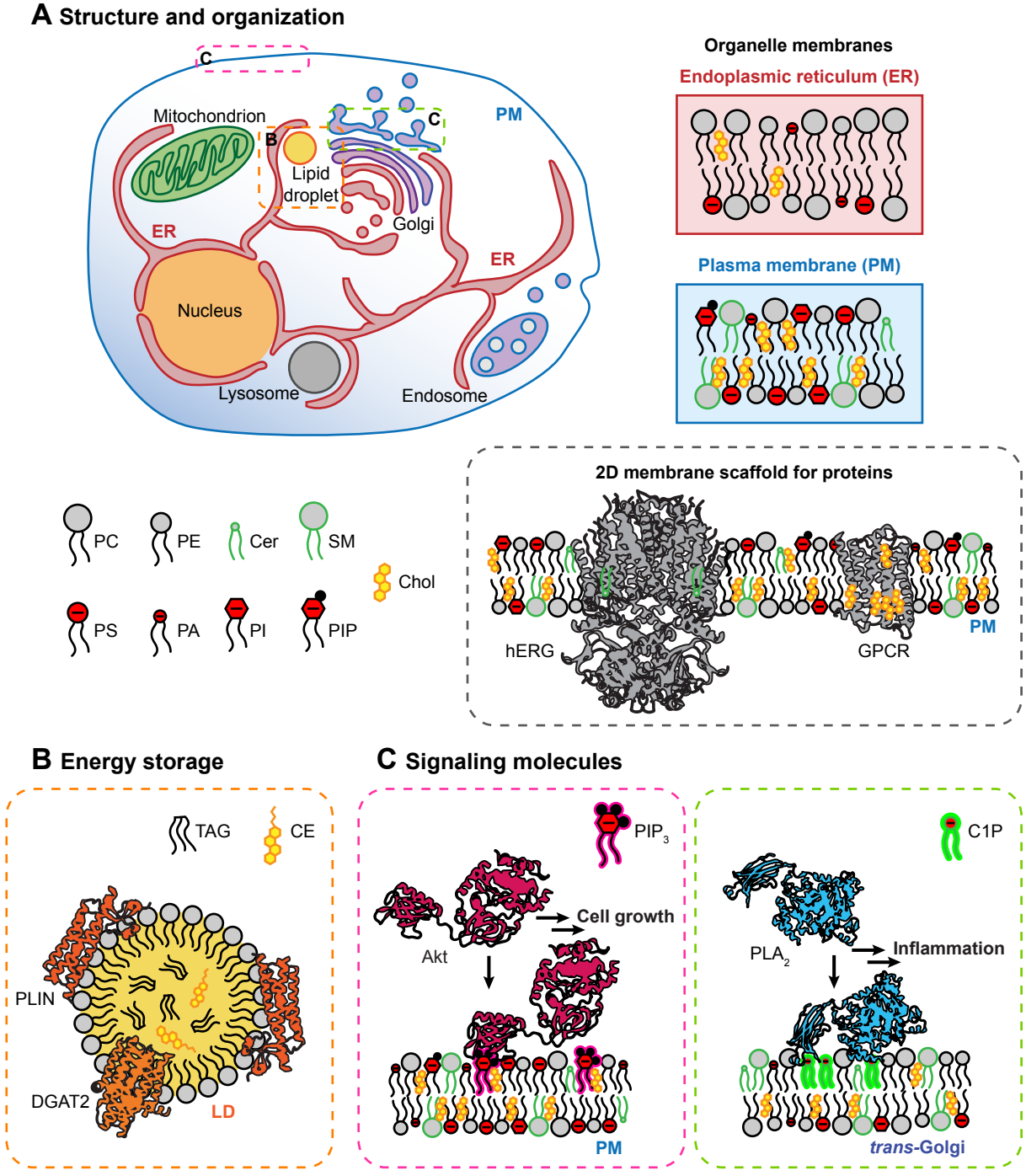


Figure 1.2: Lipids serve multiple cellular functions. (A) Lipid membranes act as physical boundaries between cells and their environment and compartmentalize cellular functions into organelles. Organelles' membrane compositions are fine-tuned to their function, and a model of two

(Continued on next page)

Figure 1.2: *(Continued from previous page)*
lipid territories can describe the starkest variations observed. The membrane of the endoplasmic reticulum (ER) epitomizes a territory characterized by loose lipid packing and neutral surfaces. The plasma membranes (PM) epitomizes a territory characterized by tight lipid packing and anionic surfaces. Illustrated are some of the common lipids found in each membrane: PC GPLs, PE GPLs, ceramide (CER), sphingomyelin (SM), PS GPLs, PA GPLs, PI GPLs, phosphatidylinositol phosphate GPLs (PIP), and cholesterol (Chol). Because membranes also serve as 2D platforms for peripheral and transmembrane proteins, lipids modulate the activity, assembly, and function of proteins, such as the human ether-a-go-go-related gene encoded K^+ channel (hERG) and G protein-coupled receptors (GPCRs). (B) Lipids are energy sources. Significant amounts of energy are stored in lipid droplets (LDs), which have a neutral core of mostly triacylglycerides (TAGs) and cholesteryl esters (CEs) surrounded by a GPL monolayer decorated by proteins such as perilipins (PLINs) and diacylglycerol acyltranferase-2 (DGAT2). (C) Lipids are signaling molecules. For example, phosphatidylinositol-(3,4,5)-trisphosphate (PIP_3) in the PM acts a secondary messenger of extracellular stimuli by activating and recruiting Akt (also called protein kinase B) to the membrane, which once activated mediates downstream pathways responsible for cell growth, proliferation, and migration. Another example is ceramide-1-phosphate (C1P). C1P activates and recruits cytosolic phospholipase A_2 ($cPLA_2$) to the *trans*-Golgi, where it produces arachidonic acid, a pro-inflammatory molecule.

ingolipids[15–17] are commonly involved in signal transduction. Lipids can be both primary and secondary messengers in signalling pathways.[18] For example, the class of lipids known as eicosanoids are primary messengers responsible for stimulating immune responses to infections and injuries.[19] Another example, phosphatidylinositol-(3,4,5)-trisphosphate (PIP_3) acts as a secondary messenger of growth factor stimulation. Bioactive lipids can also recruit cytosolic proteins to a membrane and, as a result, congregate together proteins that are responsible for further downstream signal propagation or bring enzymes that produce signaling molecules into proximity with their substrates. In fact, PIP_3 recruits Akt (also known as protein kinase B) to the plasma membrane and causes conformational changes necessary to activate Akt and further propagate downstream signals (Figure 1.2C, magenta box).[20] Another example which will be featured in Chapter 4, ceramide-1-phosphate (C1P), a bioactive sphingolipid, recruits phospholipase A_2 (PLA_2) to the Golgi membrane (Figure 1.2C, green box). When bound to the *trans*-Golgi, PLA_2 produces precursors of eicosanoids, ultimately activating inflammatory pathways.[21,22]

Thus, lipids as a class perform multiple essential cellular functions due to the fact that subtle chemical differences endow individual lipids with unique properties. Indeed, a number of genetic diseases are linked to mutations in the enzymes responsible for lipid metabolism, remodelling, and modification.[1] For example, mutations in the enzymes that catalyze the first steps of sphingolipid synthesis (serine palmitoyltransferase 1 and 2) that make them less selective for their main substrate lead to the production of 1-deoxysphingolipids. Although 1-deoxysphingolipids only differ from sphingolipids by a single hydroxyl group, they are toxic,

and buildup causes hereditary sensory and autonomic neuropathy.[23,24] Altered lipidomes have been found in a number of other diseases, including cancers,[25,26] diabetes,[27] cystic fibrosis,[28,29] and neurological ones.[30–32] In many cases, though, neither causality nor the mechanism by which aberrant lipid chemistry results in disease has been determined.

Membrane Physicochemical Properties Are Determined by their Lipid Composition

Cell membrane compositions exhibit further compositional diversity. Membrane compositions vary between species,[33,34] between tissues and cells within an organism,[35–37] between organelles within a cell[2,38] (see for example Figure 1.2A, red and blue boxes), and even between leaflets[39] and lateral domains of a single membrane.[40] Membrane compositions also vary in time, and such dynamic changes can help drive cell state transitions.[41,42]

The collective behavior of a membrane's constituent lipids determines its physicochemical properties. So, membranes with different compositions exhibit different properties. Because mesoscale membrane properties result from molecular interactions between many lipids, *in vitro* experiments using model membranes and molecular simulations have both greatly advanced the field's understanding of how lipid chemistry influences a membrane's properties.[40,43–46] While such studies lack the full complexity of cellular membranes, they form a biophysical foundation for understanding how the properties of more complex membranes arise. A subset of membrane properties with known cellular relevance (most of which are also highly relevant to Chapter 3) are discussed in the following paragraphs and illustrated in Figure 1.3.

Membrane surface electrostatics are determined by the lipid headgroups (Figure 1.3A). Membranes with a significant fraction of anionic lipids, such as PS, phosphatidic acid (PA), PIs, and PIPs, have negatively charged surfaces. In contrast, membranes with predominantly zwitterionic GPLs, such as PC and phosphatidylethanolamine (PE), and neutral sphingolipids, such as ceramide (Cer), have neutral surfaces. Membrane electrostatics vary along the secretory pathway, which is responsible for trafficking material from the ER to its final destination through the Golgi. ER and *cis*-Golgi membranes are poorly charged whereas the *trans*-Golgi and plasma membranes have a significant negative charge.[47] Because many peripheral membrane proteins bind to membranes through both specific and nonspecific electrostatic interactions, variations in cell membrane electrostatics help ensure their precise cellular localization.[48–50]

The thickness of the hydrophobic interior of a membrane is dependent on not only the length of the lipid tails but also how well lipids pack together (Figure 1.3B). Repulsion between lipid headgroups and steric hinderance among tails with kinked, unsaturated bonds can each cause looser lipid packing. Lipids with long, saturated tails, such as derived from arachidic and lignoceric acids (which have 20 and 24 carbons, respectively) assemble into thicker membranes compared to lipids with shorter tails, such as derived from lauric and

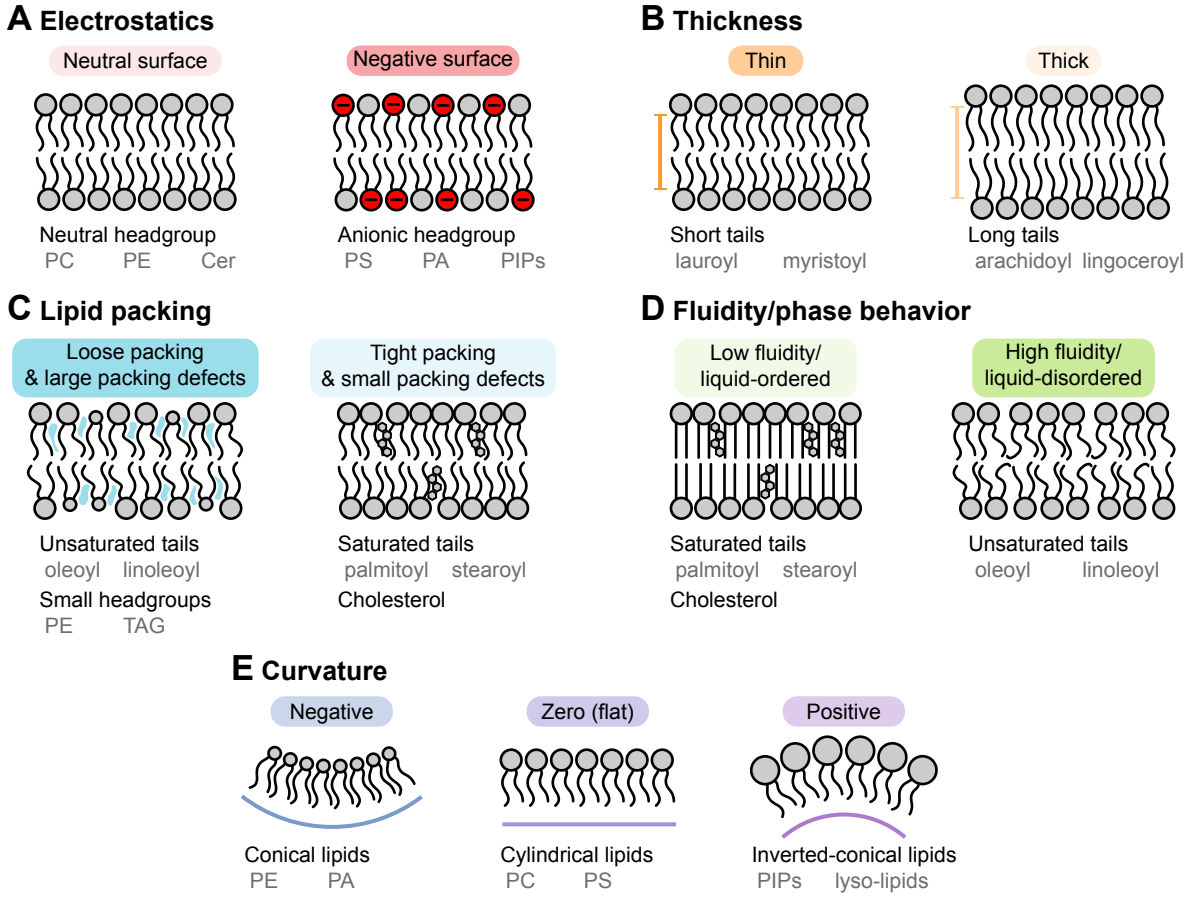


Figure 1.3: Membrane physicochemical properties are determined by their lipid composition. Examples of membrane physicochemical properties that vary between cellular membranes and influence biological processes are illustrated. Chemical features of lipids that are characteristic of a particular property are noted along with specific examples of such lipids. Because each property is determined by the collective behavior of the membrane's lipids, combinations of lipids not noted can also result in the same property. Many properties are interrelated and not mutually exclusive. (A) Membrane electrostatics are largely determined by lipid headgroup. (B) Membrane hydrophobic thickness can be increased by increasing lipid acyl chain length and also tightening lipid packing and increasing membrane order. (C) Loose lipid packing can create surface regions where the hydrophobic core is exposed to solvent, which are termed packing defects and indicated in blue. Lipids with unsaturated tails, membranes with a fraction of lipids with small headgroups, or curved membranes exhibit increased packing defects. (D) Membranes composed of lipids with saturated tails and cholesterol are ordered and commonly found in the liquid-ordered (L_o) phase if sterol is present or the gel (L_β) phase. Membranes composed of lipids with unsaturated tails are commonly found in liquid-disordered (L_d) or liquid-crystalline (L_α) phases. (E) Membrane leaflets with negative, zero, and positive curvature are preferentially formed by conical, cylindrical, and inverted-conical lipids, respectively.

myristic acid (which have 12 and 14 carbons, respectively) and lipids with unsaturated tails. Membrane thickness generally increases along the secretory pathway. Mirroring changes in organelles' membrane thickness, the length of transmembrane domains also typically increases from the ER to the plasma membrane.[51,52] By matching the length of the transmembrane domain to the membrane's thickness, no hydrophobic portions of the protein or of the membrane are exposed to solvent. Reducing such hydrophobic mismatch[53,54] can provide a driving force for proper protein localization between organelles[55–57] and lateral membrane domains[58] and also for protein clustering within a membrane.[59]

Due to imperfections in the geometrical arrangement of lipids within a membrane, packing defects arise (Figure 1.3C). Such defects expose portions of the membrane's hydrophobic core to solvent. Packing defects are more prevalent in membranes that contain high levels of unsaturated lipids,[60,61] curved membranes,[61] and membranes that include conical lipids, which have small headgroups.[62] In contrast, packing defects are sparse in membranes that contain high levels of saturated lipids and cholesterol.[60,61] Along the secretory pathway, the size and prevalence of packing defects decreases. Tight lipid packing and a minimal occurrence of defects decreases the plasma membrane's permeability and improves its barrier function. Loose lipid packing and an increased occurrence of defects may be advantageous for the ER's biosynthetic function.[47] Indeed, accumulation of saturated lipids or cholesterol, which reduce the occurrence of packing defects, triggers ER stress. Such stress can be relieved by increasing levels of saturated lipids.[63] Additionally, packing defect size and shape may help guide peripheral protein binding and cellular localization.[64,65]

Membranes are naturally fluid structures since lipids can diffuse within the plane of the membrane. The molecular structure of the membrane influences how quickly lipids diffuse: Highly ordered membranes exhibit low fluidity whereas disordered membranes exhibit high fluidity. Highly ordered membranes are typically composed of unsaturated lipids and cholesterol, which pack tightly together (Figure 1.3D). Disordered membranes are typically composed of unsaturated lipids, which loosely pack together. Even single component membranes can undergo an order-to-disorder transition from a gel phase to a liquid-cyrstalline phase when heated above their melting temperature.[66,67] Model membranes composed of a saturated lipid, unsaturated lipid, and cholesterol can separate into coexisting liquid-ordered and liquid-disordered domains.[68] Such local phase separation, in combination with specific lipid–lipid and lipid–protein interactions, may also contribute to the creation of lateral heterogeneities in cellular membranes, referred to as lipid rafts.[40,69,70] Lipid rafts may function to concentrate specific proteins together and, thus, regulate their activity.[71]

In addition to forming flat bilayers, lipids can form highly curved structures depending on the relative size of the headgroup and tails (Figure 1.3E). Cylindrical lipids, whose headgroups are roughly equal in width to their tails, such as PC and PS, yield membranes with zero spontaneous curvature. Conical lipids, whose headgroups are smaller in width than their tails, such as PE and PA, yield concave membranes with negative spontaneous curvature. Inverted-conical lipids, whose headgroups are larger in width than their tails, such as PIPs and single-tailed lyso-lipids, yield convex membranes with positive spontaneous curvature. The ability to form highly curved structures enables membrane to bud, tubulate, fuse, and

undergo fission.[72,73] Many cellular membranes, and especially those significantly involved in membrane trafficking, exhibit regions of high curvature, such as found among ER tubules and Golgi vesicles.[74]

Additional membrane phyiscochemical properties, along with those described above, have been found to regulate cellular function. These include properties such as lateral membrane pressure, membrane compressibility, and membrane stiffness.[75–77] However, as can be seen from the presented examples, many properties are interrelated, making a comprehensive list unnecessary to grasp the biological importance of maintaining proper membrane properties.

Cell Membrane Compositions are Spatiotemporally Regulated

To ensure that membrane physicochemical properties are within the ranges compatible with cellular physiology, membrane compositions are under homeostatic control.[75–77] Membranes adapt to perturbations, such as environmental changes, by altering their lipid compositions. For example, when ambient temperatures drop, tightly packing unsaturated lipids are replaced with loosely packing saturated ones in the membranes of prokaryotes and ectothermal animals to maintain membrane viscosity.[78,79] Even mammals exhibit such a 'homeoviscous adaptation' in response to dietary changes in saturated lipid content.[80] Because each organelle also relies on specific membrane physicochemical properties to function, membrane compositions are spatiotemporally regulated within a cell.

Membrane homeostasis relies on sense-and-respond systems. A protein senses changes in membrane properties and stimulates an effector to respond accordingly by altering lipidomes.[75,76] Because only a few sensors have been elucidated, their molecular mechanisms are largely uncharacterized. During a homeostatic response, lipidomes may be altered through changes in lipid metabolism, such as accomplished through transcriptional control of biosynthetic enzymes,[81,82] and also through redistribution of lipids *via* transport pathways.[83–85] The biophysical mechanisms of lipid transport are the focus of this thesis.

Lipids are transported between cellular membranes through both vesicular and non-vesicular mechanisms. Examples of both types of transport pathways are illustrated in Figure 1.4A. Vesicular transport is the predominant way that proteins and other molecules are transported along the secretory pathway and also between cells. Because lipids form vesicular transport vehicles, large amounts of lipids are non-specifically transported along with their cargoes.[77,86] Organelles that lack secretory machinery, such as mitochondria and peroxisomes, rely on non-vesicular transport to exchange lipids. Furthermore, non-vesicular transport enables rapid and specific transfer of individual lipids between all organelles.[85,87,88] Often, non-vesicular transport occurs at membrane contact sites, where organelle membranes are separated by a small (< 30 nm) cytosolic gap.[89–91] By reducing the distance that the lipid must travel, it can be transported faster, decreasing the time it takes to remodel lipidomes in response to environmental changes, acute stresses, metabolic challenges, and

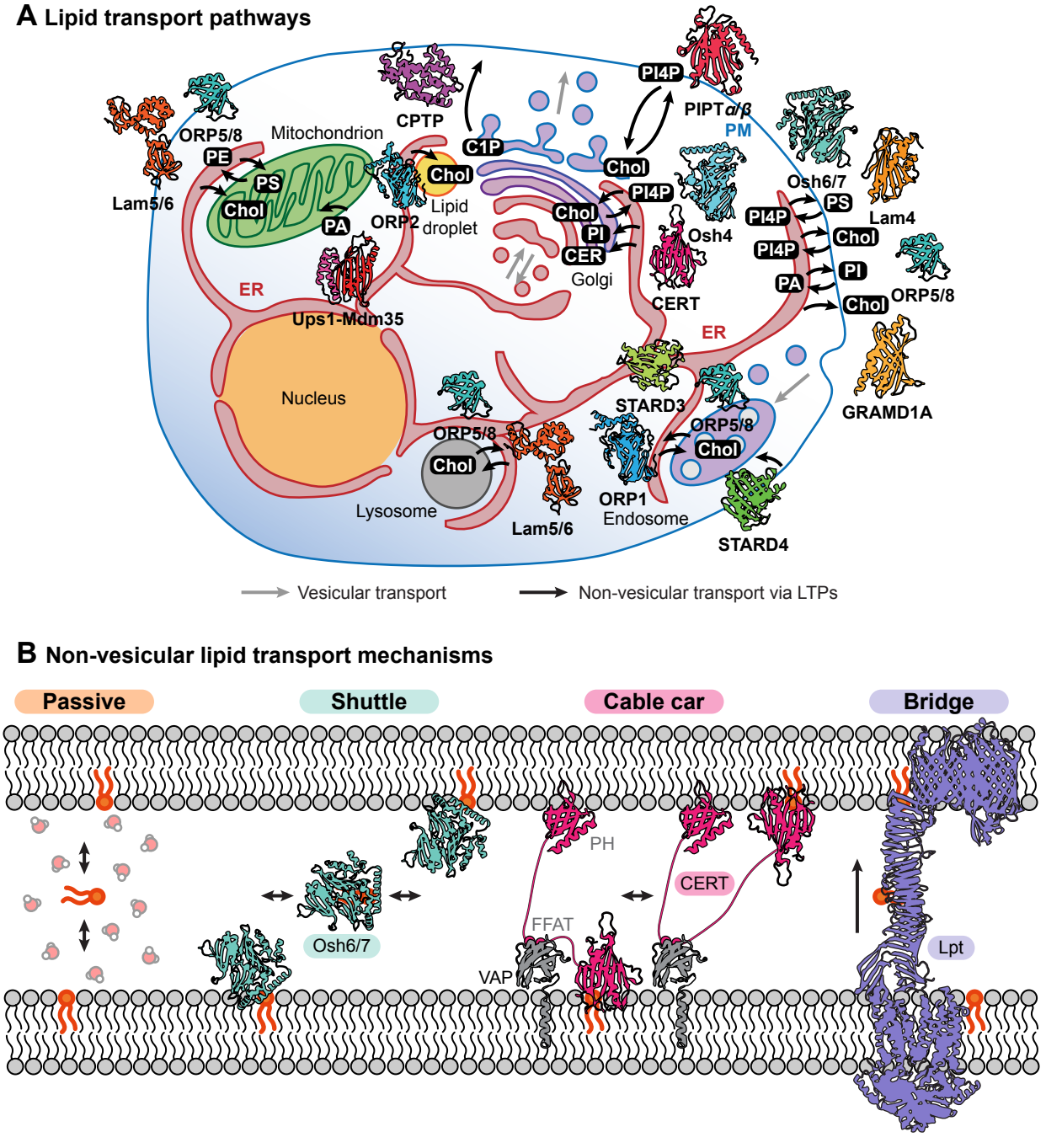


Figure 1.4: Membrane compositions can be modified through lipid transport. (A) Example vesicular (gray arrows) and non-vesicular (black arrows) lipid transport pathways between membranes are shown. Examples of lipid transfer proteins (LTPs) which facilitate non-vesicular

(Continued on next page)

Figure 1.4: *(Continued from previous page)*
transport are indicated: ceramide transfer protein (CERT), ceramide-1-phosphate transfer protein (CPTP), GRAM domain-containing protein 1A (GRAMD1A), LTP anchored at membrane contact sites (Lam), oxysterol binding protein (OSBP)-related protein (ORP), OSBP homologue (Osh), phosphatidylinositol transfer protein (PIPT), steroidogenic acute regulatory protein (StAR) domain containing protein (STARD), complex of ubiquitin–proteasome system 1 (Ups1) and mitochondrial distribution and morphology 35 (Mdm35). (B) Illustrations of non-vesicular transport mechanisms. During passive transport, a lipid traverses the cytosol alone as opposed to being escorted by a LTP. Cytosolic LTPs, such as Osh6/7, shuttle lipids between membranes. LTPs with membrane tethering domains and motifs, such as CERT with a pleckstrin homology (PH) domain that binds PIPs and with a two phenylalanines in an acidic tract (FFAT) motif that interacts with the integral membrane protein VAMP-associated protein (VAP), function like cable cars. LTPs that span the distance between two membranes, such as the lipopolysaccharide transport apparatus (Lpt), function like bridges.

cellular differentiation.

The simplest possible mechanism for non-vesicular transport is characterized by passive diffusion of a lipid between membranes (Figure 1.4B, leftmost mechanism).[85] Such a passive mechanism, involving lipid desorption from a donor membrane, diffusion through solvent, and insertion into an acceptor membrane, has been found to occur in a number of *in vitro* experiments.[92–103] However, due to the larger free energy barrier for lipid desorption, passive lipid transport occurs over hours to days, far too slowly to be the main *in vivo* mechanism. Nevertheless, a complete biophysical characterization of passive lipid exchange, as given in Chapters 2 and 3, provides a foundation to understand more complex mechanisms.

Within a cell, lipid transfer proteins (LTPs) may act as catalysts of non-vesicular lipid transport by lowering the free energy barrier for lipid desorption.[85,87,88,91,104,105] Examples of known LTPs and their transfer pathways are indicated in Figure 1.4A. While LTPs can vary in structure and utilize different transport mechanisms, all encase a lipid within their hydrophobic interiors so that the lipid's tails are shield from the cytosol during transport. A comprehensive understanding of how LTPs catalyze the transport of specific lipids between target cellular membranes, often performing the energy-intensive process of transporting lipids up concentration gradients, requires further investigation.

In vitro and computational studies have begun to elucidate some of the mechanisms used by LTPs. While some LTPs show specificity for particular lipid species, such as ceramide transfer protein (CERT), which transports ceramides with acyl chains of $14-20$ carbons,[106] others are nonspecific, such as sterol carrier protein 2 (SCP2), which transports diacyl GLPs, glycolipids, and cholesterol.[107,108] LTPs can act like cytosolic lipid shuttles (Figure 1.4B, mechanism second from the left). Such LTPs may recognize the proper donor and acceptor membranes based on their physicochemical properties.[109,110] A detailed molecular mechanism of a shuttle-like LTP, ceramide-1-phosphate transfer protein (CPTP), is described in Chapter

4. LTPs equipped with membrane tethering domains or motifs can act like lipid cable cars (Figure 1.4B, mechanism second from the right).[85,87,111,112] Both shuttle-like and cable car-like LTPs can undergo conformational changes upon membrane binding, triggering the next step in the lipid transport cycle.[113–116] In contrast to shuttle-like and cable car-like LTPs, which have a single opening to a lipid-sized cavity, bridge-like LTPs have an extended opening that runs along their length, creating a path between membranes. Lipids can slide along such a bridge, exposing only their hydrophilic heads to cytosol during their journey between membranes (Figure 1.4B, rightmost mechanism).[87,117,118]

While LTPs consistently transfer lipids between liposomes under laboratory conditions, their physiological roles may be more diverse. For example, they may sense specific lipids, present lipids to other proteins, or facilitate enzymatic remodeling of lipids.[85,87] Thus, some LTPs may act as both the sensor and effector in homeostatic sense-and-response systems.

Approaches to Model Lipid Dynamics

Non-vesicular transport of individual lipids is inherently a molecular scale process, but it occurs between membranes, which are mesoscale objects. Such a coupling of disparate scales endows membranes with many unique physicochemical properties that are essential to biological function. However, the need to access a range of length and timescales also poses challenges for modeling lipid transport.[119,120]

One approach to overcome this issue involves using models with varied resolutions to investigate different aspects of lipid transport. The most detailed model describes the system quantum mechanically and probes angstrom lengthscales and femtosecond to picosecond timescales; however, this level of detail is generally unnecessary to model lipid transport, a classical phenomena whose accurate representation does not require an explicit treatment of electrons. Furthermore, quantum mechanical calculations of cellular membranes remain too computationally expensive to perform.[120]

Classical force fields, which describe intra- and intermolecular interactions through approximate functional forms parameterized to reproduce quantum mechanical calculations or experimental data, form the basis of the most detailed models commonly used to study lipid membranes.[46,119,120] The highest resolution force fields explicitly describe all atoms in a system and are referred to as all-atom models. At a slightly lower resolution, united-atom force fields combine some atoms, such as non-polar hydrogens, into a single site and explicitly describe all other atoms. While united-atom models are somewhat computationally less expensive to simulate than all-atom models, improvements in hardware and software over the last decade have made all-atom simulations commonplace. All-atom simulations of thousands of lipids spanning lengthscales of 10s of nanometers and extending for multiple microseconds are indeed possible. As a result, all-atom simulations have been used to gain insights into numerous biological processes, including vesicle fusion,[121–123] membrane pore formation by antimicrobial and cell-penetrating peptides,[124,125] binding of peripheral proteins to membranes,[64] and modulation of transmembrane protein activity by lipids.[126,127]

Coarse-grained models group together multiple related atoms into a single bead. Thus, they trade atomistic resolution for the ability to model processes that occur over 100s of nanometers and milliseconds. The amount of detail captured by coarse-grained models varies widely. For example, the standard MARTINI model utilizes roughly $8-14$ beads to represent a lipid and a single bead to represent a water cluster[128] whereas the model of Cooke and Deserno uses 3 beads to represent a lipid and an implicit description of solvent.[129] Different approaches have been used to develop coarse-grained models, with some being developed to primarily reproduce detailed structural data from experiments or all-atom simulations and others to primarily reproduce macroscopic experimental observables.[46,119,120] Coarse-grained models have enabled simulations of biological processes otherwise too expensive to simulate at all-atom resolution. For example, coarse-grained simulations have been used to monitor lipid-mediated protein assembly and sorting between membrane domains and to probe how multiple components collectively remodel membranes.[126,127,130] While coarse-grained models of entire organelles can be built,[131,132] it is still computationally intractable to simulate their dynamics.

To reach even larger scales, including those probed with cell microscopy, even lower resolution models that do not account for chemical differences between lipid species or even explicitly model the individual lipids in a membranbe have been developed. Such models include supra-coarse-grained ones, dissipative particle dynamics models, lattice models, and continuum models.[46,119,120] Some of the largest scales are accessible with continuum models that represent a membrane mathematically based on its physical properties and otherwise ignore its chemical details.[133] For this reason, continuum models based on the elasticity theory developed by Helfrich and Canham in the 1970s continue to inform approaches used to model membranes at the micrometer scale.[132]

Models of different resolutions can be combined in multiscale approaches to capitalize on the benefits of each.[46,119,120] Consecutive multiscale schemes alternate between all-atom and coarse-grained simulations, mapping between the two models at every exchange. Long trajectories can be obtained with a coarse-grained representation to capture slow steps of a processes. By then switching to an atomistic representation, details of key fast steps are resolved. Such an approach has been used to monitor viral entry by influenza hemagglutinin[134] and is also used in Chapter 4 to investigate protein facilitated lipid transport. Concurrent multiscale schemes simulate some regions of a system at a coarse-grained level and others at an all-atom level. While potentially more powerful than a consecutive multiscale scheme, methods to interface between regions of varying resolution are still being developed.

While the choice of model significantly dictates the length and timescales that are accessible, a judicious choice of simulation method can extend the range of accessible timescales. Molecular dynamics (MD) simulations, which are widely used to study biomolecules and used in Chapters 2-4, evolve the positions of all particles in time by numerically solving classical equations of motion. Because typical integration time steps are limited by the highest frequency motions (1-2 fs in all-atom simulations to account for fast vibrations of bonds involving hydrogens), millions of integration steps must be performed to generate ns-long trajectories. Unfortunately, this is generally far too short to observe any biologically im-

portant events. Furthermore, because the space of possible system configurations is very high dimensional ($3N$-dimensional, where N is the number of atoms) as illustrated in Figure 1.5A, it is incredibly difficult to sample even the most biologically important states, let alone all of them, with a single MD trajectory. Attempts to simulate transitions between states are generally further doomed since during most (if not all) of the trajectory, the system will remain in single, free energetically favorable state waiting for a rare fluctuation to cause a rapid transition to a new state (Figure 1.5B, top panel). Unfortunately, observations of thousands of transitions are required to characterize a reaction mechanism, including that of lipid transport.

While conventional MD simulations typically generate unreactive trajectories, reactive ones can be more readily obtained from transition path sampling (TPS) simulations.[135–137] TPS simulations utilize a Monte Carlo scheme to efficiently harvest transition paths, which are trajectory segments that connect an initial state with a final state (Figure 1.5B, middle panel). By analyzing the harvested transition paths, reaction mechanisms can be deciphered. Often of prime interest is identifying the reaction coordinate, which describes the collective motion of molecules that advances a transition.[136–141] Because the reaction coordinate captures the slowest degrees of freedom that can otherwise hinder a transition, projections of a system's full configuration space onto the reaction coordinate preserve information about the transition dynamics, whereas a projection onto an order parameter that is not the reaction coordinate does not (Figure 1.5A). Furthermore, free energy profiles along the reaction coordinate resolve the rate-limiting activation barrier, whereas free energy profiles along another order parameter underestimate or even fail to resolve the activation barrier (Figure 1.5B, bottom panel).[139] Atop the activation free energy barrier is the transition state, which is the collection of configurations that are equally likely to end up in either stable state.[136–138,142] Identifying the reaction coordinate, and thus also the transition state, provides a complete description of the transition dynamics. In Chapter 2, the reaction coordinate for passive lipid transport is elucidated by analyzing transition paths.

TPS simulations can yield incredibly valuable insights into transition dynamics. Indeed they have been used to investigate enzymatic reactions,[143–145] conformational changes of proteins[146–148] and nucleic acids,[149–151] and ion transport through protein channels,[152] for example. However, TPS simulations can be ineffective at sampling long transitions with metastable intermediates and also transitions that occur through multiple, distinct paths. As an initial attempt to address some of these challenges, Chapter 5 describes the development of new path sampling approaches and their application to a simple model of non-equilibrium transport.

Once the reaction coordinate is known, approaches that do not require sampling transition paths can be used to investigate a dynamical event. Instead, free energy profiles can be efficiently calculated as a function of the reaction coordinate using enhanced sampling methods, such as umbrella sampling, metadynamics, and accelerated MD simulations.[153,154] Such free energy profiles can be combined with kinetic theories to calculate reaction rates. Additionally, they can be used to assess how subtle changes in a system, which do not change the overall reaction mechanism, affect the activation free energy and, thus, reaction rate. In

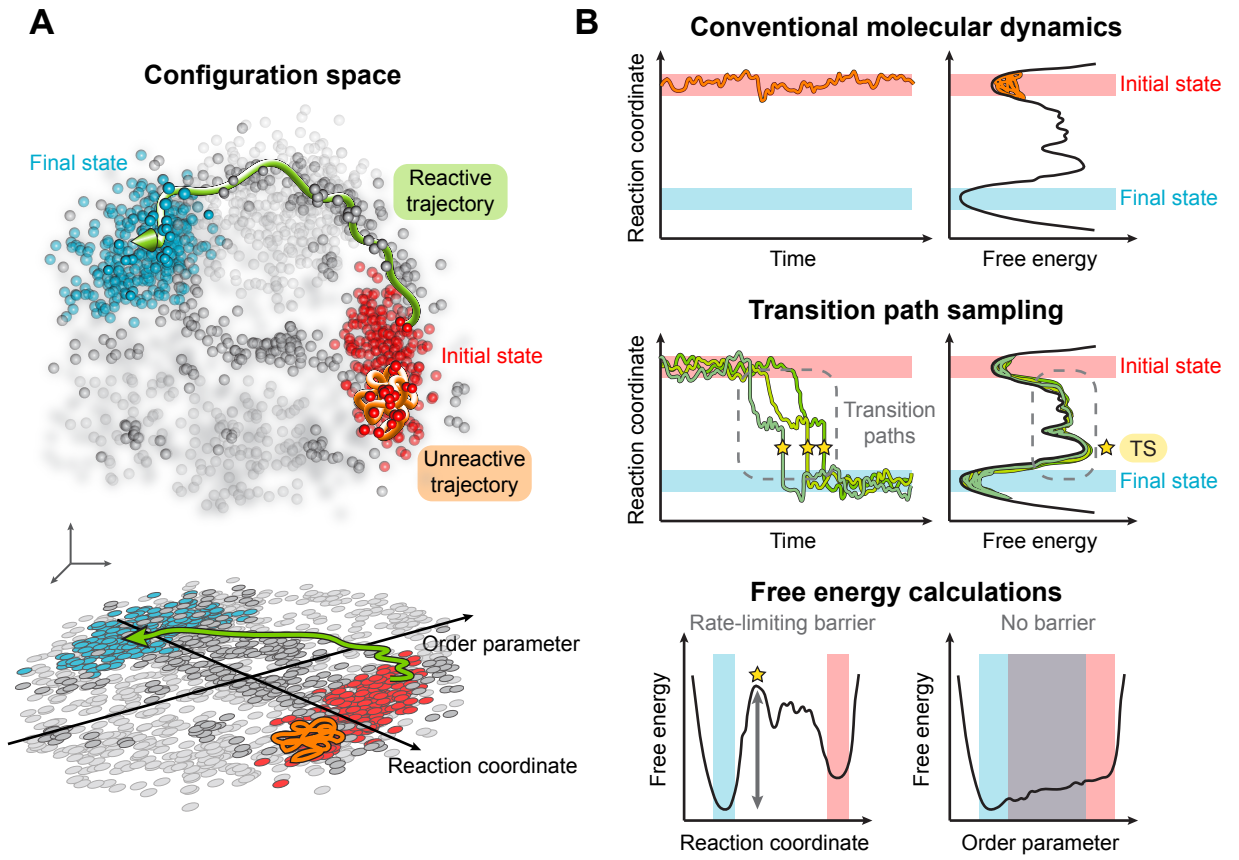


Figure 1.5: Multiple approaches can be used to overcome challenges simulating biomolecular dynamics. (A) (Top) Sampling challenges arise due to the high dimensionality of a biomolecular system's configuration space, which is illustrated in only three dimensions for visualization purposes. Each point represents a different possible configuration of the system's atoms. (Bottom) Projecting this high dimensional configuration space onto a smaller number of coordinates enables physical insights to be more readily gleaned. The choice of coordinates determines how well important aspects of the high dimensional space are maintained in the lower dimension projection. A projection onto the reaction coordinate maintains information about all degrees of freedom involved in advancing a transition between a pair of states, whereas a projection onto an order parameter that is not the reaction coordinate lacks such information. (B) Different simulation approaches can be used to study biomolecular processes. To determine the free energetics of different configurations, each point in configuration space must be sampled according to its Boltzmann probability. To investigate the kinetics of a dynamical process, or transitions between an initial state (red points) and final state (blue points), reactive trajectories (green) must be sampled. (Top) Conventional molecular dynamics (MD) simulations naturally sample biomolecular dynamics by propagating Newton's equations of motion. They typically sample fluctuations within a single stable configurational basin in the free energy landscape and generate unreactive trajectories (orange). *(Continued on next page)*

Figure 1.5: *(Continued from previous page)*
(Middle) Transition path sampling simulations sample transition paths, which are trajectory segments that connect an initial state to a final one (dashed gray box), using Monte Carlo moves performed in trajectory space. From the harvested ensemble of reactive trajectories, the transition state (TS, indicated with a star), which is composed of configurations that are equally likely to end up in either the initial or final state, can be identified, and the reaction coordinate can be determined. (Bottom) Free energy calculations can be performed using enhanced sampling techniques to obtain free energy profiles along coordinates chosen *a priori*. A free energy profile along the reaction coordinate accurately quantifies the rate-limiting activation free energy and can be combined with kinetic theories to calculate rate constants. The TS is accurately resolved in such free energy profiles. In contrast, a free energy profile along an order parameter that is not the reaction coordinate underestimates the activation free energy, does not resolve the TS, and cannot be used to accurately characterize dynamics.

Chapter 3, this approach is used to determine how subtle changes in lipid chemistry and membrane order influence the rate of passive lipid transport.

Molecular simulations are the primary method used in this thesis to gain biophysical insights into lipid transport, but this approach is not without limitations. Significantly, molecular simulations provide a picture isolated from the biological networks responsible for maintaining membrane homeostasis. In response to cellular signals and external perturbations, cell membrane compositions are altered through lipid metabolism and trafficking. However, algorithmically mimicking a cell's homeostatic response by changing membrane compositions during an MD simulation poses technical challenges. For this reason, most MD simulations are performed in thermodynamic ensembles where the number of atoms (and lipid identity) is held fixed. A recently developed framework takes initial steps to overcome this limitation by conducting simulations in the semigrand canonical ensemble and representing the influence of metabolic and transport networks on membrane composition in terms of chemical potential differences between lipid species.[155] Much future work, including the development of new simulation algorithms that explicitly account for cellular sense-and-respond systems, will be required to fully model membrane homeostasis and to understand the regulatory logic that is responsible for maintaining membrane physicochemical properties.

Chapter 2

Breakage of Hydrophobic Contacts Limits the Rate of Passive Lipid Exchange between Membranes

Introduction

Lipid membranes are essential to the structural integrity of cells. They form fluid boundaries that organize and compartmentalize cellular functions within organelles. Each organelle requires a unique membrane composition for its proper function.[2,88] To maintain such varied compositions, lipids are heterogeneously trafficked between cellular membranes through vesicular or non-vesicular transport. Vesicular transport plays a major role in trafficking lipids and proteins between organelles in the secretory pathway.[77,86] Organelles that are not connected by vesicular transport machinery rely on non-vesicular mechanisms to receive and export lipids. Even for organelles in the secretory pathway, non-vesicular transport mechanisms provide an additional way to more rapidly exchange lipids, for example, to swiftly alter membrane compositions in response to environmental changes.[85,87]

Despite its cellular importance, non-vesicular transport mechanisms *in vivo* have not been fully characterized. Non-vesicular transport predominantly involves the movement of individual lipids between membranes. Monomeric lipid transfer between membranes may occur passively, in which a lipid desorbs and freely diffuses to another membrane.[85] Alternatively, lipid transfer proteins may facilitate monomeric lipid exchange by enclosing lipids within their hydrophobic interiors during transport.[87,104] With half-times on the order of hours, passive lipid exchange is too slow to fully account for lipid transport *in vivo*,[85] indicating that lipid transfer proteins are likely catalysts of lipid exchange. Nevertheless, a

mechanistic understanding of passive lipid exchange can inform our knowledge of how lipid transfer proteins increase the rate of lipid transport.

Molecular simulations are well suited to explore the microscopic dynamics and to identify the rate limiting step of passive lipid exchange. Quantifying the free energy barrier associated with the rate limiting step, however, requires knowledge of the reaction coordinate, which characterizes the collective motion of molecules that advances a transition.[136–139] Previous computational work[156–160] on lipid transport has presumed that a lipid's displacement normal to the membrane is the reaction coordinate (Figure 2.1A) and has yielded results in conflict with experimental findings.[97–99,161] Here we show that the reaction coordinate for passive

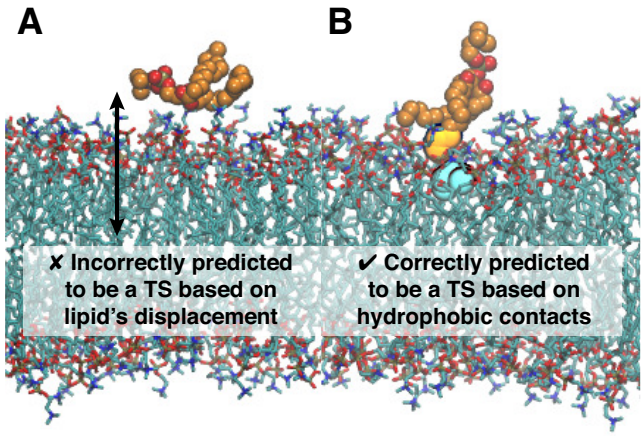


Figure 2.1: A lipid's displacement normal to the membrane has conventionally been presumed to be the reaction coordinate for passive lipid exchange; however, the lipid's displacement does not reliably identify transition state (TS) configurations. Although the lipid's displacement (black arrow) in configuration A is consistent with values observed at the transition state, this configuration is not a transition state. By contrast, configuration B is correctly predicted to be a transition state based on the extent of hydrophobic contact between the lipid and membrane (highlighted in yellow and cyan).

lipid exchange is indeed more subtle than a simple distance measurement. The reaction coordinate characterizes the creation (or disruption) of a locally hydrophobic environment around the incoming (or outgoing) lipid (Figure 2.1B). This realization resolves qualitative (but not quantitative) discrepancies between simulation and experiment and suggests that the breakage of hydrophobic contacts between a lipid and membrane limits the rate of passive lipid transport.

Experimental and Computational Background

Numerous *in vitro* studies of passive monomeric lipid exchange between membranes have demonstrated that it is a first-order process and that the rate of lipid exchange strongly

correlates with a lipid's solubility.[95–101,103,162,163] Based on these observations, the most widely accepted mechanism is characterized by aqueous diffusion.[92–103] The diffusive mechanism involves lipid desorption, which is rate limiting, followed by diffusion through solvent and insertion into another membrane. Based on experimentally measured activation energies, calculated activation free energies for lipid exchange exceed free energies for transferring a lipid from water to a membrane, indicating that there is a barrier for lipid desorption and insertion.[97–99,161] At this barrier, the desorbing lipid is hypothesized to only have the terminal carbons of its tails left within the membrane.[92,99] Thus, the activation free energy required to form the transition state has been attributed to the creation of a cavity in the membrane due to partial removal of a lipid and another cavity in the solvent to accommodate that lipid.[99,164]

However, because the experimental methods currently used to study lipid exchange have coarse temporal and spatial resolution, molecular features of the transient transition state can only be hypothesized. Molecular dynamics (MD) simulations offer an attractive means to test these hypotheses since the necessary time and length scales are accessible. Additionally, a complete free energy profile can be obtained from MD simulations. By calculating the free energy as a function of the reaction coordinate, which describes the system's dynamics during transitions between stable states,[136–138] free energies of activation can in principle be accurately quantified. If, however, the free energy is computed as a function of an order parameter that is not the reaction coordinate, then the apparent barrier generally underestimates the rate-determining free energy of activation.[139] Thus, identifying the reaction coordinate and corresponding free energy profile for passive lipid exchange can provide fundamental insights into the physical processes and work required to maintain heterogeneous cell membrane compositions.

Most previous computational studies have focused on obtaining a full free energy profile for lipid desorption and insertion.[156–160] Traditionally, free energy profiles have been computed as a function of a lipid's displacement normal to a bilayer measured from the lipid's phosphate group to the bilayer's center-of-mass (COM) (Figure 2.1A).[156–160] These free energy profiles lack a barrier for insertion,[156–160] seemingly in conflict with experimental results.[97–99,161] If the COM displacement is not the reaction coordinate for lipid exchange, the kinetically relevant barrier may not be resolved in these free energy profiles; instead, a barrier may exist along a different degree of freedom that is the reaction coordinate. Consistent with this idea, Vermaas and Tajkhorshid's MD study of lipid insertion indicated that the COM displacement is not sufficient to fully describe the microscopic dynamics of lipid insertion. They demonstrated that after the lipid associates with a bilayer, each tail of the lipid enters the bilayer successively to complete the insertion processes. The observation of splayed lipid intermediates during insertion, which are indistinguishable from other configurations based on the lipid's COM displacement, suggests that other degrees of freedom need to be considered to construct an accurate reaction coordinate for lipid exchange.[165]

Our Approach

The discrepancy between experimental[97–99,161] and computational[156–160] reports about a barrier for lipid exchange, together with the evidence suggesting that the lipid's COM displacement is a poor reaction coordinate from a MD study of lipid insertion[165] prompt two questions: (1) What is the reaction coordinate for lipid exchange? (2) What, if any, activation free energy barrier impedes the process of lipid insertion? In this article, we aim to answer these questions using molecular simulation. In doing so, we identify the reaction coordinate for passive lipid exchange in all-atom and coarse-grained lipid models. Knowledge of the reaction coordinate allows us to elucidate key biophysical details of the transition state ensemble and properly assess the free energetic cost of lipid transport.

Rather than driving the system along a presumed reaction coordinate, we instead harvest natural, unbiased trajectories in which a lipid spontaneously inserts into a membrane. Statistical analysis of this ensemble of dynamical pathways reveals the key collective motions required for lipid transport. Firstly, we find that lipid insertion is a barrier crossing event and occurs *via* three different pathways, distinguished by various splayed lipid intermediates. Secondly, we find that the reaction coordinate characterizes the formation and breakage of hydrophobic lipid–membrane contacts. The lipid's displacement normal to the bilayer, even formulated to distinguish splayed configurations, is not the reaction coordinate and obfuscates the barrier for insertion. Consistent with previous experimental results,[97–99,161] free energy profiles as a function of our reaction coordinate display a barrier for insertion and yield insertion rates calculated from Kramers theory that agree with those obtained directly from simulations. Finally, using our newfound reaction coordinate, we formulate a Smoluchoski equation for the lipid exchange rate to directly compare our simulation results with experiments. Overall, our results demonstrate that the rate limiting step for passive lipid exchange is the breakage of hydrophobic contacts, suggesting that lipid transfer proteins may catalyze lipid transport in part by lowering the associated activation free energy.

Methods

Molecular Dynamics Simulations

Spontaneous desorption of a lipid from a membrane is a very slow process occurring over minutes to hours, which is well beyond the timescale accessible in MD simulations. Most previous work has addressed this problem by introducing an external bias that allows rare configurations otherwise inaccessible in MD simulations to be readily sampled. While this approach generates configurations plausibly found along a lipid desorption trajectory, it can fail to reveal the natural, unbiased route of lipid desorption. Our approach instead exploits a fundamental statistical property of microscopic dynamics, namely its time reversibility. Natural desorption trajectories are simply the time reverse of spontaneous lipid insertion trajectories. As shown by Vermaas and Tajkhorshid,[165] the latter are straightforward to

generate in an unbiased way since lipid insertion is a rapid process. Thus, to gain insights into the dynamics of lipid exchange, we harvested trajectories of lipid insertion into a bilayer of 128 lipids using both all-atom and coarse-grained MD simulations. All simulations were performed in an isothermal-isobaric (NPT) ensemble using GROMACS 5.[166] The temperature was maintained at 320 K, ensuring that the bilayers were in the liquid crystalline phase. The pressure was maintained at 1 bar using semi-isotropic pressure coupling to allow the z dimension, which is perpendicular to the bilayer, to fluctuate separately of x and y, ensuring tensionless bilayers.

Coarse-Grained Systems. A large number of lipid insertion trajectories are required to make conclusions with high statistical accuracy. To obtain 1,000 lipid insertion trajectories in a computationally tractable manner, we performed MD simulations using the coarse-grained MARTINI force field.[167] All coarse-grained simulations used dilauroylphosphatidylcholine (DLPC) to compare to all-atom simulations of dimyristoylphosphatidylcholine (DMPC). Because MARTINI maps roughly four heavy atoms to a single coarse-grained bead, DMPC, which has 14 carbon atoms per tail, is best represented by the MARTINI model for DLPC, which has 3 beads per tail.

First, a bilayer surrounded by 3 nm thick slabs of standard MARTINI water was built using INSANE.[168] Prior to the addition of a tagged lipid into the solvent surrounding a bilayer of MARTINI DLPC lipids, the bilayer's structure was fully relaxed from its initial lattice configuration. This involved an energy minimization using the steepest descent algorithm followed by equilibration and production runs. During the first 500 ps equilibration run, a 10 fs time step and the Berendsen barostat[169] with a coupling time constant of 3 ps and an isothermal compressibility of 3×10^{-4} bar^{-1} were used. During a second 1 ns equilibration run, the time step was increased to 30 fs and the barostat was switched to the Parinello-Rahman algorithm[170] with a coupling time constant of 12 ps. A 50 ns production run using the same simulation parameters as the second equilibration run was performed to allow the bilayer's structure to fully equilibrate, as monitored by the area per lipid (Figure A.1). The temperature was maintained at 320 K with the V-rescale thermostat[171] using a coupling time constant of 1 ps. The lipids and solvent were coupled to separate thermostats to avoid the "hot solvent-cold solute" problem.[172] Dynamics were evolved according to the leapfrog algorithm.[173] As determined to yield optimal performance for simulations using MARTINI,[174] neighbor lists were updated using the Verlet neighbor searching algorithm,[175] Lennard-Jones and Coulomb interactions were truncated at 1.1 nm, and Coulomb interactions beyond the cutoff were evaluated with a reaction-field potential[176] with a relative dielectric constant of ∞.

Next, 1,000 replicate systems with a free lipid were built by inserting a tagged lipid at a random location in the solvent around the equilibrated bilayer such that the tagged lipid's COM and bilayer's COM are separated in z by at least 3.2 nm. All replicates were then energy minimized and equilibrated using the protocol described above with one modification: To ensure that the tagged lipid did not adsorb or insert into the bilayer during equilibration,

the z coordinates of its heavy atoms were restrained by a harmonic potential with a force constant of 500 kJ/mol/nm^2. Upon release of the position restraints, production runs of 1 μs were performed, and lipid insertion occurred in all replicates during this time. Reported times for MARTINI simulations are not scaled by a factor of 4, as done in other work to account for the erroneously fast solvent diffusion in MARTINI as compared to real water.[167]

All-Atom Systems. Additionally, we harvested 10 all-atom lipid insertion trajectories to compare with our results using MARTINI. The CHARMM36 force field[177] was used in combination with the CHARMM TIP3P water model[178] since it accurately reproduces many experimental observables, including the volume and area per lipid, bilayer thickness, lipid lateral diffusion coefficient, and neutron density profiles, for a liquid crystalline DMPC bilayer.[179,180] We used a simulation protocol similar to that described above for MARTINI.

First, a bilayer surrounded by 3 nm thick slabs of solvent was built using the CHARMM-GUI Membrane Builder.[181,182] The bilayer was energy minimized prior to undergoing a two-stage equilibration at 320 K and 1 bar. The first 250 ps equilibration utilized the Berendsen barostat[169] for semi-isotropic pressure coupling with a coupling time constant of 2 ps and isothermal compressibility of 4.5×10^{-5} bar^{-1}, and the second 250 ps equilibration utilized the Parinello-Rahman barostat[170] with a coupling time constant of 5 ps. A 50 ns production run was performed to allow the bilayer to fully equilibrate (Figure A.1). The lipids and solvent were coupled to separate Nosé-Hoover thermostats[183,184] using a coupling time constant of 1 ps to maintain the temperature. Dynamics were evolved according to the leapfrog algorithm[173] using a 2 fs time step. All bonds to hydrogen were constrained using the LINCS algorithm.[185] Lennard-Jones forces were smoothly switched off between 0.8 and 1.2 nm. Coulomb interactions were truncated at 1.2 nm, and long-ranged Coulomb interactions were calculated using Particle Mesh Ewald (PME) summation[186] with a Fourier spacing of 0.12 nm and an interpolation order of 4. Neighbor lists were constructed with the Verlet algorithm.[175]

Next, 10 replicate systems with a tagged lipid in solution were built as for the MARTINI systems. Each replicate was energy minimized and equilibrated using the protocol described for the CHARMM36 bilayer system with the addition of harmonic restraints on the z coordinates of all heavy atoms of the tagged lipid. After the position restraints were removed, each replicate was simulated in increments of 100 ns until the tagged lipid inserted into the bilayer.

Characterization of Transition Paths

From the harvested lipid insertion trajectories, we identified transition paths, trajectory segments that connect "reactant" and "product" states A and B. In state A, the tagged lipid is fully surrounded by solvent; in state B, the tagged lipid is fully within the bilayer. States A and B are characterized by the displacements d_{lip}, d_{sn1}, and d_{sn2} shown in Figure 2.2A. d_{lip} is the displacement in z from the COM of the tagged lipid to the COM of the closest leaflet. Similarly, d_{sn1} is the displacement in z from the terminal carbon of the sn1 tail of

the tagged lipid to the COM of the closest leaflet, and d_{sn2} is the analogous distance for the sn2 tail. Based on distributions of these three distances obtained from coarse-grained and

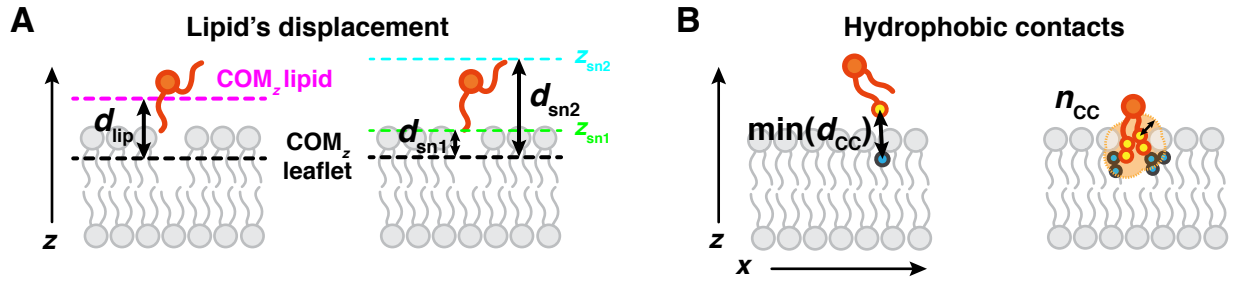


Figure 2.2: Schematics of order parameters that measure (A) the displacement of the lipid along the bilayer's normal and (B) hydrophobic lipid–bilayer contacts. The tagged lipid is colored orange. (A) d_{lip} is the displacement in z between the center-of-mass (COM_z) of the tagged lipid (magenta dashed line) and COM_z of the closest leaflet (black dashed line). d_{sn1} is the displacement in z between the terminal carbon of the sn1 tail (green dashed line) and COM_z of the closest leaflet. d_{sn2} is the displacement in z between the terminal carbon of the sn2 tail (cyan dashed line) and COM_z of the closest leaflet. (B) $\min(d_{\mathrm{CC}})$ is the minimum distance between a hydrophobic carbon of the tagged lipid and a hydrophobic carbon of the closest leaflet. The closest pair of hydrophobic carbons are drawn as circles. n_{CC} is the total number of close hydrophobic carbon contacts between the tagged lipid and closest leaflet. Any pair of lipid and membrane hydrophobic carbons within a cutoff distance of 14 Å for MARTINI and 10 Å for CHARMM36 are counted as contacts and drawn as circles. The light orange region highlights the space within a cutoff distance (black arrow) from hydrophobic carbons of the tagged lipid.

all-atom MD simulations (Figure A.2), state A is defined by $d_{\mathrm{lip}} > 24$ Å, which ensures that any configurations with the tagged lipid adsorbed onto the surface of the bilayer are not included. State B is defined by $d_{\mathrm{sn1}} < -3$ Å and $d_{\mathrm{sn2}} < -3$ Å, which ensures that the both tails of the tagged lipid are inserted into the bilayer.

In addition to d_{lip}, d_{sn1}, and d_{sn2}, we evaluated over 50 order parameters as putative reaction coordinates for lipid exchange. Hydrophobic contacts between the tagged lipid and bilayer are judged according to the distance $d_{\mathrm{CC}}^{(i)}$ between a hydrophobic carbon of the tagged lipid and a hydrophobic carbon of the closest membrane leaflet, where i indexes the many pairs of such atoms. For MARTINI lipids, hydrophobic carbons include all tail beads. For CHARMM36 lipids, hydrophobic carbons include atoms C23 – C214 and C33 – C314. More specifically, hydrophobic contacts are measured by the order parameters $\min(d_{\mathrm{CC}})$ and n_{CC} (Figure 2.2B). $\min(d_{\mathrm{CC}}) = \min_i d_{\mathrm{CC}}^{(i)}$ is the minimum distance between a hydrophobic carbon of the tagged lipid and a hydrophobic carbon of the closest leaflet. n_{CC} is the number of close contacts between hydrophobic carbons of the tagged lipid and hydrophobic carbons of the closest leaflet. The ith pair of hydrophobic carbons was counted as a close contact if $d_{\mathrm{CC}}^{(i)} \leq 14$ Å for MARTINI lipids and if $d_{\mathrm{CC}}^{(i)} \leq 10$ Å for CHARMM36 lipids. These cutoff

values encompass approximately two water solvation shells around a hydrophobic carbon of a lipid in solution and also two carbon solvation shells around a hydrophobic carbon of a lipid in a bilayer. We found smaller cutoffs to be insufficient for fully characterizing the tagged lipid's hydrophobic environment and, thus, also insufficient for constructing the reaction coordinate. Our definition of n_{CC} is similar to the order parameter developed by Lin and Grossfield to measure hydrophobic contacts between lipopeptides and phospholipids. Similar to the conclusions we make herein about phospholipid exchange, they found that a hydrophobic contact order parameter was key to accurately investigate lipopeptide insertion from a micelle into a phospholipid bilayer, whereas the COM displacement was insufficient.[187] Complete descriptions of all other order parameters are provided in the SI. The MDAnalysis Python library[188] was used to analyze all trajectories for each order parameter.

Free Energy Calculations

Based on the dynamics observed along transition paths, lipid insertion is a barrier-crossing process. To identify the physical origin of this barrier, we calculated free energy surfaces as a function of different order parameters. To obtain these free energy surfaces, we performed umbrella sampling simulations[189] using the PLUMED 2 patch[190] for GROMACS.

Coarse-Grained Systems. We computed the 2D free energy surfaces $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ and $\Delta F(d_{\mathrm{sn1}}, d_{\mathrm{sn2}})$ for the MARTINI system. To obtain $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$, we simulated 208 windows with harmonic biases centered at physically possible values of ($\min(d_{\mathrm{CC}})$, n_{CC}) for 2 μs each. Each window was initialized with a configuration, drawn from a transition path, that has a value of ($\min(d_{\mathrm{CC}})$, n_{CC}) close to the center of the window's bias. To calculate biasing forces on $\min(d_{\mathrm{CC}})$, a smooth form for the minimum function

$$\gamma \left[\log \left(\sum_i \exp(\gamma / d_{\mathrm{CC}}^{(i)}) \right) \right]^{-1}, \tag{2.1}$$

was used with $\gamma = 260$ nm. $\min(d_{\mathrm{CC}})$ calculated with Eq. 2.1 differed from the exact value by 0.007 Å on average and by at most 0.34 Å. To calculate biasing forces on n_{CC}, a switching function was used

$$\sum_i \frac{1 - \left(\frac{d_{\mathrm{CC}}^{(i)} - d_0}{r_0} \right)^6}{1 - \left(\frac{d_{\mathrm{CC}}^{(i)} - d_0}{r_0} \right)^{12}}, \tag{2.2}$$

with $d_0 = 14$ Å and $r_0 = 0.25$ Å. n_{CC} calculated with Eq. 2.2 differed from the exact value by 4 contacts on average and by at most 28 contacts.

To obtain $\Delta F(d_{\mathrm{sn1}}, d_{\mathrm{sn2}})$, we simulated 370 windows with harmonic biases centered at positions $(d_{\mathrm{sn1}}, d_{\mathrm{sn2}})$ ranging from (−10 Å, −10 Å) to (30 Å, 30 Å) for 2 μs each. To avoid unrealistically distorting the tagged lipid, we simulated only windows whose harmonic bias

centers satisfy $|d_{\mathrm{sn1}} - d_{\mathrm{sn2}}| \leq 30$ Å. Each window was initialized with a configuration, drawn from a transition path, that has a value of $(d_{\mathrm{sn1}}, d_{\mathrm{sn2}})$ close to the center of the window's bias. A force constant of 500 kJ/mol/nm^2 was used for all harmonic bias potentials.

The weighted histogram analysis method (WHAM)[191] was used to obtain both of these 2D free energy surfaces from the biased distributions, after discarding data from the first 1 μs. Error bars were calculated as the standard error of free energy surfaces estimated from five independent 200 ns blocks.

All-Atom Systems. To obtain $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ for the CHARMM36 system, 202 windows with harmonic biases centered at physically possible values of ($\min(d_{\mathrm{CC}})$, n_{CC}) were simulated for 24 ns each. Eq. 2.1 was used to calculate biasing forces on $\min(d_{\mathrm{CC}})$ with $\gamma = 200$ nm, and it differed from the exact value by 0.002 Å on average and by at most 0.27 Å. Eq. 2.2 was used to calculate biasing forces on n_{CC} with $d_0 = 10$ Å and $r_0 = 0.25$ Å, and it differed from the exact value by 68 contacts on average and by at most 175 contacts. Each window was initialized with a configuration, drawn from a transition path, that has a value of ($\min(d_{\mathrm{CC}})$, n_{CC}) close to the center of the window's bias. The first 4 ns of data from these windows was discarded to account for equilibration. Finally, data from all windows was combined with WHAM to obtain a free energy surface as a function of $\min(d_{\mathrm{CC}})$ and n_{CC}. Error bars were calculated as the standard error of free energy surfaces estimated from five independent 5 ns blocks.

Calculation of 1D Free Energy Profiles from 2D Free Energy Surfaces. We calculated 1D free energy profiles $\Delta F(\min(d_{\mathrm{CC}}))$ and $\Delta F(n_{\mathrm{CC}})$ by numerically integrating $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ over one of its variables. Denoting the two variables as q and q' (in either order), the free energy profile $\Delta F(q)$ is obtained from the free energy surface $\Delta F(q, q')$ according to

$$\Delta F(q) = -\beta^{-1} \ln \left[\int dq' \; e^{-\beta \Delta F(q,q')} \right], \tag{2.3}$$

where $\beta = (k_{\mathrm{B}}T)^{-1}$ is the inverse of Boltzmann's constant, k_{B}, multiplied by temperature.

Committor Analysis

To identify the reaction coordinate for passive lipid exchange and determine if transition states are sampled in umbrella sampling simulations, we characterized configurations according to their tendency to proceed to state B using committor analysis.[136–138] The committor, p_{B}, is the probability that a configuration will reach state B prior to state A when its momenta are chosen randomly from a Maxwell-Boltzmann distribution. By construction, the committor distinguishes transition states, which have $p_{\mathrm{B}} = 0.5$, from stable state A and B configurations, which have $p_{\mathrm{B}} = 0$ and 1, respectively. Thus, p_{B} is the true reaction coordinate. Through committor analysis of configurations found along transition paths, we identified order parameters that are strongly correlated with the committor and, thus, can

be used as approximate reaction coordinates. These order parameters have the advantage of being more physically descriptive and, thus, more easily interpreted than the committor. Henceforth we refer to order parameters that correlate strongly with p_{B} as the reaction coordinate. We calculated committor values for 98,094 MARTINI and 138 CHARMM36 configurations sampled along transition paths. Additionally, we calculated the committor for 500 MARTINI and 100 CHARMM36 configurations from each umbrella sampling simulation. For each MARTINI configuration, the outcome of 50 trajectories, each 3 ns long and initialized with random velocities sampled from a Maxwell-Boltzmann distribution, were used to calculate its committor value. For each CHARMM36 configuration, the outcome of 20 trajectories, each 12 ns long, were used to calculate its committor value.

Rate Calculations

To further assess how well the dynamics of lipid transport are captured by monitoring lipid–membrane hydrophobic contacts, we calculated the rate constant for lipid insertion, k_{ins}, using Kramers theory[192] coupled with thermodynamic information about hydrophobic contacts. The resulting value was compared with the insertion rate calculated from the mean first passage time in MD simulations. Based on Kramers theory,

$$k_{\mathrm{ins}} = D \left[\int_{x_B}^{L} dx \; e^{-\beta U(x)} \int_{x_B}^{x} dx' \; e^{\beta U(x')} \right]^{-1}, \tag{2.4}$$

where D is the diffusion coefficient along a reaction coordinate x for lipid insertion, x_B is the value of x in state B, L is the maximal value of x in state A sampled in our simulations, and $U(x)$ is the effective interaction potential biasing the dynamics of x. We have taken x to be $\min(d_{\mathrm{CC}})$ since $\min(d_{\mathrm{CC}})$ describes the spatial motion of the lipid during insertion in addition to hydrophobic contact formation. For the same reason we set D to be the diffusion coefficient of a freely diffusing lipid in solution. Note that $\min(d_{\mathrm{CC}})$ alone is not the reaction coordinate for lipid exchange (a linear combination of $\min(d_{\mathrm{CC}})$ and n_{CC} is the reaction coordinate), but it is more simply interpretable to utilize a single order parameter for these calculations. $U(x)$ is taken to be the free energy profile $\Delta F(\min(d_{\mathrm{CC}}))$. $U(x)$ was obtained from the free energy surface that depends jointly on $\min(d_{\mathrm{CC}})$ and n_{CC} using Eq. 2.3. We evaluated Eq. 2.4 with numerical integration using the trapezoidal rule. The diffusion coefficient was calculated from an additional MD simulation of a single lipid solvated in a cubic box with the same area and simulation protocol as the bilayer systems, but with isotropic pressure coupling. D was calculated according to Einstein's relation from the mean squared displacement of the lipid's COM obtained from 1 μs and 100 ns trajectories of a MARTINI and CHARMM36 lipid, respectively.

Results and Discussion

Lipid Insertion Is a Barrier Crossing Process that Occurs *via* Multiple Pathways

We first investigated the dynamics of lipid insertion by harvesting 1,000 MARTINI DLPC and 10 CHARMM36 DMPC insertion trajectories from MD simulations. A lipid insertion event is classified by a transition from state A, in which the tagged lipid is fully in the solvent, to state B, in which the tagged lipid resides within the bilayer. State A is distinguished by d_{lip}, which measures the COM displacement of the tagged lipid along the bilayer's normal; state B is distinguished by d_{sn1} and d_{sn2}, which measure the displacement of each tail of the tagged lipid along the bilayer's normal (Figure 2.2A). Precise definitions are given above in the **Methods** section. The fact that state B configurations cannot be reliably identified by d_{lip} alone suggests that the COM displacement is not the reaction coordinate for passive lipid exchange.

Free energy profiles as a function of the lipid's displacement along the bilayer normal obtained from previous computational studies[156–160] give the impression that lipid insertion is a barrier-less process. If that were true, insertion should occur immediately once the tagged lipid reaches the bilayer. However, as seen in snapshots from a MARTINI and a CHARMM36 trajectory shown in Figure 2.3, the tagged lipid repeatedly arrives at the bilayer and adheres to its surface without inserting. Instead, it detaches from the bilayer's surface and returns to the solvent. In typical trajectories, many such unproductive encounters occur before the lipid inserts into the bilayer. Indeed, as seen in time traces of d_{lip} in Figure 2.4A, many adsorption events commonly precede insertion. Similar adsorption events have been observed during simulations of 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) insertion into a bilayer.[165]

When lipid insertion eventually occurs in these trajectories, it does so suddenly. Characteristic of barrier-crossing dynamics, d_{lip}, d_{sn1}, and d_{sn2} change sharply from values of state A to those of state B (Figure 2.4A). The fact that the transition times are much faster than the inverse rate constant for lipid insertion, $1/k_{\mathrm{ins}}$, points to a substantial free energy barrier for insertion. Free energy profiles as a function of the lipid's displacement simply do not resolve this barrier.[156–160] Thus, the barrier must exist along a different degree of freedom that captures other important features of the dynamics.

In fact, lipid insertion occurs *via* three different pathways which cannot be differentiated by d_{lip}. Each pathway is characterized by a distinct lipid configuration, which is distinguished by d_{sn1} and d_{sn2}, near the bilayer's surface: (1) In the sliding pathway, the two tails enter the bilayer almost simultaneously as the tagged lipid slides into the bilayer (Figure 2.3, CHARMM36 trajectory and Figure 2.4A, magenta time traces). Near the bilayer's surface, both tails are a similar distance from the bilayer (Figure 2.4B, magenta region). (2) In the sn1 splayed pathway, the sn1 tail enters the bilayer first (Figure 2.4A, green time traces), creating a splayed intermediate with the sn1 tail anchored in the bilayer (Figure 2.4B, green region). (3) In the sn2 splayed pathway, the sn2 tail enters the bilayer first (Figure 2.3,

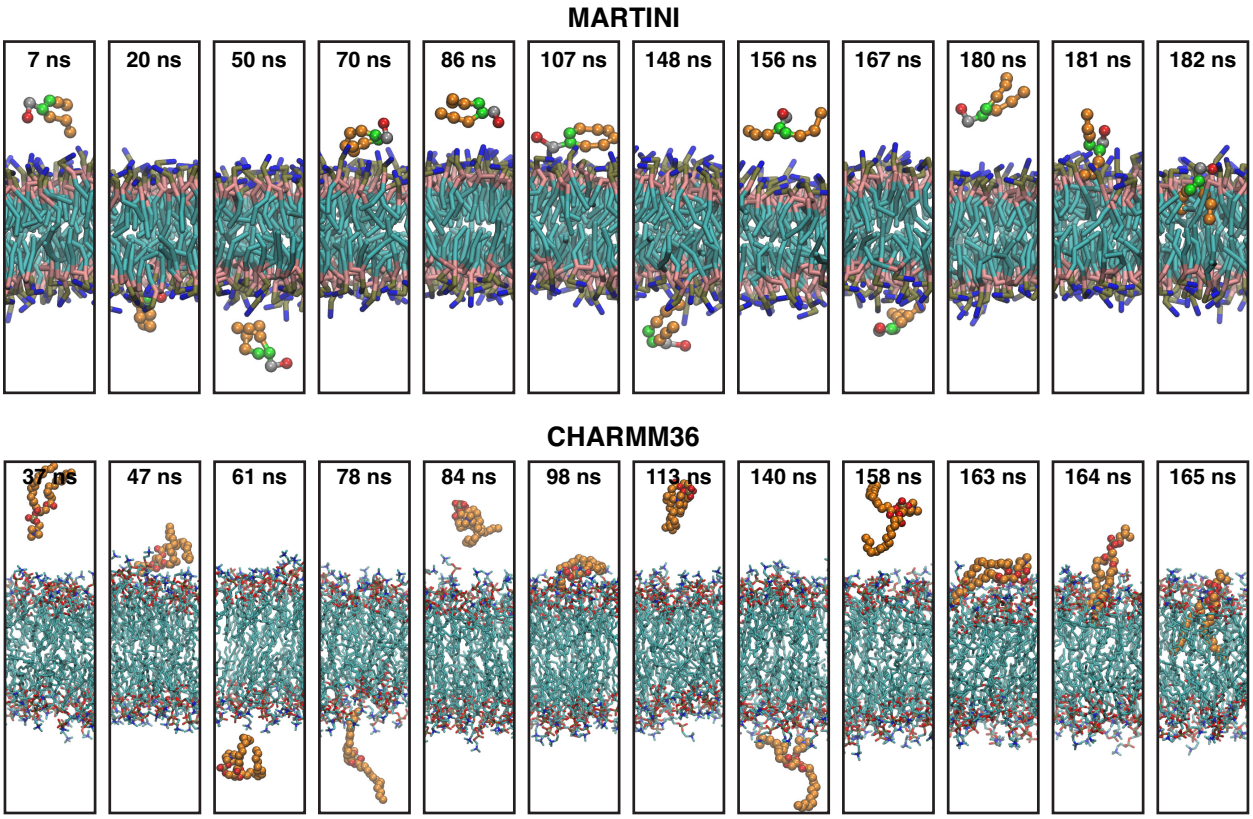


Figure 2.3: Snapshots along MD trajectories of a tagged lipid inserting into a bilayer illustrate that a membrane-adsorbed lipid does not immediately insert into a bilayer, suggesting that there is a barrier for insertion. For clarity, solvent is not shown in the snapshots. The tagged lipid is rendered with van der Waals spheres. For the MARTINI simulation, the headgroup, phosphate, glycerol, and tail beads of the tagged lipid (bilayer lipids) are colored red (blue), gray (brown), green (pink), and orange (cyan), respectively. For the CHARMM36 simulation, the carbon atoms of the tagged lipid (bilayer lipids) are colored orange (cyan).

MARTINI trajectory and Figure 2.4A, cyan time traces), creating a splayed intermediate with the sn2 tail anchored in the bilayer (Figure 2.4B, blue region). Table A.1 reports the frequency of each pathway in our simulations. Splayed intermediates have also been observed in a previous study of lipid insertion [165] and postulated as transition states for stalk formation during membrane fusion, [122,193–197] a key step in vesicular transport. The existence of distinct insertion pathways might suggest that the displacements of individual tails along the bilayer's normal could serve as reaction coordinates since they encapsulate dynamically relevant information that is not contained in the COM displacement. We demonstrate below, however, that the displacements of individual tails are not the reaction coordinates for lipid exchange.

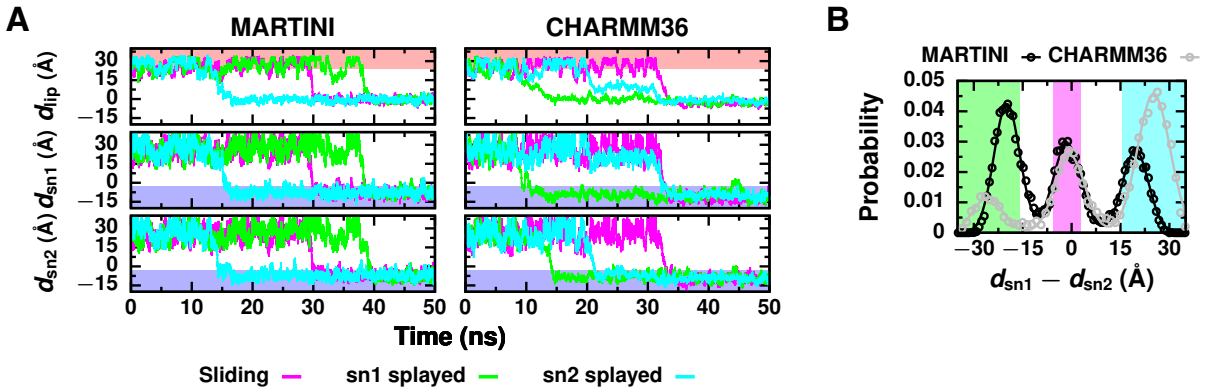


Figure 2.4: Time evolution of d_{lip}, d_{sn1}, and d_{sn2} during MD simulations of lipid insertion is indicative of barrier crossing dynamics. (A) State A configurations ($d_{lip} > 24$ Å) are located in the red region. State B configurations ($d_{sn1} < -3$ Å and $d_{sn2} < -3$ Å) are located in the blue region. Dips in the distances d_{lip}, d_{sn1}, and d_{sn2} out of the state A region without entering the state B region occur when the tagged lipid adsorbs to the surface of the bilayer. Examples for the sliding, sn1 splayed, and sn2 splayed pathways are shown for both MARTINI and CHARMM36. The initial times have been shifted so that all trajectories can be plotted together. (B) Probability distributions of $d_{sn1} - d_{sn2}$ when the tagged lipid is near the surface of the bilayer (specifically, MARTINI configurations with $7 \leq d_{lip} \leq 9$ Å and CHARMM36 configurations with $2.5 \leq d_{lip} \leq 15$ Å), which were used to identify splayed tail configurations. Distributions from MD simulations are plotted with open circles, and the fits to a sum of three Gaussians are plotted with solid lines. Transition paths that follow the sn1 splayed pathway sample configurations found in the green region, and those that follow the sn2 splayed pathway sample configurations found in the cyan region. Transition paths that follow the sliding pathway sample configurations in the magenta region and do not sample splayed configurations in either the cyan or green regions.

Reaction Coordinate Characterizes Hydrophobic Contacts Between the Lipid and Membrane

Based on experiments, a cavity model has been proposed to describe the transition state.[99,164] According to the cavity model, solvent is evacuated above the desorbing lipid and a void forms in the membrane below. Such a focus on cavities is reminiscent of modern theories of the hydrophobic effect, which characterize hydrophobicity in terms of the statistics of solvent density fluctuations.[198] Although no true cavities are observed at transition states sampled in our simulations (Figure A.3), the importance of hydrophobicity in lipid transport is evident from our analysis of over 50 order parameters. For example, the number of water molecules solvating the tagged lipid steadily decreases during insertion (Figure A.4). The density of hydrophobic molecular fragments below the tagged lipid gradually increases while the number of hydrophilic ones decreases (Figure A.3). Defects in the polar head group region of the bilayer that expose hydrophobic membrane patches[60,61] are also observed in

the transition state ensemble (Figures A.5 and A.6). Based on these results, we hypothesized that the reaction coordinate for passive lipid exchange monitors the formation and breakage of hydrophobic contacts between the tagged lipid and membrane (Figure 2.1B).

To rigorously test this hypothesis, we employed committor analysis.[136–138] The committor, p_B, is the probability that a trajectory initiated at a given configuration will reach state B prior to state A when initial momenta are chosen randomly from a Maxwell-Boltzmann distribution. Configurations within stable states A and B have $p_B = 0$ and 1, respectively. Transition states are equally likely to advance to states B and A, such that $p_B = 0.5$.[136–138,142] Since p_B directly measures the progress of a reaction, p_B is the true reaction coordinate.[136,137,199] However, p_B is a complicated function of the system's microscopic configuration — a very large set of variables in the case of biomolecular systems. The complete functional form for p_B is practically unobtainable and would provide little physical insight into reaction mechanisms. Instead, it is more informative to identify an order parameter, q, that is strongly correlated with p_B and closely approximates the true reaction coordinate.[141,199] Henceforth we refer to such order parameters as the reaction coordinate.

As one assessment of correlation between an order parameter q and the true reaction coordinate p_B, we compare the probability distribution of q from transition states to distributions of q from states A and B. The typical values of q at the transition state should differ from its values in states A and B such that these probability distributions do not overlap. Otherwise, q cannot reliably distinguish transition states from stable reactant and product states and, thus, poorly recapitulates the true reaction coordinate. Additionally, we compare the probability distribution of q from transition states to distributions from pre- and post-transition states. Because pre-transition states are intermediates between state A and the transition state and post-transition states are intermediates between the transition state and state B, it is often challenging to devise an order parameter q that distinguishes transition states from them, as p_B naturally does. A good approximation to the true reaction coordinate should reliably distinguish transition states from pre- and post-transition states in addition to state A and B configurations.

A more stringent test of a putative reaction coordinate examines a histogram of p_B values for configurations with a particular value of q. If q is the reaction coordinate, then a histogram of p_B for configurations with a given value of q will be peaked at the corresponding value of p_B. This histogram test is useful to determine if the mapping from q to p_B is approximately one-to-one, a requirement for q to be the reaction coordinate. Bimodal histograms of p_B are clear indicators that q is not the reaction coordinate. By contrast, a histogram sharply peaked at $p_B = 0.5$ for values of q characteristic of transition states clearly indicates that q accurately describes the true reaction coordinate.[136,137]

Using the first criterion that q must reliably distinguish transition states from all other configurations to be the reaction coordinate, we assessed measures of hydrophobic lipid–membrane contacts as approximations to the true reaction coordinate. Specifically, we characterize hydrophobic contacts by the minimum distance between hydrophobic carbons of the tagged lipid and hydrophobic carbons of the bilayer, $\min(d_{CC})$, and the number of close hydrophobic carbon–carbon contacts between the tagged lipid and bilayer, n_{CC} (Figure 2.2B).

Precise definitions are given above in the **Methods** section. To determine if this criterion was satisfied, we compared the probability distributions of $\min(d_{CC})$ and n_{CC} from several ensembles: equilibrium configurations representative of (1) state A and (2) state B (as defined above in **Methods**); and three ensembles drawn from transition paths of (3) pre-transition state configurations identified by $p_B = 0$, (4) transition states identified by $p_B \approx 0.5$ (specifically, $0.45 \leq p_B \leq 0.55$ for MARTINI and $0.4 \leq p_B \leq 0.6$ for CHARMM36 configurations), and (5) post-transition state configurations identified by $p_B = 1$. Joint distributions of $\min(d_{CC})$ and n_{CC} in these five different ensembles are shown in Figure 2.5. Corresponding 1D probability distributions of $\min(d_{CC})$ and n_{CC} are shown in Figures A.7 and A.8. The distributions from MARTINI and CHARMM36 configurations exhibit similar

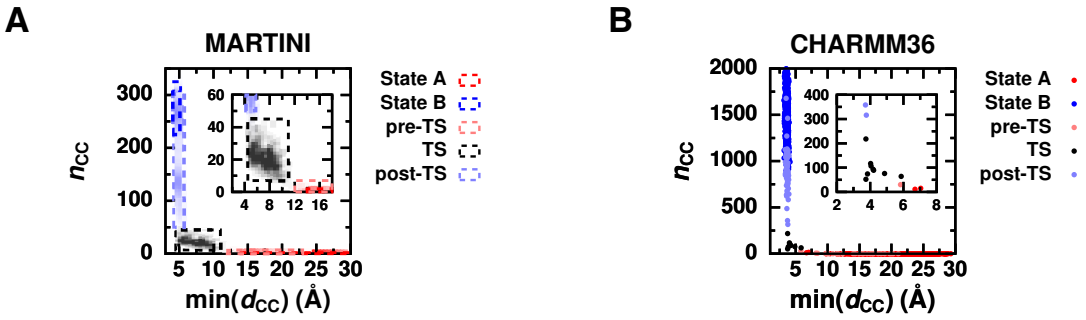


Figure 2.5: Distributions of $\min(d_{CC})$ and n_{CC} demonstrate that hydrophobic lipid–membrane contacts reliably identify transition states. (A) 2D probability distribution of $\min(d_{CC})$ and n_{CC} from MARTINI simulations. Individual 2D distributions for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations, transition states (TS), and post-transition state (post-TS) configurations are each plotted in a single color and outlined in a dashed rectangle. (B) Scatter plot of $\min(d_{CC})$ and n_{CC} from CHARMM36 simulations. Magnified views around transition states are shown in the insets.

features. In state A, $\min(d_{CC})$ effectively measures the separation between the tagged lipid and the distant bilayer. Large values of $\min(d_{CC})$ are therefore typical. No close hydrophobic contacts are formed since the tagged lipid is fully solvated. As the tagged lipid progresses from state A towards the transition state, it becomes a pre-transistion state configuration. Pre-transition state configurations are closer to the bilayer than state A configurations, resulting in decreased values of $\min(d_{CC})$, but hydrophobic contacts between the tagged lipid and bilayer still scarcely exist. At the transition state, the tagged lipid has made only a few initial hydrophobic contacts with the bilayer. The transition state distribution is centered at values of $\min(d_{CC})$ and n_{CC} intermediate between those of states A and B. As the tagged lipid progress from the transition state to state B, it becomes a post-transition state configuration. Post-transition state configurations, which include splayed lipid configurations, have a substantial number of hydrophobic contacts between the tagged lipid and bilayer compared to transition states but on average half as many as state B configurations. Both

the post-transition state and state B distributions of $\min(d_{\mathrm{CC}})$ are sharply peaked at a value consistent with the minimum of the Lennard-Jones potential for a carbon–carbon interaction (or a bead–bead interaction in the MARTINI model). In state B, a maximal number of hydrophobic contacts exist between the tagged lipid and adjacent lipids in the bilayer. Importantly, the transition state distribution overlaps negligibly with the other four distributions. A combination of $\min(d_{\mathrm{CC}})$ and n_{CC} distinguishes transition states from not only stable state A and B configurations but also pre- and post-transition state configurations, and, therefore, may serve as the reaction coordinate for lipid exchange.

If the true reaction coordinate is well described by a combination of $\min(d_{\mathrm{CC}})$ and n_{CC} — in other words if the formation and breakage of hydrophobic contacts are the essential processes required for lipid exchange — the free energy surface $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ should exhibit a barrier for insertion and desorption. These free energy surfaces for both MARTINI and CHARMM36 are shown in Figure 2.6A. Corresponding 1D free energy profiles along $\min(d_{\mathrm{CC}})$ and n_{CC} are shown in Figure A.10. A deep free energy minimum exists at values of $\min(d_{\mathrm{CC}})$ and n_{CC} characteristic of state B due to the formation of many favorable hydrophobic contacts between the tagged lipid and membrane lipids. At larger values of $\min(d_{\mathrm{CC}})$ characteristic of state A, the free energy surface flattens; when the lipid is far away from the membrane, the free energy is no longer sensitive to $\min(d_{\mathrm{CC}})$. The surface plateaus at larger free energy values for MARTINI compared to CHARMM36. This discrepancy is consistent with the fact that compared to atomistic models, the MARTINI model overestimates the free energy difference between a lipid in solution and in the bilayer.[167] In qualitative agreement with experimental findings,[97–99,161] there is a free energy barrier for lipid insertion (Figure 2.6A, outlined in dashes). The activation free energy for the formation of hydrophobic contacts is roughly 5 $k_{\mathrm{B}}T$. Additionally, transition paths closely follow the minimum free energy path in the space of $\min(d_{\mathrm{CC}})$ and n_{CC} (Figure A.9), indicating that the dynamics of lipid exchange are well described in terms of hydrophobic contacts.

Most importantly, a combination of $\min(d_{\mathrm{CC}})$ and n_{CC} can locate transition states with high fidelity. We find that a linear combination suffices for this purpose ($r_{\mathrm{LxS}} = \alpha_1 \min(d_{\mathrm{CC}}) + \alpha_2 n_{\mathrm{CC}} + \alpha_0$ with coefficients α_i determined using a maximum likelihood approach[140,200] as detailed in Appendix A). By construction, $r_{\mathrm{LxS}} = 0$ at the dividing surface between states A and B where transition states are located. The dashed lines in Figure 2.6A outline a narrow range around $r_{\mathrm{LxS}} = 0$, which roughly traces the ridgeline of $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ between states A and B. To definitively test if r_{LxS} is the reaction coordinate reaction coordinate for lipid (L) transport via solvent (xS), we performed committor analysis of configurations drawn from a narrow range around $r_{\mathrm{LxS}} = 0$. Figure 2.6B shows that the ensemble defined by $r_{\mathrm{LxS}} \approx 0$ predominantly includes transition states for both MARTINI and CHARMM36. Thus, a measure of hydrophobic contacts between the tagged lipid and bilayer is the reaction coordinate for lipid exchange.

Together, $\min(d_{\mathrm{CC}})$ and n_{CC} capture many different aspects of the lipid's hydrophobic environment, underlying the ability of r_{LxS} to precisely describe the process of lipid exchange. Hydrophobicity can be quantified in other ways as well. For example, we constructed a more complicated reaction coordinate, $r_c^{(M=48)}$, that is a linear combination of 48 order parameters

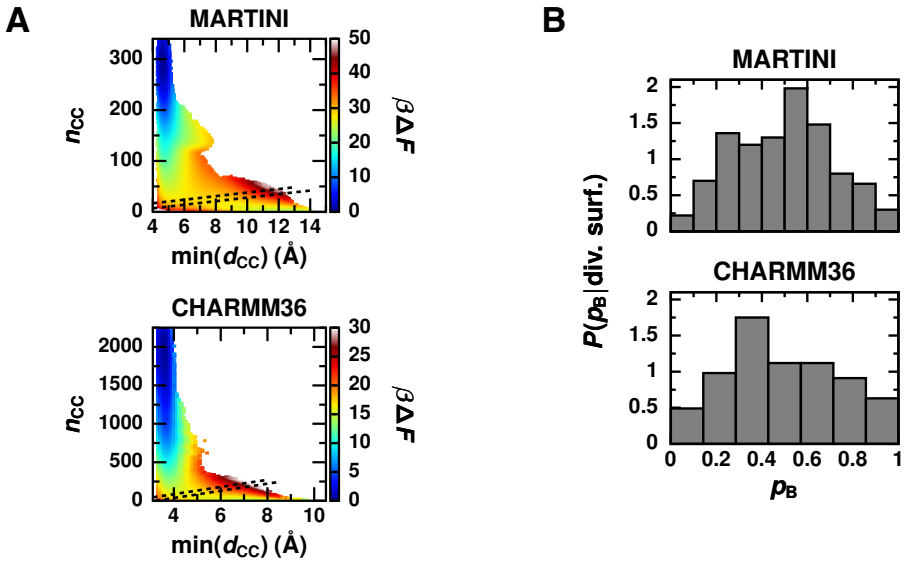


Figure 2.6: A free energy barrier for hydrophobic lipid–membrane contact breakage and formation limits the rate of lipid desorption and insertion, respectively. (A) Free energy surfaces as a function of $\min(d_{CC})$ and n_{CC} for the MARTINI and CHARMM36 force fields exhibit a barrier for insertion, which separates states A (bottom right corner) and B (top left corner). In each free energy profile, the dashed lines outline a narrow range about the dividing surface between states A and B. Specifically, the dashed lines outline the region $-0.2 \leq r_{\mathrm{LxS}} \leq 0.2$ for MARTINI and $-0.12 \leq r_{\mathrm{LxS}} \leq 0.12$ for CHARMM36, where $r_{\mathrm{LxS}} = \alpha_1 \min(d_{CC}) + \alpha_2 n_{CC} + \alpha_0$ with coefficients α_i determined using a maximum likelihood approach[140,200] (Table A.3). Statistical errors in the free energy profiles are shown in Figure A.9. (B) Histogram of committor values, p_B, for configurations drawn from umbrella sampling simulations within a narrow range about the dividing surface, demonstrate that a linear combination of $\min(d_{CC})$ and n_{CC} is indeed the reaction coordinate.

excluding $\min(d_{CC})$ and n_{CC}, with coefficients again determined using a maximum likelihood approach[140,200] (Table A.2). As fully described in Appendix A, each of these 48 order parameters measure different details about the lipid's environment, including the number of water molecules solvating the tagged lipid and the size of exposed hydrophobic membrane defects near the tagged lipid. This more complex reaction coordinate identifies transition states almost as faithfully as r_{LxS} does (Figure A.11A), but $r_c^{(M=48)}$ is difficult to interpret physically. The reaction coordinates r_{LxS} and $r_c^{(M=48)}$ are highly correlated (Figure A.11B), demonstrating that detailed information about the lipid's environment is well represented by a simple combination of $\min(d_{CC})$ and n_{CC}. Not only is a linear combination of $\min(d_{CC})$ and n_{CC} the most accurate reaction coordinate out of all tested (Table A.3), but it has the advantage of providing physical insight into lipid exchange: The rate limiting step for

desorption is the breakage of hydrophobic contacts between the lipid and membrane.

Lipid's Displacement Normal to the Bilayer Is Not the Reaction Coordinate

Previous computational studies presumed that the lipid's displacement normal to the bilayer is the reaction coordinate for lipid desorption and insertion.[156–160] However, d_{lip} is not the reaction coordinate, a fact we confirm by performing committor analysis. The probability distributions of d_{lip} from transition states are compared to the distributions from state A, state B, pre-transition state, and post-transition state configurations in Figure 2.7. The

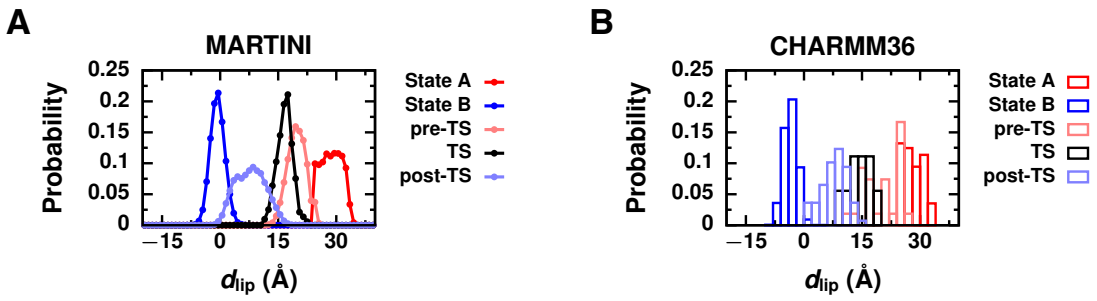


Figure 2.7: Probability distributions of d_{lip} from (A) MARTINI and (B) CHARMM36 simulations indicate that d_{lip} is not the reaction coordinate for lipid exchange. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations, transition states (TS), and post-transition state (post-TS) configurations.

distributions from MARTINI and CHARMM36 configurations are quite similar. In state A, where the tagged lipid is fully solvated and away from the bilayer, the typical value of d_{lip} is large. As the lipid enters the bilayer, it progresses from being in state A to being a pre-transition state configuration, a transition state, a post-transition state, and finally in state B; correspondingly, the centers of each distribution shift to smaller values of d_{lip}. In state B, the tagged lipid is in register with the other lipids in the membrane such that the distribution is centered near zero. For both MARTINI and CHARMM36, distributions from transition states overlap significantly with those from the other ensembles, indicating that transition states cannot be reliably identified by d_{lip} (Figure 2.1A). Furthermore, histograms of p_{B} for configurations sampled along transition paths with values of d_{lip} typical of transition states are bimodal and lack a peak at 0.5 (Figure A.12), indicating that d_{lip} is not strongly correlated with p_{B}. Therefore, d_{lip} is not the reaction coordinate for lipid exchange.

Based on the observation of multiple pathways for lipid insertion characterized by splayed intermediates (Figure 2.4), combinations of the tail displacements d_{sn1} and d_{sn2} might have been considered as pathway-specific reaction coordinates. For transitions *via* the sliding

pathway, $\frac{1}{2}d_{\mathrm{sn1}} + \frac{1}{2}d_{\mathrm{sn2}}$ is a plausible reaction coordinate since both tails enter the bilayer at approximately the same time. For transitions *via* the sn1 splayed and sn2 splayed pathways, d_{sn1} and d_{sn2} are plausible reaction coordinates for each pathway, respectively, since the lipid is committed to fully insert after the first tail has entered the bilayer. Probability distributions of these putative reaction coordinates from MARTINI transition states are compared to the distributions from state A, state B, pre-transition state, and post-transition state configurations in Figure 2.8 for each pathway individually. With only 10 CHARMM36

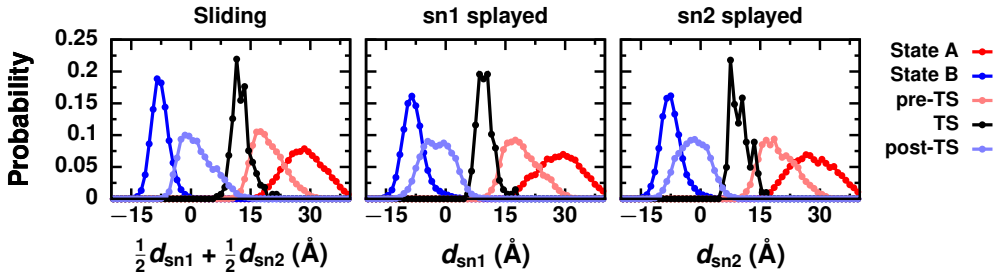


Figure 2.8: Probability distributions of $\frac{1}{2}d_{\mathrm{sn1}} + \frac{1}{2}d_{\mathrm{sn2}}$, d_{sn1}, and d_{sn2} from MARTINI simulations that follow a given lipid insertion pathway demonstrate that the tail displacements can reliably identify transition states. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations, transition states (TS), and post-transition state (post-TS) configurations.

insertion trajectories, there is insufficient data to reliably examine distributions for each pathway individually (data from all CHARMM36 transitions is shown in Figure A.13). With limited overlap between the transition state distribution and distributions from all other ensembles (Figure 2.8), $\frac{1}{2}d_{\mathrm{sn1}}+\frac{1}{2}d_{\mathrm{sn2}}$, d_{sn1}, and d_{sn2} appear to be potential reaction coordinates that are specific to each pathway. d_{sn1} cannot be used to accurately identify transition states along the sn2 splayed pathway and vice versa (Figure A.14). While this indicates that transition states have values of d_{sn1} and d_{sn2} distinct from configurations in the other ensembles, it does not guarantee that $\frac{1}{2}d_{\mathrm{sn1}} + \frac{1}{2}d_{\mathrm{sn2}}$, d_{sn1}, and d_{sn2} are strongly correlated with p_{B}, as required for them to be the reaction coordinates for each lipid exchange pathway.

Furthermore, if $\frac{1}{2}d_{\mathrm{sn1}}+\frac{1}{2}d_{\mathrm{sn2}}$ is the reaction coordinate for the sliding pathway, d_{sn1} for the sn1 splayed pathway, and d_{sn2} for the sn2 splayed pathway, then they should fully describe the dynamics of lipid exchange. In that case, the free energy surface $\Delta F(d_{\mathrm{sn1}}, d_{\mathrm{sn2}})$ should have a barrier for lipid insertion and desorption to be consistent with the observed barrier crossing dynamics (Figure 2.4A). This free energy surface is shown in Figure 2.9A for the MARTINI model. Between states A (top right corner) and B (bottom left corner), slight plateaus in the free energy profile occur at combinations of d_{sn1} and d_{sn2} characteristic of splayed configurations (top left and bottom right corners). The free energy profile has a saddle point near $d_{\mathrm{sn1}} \approx d_{\mathrm{sn2}} \approx 15$ Å with a barrier for insertion of approximately 1 $k_{\mathrm{B}}T$ (outlined in dashes), which is only slightly larger than the statistical error in the calculated

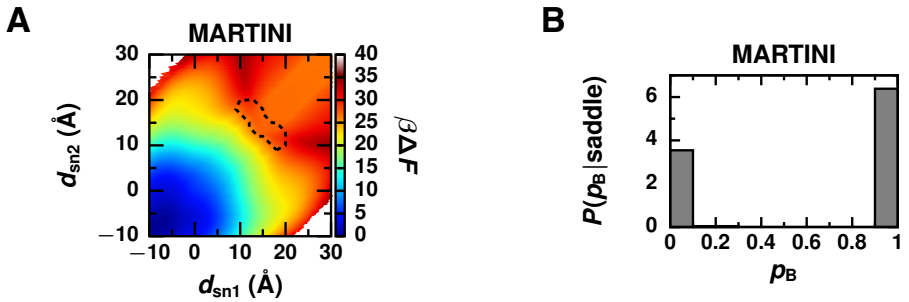


Figure 2.9: Lipid tail displacements from the membrane are not the reaction coordinate. (A) Free energy surface as a function of d_{sn1} and d_{sn2} obtained from umbrella sampling simulations using the MARTINI force field lacks a significant barrier for insertion. The dashed line outlines a saddle point separating states A (top right corner) and B (bottom left corner). Statistical error in the free energy profile is shown in Figure A.15. (B) Histogram of committor values, p_B, for configurations located at the saddle point during umbrella sampling simulations demonstrate that the reaction coordinate is not simply a function of d_{sn1} and d_{sn2}.

free energy (Figure A.15). This small barrier would hardly impede the dynamics of insertion. Indeed, it is significantly lower than the barrier for hydrophobic contact formation (Figure 2.6A), demonstrating that decreases in the tail displacments do not limit the rate of lipid insertion.

Finally, to determine if the tail displacements are strongly correlated with the true reaction coordinate, we constructed a histogram of p_B values for configurations found near the saddle point during umbrella sampling simulations. The histogram of p_B for configurations that are predicted to be transition states based on the free energy surface is strongly bimodal (Figure 2.9B). Almost all of the tested saddle point configurations are committed to state A or B, and transitions states are seldom sampled in our simulations. Thus, the lipid's displacement normal to the bilayer, even reformulated to account for the three different insertion pathways, is not the reaction coordinate for lipid exchange.

Calculation of Lipid Exchange Rate Enables Comparison to Experiment

Having identified the reaction coordinate, which measures hydrophobic contact formation and breakage between the lipid and membrane, we utilized this knowledge to calculate the kinetic parameters of lipid exchange. First, we calculated the lipid insertion rate, k_{ins}, using Kramers theory (Eq. 2.4) as detailed in the **Methods** section and compared it to the rate obtained directly from our unbiased trajectories. Kramers theory provides an expression for the reaction rate of a process whose dynamics are diffusive and well described by an overdamped Langevin equation.[192] During insertion, the lipid is buffeted by solvent molecules

and other membrane lipids while crossing the broad free energy barrier for hydrophobic contact formation (Figure 2.6A), resulting in diffusive barrier crossing dynamics (Figure A.9) and making Kramers theory appropriate. The value of k_{ins} calculated with Kramers theory is 7.0 μs^{-1} for MARTINI and 21.0 μs^{-1} CHARMM36; k_{ins} obtained from the mean first passage time in the spontaneous insertion simulations is 8.0 μs^{-1} for MARTINI and 5.7 μs^{-1} for CHARMM36. The insertion rates calculated based on hydrophobic contact formation differ from those obtained from dynamical simulations by a factor of 4 or less. Such good agreement demonstrates that hydrophobic contacts describe the dynamics of lipid insertion, an elementary step of lipid exchange, with quantitative accuracy. Unfortunately, the timescale of insertion is too fast to measure with straightforward experimental methods.

To compare our results to experiments, we sought to calculate instead the much slower lipid exchange rate, k_{ex}. Although we obtained k_{ins} directly from simulations, k_{ex} cannot be obtained in the same manner. Complete exchange events, which involve lipid desorption and diffusion to another membrane in addition to insertion, could not be simulated due to their prohibitively long waiting times. Given the good agreement between k_{ins} calculated directly from simulation and from Kramers theory, we instead calculated k_{ex} from a Smoluchowski equation that contains information about the free energetics of hydrophobic lipid–membrane contacts.

To experimentally probe passive lipid exchange, two populations of vesicles, one initially composed of labeled lipids and the other initially of unlabeled ones, are combined in solution. k_{ex} is then determined by monitoring the rate of mixing of labeled lipids between these two populations of vesicles.[98,101,103,201–203] We can obtain an expression for the rate of mixing by first considering how the concentration of labeled lipids in a single vesicle changes over time. The concentration of labeled lipids in each vesicle will adopt a steady state with the concentration in the solution surrounding that vesicle. At steady state, the flux of lipids over any closed surface around a vesicle is constant. Importantly, since lipids that desorb may rapidly return to their original vesicle given the small barrier for insertion compared to desorption (Figure 2.6A), the flux contains contributions from lipids both leaving and entering a vesicle. According to the Smoluchowski equation, the flux, J, of lipids radially into a vesicle is

$$-J = \frac{\partial n_{\text{B}}}{\partial t} = 4\pi r^2 D e^{-\beta U(r)} \frac{\partial}{\partial r} \left(e^{\beta U(r)} \rho(r) \right), \tag{2.5}$$

where n_{B} is the number of labeled lipids in a vesicle, r is the radial distance from the center of a vesicle, D is the diffusion coefficient of a lipid in solution, $U(r)$ is the effective interaction potential between a lipid and a vesicle, and $\rho(r)$ is the concentration of labeled lipids at a distance r. As for the calculations using Kramers theory, we use $\min(d_{\text{CC}})$ to describe interactions between a lipid and a vesicle in 3D space. $U(r)$ is then the free energy profile as a function of $\min(d_{\text{CC}})$ (Figure A.10). We set $U(r) = 0$ far away from a vesicle and shift the peak of the free energy barrier to $r = 50$ nm, a typical radius of large unilamellar vesicles (LUVs) used in experimental studies.[98,101,103,201–203] Assuming that the concentration profile of labeled lipids reaches steady state much faster than n_B is varying, the time dependence of the bulk concentration of labeled lipids, $\bar{\rho}$, can be regarded as constant while calculating

the steady-state profile $\rho(r)/\bar{\rho}$. Assuming as well that equilibration within state B is fast, the concentration of labeled lipids within a vesicle obeys equilibrium statistics,

$$\rho(r_{\mathrm{B}}) = n_{\mathrm{B}}^{(i)} \frac{e^{-\beta U(r_{\mathrm{B}})}}{q_{\mathrm{B}}} \tag{2.6}$$
$$q_{\mathrm{B}} = \int_{\mathrm{B}} dr\ 4\pi r^2 e^{-\beta U(r)},$$

where i indicates whether the vesicle was initially composed of labeled ($i = 1$) or unlabeled ($i = 2$) lipids, r_{B} is a particular location within a vesicle, q_{B} is the partition function for labeled lipids in a vesicle, and the integral is performed over the range of r designated as state B. With these two assumptions, Eq. 2.5 integrates to

$$-J^{(i)} = 4\pi D \left(\bar{\rho} - \frac{n_{\mathrm{B}}^{(i)}}{q_{\mathrm{B}}}\right) \left[\int_{r_{\mathrm{B}}}^{r_{\mathrm{c}}} dr\ \frac{e^{\beta U(r)}}{r^2} + \frac{1}{r_{\mathrm{c}}}\right]^{-1}, \tag{2.7}$$

where r_{c} is the distance beyond which $U(r)$ is zero. Finally, the rate of mixing is obtained by considering how the difference between the number of labeled lipids in type 1 vesicles and in type 2 changes over time:

$$\frac{\partial}{\partial t}\left(n_{\mathrm{B}}^{(1)} - n_{\mathrm{B}}^{(2)}\right) = -k_{\mathrm{ex}}\left(n_{\mathrm{B}}^{(1)} - n_{\mathrm{B}}^{(2)}\right) \tag{2.8}$$
$$k_{\mathrm{ex}} = 4\pi D \left[q_{\mathrm{B}}\left(\int_{r_{\mathrm{B}}}^{r_{\mathrm{c}}} dr\ \frac{e^{\beta U(r)}}{r^2} + \frac{1}{r_{\mathrm{c}}}\right)\right]^{-1}. \tag{2.9}$$

Eq. 2.8 correctly reflects the fact that lipid exchange is a first order process.[92–103] Both the first term in the denominator of Eq. 2.9, which dominates when the free energy barrier is very large, and the second term, which dominates when diffusion is rate limiting, are key to the rate of lipid exchange.

The values of k_{ex} calculated with Eq. 2.9 for MARTINI DLPC and CHARMM36 DMPC are compared to experimental values in Table 2.1. In the case of MARTINI DLPC, our

Table 2.1: Rates of lipid exchange.

	k_{ex} (s^{-1})	
	Simulation	**Experiment**[98,101,103,201–203]
MARTINI	1.2×10^{-2}	$1.2 \times 10^{-3} - 4.3 \times 10^{-2}$
CHARMM36	5.7×10^{2}	$4.5 \times 10^{-5} - 6.6 \times 10^{-4}$

calculated k_{ex} is in agreement with experimental values. However, in the case of CHARMM36 DMPC, our calculated k_{ex} differs from experimental values by six orders of magnitude.

Some discrepancy between experiment and simulation/theory is to be expected because small errors in the free energy profile are magnified exponentially in k_{ex}. Yet, a six-order-of-magnitude difference between theoretical and experimental rates corresponds to a significant underestimation of the free energy barrier of approximately 12 k_BT for CHARMM36. This large discrepancy could be due to any number of the following reasons:

1. Sampling errors, which are within about 1 k_BT throughout the free energy surface (Figure A.9), may influence our rate calculations.

2. Inaccuracies in the force field may cause errors in the calculated free energy surface. Based on reported differences between permeabilities and partition coefficients for water, alkanes, and other small molecules obtained from simulations with CHARMM36 and from experiments,[204,205] we estimate that errors due to the force field are approximately $1 - 2$ k_BT. We also repeated our calculations using a different force field (Stockholm lipids, also known as Slipids[206,207]) as detailed in Appendix A. The height of the free energy barrier is approximately 2 k_BT higher for Slipids compared to CHARMM36 (Figure A.16), resulting in an estimate of k_{ex} that still differs from experiment by about five orders of magnitude.

3. The ionic strength of the solution may influence the free energy surface. Our simulations are performed with neat water, but experimental systems are conducted in buffer, which contains significant amounts of monovalent salts, such as NaCl or KCl.[98,201–203] The addition of NaCl or KCl could salt out the free lipids in solution, increasing the difference in free energy between state A and B. Additionally, monovalent cations bind to the carbonyl region of phosphatidylcholine membranes, causing the thickness of the bilayer and order of the tails to increase.[208–210] These structural changes are expected to increase the free energy barrier for hydrophobic contact formation.

4. The assumptions we made to obtain Eq. 2.7 may not be consistent with what occurs in experimental systems, causing discrepancies between our calculated k_{ex} and experimental measurements. In writing down a Smoluchowski equation (Eq. 2.5), we assumed that lipids exchange between stable, spherical vesicles of uniform size. This may be justified since vesicles composed of biological phospholipids are kinetically-trapped, metastable states, which by definition do not quickly relax into equilibrium lamellar structures.[211–214] But, we lack the requisite knowledge about the processes and associated timescales by which vesicles relax into equilibrium structures to confirm the validity of this assumption. Other relaxation processes, such as uncatalyzed vesicle fusion, which has an experimentally measured rate similar to lipid exchange,[215] could occur simultaneously. We note that lipid flip-flop, the process by which a lipid moves between leaflets of the same bilayer,[2,85] also has a rate measured under some experimental conditions that is similar to the exchange rate.[201,216] Unlike vesicle fusion, lipid flip-flop is accounted for in experimental determination of lipid exchange rates.

5. Similar issues of random error and unjustified kinetic assumptions could plague the inference of exchange rates from laboratory measurements.

Conclusions

The breakage of hydrophobic contacts limits the rate of passive lipid transport. To reach this conclusion, we investigated the elementary steps of lipid exchange, lipid insertion into and desorption from a membrane, using molecular dynamics simulations of the coarse-grained MARTINI and all-atom CHARMM36 lipid models. Results from MARTINI and CHARMM36 provide consistent pictures; even with a coarse description of lipids and water, the MARTINI model captures the essential features of lipid exchange exhibited in an all-atom model. We discovered that the reaction coordinate for passive lipid exchange measures the formation and breakage of hydrophobic lipid–membrane contacts, which gives rise to a free energy barrier for both lipid desorption and insertion.

Thus, knowledge of the reaction coordinate resolves previous qualitative discrepancies between simulations, which predicted that there is no barrier for lipid insertion, and experiments, which indicated that there is a barrier. This barrier likely plays an important biological role: A barrier for lipid insertion ensures that membrane compositions, which are spateotemporally regulated to maintain cell homeostasis,[2,88] are not easily disrupted. We suspect that the formation of hydrophobic contacts may generally give rise to a free energy barrier for transporting amphiphiles, including synthetic surfactants[217] and lipopeptides,[187,218] which is not resolved by monitoring the amphiphile's center-of-mass displacement from a membrane.

Additionally, knowledge of the reaction coordinate allowed us to formulate a Smoluchowski equation to model lipid exchange between vesicles, which occurs over time and length scales inaccessible in MD simulations, and calculate the rate of lipid exchange. Differences between our calculated lipid exchange rate and experimental measurements indicate that considerable quantitative discrepancies between simulation and experiment still exist. Future studies will be performed to better assess the sources of these discrepancies.

Finally, this knowledge provides a foundation to understand how catalysts of lipid exchange work at a molecular level. Lipid transfer proteins may efficiently extract lipids from membranes by lowering the activation free energy barrier for hydrophobic contact breakage. Interestingly, catalysts of vesicle fusion, the key step in vesicular lipid transport, may function in a similar way. For example, carbon nanotubes aid vesicle fusion by facilitating the formation of hydrophobic contacts between two vesicles,[219] and viral fusion peptides are thought to catalyze fusion by promoting hydrophobic lipid tail protrusions and contacts.[122] Thus, common physical properties of lipids, specifically their hydrophobicity, may be exploited *in vivo* to precisely control both non-vesicular and vesicular lipid transport.

Chapter 3

Membrane Hydrophobicity Determines the Activation Free Energy of Passive Lipid Transport

Reproduced with permission from Rogers, J. R.; Espinoza Garcia, G.; Geissler, P. L. bioRxiv doi: 10.1101/2021.03.17.435885.

Introduction

Over a thousand chemically diverse lipid species are heterogeneously distributed among eukaryotic cell membranes.[1,2] Even within a single membrane, compositional differences exist between leaflets[39] and laterally between nanodomains.[40] Because the collective behavior of a membrane's constituent lipids determines its physical properties, such as fluidity, thickness, and curvature, membrane compositions are under homeostatic control.[75–77] One way that proper lipid distributions are maintained is through non-vesicular transport of individual lipids between membranes. Non-vesicular transport enables rapid and specific alteration of membrane compositions, such as required to withstand cellular stress.[85,88]

Despite the recognized importance of lipid chemistry in determining membrane physical properties, how lipid physicochemical properties influence the dynamic processes that maintain precise membrane compositions is poorly understood. *In vivo*, lipid transfer proteins may be largely responsible for selectively transferring lipids recognized through specific protein–lipid interactions. Lipid transfer proteins are also equipped with membrane binding domains or motifs that may enable them to target donor and acceptor membranes with particular compositions and biophysical properties.[87,105] As demonstrated by *in vitro* experiments, lipids with subtle chemical differences are also passively exchanged between membranes at different rates.[95–101,103,161–163,202,220,221] For example, diacyl phosphatidylcholine (PC) exchanges much more slowly than lysoPC, its single-tailed counterpart.[98,100] Even the addition of just two carbons to an acyl chain of a phospholipid reduces its exchange rate by

roughly 10-fold.[96,98–101,103,161,163,220,221] The physical properties and chemical composition of the donor and acceptor membranes additionally influence the rate at which a lipid is passively transported.[95,96,98,101,161,202,221,222] For example, dimristoylphosphatidylcholine (DMPC) exchanges more rapidly between liquid-crystalline (L_α) phase vesicles than more ordered gel (L_β) phase ones.[98,222] Therefore, the underlying free energy barriers for transport also depend on a lipid's chemical structure and properties of the donor and acceptor membranes.

During passive lipid transport, a lipid first desorbs from a donor membrane, diffuses through solvent, and then inserts into an acceptor membrane. Due to the large free energetic cost of disrupting a lipid's local hydrophobic environment in a membrane, lipid desorption is the rate-limiting step and, thus, determines the activation free energy of lipid transport. A smaller free energy barrier exists for lipid insertion.[97–99,161,223] In Chapter 2, we elucidated the molecular origins of both free energy barriers by identifying the reaction coordinate for passive DMPC transport: The reaction coordinate measures the extent of hydrophobic contact between the transferring DMPC and membrane. The free energy barrier for DMPC desorption reflects the thermodynamic cost of breaking hydrophobic lipid–membrane contacts, and the free energy barrier for DMPC insertion reflects the cost of disrupting the membrane–solvent interface during the formation of initial hydrophobic lipid–membrane contacts.[223]

Here, we systematically investigate how the free energy barriers for both lipid desorption and insertion vary with lipid chemistry and membrane phase. By focusing our analysis on glycerophospholipids, we determine how the chemical diversity found among structural lipids in eukaryotic lipidomes[1,2] influences the rate of passive lipid transport. Since all glycerophospholipids possess the same general physical properties, we assume at the outset that the reaction coordinate for DMPC transport discovered in Chapter 2[223] also captures the salient features of transporting other glycerophospholipids. Working under this assumption allows us to efficiently quantify the free energy barriers for transporting lipids between 14 different membranes using all-atom molecular dynamics (MD) simulations. The results reported herein indicate that the biophysical mechanism of lipid transport is indeed invariant to glycerophospholipid chemistry, supporting our initial assumption. We find that the activation free energy for lipid transport increases as lipid acyl chain length increases and as membrane order increases such that L_β phase membranes pose the largest barriers for lipid transport. Additionally, we find the activation free energy to be strongly dependent on the identity of the lipid headgroups. Our results are consistent with *in vitro* measurements and provide a biophysical rationale for previously unexplained experimental observations. Furthermore, the atomistic detail provided by molecular simulations allows us to correlate the free energy barriers for lipid desorption and insertion with the molecular properties of a membrane and, thus, to identify the key physicochemical properties that control transport rates: Membrane hydrophobicity, which can be tuned through subtle changes in lipid chemistry and membrane order, determines the rate of passive lipid transport.

Methods

The rate-limiting step of passive lipid transport – spontaneous desorption of a lipid from a membrane – occurs over minutes to hours. As a result, calculating the transport rate directly from unbiased MD simulations is computationally intractable for even one membrane system. Instead, our approach to quantitatively assess how the rate of passive lipid transport varies with lipid chemistry and membrane phase capitalizes on knowledge of the reaction coordinate,[223] which we assume to be the same for all lipids investigated.

Free energy profiles along order parameters other than the reaction coordinate generally yield barriers that underestimate the rate-determining activation free energy, and therefore do not provide sufficient thermodynamic information to accurately determine barrier-crossing rates.[139] A barrier can even be entirely absent from such a free energy profile, in which case even a qualitative inference of the transition dynamics is erroneous. As shown in Chapter 2, lipid insertion is a prime example of this pathology: Free energy profiles calculated as a function of the center of mass displacement of the lipid from the membrane incorrectly suggest that lipid insertion is a barrier-less process.[223]

By calculating free energy profiles along the reaction coordinate, we efficiently determined the rate-limiting free energy barrier for transporting 11 different lipids between membranes composed of the same lipid species as the one being transported. For high melting temperature lipids, free energy barriers for lipid desorption from and insertion into both L_α and L_β phase membranes were determined (Figure 3.1).

Simulated Systems

To systematically assess how the activation free energy of lipid transport depends on lipid physicochemical properties, we simulated membrane systems composed of one of eleven different glycerophospholipid species. This included a series of PC lipids with increasing acyl chain lengths and various degrees of saturation (Figure 3.1C): 1,2-dimyristoyl-*sn*-glycero-3-PC (DMPC), 1,2-dipalmitoyl-*sn*-glycero-3-PC (DPPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-PC (POPC), 1,2-dioleoyl-*sn*-glycero-3-PC (DOPC), 1-stearoyl-2-oleoyl-*sn*-glycero-3-PC (SOPC), and 1,2-distearoyl-*sn*-glycero-3-PC (DSPC). This also included a series of 1-palmitoyl-2-oleoyl (PO) lipids with both zwitterionic and anionic headgroups (Figure 3.1D): PO-*sn*-glycero-3-phosphoethanolamine (POPE), POPC, PO-*sn*-glycero-3-phosphate (POPA), PO-*sn*-glycero-3-phospho-L-serine (POPS), PO-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (POPG), and PO-*sn*-glycero-3-phosphoinositol (POPI). Membranes composed of each lipid were simulated in the L_α phase at 320 K with the exception of DSPC, which was simulated at 350 K due to its higher melting temperature. Membranes composed of the saturated lipids, DMPC, DPPC, and DSPC, were additionally simulated in the L_β phase (Figure 3.1E) at 275 K, 295 K, and 320 K, respectively.

For all lipid species, initial L_α phase bilayers of 128 lipids surrounded by 3.2 nm thick slabs of solvent were built using the CHARMM-GUI Membrane Builder.[181,182] Neutralizing

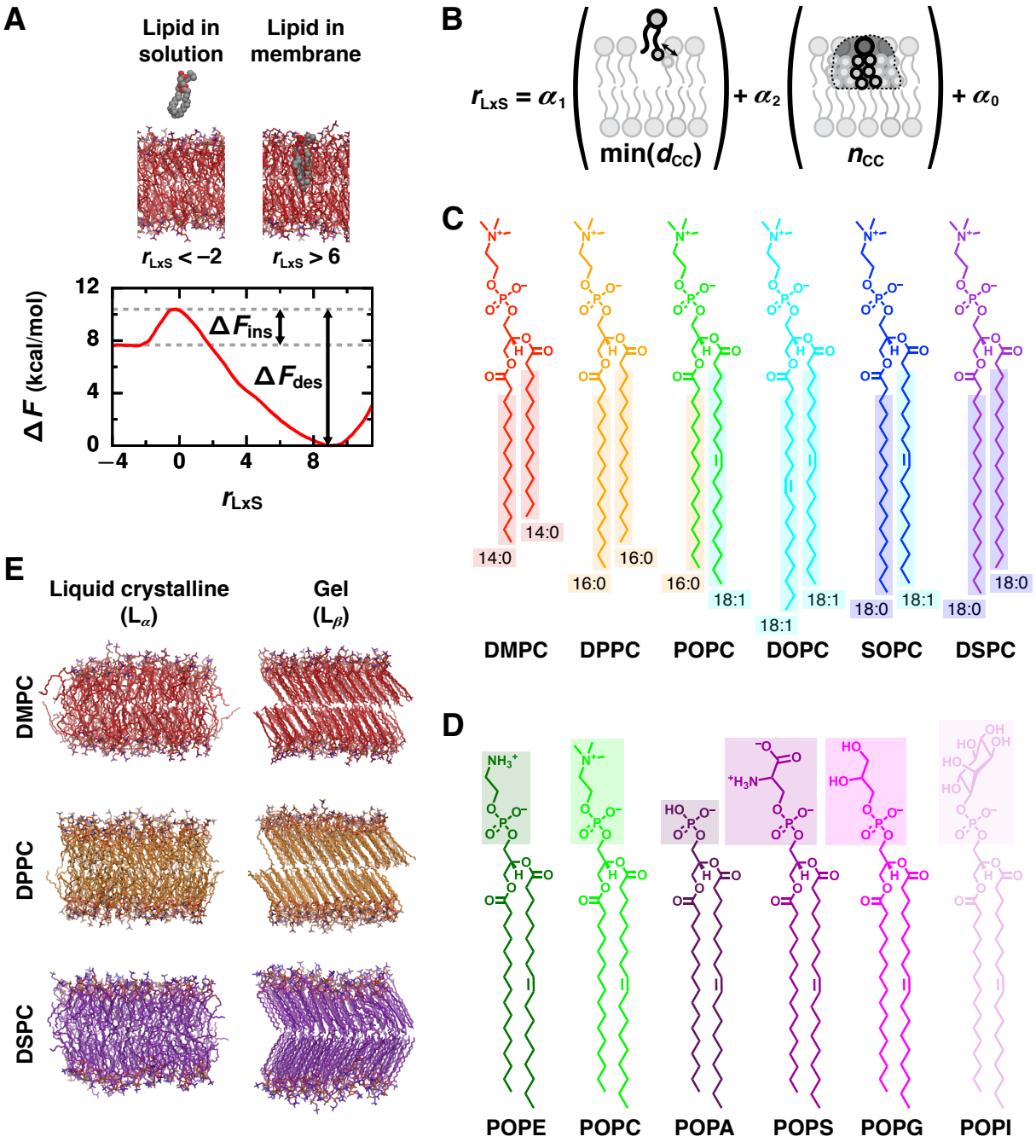


Figure 3.1: Physicochemical properties that determine the rate of lipid transport were assessed for a variety of lipid chemistries and phases by calculating the free energy barriers for lipid desorption and insertion. (A) An illustrative free energy profile, calculated as a function of the reaction coordinate, r_{LxS}, for a L_α phase DMPC membrane, reveals barriers

(Continued on next page)

Figure 3.1: *(Continued from previous page)*
for lipid desorption (ΔF_{des}) and lipid insertion (ΔF_{ins}). Representative configurations are shown for negative values of r_{LxS} (where the lipid is in solution) and for positive values of r_{LxS} (where the lipid is in the membrane). Solvent is not rendered. (B) r_{LxS} is a linear combination of $\min(d_{\mathrm{CC}})$ (the minimum distance between any hydrophobic carbon of the lipid and of the closest membrane leaflet) and n_{CC} (and the number of close hydrophobic carbon contacts between the lipid and closest leaflet). (C and D) Chemical structures of the lipids investigated with varied (C) acyl chain lengths and degrees of saturation, and (D) headgroups. In (C), hydrophobic carbons are boxed. (E) Representative configurations of membranes composed of DMPC, DPPC, or DSPC in both L_α and L_β phases.

sodium ions were added to each anionic membrane. The L_α phase DSPC bilayer was also used to initialize a simulation at 320 K that yielded a L_β phase bilayer within 100 ns. Configurations of L_β phase DPPC and DMPC bilayers were instead used to initialize simulations at temperatures consistent with a L_β phase to avoid these systems getting trapped in the ripple phase upon cooling, which can occur in simulations of shorter chain lipids.[224] The initial L_β phase DPPC bilayer was obtained from the LipidBook repository[225] and also used to construct an initial L_β phase DMPC bilayer. The CHARMM36 force field[177] was used to model all lipids in combination with the CHARMM TIP3P water model.[178]

Molecular Dynamics Simulations

For each different lipid species and membrane phase, solvated bilayers were simulated to characterize the physical properties of each membrane at equilibrium. All simulations were performed in an isothermal–isobaric (NPT) ensemble using GROMACS 2019.[166] The pressure was maintained at 1 bar using semi-isotropic pressure coupling with an isothermal compressibility of $4.5 \times 10 - 5$ bar^{-1}, and the temperature was maintained using the Nosé-Hoover thermostat[183,184] with a coupling time constant of 1 ps. The lipids and solvent were coupled to separate thermostats. Dynamics were evolved using the leapfrog algorithm[173] and a 2 fs time step. All bonds to hydrogen were constrained using the LINCS algorithm.[185] Lennard-Jones forces were smoothly switched off between 0.8 and 1.2 nm. Coulomb interactions were truncated at 1.2 nm, and long-ranged Coulomb interactions were calculated using Particle Mesh Ewald (PME) summation.[186] Neighbor lists were constructed using the Verlet list cut-off scheme.[175]

Each initial bilayer configuration was first energy minimized using the steepest descent algorithm and then equilibrated in two steps: The first 250 ps equilibration used the Berendsen barostat[169] to maintain the pressure with a coupling time constant of 2 ps, and the second 250 ps equilibration used the Parinello-Rahman barostat[170] with a coupling time constant of 5 ps. To allow the bilayers' structures to fully equilibrate, runs of 50 ns for the L_α and of 100 ns for the L_β phase membranes were performed using the same parameters as the

second equilibration step. The final configurations from these runs were used to construct initial configurations for enhanced sampling simulations that had an additional lipid free in solution. Simulations of each bilayer were extended for an additional 300 ns, which was analyzed to calculate average properties of each membrane.

Free Energy Calculations

To determine the activation free energy for passively transporting a lipid between membranes, we calculated free energy profiles as a function of the reaction coordinate (Figure 3.1A). As determined in Chapter 2,[223] the reaction coordinate, r_{LxS}, is a linear combination of two order parameters that measure hydrophobic lipid–membrane contacts between the transferring lipid and closest membrane leaflet (Figure 3.1B). Both parameters are based on the collection of distances $d_{\mathrm{CC}}^{(i)}$ between a hydrophobic carbon of the lipid and a hydrophobic carbon of the membrane, where i labels a particular carbon-carbon (CC) pair: (1) $\min(d_{\mathrm{CC}}) = \min_i d_{CC}^{(i)}$ is the minimum of these CC distances; and (2) n_{CC} is the number of CC pairs that satisfy $d_{\mathrm{CC}}^{(i)} \leq 1$ nm. The hydrophobic carbons that are considered to calculate n_{CC} for each lipid are boxed in Figure 3.1C. In detail,

$$r_{\mathrm{LxS}} = \alpha_1 \min(d_{\mathrm{CC}}) + \alpha_2 n_{\mathrm{CC}} + \alpha_0 \tag{3.1}$$

with coefficients $\alpha_1 = -2.247$ nm^{-1}, $\alpha_2 = 0.004828$, and $\alpha_0 = 0.6014$ such that r_{LxS} is a unitless quantity. By construction, $r_{\mathrm{LxS}} = 0$ at the free energy barrier that separates configurations with the transferring lipid fully in solution, which have negative values of r_{LxS}, and configurations with the transferring lipid fully in the membrane, which have positive values of r_{LxS} (Figure 3.1A). Regardless of its total number of hydrophobic carbons, a lipid forms only a few hydrophobic contacts with the membrane at the free energy barrier.[223] The largest free energy barrier therefore occurs at $r_{\mathrm{LxS}} \approx 0$ for all lipid species investigated, even though $\{\alpha_i\}$ were determined specifically for DMPC. During all enhanced sampling simulations, r_{LxS} was calculated using a differentiable form for $\min(d_{\mathrm{CC}})$,

$$\gamma \left[\log \left(\sum_i \exp(\gamma / d_{\mathrm{CC}}^{(i)}) \right) \right]^{-1} \tag{3.2}$$

with $\gamma = 200$ nm, and for n_{CC},

$$\sum_i \frac{1 - \left(\frac{d_{\mathrm{CC}}^{(i)} - d_0}{r_0} \right)^6}{1 - \left(\frac{d_{\mathrm{CC}}^{(i)} - d_0}{r_0} \right)^{12}} \tag{3.3}$$

with $d_0 = 1$ nm and $r_0 = 0.025$ nm.

We performed umbrella sampling simulations[189] using the PLUMED 2 patch[190] for GROMACS to obtain the free energy profiles $\Delta F(r_{\mathrm{LxS}})$ for each system. To generate initial configurations, a tagged lipid of the same species as the membrane lipids was randomly inserted

into the solvent around each equilibrated bilayer such that the center of mass of the tagged lipid and of the bilayer were separated by at least 3.2 nm along z, which is the axis perpendicular to the bilayer. Each system was then energy minimized and equilibrated using the same two-step procedure as used for the bilayers with the addition of harmonic position restraints on the z coordinates of the tagged lipid's heavy atoms with a force constant of 500 kJ/mol/nm^2. Next, to generate initial configurations for each umbrella sampling window, a steered MD simulation was performed using a harmonic bias on r_{LxS} with a spring constant of 500 kJ/mol. Each umbrella sampling window was initialized with a configuration from the steered MD simulation that has a value of r_{LxS} within 0.1 of the window's bias center. Bias parameters used for umbrella sampling are tabulated in Table B.1. All windows were run for 52 ns. After discarding 20 ns to account for equilibration, data from all windows was combined with the weighted histogram analysis method (WHAM)[191] to obtain $\Delta F(r_{\mathrm{LxS}})$. Error bars were calculated as the standard error of $\Delta F(r_{\mathrm{LxS}})$ estimated from 4 independent 8 ns blocks. As indicated in Figure 3.1A, by setting $\Delta F(r_{\mathrm{LxS}})$ to zero at the global free energy minimum, the barriers for lipid desorption and insertion are defined as

$$\Delta F_{\mathrm{des}} = \max_{-3<r_{\mathrm{LxS}}<3} \Delta F(r_{\mathrm{LxS}}) \tag{3.4}$$

$$\Delta F_{\mathrm{ins}} = \Delta F_{\mathrm{des}} - \Delta F(r_{\mathrm{LxS}} = -3.45). \tag{3.5}$$

Analysis of Membrane Properties

We analyzed trajectories of each bilayer system to assess how lipid chemistry and membrane phase influence the molecular structure of a membrane, and to provide a reference for rationalizing activation free energies for passive lipid transport in terms of physicochemical properties. In addition to the measures of hydrophobic lipid–membrane contacts defined above, we computed the membrane–solvent interaction energy, area per lipid, area of interfacial packing defects, membrane thickness, density profiles along the membrane normal, carbon–deuterium order parameters of the lipid tails, and radial distribution functions that characterize the intermolecular structure of the membrane–solvent interface. A combination of MDAnalysis[188] and NumPy[226] Python libraries in addition to GROMACS tools were used to calculate all properties. The membrane–solvent interaction energy, $E_{\mathrm{mem-solv}}$, was calculated as the sum of short-ranged Lennard-Jones and Coulomb interaction energy terms between the membrane and solvent. The average area per lipid, $\langle A_{\mathrm{lip}} \rangle$, was calculated as the area of the box in the xy plane divided by the total number of lipids in a leaflet. Lipid packing defects, which are interfacial voids between polar lipid heads that expose aliphatic atoms to solvent, were identified with PackMem.[61] The packing defect size constant, π_{defect}, was obtained by fitting the distribution of defect areas, A_{defect}, to a monoexponential decay, $P(A_{\mathrm{defect}}) \propto e^{-A_{\mathrm{defect}}/\pi_{\mathrm{defect}}}$. Fits were performed on $P(A_{\mathrm{defect}}) \geq 10^{-4}$ and $A_{\mathrm{defect}} \geq 1.5$ nm^2 for L_α phase membranes or $A_{\mathrm{defect}} \geq 0.5$ nm^2 for L_β phase membranes. Definitions of all other properties are provided in Appendix B. All analysis was performed on the final 300 ns of each bilayer trajectory split into 100 ns intervals for block averaging. Reported error is the standard error calculated from the three 100 ns intervals.

Results

For a series of 14 different single-component membranes, we quantified the free energy barriers that limit the rates of lipid desorption from and insertion into a membrane (Figure 3.1). By calculating free energy profiles $\Delta F(r_{\mathrm{LxS}})$ as a function of the reaction coordinate, which captures the collective motion of molecules that advances a transition,[136–139] we are able to extract dynamical information from them. Lipid desorption, which limits the rate of passive lipid transport, occurs when a lipid transitions from the membrane to the solvent, breaking hydrophobic contacts with the membrane along the way. Lipid insertion, which is the final step of lipid transfer, is the reverse process. The reaction coordinate for lipid (L) transport via solvent (xS), r_{LxS}, measures the breakage and formation of hydrophobic lipid–membrane contacts through a linear combination of the minimum distance, $\min(d_{\mathrm{CC}})$, and number of close contacts, n_{CC}, between hydrophobic carbons of the transferring lipid and membrane (Figure 3.1B).[223] We assume that r_{LxS} serves as a good reaction coordinate for all lipids investigated as long as all hydrophobic carbons of each lipid species (Figure 3.1C) are used to calculate n_{CC}. r_{LxS} describes progress from configurations with the lipid in the membrane (which exhibit many hydrophobic lipid–membrane contacts and large positive values of r_{LxS}) to configurations with the lipid in solution (which have very few if any hydrophobic lipid–membrane contacts and negative values of r_{LxS}). Correspondingly, the rate of lipid desorption is limited by the free energy barrier ΔF_{des}, which is the difference between $\Delta F(r_{\mathrm{LxS}} \approx 0)$ and the global free energy minimum found at large positive values of r_{LxS}; the rate of lipid insertion is limited by the free energy barrier ΔF_{ins}, which is the difference between $\Delta F(r_{\mathrm{LxS}} \approx 0)$ and the free energy at very negative values of r_{LxS} (Figure 3.1A).

Increasing Lipid Acyl Chain Length Increases the Desorption Barrier

To investigate how the chemical structure of the lipid tails influences the rate of lipid transport, we simulated a series of L_α phase membranes composed of PC lipids with acyl chain lengths ranging from 14 to 18 carbons and degrees of unsaturation ranging from zero to two (Figure 3.1C). This list included fully saturated lipids DMPC, DPPC, and DSPC, that only differ in chain lengths by two carbons; lipids with acyl chains of 18 carbons each DOPC, SOPC, and DSPC, that only differ by one degree of unsaturation; and the mixed-chain lipid POPC, which has an *sn1* chain of 16 carbons and *sn2* chain of 18 carbons with one double bond.

The free energy profiles of each lipid exhibit the same qualitative features (Figure 3.2A). This commonality is consistent with our assumption that PC lipids with different tails share the same biophysical mechanism of lipid transport, and that the generic transition pathway is well characterized by r_{LxS}. Importantly, all of the free energy profiles exhibit a barrier at $r_{\mathrm{LxS}} \approx 0$. However, the free energy profiles are quantitatively different.

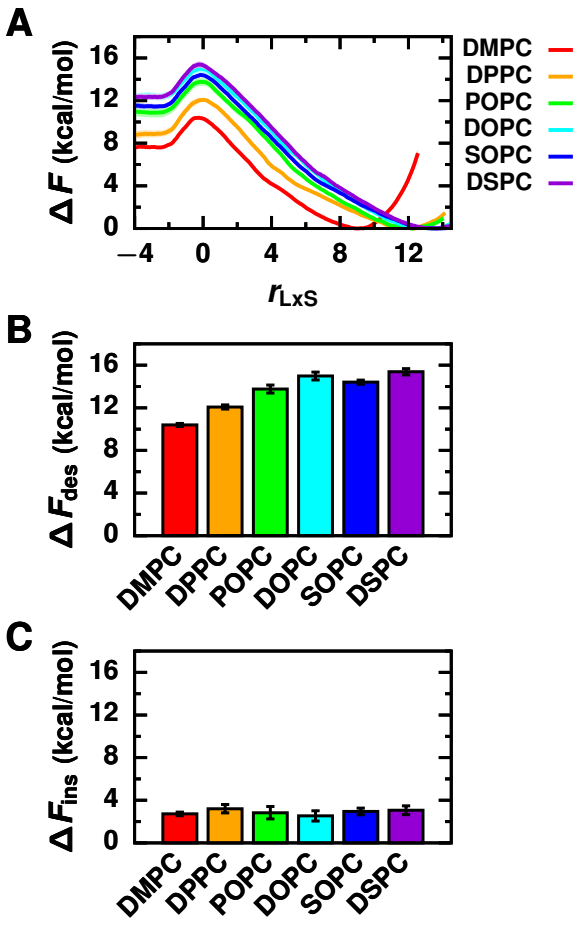


Figure 3.2: Increasing lipid acyl chain length increases the desorption barrier and does not substantially influence the insertion barrier. (A) Free energy profiles as a function of the reaction coordinate for L_α phase membranes composed of lipids with different tail chemistries. Corresponding free energy barriers are shown for (B) lipid desorption and (C) lipid insertion.

As shown in Figure 3.2B, ΔF_{des} varies significantly with lipid tail chemistry. Increasing the total number of carbons in the acyl chains increases ΔF_{des}. Lipids with longer tails have more hydrophobic carbons (Figure 3.1C) that can form hydrophobic contacts with surrounding lipids in the membrane (Table B.2). Consequently, more hydrophobic lipid–membrane contacts must break for lipids with longer tails to desorb, explaining why (a) ΔF_{des} increases for the fully saturated lipids in the order DMPC $<$ DPPC $<$ DSPC, and (b) ΔF_{des} increases for the unsaturated lipids that differ in the length of a single acyl chain in the order POPC $<$ SOPC. ΔF_{des} is smaller for lipids with unsaturated bonds than for fully saturated lipids with the same acyl chain lengths (Figure 3.2B); both DOPC and SOPC have slightly reduced ΔF_{des} compared to DSPC. Membranes composed of lipids with

increased degrees of unsaturation are more disordered (Figure B.1), with decreased bilayer thicknesses (Figure B.2 and Table B.2) and increased areas per lipid (Table B.2). As a result, hydrophobic contacts in these membranes can be more easily disrupted during lipid desorption.

ΔF_{ins} is substantially smaller than ΔF_{des}, again demonstrating that lipid desorption limits the rate of lipid transport. In contrast to ΔF_{des}, ΔF_{ins} does not vary significantly with lipid tail chemistry (Figure 3.2C). During lipid insertion, a lipid must break through the membrane's interface to form an initial hydrophobic lipid–membrane contact. Since all the lipids that we investigated with different tail chemistries have the same PC headgroup, the interfacial structure and chemistry of the membranes are quite similar (Figure B.2-B.4). Consequently, the free energetic cost for a lipid to cross the membrane's interface during insertion is roughly the same.

Lipid Headgroup Influences Both Desorption and Insertion Barriers

To investigate how the chemical structure of the lipid headgroups influence lipid transport rates, we simulated a series of L_α phase membranes composed of PO lipids with both zwitterionic and anionic headgroups (Figure 3.1D). This list included neutral lipids POPE, which has a terminal amine group, and POPC, which has a bulkier choline group; and the anionic lipids POPA and POPS, which have terminal acidic groups, and POPG and POPI, which have terminal polar groups.

The free energy profiles of the PO lipids with different headgroups exhibit the same general features as the profiles for the PC lipids, including a free energy barrier at $r_{\text{LxS}} \approx 0$ (Figure 3.3A). This commonality further supports our assumption that r_{LxS} is a generic reaction coordinate among glycerophospholipids. It additionally suggests that lipids with different headgroups and tails are transferred through the same biophysical mechanism, which is characterized by the breakage and formation of hydrophobic lipid–membrane contacts.

Quantitative differences in ΔF_{des} indicate that lipids with different headgroups are transported at different rates (Figure 3.3B). It is difficult to rationalize why certain headgroups result in smaller or larger ΔF_{des} based on their chemical structures alone. But, differences in ΔF_{des} can be explained based on each headgroup's influence on general physical properties of a membrane,[62,227,228] including lipid packing as measured by the average area per lipid (Table B.2), membrane thickness (Figure B.2 and Table B.2), and acyl chain order parameters (Figure B.1). Crucial to lipid transport, the headgroup impacts the number of hydrophobic contacts that a lipid forms with surrounding lipids in the membrane. A greater number of hydrophobic contacts between lipids in PE, PS, and PA membranes must be disrupted on average than in PI, PC, and PG membranes (Table B.2) during lipid desorption. ΔF_{des} increases accordingly in the order POPG < POPC < POPI < POPA < POPS < POPE.

The interfacial structure and chemistry of a membrane is largely determined by the lipid headgroup (Figure B.2-B.4). ΔF_{ins}, which reflects the cost for a hydrophobic lipid tail to cross the membrane's interface, thus varies with lipid headgroup (Figure 3.3C). On average, zwitterionic lipids have smaller ΔF_{ins} compared to anionic lipids; when comparing lipids

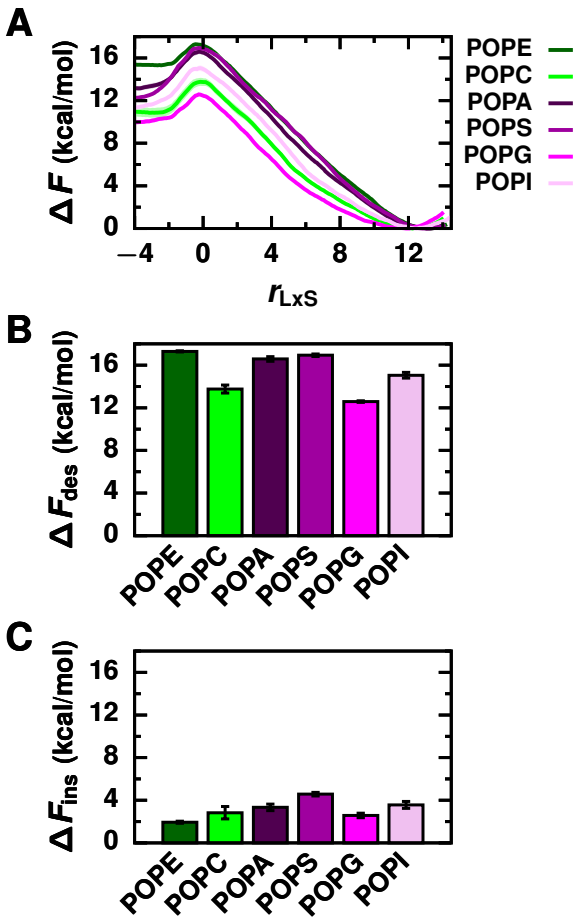


Figure 3.3: Lipid headgroup chemistry influences both desorption and insertion barriers. (A) Free energy profiles as a function of the reaction coordinate for L_α phase membranes composed of lipids with different headgroups. Corresponding free energy barriers are shown for (B) lipid desorption and (C) lipid insertion.

with chemically similar headgroups (PE and PC; PA and PS; PG and PI), ΔF_{ins} decreases as the molecular size of the headgroup decreases. Decreasing the net charge of the membrane surface and decreasing the surface density of polar functional groups both reduce the free energetic cost of disrupting the membrane's interfacial structure during the formation of an initial hydrophobic contact with an incoming lipid.

Increasing Membrane Order Increases Both Desorption and Insertion Barriers

Finally, we investigated how the rate of lipid transport depends on membrane phase. We simulated L_β phase membranes composed of the high melting temperature lipids DMPC, DPPC, and DSPC. Because the other lipid species investigated have very low melting temperatures, it would have been intractable to simulate them in a L_β phase and, furthermore, they are unlikely to be dominant components of highly ordered membrane domains at physiological temperatures.

As with L_α phase membranes, the free energy profiles of L_β phase membranes exhibit a rate-limiting free energy barrier at $r_{\mathrm{LxS}} \approx 0$, albeit broader and more rugged (Figure 3.4A). Thus, for both L_α and L_β phase membranes, the rate of lipid transport critically depends on the breakage and formation of hydrophobic lipid–membrane contacts. In contrast to free energy profiles of L_α phase membranes, profiles of L_β phase membranes exhibit a local free energy minimum at positive values of r_{LxS}. At this local minimum, the transferring lipid adopts a splayed configuration with one tail anchored in the membrane and the other exposed to solvent (Figure 3.4A and B.5). While splayed intermediates have been observed in trajectories of lipid insertion into L_α phase membranes,[165,223] they are not locally thermodynamically stable in those cases. Given that splayed lipids form the transition state for vesicle fusion,[122,193–197] the enhanced stability we have observed in more ordered membranes may help explain why ripple phase membranes fuse faster than L_α phase membranes in *in vitro* assays[229] and why many viral fusion proteins localize to lipid rafts.[230–235] In L_β phase membranes, splayed lipids persist after hydrophobic contacts between a single lipid tail and the membrane are broken since a second free energy barrier must be crossed to break contacts with the other lipid tail. The free energy maximum at $r_{\mathrm{LxS}} \approx 0$ ultimately limits the rate of lipid transport; its value relative to the global minimum is reported as ΔF_{des} below.

Both ΔF_{des} and ΔF_{ins} are substantially larger for L_β compared to L_α phase membranes (Figure 3.4B and 3.4C). Lipids in highly ordered and tightly packed membranes, epitomized by L_β phase membranes, form a greater number of hydrophobic contacts with surrounding membrane lipids (Table B.2). More contacts must be broken for a lipid to desorb from a L_β phase membrane compared to a L_α phase membrane, thus increasing ΔF_{des}. Consistent with the trend observed for L_α phase membranes, L_β phase membranes composed of lipids with longer acyl chains also have larger ΔF_{des}. The tight packing of lipids in L_β compared to L_α phase membranes also increases the density of polar lipid headgroups at the membrane's interface (Figure B.2). Consequently, the free energetic cost for a lipid to traverse the membrane's interface to form an initial hydrophobic lipid–membrane contact increases, explaining the overall increase in ΔF_{ins} for L_β phase membranes. Lipids pack more tightly in L_β phase DPPC and DMPC membranes than in L_β phase DSPC membranes, as indicated by their smaller areas (Table B.2) and more ordered tails (Figure B.1). The resulting increased surface density of headgroups further hinders the disruption of the membrane's interfacial structure during lipid insertion and, thus, ΔF_{ins} increases accordingly in the order DSPC $<$ DMPC $<$ DPPC for L_β phase membranes.

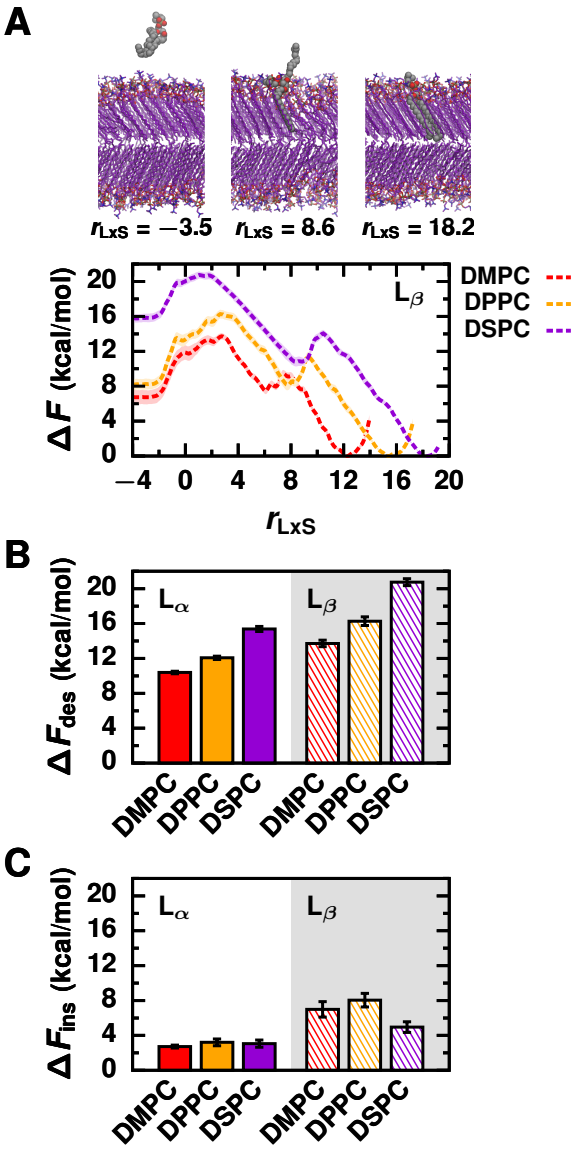


Figure 3.4: L_β phase membranes exhibit increased barriers for both desorption and insertion compared to L_α phase membranes. (A) Free energy profiles as a function of the reaction coordinate for L_β phase membranes, together with configurations of DSPC representating all three local minima of $\Delta F(r_{LxS})$: the lipid in solution, splayed lipid intermediate, and lipid in the membrane. (B and C) Free energy barriers of (B) lipid desorption and (C) lipid insertion for membranes simulated in both L_α and L_β phases.

Desorption Barrier Depends on a Lipid's Local Hydrophobic Environment in a Membrane

We attribute the main free energetic cost for lipid desorption to the disruption of a lipid's locally hydrophobic environment in a membrane. A lipid's hydrophobic environment is

quantified by the average number of close hydrophobic contacts that a lipid in a membrane makes with surrounding lipids, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. As shown in Figure 3.5, a general trend between $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ and ΔF_{des} exists: As $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ increases, ΔF_{des} increases since more hydrophobic

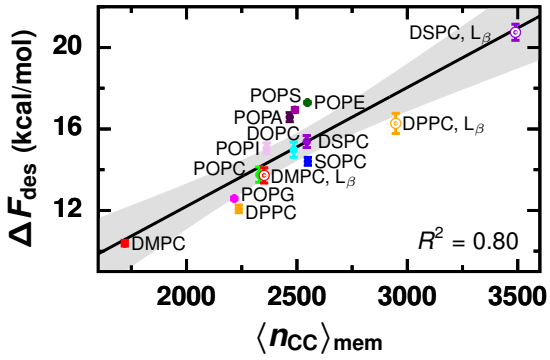


Figure 3.5: The barrier for desorption of a lipid is determined by its local hydrophobic membrane environment, which we quantify by the average number of close hydrophobic contacts between membrane lipids, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. Black line is the best linear fit: $\Delta F_{\mathrm{des}} = 0.0058 \langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}} + 0.53$. Gray region indicates 95% confidence interval.

contacts must be broken to displace a lipid from the membrane. ΔF_{des} is roughly linearly correlated with $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, indicating that the cost of breaking each hydrophobic contact is approximately the same (roughly 5.8 cal/mol/contact). Because $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ is sensitive to both lipid chemistry and membrane phase, the observations made above about how lipid tail chemistry, headgroup, and membrane phase individually affect ΔF_{des} are all attributable to differences in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. In other words, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ integrates the contributions to ΔF_{des} from all chemical and physical features of a membrane. Differences in ΔF_{des} between membranes that vary in both lipid chemistry and membrane phase, for example a L_β phase DMPC membrane compared to a L_α phase POPE membrane, are indeed explained by differences in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. Thus, the hydrophobicity of a membrane's core, as measured by $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, predominantly determines ΔF_{des} and, consequently, the rate of passive lipid transport.

Insertion Barrier Depends on a Membrane's Interfacial Hydrophobicity

We attribute the main free energetic cost for lipid insertion to the disruption of the membrane's stratified chemical organization when the lipid breaches the membrane's interface to form an initial hydrophobic lipid–membrane contact. The price of such a "mixing" of lipid headgroups and tails is greatest when the two groups are most chemically dissimilar; there is more resistance for lipid insertion into membranes with more hydrophilic surfaces. However,

this picture is complicated by the fact that a membrane's interface is chemically heterogeneous at the molecular scale, exhibiting regions dominated by headgroups and regions where tails are partially exposed to solvent. Contributions from both types of regions must be considered to explain all observed variations in ΔF_{ins}. To compare different membranes, we characterize the net chemical character of their surfaces by an interfacial hydrophobicity. (The textbook description of lipid membranes' surfaces is strongly hydrophilic, corresponding to very low interfacial hydrophobicity.) We quantify contributions from the headgroups by the average membrane–solvent interaction energy, $\langle E_{\text{mem-solv}} \rangle$, and we consider membranes with less favorable $\langle E_{\text{mem-solv}} \rangle$ to have more hydrophobic interfaces. As shown in Figure 3.6A, for membranes with different headgroups, ΔF_{ins} decreases as $\langle E_{\text{mem-solv}} \rangle$ becomes less

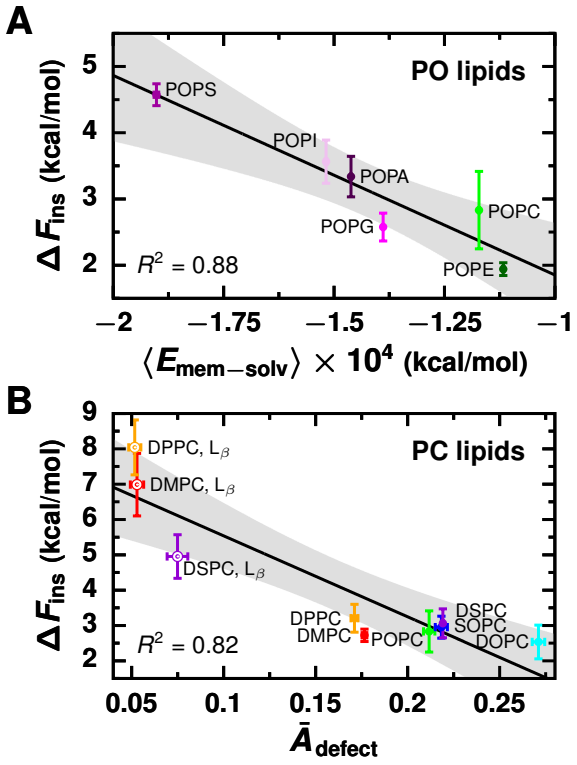


Figure 3.6: The barrier for lipid insertion is determined by a membrane's interfacial hydrophobicity, which we quantify by the average membrane–solvent interaction energy ($\langle E_{\text{mem-solv}} \rangle$) and the relative area of membrane packing defects ($\bar{A}_{\text{defect}}$). (A) Changes in ΔF_{ins} due to lipid headgroup chemistry are attributed to differences in $\langle E_{\text{mem-solv}} \rangle$. Black line is the best linear fit: $\Delta F_{\text{ins}} = -0.00030 \langle E_{\text{mem-solv}} \rangle - 1.2$. (B) Changes in ΔF_{ins} due to lipid tail chemistry and membrane phase are attributed to differences in $\bar{A}_{\text{defect}}$. Black line is the best linear fit: $\Delta F_{\text{ins}} = -22.9 \bar{A}_{\text{defect}} + 7.8$. Gray regions indicate 95% confidence intervals.

favorable since the cost to compromise interactions between the headgroups and solvent dur-

ing lipid insertion is reduced. For membranes with the same headgroup, differences in ΔF_{ins} arise from variations in the contributions to the interfacial hydrophobicity from regions where lipid tails are partially exposed. Because these regions result from lipid packing defects, we quantify these contributions by the relative area of packing defects, $\bar{A}_{\text{defect}}$, defined as the ratio of the characteristic packing defect size π_{defect}[61] (Figure B.6) to the average area of a lipid in the membrane, $\langle A_{\text{lip}} \rangle$. We consider membranes with increased $\bar{A}_{\text{defect}}$ to have more hydrophobic interfaces. As shown in Figure 3.6B, for membranes with the same headgroup but different tails or phases, ΔF_{ins} decreases as $\bar{A}_{\text{defect}}$ increases since the cost to locally expose the membrane's hydrophobic core during lipid insertion is reduced.

Overall, a membrane's interfacial hydrophobicity determines ΔF_{ins}, and, consequently, the rate of lipid insertion. To identify a relationship between ΔF_{ins} and quantitative metrics of a membrane's interfacial hydrophobicity, we have considered variations due to lipid headgroup, which modulate interfacial hydrophobicity through changes in $\langle E_{\text{mem-solv}} \rangle$, separately from variations due to tail chemistry and membrane phase, which tune interfacial hydrophobicity through changes in $\bar{A}_{\text{defect}}$. More sophisticated and precise measures of a membrane's interfacial hydrophobicity, such as have been developed to characterize protein surfaces,[236–239] may abrogate the need to consider contributions from headgroups separately from lipid tails and membrane phase. However, the two simplistic measures that we identify, $\langle E_{\text{mem-solv}} \rangle$ and $\bar{A}_{\text{defect}}$, reasonably explain variations in ΔF_{ins} for different membranes relative to the errors in calculated values of ΔF_{ins}.

Discussion

In order to decipher the physicochemical properties that control the rate of passive lipid transport between membranes, we systematically investigated how the free energy barriers for lipid desorption and insertion vary with lipid chemistry and membrane phase. Specifically, we quantified the barriers for transporting glycerophospholipids, the major structural component of cell membranes,[1,2] with a diversity of acyl chains and headgroups. To do so efficiently, we calculated free energy profiles as a function of the reaction coordinate r_{LxS}, which we first identified for DMPC transport between L_α phase membranes in Chapter 2[223] and here assumed applicable to similar systems. All of the free energy profiles exhibit a rate-limiting barrier at nearly the same value of the reaction coordinate (Figure 3.2A, 3.3A, and 3.4A), supporting our assumption that r_{LxS} is a good reaction coordinate for all glycerophospholipids; regardless of glycerophospholipid chemistry or membrane phase, the dynamics of lipid transport are best described in terms of the breakage and formation of hydrophobic lipid–membrane contacts. In all cases, the free energy barrier for lipid desorption is at least twice the barrier for lipid insertion, confirming that desorption limits the rate of lipid transport and that the free energy barrier for desorption is equivalent to the activation free energy barrier for lipid transport.

We have revealed a correlation between the activation free energy and the hydrophobicity of a lipid's local membrane environment, as measured by the average number of close hy-

drophobic contacts that a lipid forms with surrounding membrane lipids, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ (Figure 3.5). $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, and hence the activation free energy, is sensitive to subtle changes in a membrane's chemistry and structure. In general, lipids form more hydrophobic contacts in thick, tightly packed, and ordered membranes (Figure B.7). Membranes with these characteristics are typically composed of lipids with long, saturated acyl chains, lipids with headgroups that facilitate tight lipid packing, or lipids in ordered phases. Indeed, lipids with long, saturated acyl chains exhibit increased activation free energies (Figure 3.2B) and slower experimentally measured transfer rates[95–101,103,161–163,220,221] than counterparts with unsaturated and shorter acyl chains; conical lipids with smaller headgroups that pack well, for example PE,[47,62,240] also have increased activation free energies (Figure 3.3B) and slower experimentally measured transfer rates[95,163] compared to cylindrical lipids with bulkier headgroups such as PC; and lipids in highly ordered L_β phase membranes exhibit the largest activation free energies (Figure 3.4B) and slowest experimentally measured transfer rates.[96,98,222]

We report an average increase in the activation free energy of 0.62 kcal/mol per methylene unit added to the tails of a saturated lipid due to an average increase in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ of 103 contacts per methylene unit. Previous *in vitro* experiments have demonstrated that lipid transfer rates decrease with increasing acyl chain length,[95–101,103,161–163,220,221] estimating an increase in the activation free energy of $0.32 - 0.65$ kcal/mol per methylene unit.[163,220,221] Our results are in quantitative agreement with these values. Furthermore, our finding that the activation free energy depends on the hydrophobic environment around the desorbing lipid may explain why different values are reported from experiments using different donor membranes. Changing the desorbing lipid's tail length alters the average number of hydrophobic contacts it makes with surrounding lipids, $\langle n_{\mathrm{CC}} \rangle$, and correspondingly alters its activation free energy. However, if donor membranes contain components capable of regulating $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, then desorbing lipids with different tail lengths may not exhibit substantially different $\langle n_{\mathrm{CC}} \rangle$, resulting in similar activation free energies. Proteins, for example, could act as such regulating components. Consistent with this hypothesis, smaller changes in the activation free energy were determined from experiments that utilized model high-density lipoproteins (HDL) composed of apolipoprotein A-I (apo A-I) and POPC as donors[220] compared to those using large unilamellar vesicles composed of only PC lipids.[163,221] Apo A-I wraps around lipids in model HDL to create a discoidal morphology,[241] suggesting that it may regulate $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ by controlling the membrane's overall structure. Alternatively, changes in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ could be directly compensated by specific protein–lipid interactions.

We report an average increase of 4.3 kcal/mol in the activation free energy for a L_β phase membrane compared to a L_α one due to an average increase in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ of 762 contacts. Due to the exponential dependence of the rate on the activation free energy, this increase corresponds to a lipid transport rate that is roughly 3 orders of magnitude slower for L_β compared to L_α phase membranes. In agreement with our results, *in vitro* assays have demonstrated that lipid transport between L_β phase membranes is substantially slowed compared to transport between L_α phase membranes, with an increased activation free energy of $1.6 - 12.7$ kcal/mol.[96,98,222] We suggest that the wide range of activation free energies reported likely reflects differences in the hydrophobicity of the donor membranes

used in the experiments. The smallest difference is reported for lipid transport between highly curved small unilamellar vesicles,[98] which are more disordered than planar membranes.[242,243] Thus, they likely have reduced $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, and correspondingly reduced activation free energy, compared to other experimental L_β phase systems.[96,222]

We report an increase in the activation free energy for lipids with different headgroups in the order PG < PC < PA < PS < PE due to an increase in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ in the same order. We cannot directly compare our results to experiments since significantly different donor membrane compositions were used. While we used donor membranes composed of the same lipid species as the one being transferred, donor membranes composed of predominantly PC lipids were used in experiments.[95,103,162,163] Nevertheless, we note that in multiple experiments, PE lipids are transferred at slower rates than PC lipids,[95,163] consistent with our results. However, no consensus ranking of transfer rates for lipids with other headgroups has been reached based on experiments.[95,103,162,163] Results from even the same experimental setup have often been difficult to interpret,[95,103,163] most likely because no clear trend in the activation energy can be identified based on the chemical structure of the headgroup alone. Our finding that the activation free energy depends on $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, which depends on lipid headgroup chemistry in a nontrivial way, may help guide the design of experiments to conclusively assess the impact of headgroup on transfer rates.

In contrast to the activation free energy barrier for lipid desorption, trends in the free energy barrier for lipid insertion can be explained by chemical features of the headgroup, especially net charge, presence of terminal acidic groups, and molecular size (Figure 3.3C). Changing any of these chemical features modulates the membrane's interfacial hydrophobicity by altering the strength of membrane–solvent interactions (Table B.2). A membrane's interfacial hydrophobicity can also be tuned by changing lipid tail chemistry or membrane phase so that packing defects expose more of the hydrophobic core (Figure B.6 and Table B.2). Membranes with increasingly hydrophobic surfaces, characterized by weaker headgroup–solvent interactions and larger packing defects, exhibit reduced insertion free energy barriers (Figure 3.6) and correspondingly faster lipid insertion rates. Unfortunately, the rate of lipid insertion is too fast to monitor with standard experimental techniques, so we cannot compare our calculated free energy barriers for insertion to experimental ones. Nevertheless, one of the parameters that we find controls the rate of lipid insertion is similar to the membrane property that is predictive of experimentally measured rates of nanoparticle insertion into PC membranes: Nanoparticle insertion rates depend on the prevalence of low density areas at the membrane surface, which are physically similar to packing defects.[244] Given this similarity, we suspect that accounting for changes in a membrane's interfacial hydrophobicity through additional measures such as the membrane–solvent interaction energy may be necessary to predict nanoparticle insertion rates into more complex membranes that contain lipids with diverse headgroups and better mimic cell membrane compositions.

In summary, membrane hydrophobicity determines the rate of passive lipid transport. Our discovery that the rate of lipid transport depends on the physicochemical properties of the donor membrane within 1 nm of the transferring lipid suggests that transport rates can be regulated, for example, by membrane proteins and local changes in membrane composi-

tion that alter the hydrophobicity of the membrane core. Sequestering lipids into ordered, tightly packed lipid rafts,[40] where a significant expenditure of energy is required to extract a lipid, may create pools of non-exchangeable lipids within a single membrane. Membrane physicochemical properties vary even more substantially between organelles, creating two cell membrane territories: Loose lipid packing and minimally charged membrane surfaces define the territory found among membranes of the early secretory pathway, whereas tight lipid packing and highly anionic membrane surfaces define the territory found among late secretory membranes.[47,77] Given that tight lipid packing increases the hydrophobicity of the membrane core and that stronger membrane electrostatics decrease interfacial hydrophobicity, we suggest that the free energy barriers for lipid transport may transition from low to high between these two territories.

Within cells, lipid transfer proteins are largely responsible for specifically transporting individual lipids between membranes.[85,87,104] Lipid transfer proteins may generally enhance the rate of lipid desorption by disrupting the hydrophobic environment around a target lipid. Those that transport lipids from membranes with highly hydrophobic cores, such as may be found among late secretory membranes, may have acquired additional mechanisms to overcome increased activation free energies. Similarly, lipid transfer proteins may aid lipid insertion by increasing the surface hydrophobicity of the membrane. For example, multiple lipid transfer proteins contain basic surface regions that enhance binding to anionic membranes[110,114,245] and that could also compensate for the disruption of especially favorable interactions between solvent and anionic lipids during insertion. Thus, the membrane physicochemical properties that control the rate of passive lipid transport may also be exploited to direct non-vesicular lipid traffic *via* lipid transfer proteins.

Chapter 4

Ceramide-1-phosphate Transfer Protein Enhances Lipid Transport by Disrupting Hydrophobic Lipid–Membrane Contacts

Introduction

Bioactive sphingolipids play critical roles in most cellular processes, including cell growth, differentiation, migration, and survival.[15,246] Normal biological function requires spatiotemporal regulation of sphingolipid levels, and altered distributions are linked to the pathophysiology of numerous diseases, including cancers, inflammatory diseases, metabolic syndromes, and neurological disorders.[15–17] Precise sphingolipid distributions are maintained, in part, through non-vesicular transport of specific lipids by lipid transfer proteins.[85,88] One such protein, ceramide-1-phosphate transfer protein (CPTP) transports ceramide-1-phosphate (C1P) from the *trans*-Golgi to target membranes.[247] Elevated levels of C1P in the *trans*-Golgi overstimulate the production of pro-inflammatory eicosanoids by group IVA cytosolic phospholipase A_2. By selectively trafficking C1P from the *trans*-Golgi, CPTP modulates eicosanoid production and, thus, helps reduce the risk of chronic inflammation.[247,248]

CPTP acts as a cytosolic shuttle that transports C1P from the *trans*-Golgi membrane to nuclear and plasma membranes.[247] Experiments suggest that lipid transfer occurs through the following steps: CPTP binds the donor membrane in its apo form to initiate a transfer event. CPTP then extracts C1P from the membrane. Once C1P is housed inside its hydrophobic cavity, CPTP unbinds from the membrane and diffuses through solution to reach an acceptor membrane. After the C1P-bound form binds the acceptor membrane, C1P exits the hydrophobic cavity and inserts into the membrane. Unbinding of CPTP, which is now in its apo form, restarts the transfer cycle.

Structural and biochemical studies have elucidated further details about how CPTP functions as a C1P transfer protein.[247] CPTP has a two-layered, all α-helical topology that is homologous with the glycolipid transfer protein fold. At the interior of its sandwich-like

structure is a hydrophobic cavity that shields the hydrophobic tails of a C1P from unfavorable exposure to the cytosol during transport. Although the hydrophobic cavity adaptively expands to accommodate C1P species with acyl chain lengths ranging from 2 to 18 carbons, CPTP preferentially transports species with chain lengths of 16 or 18 carbons. CPTP selectively targets C1P by recognizing its phosphate headgroup through a triad of cationic residues located within a positively-charged surface cavity at the entrance to the hydrophobic cavity and by also recognizing its acyl-amide moiety through a bifurcated hydrogen bond. Based on crystallographic measurements, a cleft-like gating mechanism has been proposed to facilitate C1P entry into and exit from the hydrophobic cavity. Furthermore, the observation of two C1P-binding modes suggests that C1P entry and exit occur through a two-step mechanism whereby the acyl chain enters into the cavity before the sphingosine during entry and leaves after the sphingosine during exit.

However, because structures were solved in the absence of a membrane in the aforementioned studies, it is unclear if CPTP uses the proposed mechanism to efficiently extract (or insert) C1P from (or into) a membrane. Furthermore, how CPTP overcomes the significant free energy barrier for passive lipid desorption from a membrane, which limits the rate of passive lipid transport,[97–99,161,223] is unknown. In Chapters 2 and 3, we demonstrated that the free energy barrier for passive lipid desorption reflects the thermodynamic cost of breaking hydrophobic lipid–membrane contacts,[223] and that it can be lowered by decreasing the hydrophobicity of the transferring lipid's local membrane environment.[249] Thus, we hypothesize that CPTP rapidly extracts C1P from a membrane by disrupting C1P–membrane hydrophobic contacts, thereby lowering the free energy barrier for lipid desorption and catalyzing C1P transport.

To test this hypothesis, we investigate the molecular mechanism by which CPTP extracts and inserts C1P into a membrane using molecular dynamics simulations. Significant structural differences between the apo and C1P-bound forms, observed in both published crystal structures[247] and our solution simulations, indicate that conformational changes are required to accommodate C1P in CPTP's hydrophobic cavity. We find that CPTP exhibits a single favorable membrane binding pose, which aptly positions the entrance to its hydrophobic cavity at the membrane interface. Membrane binding promotes rearrangements of CPTP's helices much like that expected for a cleft-like gating mechanism to facilitate C1P entry and exit into its hydrophobic cavity. When closed, a gating helix anchors the apo form deeply into the membrane through strong electrostatic interactions. Furthermore, the apo form substantially disrupts a lipid's local hydrophobic environment in the membrane below. As a result, CPTP lowers the free energy barrier for C1P extraction. Uptake of C1P may then trigger conformational changes that prompt membrane unbinding through an electrostatic switching mechanism such as used by other lipid transfer proteins.[109] Our proposed molecular mechanism not only sheds light on how CPTP efficiently traffics C1P between membranes but also offers explanations for previous experimental findings. Significantly, we provide, to our knowledge, the first biophysical characterization of an essential catalytic step of all lipid transfer proteins at atomic resolution.

Results and Discussion

We use a multiscale approach that builds upon previous structural studies and our earlier studies of passive lipid transport to efficiently and accurately characterize key elementary steps of C1P transport by CPTP. Specifically, we model the transport of C1P with a saturated, 16-carbon acyl chain, which is CPTP's most likely *in vivo* substrate,[247] between membranes composed of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-PC (POPC) lipids, which have been used in experimental transfer assays,[250] by human CPTP. We use atomistic simulations to resolve conformational changes that occur upon membrane binding and to capture C1P extraction from and insertion into a membrane, whereas we use coarse-grained simulations to capture membrane binding events, which require multiple microsecond-long simulations.

CPTP's Hydrophobic Cavity Must Expand to Accommodate C1P

First, we examine how the structures of the apo and C1P-bound forms of CPTP differ in solution. To initialize solution simulations, we use the most complete crystal structure of CPTP with C1P fully bound in its hydrophobic cavity (PDB 4K85). Final structures from our solution simulations of the apo and C1P-bound forms are shown in Figures 4.1A-4.1C. Upon artificial removal of C1P, CPTP's helices reorient themselves, closing off the entrance to the hydrophobic cavity and barring the entry of solvent molecules. In particular, helices $\alpha 2$ and $\alpha 4$, which are located on opposite layers of CPTP's sandwich topology, angle into the hydrophobic cavity and fill the space otherwise occupied by C1P (Figure 4.1A). Due to the closer packing of multiple helices in the apo form, the hydrophobic cavity collapses (Figure 4.1B and 4.1C). The total volume of the hydrophobic cavity in the apo form is on average only half the volume of the cavity in the C1P-bound form (Figure 4.1D). Consistent with conclusions made based on crystal structures,[247] our results indicate that CPTP must undergo significant conformational changes to accommodate C1P in its hydrophobic cavity.

Membrane Binding Poises CPTP to Extract C1P

We next characterize how CPTP, in both apo and C1P-bound forms, binds to a POPC membrane. To observe spontaneous binding events, we performed coarse-grained simulations. While our coarse-grained model accurately recapitulates the structure and internal dynamics of CPTP observed in our atomistic solution simulations (Figures C.1 and C.2), it cannot capture conformational changes that may occur upon membrane binding. Thus, we converted membrane-bound, coarse-grained configurations back into all-atom representations in order to resolve conformational changes that may result from CPTP–membrane interactions.

A single binding pose is observed in coarse-grained simulations, and CPTP remains favorably bound to the membrane in the same overall pose in all-atom simulations (Figure 4.2). In both apo and C1P-bound forms, the loop between helices $\alpha 1$ and $\alpha 2$ and helix $\alpha 6$ inserts into the membrane's hydrophobic core (Figure 4.2A). A combination of hydrophobic

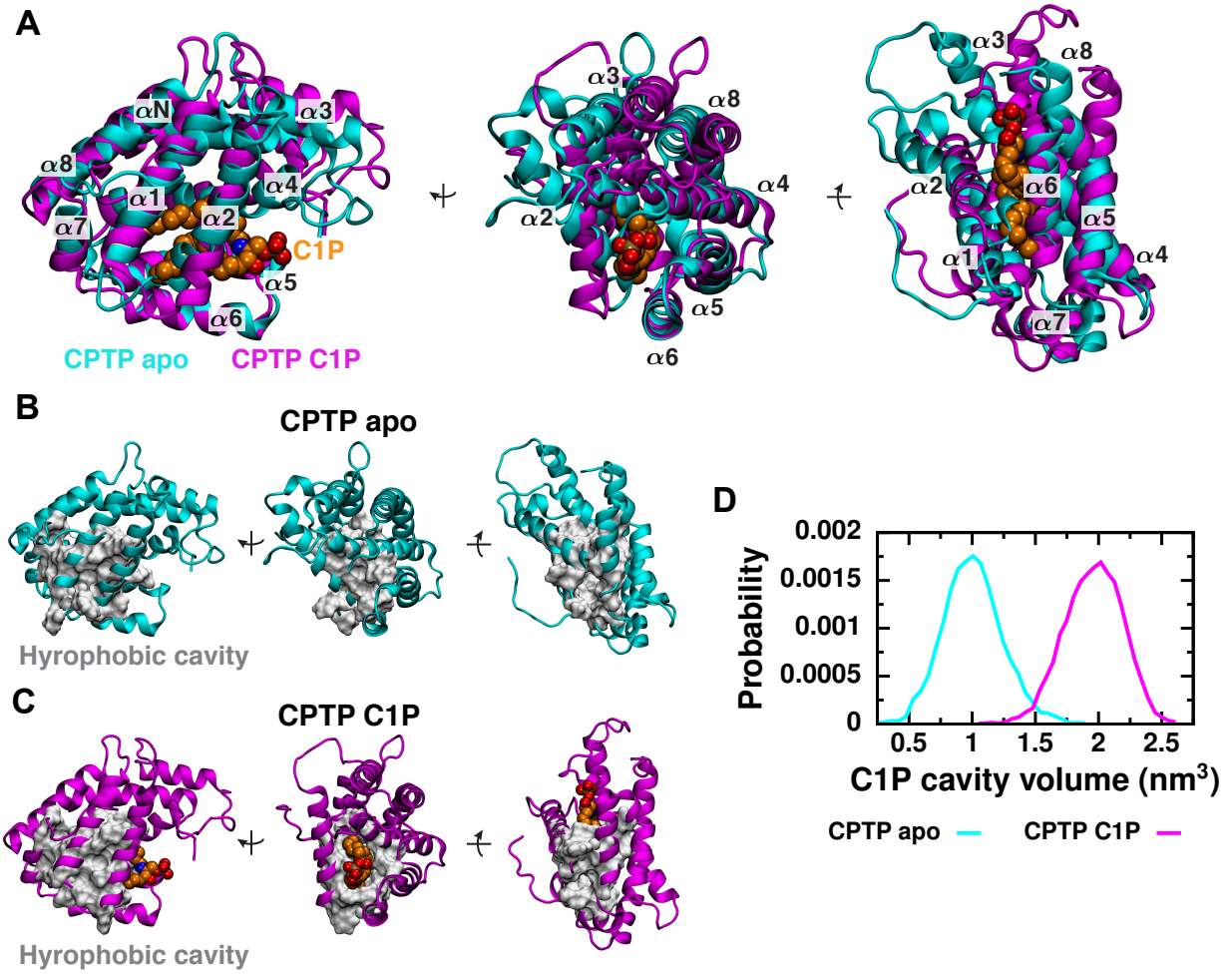


Figure 4.1: In solution, conformational differences between the apo and C1P-bound forms of CPTP tune the accessibility and volume of the hydrophobic cavity. (A) Superimposed structures of the apo form (cyan) and C1P-bound form (magenta) obtained from solution-phase simulations are shown from three views. Structures were aligned to minimize the root-mean-square deviation (RMSD) of Cα atoms of helix α6. α-helices (αN and $\alpha 1 - \alpha 8$) are numbered from amino (N) to carboxy (C) termini. C1P is rendered as van der Waals spheres and colored orange. (B and C) Residues lining the hydrophobic cavity are shown as a gray surface on the (B) apo and (C) C1P-bound structures. (D) Distributions of the total volume of the hydrophobic cavity, including volume occupied by C1P, during solution simulations of the apo and C1P-bound forms.

and aromatic residues, specifically I49, F50, and F52 of the α1-α2 loop and W152, V153, and V160 of helix α6, anchor CPTP to the membrane as predicted from cystral structures.[247] When bound to the membrane, the entrance to CPTP's hydrophobic cavity is aptly positioned at the membrane–solvent interface (Figure 4.2B and 4.2C). Additionally, the cleft

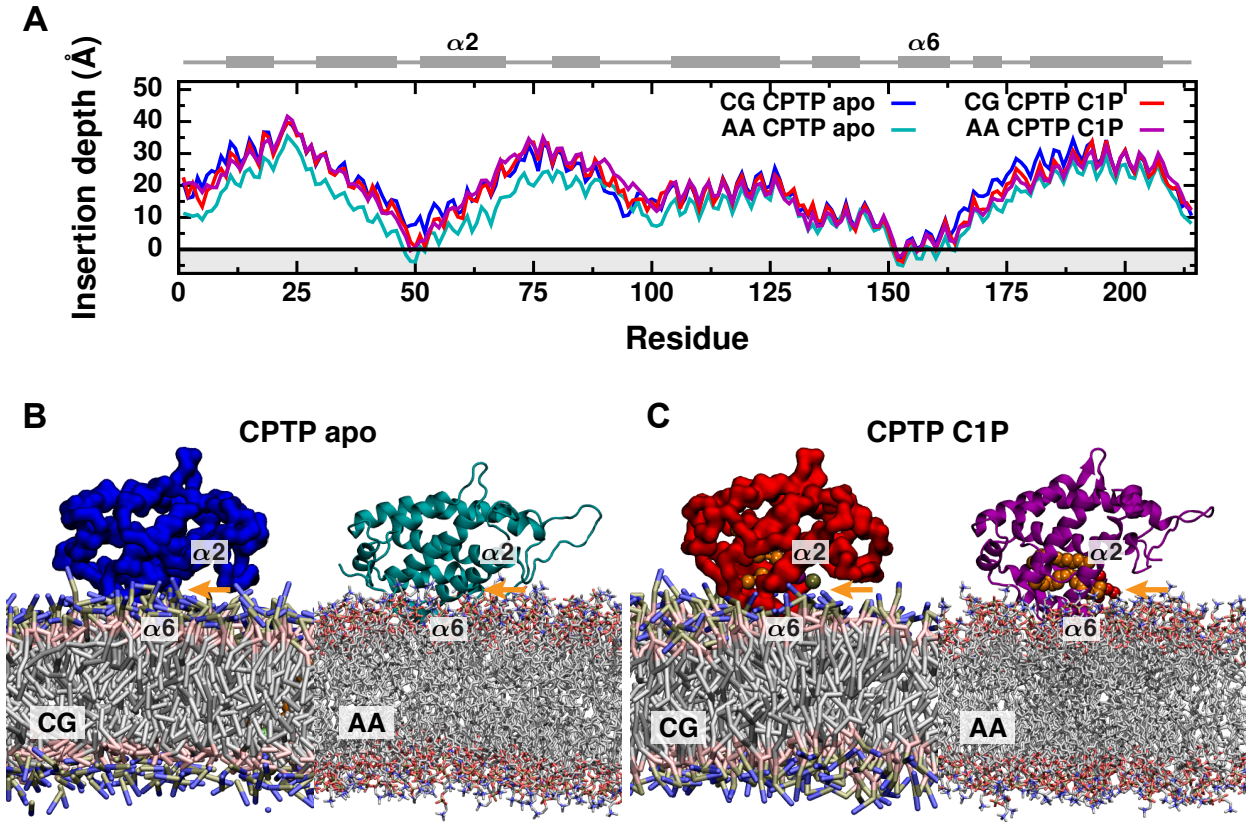


Figure 4.2: Both apo and C1P-bound forms of CPTP bind a membrane poised to extract or insert C1P. (A) Average insertion depth of each residue measured relative to the average position of lipid phosphate groups (black line) from coarse-grained (CG) and all-atom (AA) simulations. The gray region highlights the location of the membrane's hydrophobic core. CPTP's secondary structure is schematically illustrated above with helices represented as rectangles and unstructured loop regions as lines. (B and C) Representative configurations of the (B) apo form and (C) C1P-bound form of CPTP bound to a POPC membrane for both coarse-grained and all-atom representations. The entrance to CPTP's hydrophobic cavity is indicated with an orange arrow. C1P is rendered as van der Waals spheres and colored orange.

between CPTP's two α-helical layers faces into the membrane's hydrophobic core, providing an avenue for C1P to move from the membrane into CPTP's hydrophobic cavity without exposing its tails to solvent.

Subtle differences in the binding pose of the apo form observed in all-atom versus coarse-grained representations (Figures 4.2A and 4.2B) suggest that conformational changes may be coupled to membrane binding. Indeed, membrane binding promotes a widening of the entrance to CPTP's hydrophobic cavity. Helix α2, which partially gates the entrance to the hydrophobic cavity (Figure 4.1A), is rotated outward in membrane bound structures of

both the apo and C1P-bound forms compared to their respective solution structures (Figure 4.3A). We quantify this conformational change by the angle $\theta_{\alpha 2}$ between helix $\alpha 2$ and the

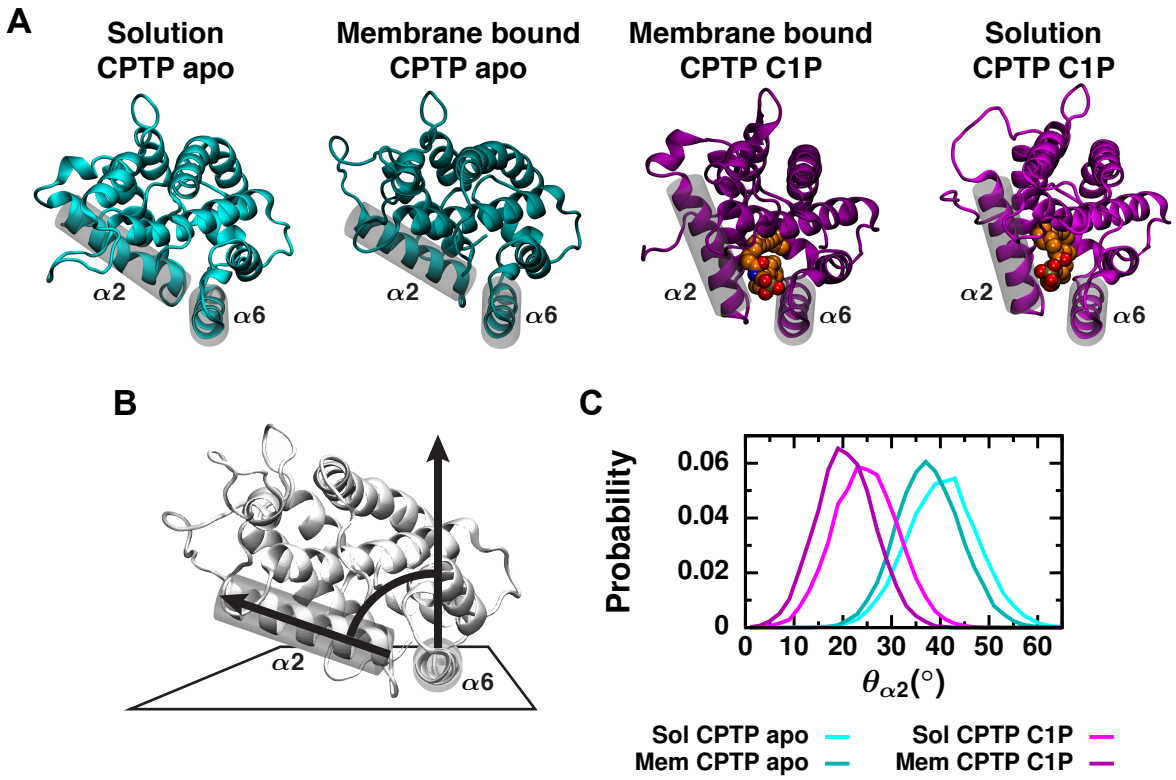


Figure 4.3: Conformational changes that promote the opening of gating helix $\alpha 2$ are coupled to membrane binding. (A) Orientation of helix $\alpha 2$ relative to helix $\alpha 6$ is highlighted on structures of both the apo and C1P-bound forms in both solution and bound to the membrane. Both helices are outlined by a gray cylinder. C1P is rendered as van der Waals spheres and colored orange. (B) Definition of angle $\theta_{\alpha 2}$ used to characterize gate opening. $\theta_{\alpha 2}$ is the angle between the helical axis of $\alpha 2$ and a vector normal to the plane of the top surface of helix $\alpha 6$ when CPTP is membrane bound. This plane nearly corresponds to the location of the membrane–solvent interface when CPTP is membrane bound. (C) Distributions of $\theta_{\alpha 2}$ from all-atom simulations of the apo and C1P-bound forms of CPTP both in solution and bound to the membrane.

axis normal to the membrane when CPTP is bound. In order to calculate $\theta_{\alpha 2}$ for both solution and membrane-bound structures, we define the latter vector as the vector normal to the plane that contains the top surface of helix $\alpha 6$ when CPTP is membrane bound (Figure 4.3B). We use helix $\alpha 6$ as an internal reference point since helix $\alpha 6$ is approximately parallel to the membrane surface when CPTP is membrane bound (Figure 4.2) and has a rather rigid structure. A decrease in $\theta_{\alpha 2}$ indicates opening of the gate created by helix $\alpha 2$, whereas an increase in $\theta_{\alpha 2}$ indicates closing. As shown in Figure 4.3C, membrane bound structures of CPTP indeed have reduced $\theta_{\alpha 2}$ compared to solution structures on average. By locking

down the gate created by helix $\alpha 2$ when in solution, the hydrophobic cavity, and any lipid inside, remains sealed off from solvent. Widening of the entrance to the hydrophobic cavity upon membrane binding then readies CPTP to extract or insert C1P and is consistent with the proposed cleft-like gating mechanism for C1P entry and exit into the hydrophobic cavity.

Both the apo and C1P-bound forms of CPTP bind the membrane in the same overall pose and exhibit conformational changes of similar magnitude upon binding; however, such conformational changes enable the apo form to insert more deeply into the membrane's hydrophobic core (Figure 4.2A and 4.2B). In the apo form, a slight widening of the entrance to the hydrophobic cavity brings helix $\alpha 2$ closer to the membrane surface. As a result, energetically favorable interactions between residues on helix $\alpha 2$ and the membrane are present in the apo form but not the C1P-bound form (Figure 4.4). In particular, two positively charged residues, R62 and R66, form significantly favorable electrostatic interactions with the membrane. By anchoring helix $\alpha 2$ to the membrane, such favorable interactions may facilitate the opening of the gate created by helix $\alpha 2$ in the apo form. Although our simulated membranes are composed of zwitterionic POPC lipids, the presence of anionic lipids may further strengthen such interactions and thereby facilitate gate opening. Consistent with this idea, including anionic phosphatidylserine lipids in donor membranes has been shown to enhance C1P transport in *in vitro* assays.[250] Complete opening of the gate created by helix $\alpha 2$ to accommodate C1P inside CPTP's hydrophobic cavity breaks these favorable electrostatic interactions with membrane. Thus, gate opening may facilitate membrane unbinding in a manner reminiscent of the electrostatic switching mechanism used by other cytosolic lipid transfer proteins that have well-defined lids to their hydrophobic cavities.[109]

When bound to the membrane, CPTP also influences the molecular structure and physical properties of the membrane. Especially relevant to lipid transport is how CPTP impacts nearby lipids' hydrophobic membrane environments. In Chapter 3, we demonstrated that the activation free energy for passive lipid transport correlates with the average number of hydrophobic contacts that a lipid makes with surrounding membrane lipids, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. As $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ decreases, the rate of passive lipid desorption from a membrane increases.[249] Figure 4.5 demonstrates how the apo and C1P-bound forms of CPTP impact $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. The C1P-bound form only minimally reduces the number of hydrophobic contacts that nearby lipids make with other membrane lipids. In contrast, the apo form significantly disrupts hydrophobic contacts that lipids within roughly $1-2$ nm of its binding location make with other membrane lipids. The membrane lipids closest to the apo form of CPTP make only 1,650 contacts with other membrane lipids, whereas when CPTP is not bound to the membrane, POPC lipids form an average of 2,340 contacts with each other.[249] Based on the relationship between $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ and the free energy barrier for passive lipid desorption provided in Chapter 3 (Figure 3.5),[249] this reduction in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ by 690 contacts corresponds to a 4 kcal/mol decrease in the free energy barrier. Thus, the apo form of CPTP may efficiently extract C1P by reducing the barrier for hydrophobic C1P–membrane contact breakage.

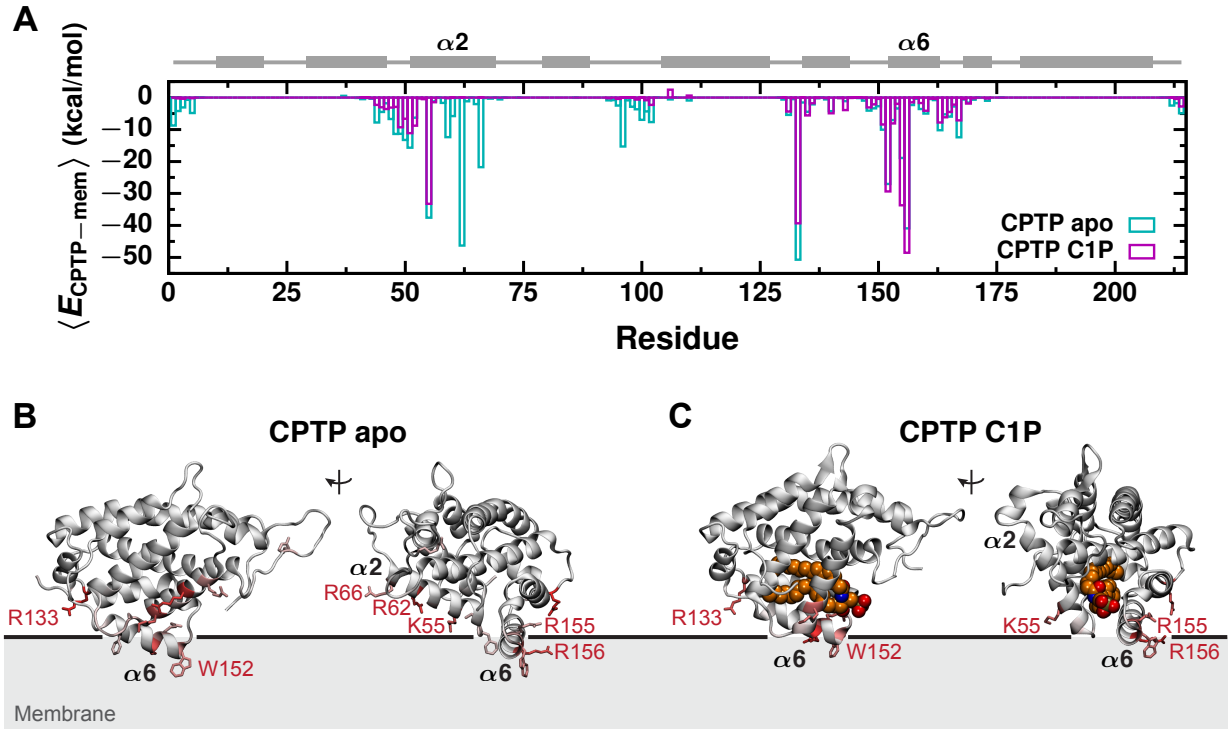


Figure 4.4: Favorable electrostatic interactions anchor the apo form deeply into the membrane. (A) Average interaction energy between each protein residue and the membrane, $\langle E_{\mathrm{CPTP-mem}} \rangle$, from all-atom simulations of the apo and C1P-bound forms of CPTP bound to a POPC membrane. CPTP's secondary structure is schematically illustrated above with helices represented as rectangles and unstructured loop regions as lines. (B and C) $\langle E_{\mathrm{CPTP-mem}} \rangle$ mapped onto the structures of the (B) apo and (C) C1P-bound forms of CPTP bound to the membrane. $\langle E_{\mathrm{CPTP-mem}} \rangle$ is indicated by the color of the residue. Residues that most favorably interact with the membrane are colored dark red, those that weakly or do not interact with the membrane are colored gray. Side chains of residues with $\langle E_{\mathrm{CPTP-mem}} \rangle < -12$ kcal/mol are rendered. Residues with $\langle E_{\mathrm{CPTP-mem}} \rangle < -18$ kcal/mol are labeled. C1P is rendered as van der Waals spheres and colored orange. The black line indicates the average position of phosphate groups of membrane lipids, and the membrane's hydrophobic core is shown in gray.

CPTP Efficiently Extracts C1P by Lowering the Free Energy Barrier for Breaking Hydrophobic C1P–Membrane Contacts

To determine if CPTP lowers the free energy barrier for hydrophobic C1P–membrane contact breakage, we performed free energy calculations. Specifically, we calculated free energy landscapes along two coordinates that may capture key intermediates of CPTP facilitated C1P transport: (1) The reaction coordinate for passive lipid (L) transport via solvent (xS), r_{LxS}, defined in Eq. 3.1 of Chapter 3 describes interactions between C1P and the membrane

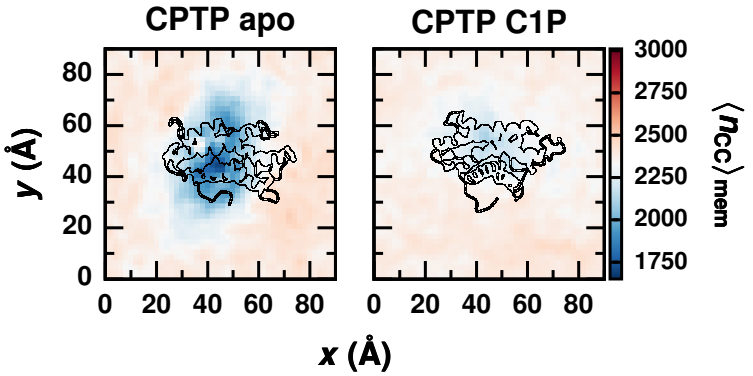


Figure 4.5: The apo form of CPTP significantly disrupts the local hydrophobic membrane environment of lipids below. Average number of hydrophobic contacts that a lipid in the membrane makes with surrounding lipids, $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$, as a function of its displacement in the xy plane from CPTP. Color scale is set relative to $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ for a POPC membrane in the absence of CPTP; membrane regions where lipids make more contacts than an unperturbed POPC membrane are colored red, and regions where lipids make less are colored blue. The structure of the apo and C1P-bound forms of CPTP bound to the membrane are outlined for reference.

and monitors progress from configurations with C1P in the membrane (which exhibit many hydrophobic lipid–membrane contacts and large values of r_{LxS}) to configurations with the lipid outside the membrane (which lack hydrophobic lipid–membrane contacts and have negative values of r_{LxS});[223,249] (2) The fraction of contacts C1P makes with CPTP when fully inside its hydrophobic cavity, Q, monitors progress from configurations with C1P outside CPTP (which have values of Q close to zero) to configurations with C1P inside CPTP (which have values of Q close to one). Specifically, Q is calculated according to

$$Q = \frac{1}{N} \sum_{(i,j)} \left[1 + \exp\left[\beta (r_{ij} - \lambda r_{ij}^{(0)}) \right] \right]^{-1}, \tag{4.1}$$

where N is the total number of contact pairs (i, j) between an atom i of C1P and atom j of CPTP, r_{ij} is the distance between atoms i and j, $r_{ij}^{(0)}$ is a reference distance between atoms i and j, which we take to be the average distance between the atom pair in solution simulations of the C1P-bound form, $\beta = 50\ \mathrm{nm}^{-1}$, and $\lambda = 1.8$. Contact pairs used to calculate Q include: (1) hydrophobic contacts between C1P and residues that line CPTP's hydrophobic cavity and (2) nonpolar contacts between C1P's headgroup and sphingoid backbone and residues responsible for C1P recognition.

A preliminary free energy landscape is shown in Figure 4.6A. At the global free energy minimum, C1P resides in the membrane where it makes many hydrophobic contacts with surrounding membrane lipids but does not substantially interact with CPTP. An example of a configuration found within the global free energy minimum is shown in Figure 4.6D1.

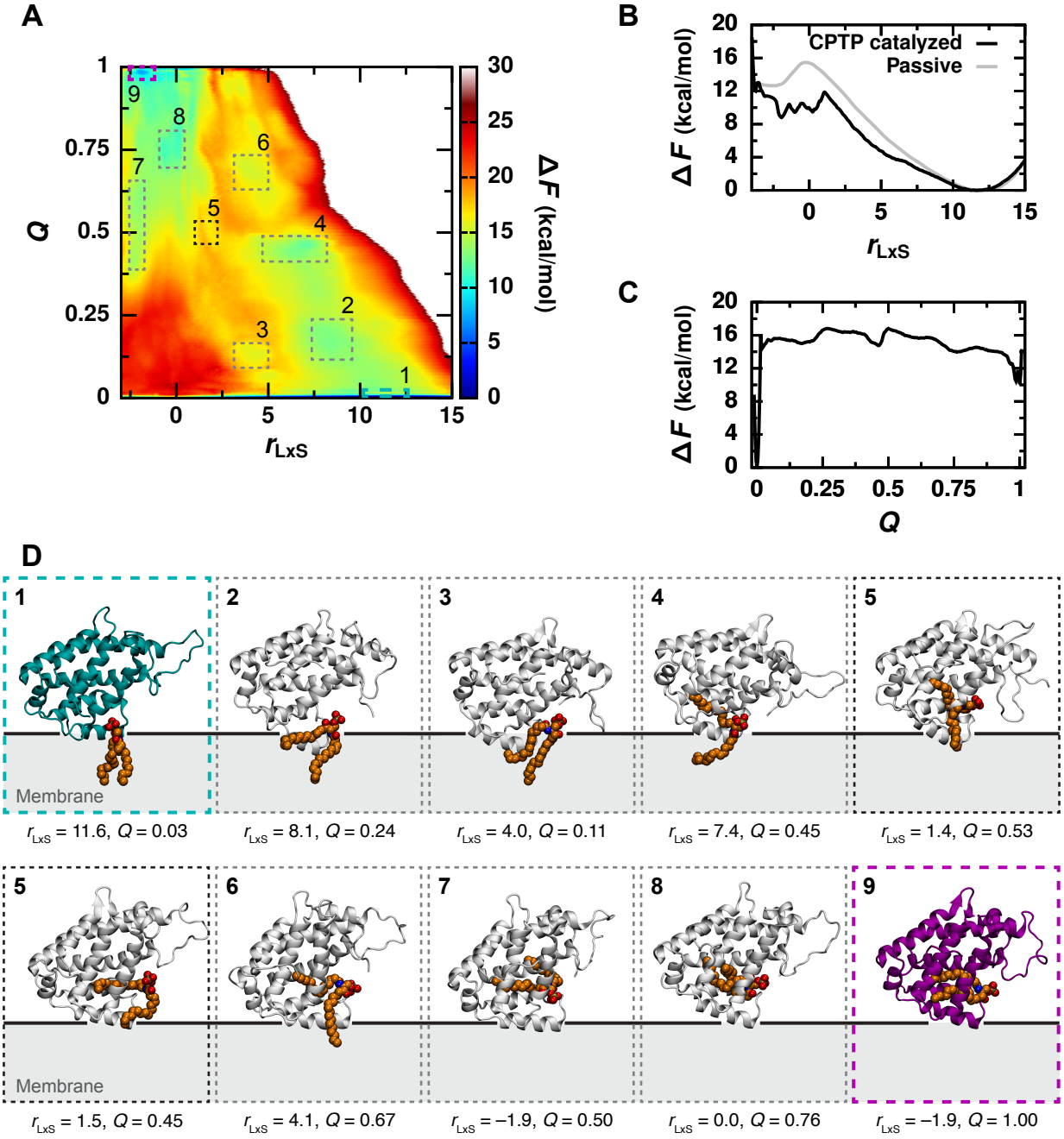


Figure 4.6: CPTP lowers the free energy barrier for hydrophobic C1P–membrane contact breakage. (A) Preliminary free energy landscape along r_{LxS} and Q. (B and C) Free energy profiles along (B) r_{LxS} or (C) Q obtained from the free energy landscape in (A). (D) Example configurations from different regions of the free energy landscape outlined in (A). C1P is rendered

(Continued on next page)

Figure 4.6: *(Continued from previous page)*
as van der Waals spheres and colored orange. The apo form of CPTP in configuration 1 is colored dark cyan, and the C1P-bound form in configuration 9 is colored dark magenta. The black line indicates the average position of phosphate groups of membrane lipids, and the membrane's hydrophobic core is shown in gray.

At the second most favorable minimum in the free energy landscape, C1P resides within CPTP's hydrophobic cavity but does not form hydrophobic contacts with membrane lipids. An example of a configuration found at this local minimum is shown in Figure 4.6D9. This free energy profile is consistent with CPTP's observed function as a C1P transporter; CPTP serves as a free energetically favorable vehicle for C1P transport between even more favorable states in which C1P is part of a membrane.

Figure 4.6B shows preliminary free energy profiles along r_{LxS} both in the presence and absence of CPTP. The free energy profile for passive C1P transport exhibits a rate-limiting activation free energy barrier of 15.5 kcal/mol at $r_{\mathrm{LxS}} \approx 0$. When CPTP is present, the free energy barrier for C1P desorption from the membrane is reduced to 11.9 kcal/mol and shifted to slightly larger values of r_{LxS}. CPTP lowers the free energy barrier by 3.6 kcal/mol, corresponding to a roughly 500 fold speedup in the rate of lipid transport at room temperature. Furthermore, this value in reasonable agreement with the one predicted above based on the maximal reduction in $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ caused by the apo form. By disrupting C1P's local hydrophobic environment in the membrane (Figure 4.5), CPTP lowers the free energy barrier that limits the rate of C1P desorption from a membrane and, thus, catalyzes lipid extraction. Because CPTP provides C1P with another favorable hydrophobic environment, C1P enters CPTP's hydrophobic cavity after fewer hydrophobic contacts with the membrane have been broken than required for C1P desorption into solvent, which offers no favorable hydrophobic environment. As a result, the free energy barrier for lipid desorption is shifted to slightly larger values of r_{LxS}, which are indicative of more C1P–membrane contacts, when CPTP is present. Indeed, as illustrated by configurations around this free energy barrier in Figure 4.6D5, C1P maintains a few hydrophobic contacts with the membrane while also forming contacts with CPTP.

A preliminary free energy profile along Q (Figure 4.6C) suggests that C1P entry into CPTP's hydrophobic cavity may require passage through multiple metastable intermediates. Configurations representative of potential intermediates are shown in Figure 4.6D. Consistent with the two C1P-binding modes to CPTP resolved by crystallography, we observe metastable intermediates with one tail of C1P inside CPTP's hydrophobic cavity and the other still within the membrane. Thus, our findings support a two-step mechanism whereby each tail of C1P enters into and exits from CPTP's hydrophobic cavity individually.

Conclusions

Based on atomistic simulations of CPTP in the presence of a membrane, we propose a molecular mechanism for how CPTP functions as a C1P transfer protein that is consistent with the previously suggested cleft-like gating mechanism:[247] To initiate a transfer event, the apo form of CPTP, whose collapsed hydrophobic cavity is closed off to C1P entry, binds a donor membrane. Interactions between the membrane and gating helix α2 and also helix α6 anchor CPTP deeply into the membrane and aptly position the entrance to its hydrophobic cavity at the membrane-solvent interface. Strong electrostatic interactions between positively charged residues on gating helix α2 facilitate cleft widening and gate opening. Poised for C1P entry into its hydrophobic cavity, CPTP significantly disrupts C1P's local membrane environment thereby lowering the free energy barrier for C1P extraction from the membrane. Complete uptake of C1P into CPTP's hydrophobic cavity further opens gating helix α2. Such conformational changes break strong electrostatic interactions with the membrane that had previously anchored the apo form, potentially facilitating membrane unbinding through an electrostatic-switching-like mechanism.[109] Having extracted C1P from a donor membrane, CPTP then inserts C1P into a target membrane, a free energetically downhill event, to complete a transfer event. While this molecular mechanism is partially based on preliminary free energy calculations, simulations to more rigorously evaluate its potential validity are ongoing.

Significantly, we demonstrate how CPTP enhances the rate of C1P transport by disrupting hydrophobic C1P–membrane contacts and lowering the free energy barrier for lipid extraction. To our knowledge, we provide the first molecular evidence that lipid transfer proteins catalyze lipid extraction from and insertion into a membrane. Because the catalytic mechanism described here is closely related to that of passive lipid transport,[223,249] we suspect it may be generally used by lipid transfer proteins, not just CPTP, to increase the rate of lipid transport between cell membranes.

Methods

Molecular Dynamics Simulations

We performed molecular dynamics simulations of CPTP both in solution and in the presence of a POPC membrane at both all-atom and coarse-grained resolutions using GROMACS 2019.[166] The CHARMM36 force field[177] in combination with the CHARMM TIP3P water model[178] was used for the all-atom simulations. CHARMM36 force field parameters developed for sphingolipids[251] were used for C1P with a saturated acyl chain of 16 carbons (16:0-C1P) and are provided in Figure C.3. The MARTINI 2.3P force field[252] in combination with polarizable MARTINI water model[253] was used for all coarse-grained simulations since it accurately reproduces the binding poses of peripheral membrane proteins. An ElNeDyn elastic network[254] with harmonic springs connecting all backbone beads within 0.9 nm of

each other and a spring constant of 500 kJ/mol/nm^2 was used to maintain the overall shape of CPTP in its apo and C1P-bound forms. Parameters for 16:0-C1P were chosen according to the standard MARTINI 2.0 lipid definitions and building block rules[167] and are provided in Figure C.4.

All-Atom Solution Simulations. The crystal structure of CPTP provided in PDB 4K85[247] was used to initialize all-atom simulations since it has the most residues fully resolved out of all structures of CPTP with C1P fully bound. The first six N-terminal residues that were not resolved in PDB 4K85 were built using MODELLER.[255,256] To investigate the transfer of CPTP's most likely *in vivo* substrate, the structure of 16:0-C1P fully bound inside CPTP provided in PDB 4K84 was aligned onto the structure of C1P with a saturated acyl chain of 12 carbons (12:0-C1P) present in PDB 4K85 to generate an initial configuration of CPTP in its C1P-bound form. C1P was removed from this structure to generate an initial configuration of the apo form. Both initial structures of the apo and C1P-bound forms were solvated in triclinic boxes of pre-equilibrated water molecules with dimensions chosen such that the minimum distance between any protein atom and the box walls was at least 2 nm. Sodium and chloride ions were added to achieve a salt concentration of 150 mM, matching the conditions used in *in vitro* experiments of C1P transfer by CPTP,[250] and to fully neutralize each system. Both systems were then energy minimized using the steepest descent algorithm and equilibirated using a three-step protocol: (1) a 200 ps simulation in an isothermal–isochoric (NVT) ensemble with the positions of CPTP's heavy atoms restrained by harmonic potentials with a force constant of 1000 kJ/mol/nm^2 to equilibrate the solvent molecules and adjust the system density; (2) a 200 ps simulation in an isothermal–isobaric (NPT) ensemble using the Berendsen barostat[169] to maintain the pressure with isotropic pressure coupling and a coupling time constant of 2 ps; and (3) a 200 ps simulation in an NPT ensemble using the Parinello-Rahman barostat[170] with a coupling time constant of 5 ps. Subsequently, production simulations in an NPT ensemble were run for 300 ns using the parameters given below. The final 200 ns of each of these simulations was analyzed to characterize the solution structure of the apo and C1P-bound forms of CPTP. The final configurations from these simulations were used to initialize simulations with a POPC bilayer present.

Coarse-Grained Simulations of CPTP in the presence of a Membrane Coarse-grained bilayers of 264 lipids surrounded by 3.2 nm slabs of solvent were built using INSANE[168] for simulations of CPTP in the presence of a membrane. For simulations of the apo form, the bilayer consisted of 262 POPC lipids and two 16:0-C1P lipids, one in each leaflet. For simulations of the C1P-bound form, the bilayer consisted of 264 POPC lipids. Prior to the addition of CPTP into the solvent surrounding the bilayer, the structures of both bilayers were fully relaxed from their initial lattice configurations. This involved an energy minimization using the steepest descent algorithm followed by equilibration using a two-step protocol: (1) a 500 ps simulation using a 10 fs time step and the Berendsen

barostat[169] with semi-isotropic presssure coupling and a coupling time constant of 3 ps; and (2) a 1 ns simulation using a 20 ps time step and the Parinello-Rahman barostat[170] with a coupling time constant of 12 ps. Subsequently, a 50 ns production run was performed to obtain bilayer structures that were used to construct initial configurations of CPTP in the presence of a membrane. The martinize.py script[257] was used to construct coarse-grained protein topologies and structures based on the final configurations of the apo and C1P-bound forms of CPTP obtained from all-atom solution simulations. Both coarse-grained structures were then fully equilibrated in solution through an energy minimization and a three-step protocol: (1) a 500 ps simulation in an NVT ensemble using a 10 fs time step and with the positions of CPTP's backbone beads restrained by harmonic potentials with a force constant of 1000 kJ/mol/nm^2 to equilibrate solvent molecules and adjust the system density; (2) a 500 ps simulation in an NPT ensemble using a 20 fs time step and the Berendsen barostat[169] with isotropic pressure coupling and a coupling time constant of 3 ps; and (3) a 500 ps simulation using a 20 fs time step and the Parinello-Rahman barostat[170] with a coupling time constant of 12 ps. Subsequently, 300 ns production runs were performed to obtain structures of CPTP used to initialize simulations that probe membrane binding. For simulations of both forms of CPTP in the presence of a membrane, four different initial coarse-grained configurations were built by inserting a structure of CPTP sampled during the last 100 ns of the coarse-grained solution simulation in a random location and orientation in the solvent around the equilibrated bilayer such that the minimum distance between any atom of CPTP and the membrane was 6 Å. Each replicate was then energy minimized and equilibrated using a three-step protocol: (1) a 500 ps simulation in an NVT ensemble using a 10 fs time step and with the positions of CPTP's backbone beads restrained by harmonic potentials with a force constant of 1000 kJ/mol/nm^2 to equilibrate solvent molecules and adjust the system density; (2) a 500 ps simulation in an NPT ensemble using a 20 fs time step and the Berendsen barostat[169] with semi-isotropic pressure coupling and a coupling time constant of 3 ps; and (3) a 1 ns simulation using a 20 fs time step and the Parinello-Rahman barostat[170] with a coupling time constant of 12 ps. During production runs of 2 μs, CPTP spontaneously bound to the membrane in all but one replicate of the apo form (Figure C.5), which was excluded from analysis of CPTP's membrane binding pose in coarse-grained simulations.

All-Atom Simulations of CPTP bound to a Membrane Membrane bound coarse-grained configurations of the apo and C1P-bound forms of CPTP were then backmapped into all-atom representations using the backward method.[258] The systems then underwent a three-step equilibration using the same protocol as used for the all-atom solution simulations but with semi-isotropic pressure coupling. Subsequently, 300 ns production runs were performed. The final 200 ns from each of these runs was used to analyze all-atom configurations of CPTP bound to a membrane.

All-Atom Production Simulation Parameters. All simulations were performed in an NPT ensemble at a pressure of 1 bar and temperature of 310 K. Separate Nosé-Hoover

thermostats[183,184] were used for the solvent and solute molecules. The pressure was maintained using the Parinello-Rahman barostat[170] with a compressibility of 4.5×10^{-5} bar^{-1} and coupling time constant of 1 ps. Isotropic pressure coupling was used in solution simulations, and semi-isotropic coupling was used in simulations with a bilayer. Dynamics were evolved using the leapfrog algorithm[173] and a 2 fs time step. All bonds to hydrogen were constrained using the LINCS algorithm.[185] Lennard-Jones forces were smoothly switched off between 0.8 and 1.2 nm. Real-space Coulomb interactions were truncated at 1.2 nm, and long-ranged Coulomb interactions were calculated using Particle Mesh Ewald (PME) summation.[186] Neighbor lists were constructed using the Verlet list cut-off scheme.[175]

Coarse-Grained Production Simulation Parameters. All simulations were performed in an NPT ensemble at a pressure of 1 bar and temperature of 310 K. To avoid the "hot solvent–cold solvent" problem,[172] separate V-rescale thermostats[171] were used for the solvent and solute molecules. The pressure was maintained using the Parinello-Rahman barostat[170] with a compressibility of 3×10^{-4} bar^{-1} and coupling time constant of 12 ps. Isotropic pressure coupling was used in solution simulations, and semi-isotropic coupling was used in simulations with a bilayer. Dynamics were evolved using the leapfrog algorithm[173] and a 20 fs time step. Lennard-Jones and Coulomb interactions were truncated at 1.1 nm, and long-ranged Coulomb interactions were calculated using PME summation.[186] A relative dielectric constant of 2.5 was used. Neighbor lists were constructed using the Verlet list cut-off scheme.[175]

Analysis of CPTP's Structure and Interactions with the Membrane

We analyzed the structure of apo and C1P-bound forms of CPTP both in solution and bound to a membrane. We quantified the volume of CPTP's hydrophobic cavity using MDpocket.[259] All other analysis was performed using a combination of MDAnalysis[188] and NumPy[226] Python libraries in addition to GROMACS tools. The insertion depth of each residue was calculated as the average signed distance in z, which is the axis normal to the membrane, between the center of mass (COM) of each residue and the average position of phosphate atoms (beads) in all-atom (coarse-grained) simulations. To calculate the angle $\theta_{\alpha 2}$ (Figure 4.3B), the Cα atoms of residues 54 and 65 were used to define the vector along the axis of helix α2, and the Cα atoms of residues of residues 155, 158, and 162 were used to define the plane that contains the top surface of helix α6 when CPTP is membrane bound. The interaction energy between each residue of CPTP and the membrane, $E_{\mathrm{CPTP-mem}}$, was calculated as the sum of short-ranged Lennard-Jones and Coulomb interaction energy terms between each residue and the membrane. The number of hydrophobic contacts between lipids was calculated based on the distances $d_{\mathrm{CC}}^{(i)}$ between hydrophobic carbons of two different lipids, where i labels a particular carbon-carbon (CC) pair. For all-atom POPC lipids, hydrophobic carbons include atoms C23−C216 and C33−C316. For all-atom 16:0-C1P, hydrophobic carbons include atoms C3F−C16F and C5S−C18S. The number of hydrophobic contacts, n_{CC}, is the number of CC pairs that satisfy $d_{\mathrm{CC}}^{(i)} \leq 1$ nm.

Free Energy Calculations

CPTP Catalyzed C1P Transport. We calculated free energy landscapes along the reaction coordinate for passive lipid transport, r_{LxS}, and fraction of contacts between C1P and CPTP, Q. r_{LxS} is defined in Eq. 3.1 of Chapter 3 and was calculated during all enhanced sampling simulations using the differential forms given in Eqs. 3.2 and 3.3 of Chapter 3. Q is calculated using the functional form developed to calculate the fraction of native protein contacts[260] given in Eq. 4.1 above. Contact pairs used to calculate Q include: (1) pairs of hydrophobic carbons of C1P and carbons of the residues that line CPTP's hydrophobic cavity (residues 10, 14, 36, 39, 40, 42, 43, 46, 48, 50, 52, 53, 57, 110, 111, 114, 117, 118, 121, 122, 146, 150, 153, 154, 158, 162, 165, 171, and 175) that have $\langle r_{ij} \rangle \leq 7.8$ Å during solution simulations of the C1P-bound form of CPTP; and (2) pairs of heavy atoms of C1P's headgroup and sphingoid backbone and heavy atoms of the residues responsible for C1P recognition (residues 52, 53, 56, 60, 96−102, 106, 110, 113, 114, 146−151, 154, and 214) that have have $\langle r_{ij} \rangle \leq 5.5$ Å during solution simulations of the C1P-bound form.

To calculate the free energy landscape, we performed multi-walker[261] well-tempered metadynamics simulations[262] using the PLUMED 2 patch[190] for GROMACS. A total of 20 walkers, 10 initialized from all-atom configurations of the apo form of CPTP bound to the membrane and 10 initialized from all-atom configurations of the C1P-bound form bound to the membrane, were used. Gaussian hills with $\sigma_{r_{\mathrm{LxS}}} = 0.1$ and $\sigma_Q = 0.01$ and initial heights of 2.5 kJ/mol were added by each walker every 4 ps with a bias factor of 30. To avoid pushing the system into regions with very negative values of r_{LxS}, where C1P is free in solution far away from either the membrane or CPTP, and at very large positive values of r_{LxS}, which are unphysical, harmonic wall potentials with a spring constant of 500 kJ/mol were applied at $r_{LxS} = -3$ and 15. The simulation was run for 560 ns (cumulatively 11.2 μs). To facilitate sampling of configurations with C1P inside CPTP by walkers initialized with the apo form, and vice versa for walkers initialized with the C1P-bound form, moving restraints were applied to each walker to pull it to ($r_{\mathrm{LxS}} = -2.5, Q = 1$), if it was initialized from the apo form, and to ($r_{\mathrm{LxS}} = 12, Q = 0$), if it was initialized from the C1P-bound form, over the course of 20 ns. The moving restraint applied a harmonic bias on r_{LxS} with a spring constant of 75 kJ/mol and a harmonic bias on Q with a spring constant of 7500 kJ/mol. During this time, no Gaussians were deposited, but the bias due to previously accumulated Gaussians was applied. Subsequently, the multi-walker well-tempered metadynamics simulation was continued for an additional 20 ns without the moving restraints. After discarding the data from the first 100 ns of each walker and from the time period when moving restraints were applied, a free energy landscape and profiles along r_{LxS} and Q were obtained using the time-independent free energy estimator of Tiwary and Parrinello.[263] As the simulations have not yet converged, these free energy landscape and profiles are considered preliminary.

Passive C1P Transport. We calculated a free energy profile along r_{LxS} for passive transport of C1P between membranes using umbrella sampling simulations.[189] Simulations were performed using two different membrane compositions that matched the membrane compo-

sitions used for simulations with both forms of CPTP. An initial bilayer of 264 POPC lipids surrounded by 3.2 nm thick slabs of solvent was built using the CHARMM-GUI Membrane Builder.[181,182] A lipid in each leaflet was replaced with a 16:0-C1P molecule to construct a second membrane. Sodium and chloride ions were added to achieve a salt concentration of 150 mM and to fully neutralize each system. Both membranes were energy minimized and equilibrated using a two-step protocol: (1) a 250 ps simulation using the Berendsen barostat;[169] and (2) a 250 ps simulation using the Parinello-Rahman barostat.[170] All other simulation parameters matched those given above for all-atom simulations of CPTP bound to a membrane. 50 ns production runs were performed to fully equilibrate the structure of each bilayer. The final configuration from the simulation of the membrane composed of both POPC and 16:0-C1P was used to initialize one enhanced sampling simulation. Another initial configuration was constructed by inserting a 16:0-C1P lipid randomly into the solution above the POPC membrane such that its COM was at least 3.2 nm above the COM of the membrane, energy minimizing the system, and equilibrating it using the two-step protocol used for the membrane only system with the addition of harmonic restraints with a force constant of 500 kJ/mol/nm^2 on the z coordinates of all heavy atoms of 16:0-C1P. Next, to generate configurations to initialize each umbrella sampling window, 20 ns steered MD simulations were performed using a harmonic bias on r_{LxS} with a spring constant of 75 kJ/mol. During one steered simulation, a tagged 16:0-C1P was pulled from $r_{\mathrm{LxS}} = 14$ to $r_{\mathrm{LxS}} = -5$ to extract it from the membrane. In the other, the 16:0-C1P was pulled from $r_{\mathrm{LxS}} = -5$ to $r_{\mathrm{LxS}} = 14$ to insert it into the membrane. A total of 66 umbrella sampling windows were simulated with equally spaced harmonic bias centers ranging from $r_{\mathrm{LxS}} = -5$ to $r_{\mathrm{LxS}} = 15$ and a spring constant of 40 kJ/mol. All windows were run for 52 ns, and the first 20 ns was discarded to account for equilibration. Data from all windows of both systems was combined with the weighted histogram analysis method (WHAM)[191] to obtain the free energy profile. Error bars were calculated as the standard error estimated from the two independent systems.

Chapter 5

Development of Brownian Bridge Based Path Sampling Methods for a Nonequilibrium Transport Model

The work presented in this chapter was conducted in collaboration with Prof. Juan P. Garrahan from the University of Nottingham.

Introduction

In the previous chapters, lipid transport has been assumed to occur in thermal equilibrium. Passive lipid transport is, indeed, an equilibrium process, and many lipid transfer proteins, including the one highlighted in Chapter 4, can function under equilibrium conditions used in *in vitro* experiments. However, within a cell, some lipid transfer proteins require a direct expenditure of energy to transport lipids up concentration gradients.[85,87] For example, the lipopolysaccharide transport apparatus, which is the bridge-like lipid transfer protein shown in Figure 1.4B in Chapter 1, utilizes ATP hydrolysis to extract lipopolysaccharides from the inner membrane of Gram-negative bacteria and drive their transport to the outer membrane.[117,118] In this case, lipid transport is an inherently nonequilibrium process.

Much as equilibrium systems are described by free energy functions, nonequilibrium states are characterized by large deviation functions, or generalized free energies. Both equilibrium free energy functions and large deviation functions encode information about the stability of different states of a system.[264,265] Thus, a wealth of information about nonequilibrium transport of lipids between membranes is contained in large deviation functions. However, large deviation functions are difficult to calculate using simulations for complex nonequilibrium systems. Here, we attempt to address some of the challenges that plague current simulation methods used to calculate large deviation functions.

Because large deviation functions are averages taken over trajectories ensembles, they can be calculated from transition path sampling (TPS) simulations.[266–269] TPS simulations

sample trajectory ensembles by performing Monte Carlo moves in trajectory space.[136,137] To accurately calculate a large deviation function, rare trajectories must be harvested. By biasing the system appropriately, rare trajectories can be more easily sampled. Unfortunately, standard TPS moves perform poorly at high bias.[266] Standard TPS moves generate new trajectory segments by shooting off dynamics from a random time point of an original trajectory.[270] At high bias, only very short shooting moves at either end of the trajectory are accepted, leaving the bulk of the trajectory in the middle unchanged.[266] To overcome this limitation, we developed new TPS moves that utilize Brownian bridges, which are stochastic trajectories constrained to start and end at specified configurations, to readily modify the middle of trajectories even at high bias. Specifically we developed Brownian bridge moves to sample trajectories of periodic asymmetric simple exclusion processes (ASEPs), a paradigmatic nonequilibrium transport model[271,272] that exhibits a dynamical phase transition.[273,274] ASEPs have been previously used to model various biological phenomena,[271] such as ribosome scanning on mRNA[275,276] and proton transport through channels,[277] and could foreseeably be used to model lipid transport via bridge-like lipid transfer proteins.

The remainder of this chapter is organized as follows: First, we provide a brief theoretical overview of trajectory ensembles and large deviation functions. Second, because we have developed new TPS moves to specifically sample rare trajectories of the periodic ASEP, we describe the details of the model and its behavior. Next, we discuss the details of shooting moves before presenting our new Brownian bridge moves. In the final section, we demonstrate how Brownian bridge moves enable large deviation functions to be efficiently calculated from TPS simulations. We conclude by discussing potential ways to further improve our newly developed Brownian bridge moves in addition to other types of sampling problems that could be addressed with Brownian bridges moves.

Theory

Trajectory Ensembles

A trajectory is a time-ordered sequence of system configurations. We denote a trajectory, $\mathscr{X}(T)$, that extends for a total time T as $\mathscr{X}(T) = [\boldsymbol{x}(0), \boldsymbol{x}(t_1), \boldsymbol{x}(t_2), \ldots, \boldsymbol{x}(T)]$, where $\boldsymbol{x}(t_j)$ is the system configuration at time t_j. We focus on sampling trajectories of stochastic Markov processes, which generally obey a continuity equation of the form

$$\frac{\partial p(\boldsymbol{x}(t))}{\partial t} = W p(\boldsymbol{x}(t)), \tag{5.1}$$

where $p(\boldsymbol{x}(t))$ is the probability of observing the system in configuration $\boldsymbol{x}$ at time t and W is the linear operator in configuration space that propagates the system's dynamics. For discrete configuration space, W is a transition rate matrix and the continuity equation is a master equation. For continuous configuration space, W is a Liouvillian and the continuity equation is a Fokker-Planck equation. If W obeys detailed balance, then the steady state

of the system is an equilibrium distribution that follows Boltzmann statistics. If, however, W does not obey detailed balance but is irreducible, unique nonequilibrium steady states with non-vanishing currents are generated in the long time limit.[278] The set of trajectories that collectively characterize a nonequilibrium steady state form an ensemble. In general, for both equilibrium and nonequilibrium trajectories, the probability of a trajectory is

$$P[\mathscr{X}(T)] = \frac{\rho(\boldsymbol{x}(0))}{Z} \prod_j p(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1})) \tag{5.2a}$$

$$Z = \sum_{\mathscr{X}(T)} \rho(\boldsymbol{x}(0)) \prod_j p(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1})), \tag{5.2b}$$

where $\rho(\boldsymbol{x}(0))$ is the stationary distribution of initial system configurations, $p(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1}))$ is the transition probability from configuration $\boldsymbol{x}(t_j)$ to $\boldsymbol{x}(t_{j+1})$, and Z is the path partition function. A trajectory ensemble affords the calculation of probability distributions of time extensive observables and, furthermore, the large deviation function.

Large Deviation Functions

Nonequilibrium steady states can be characterized by time extensive observables $\mathcal{O}[\mathscr{X}(T)]$. In a nonequilibrium steady state, a time extensive observable has a probability distribution that can be constructed as a marginal over an ensemble of trajectories.[264,265] The probability of observing the value $\mathcal{O}$ is

$$P(\mathcal{O}) = \langle \delta(\mathcal{O} - \mathcal{O}[\mathscr{X}(T)]) \rangle = \sum_{\mathscr{X}(T)} P[\mathscr{X}(T)] \delta(\mathcal{O} - \mathcal{O}[\mathscr{X}(T)]). \tag{5.3}$$

In the long time limit, the probability distributions of time extensive observables that are correlated over only a finite amount of time have a time intensive form,

$$-T^{-1} \ln P(\mathcal{O}) \sim I(\mathcal{O}/T), \tag{5.4}$$

where $I(\mathcal{O}/T)$ is the rate function.[264,265] Additionally, the Laplace transform of Eq. 5.3 yields the generating function

$$\langle e^{-\zeta \mathcal{O}} \rangle = \sum_{\mathscr{X}(T)} P[\mathscr{X}(T)] e^{-\zeta \mathcal{O}} \sim e^{\psi(\zeta) T}, \tag{5.5}$$

where $\psi(\zeta)$ is the large deviation function and ζ is a (generally nonphysical) field conjugate to the observable. Like an equilibrium free energy, the large deviation function is also a scaled cumulant generating function. When the asymptotic, long time limit exists and the rate function is smooth and convex, then the large deviation function is obtained from the rate function through a Legendre transform,

$$\psi(\zeta) = \inf_{\mathcal{O}} \left[I(\mathcal{O}/T) + \zeta \mathcal{O}/T \right], \tag{5.6}$$

where inf is the largest lower bound over $\mathcal{O}$.[264,265] As in equilibrium thermodynamics, the Legendre transform switches between ensembles in which either of a pair of conjugate variables is held fixed. Ensembles with fixed values of $\mathcal{O}$ are typically equivalent to ensembles with fixed values of the field ζ in the long time limit.[279]

Calculating large deviation functions from simulations requires sampling trajectories with rare values of the observable, which are typical of large values of ζ. Conventional simulations are ill-suited to this task, just as they are for calculating equilibrium free energies. Instead, the trajectory ensemble is biased, or tilted, to enhance the sampling of rare trajectories. In the tilted ensemble, the probability of a trajectory is

$$P_\zeta[\mathscr{X}(T)] = P[\mathscr{X}(T)]e^{-\zeta\mathcal{O}[\mathscr{X}(T)]-T\psi(\zeta)}, \tag{5.7}$$

such that the trajectories which contribute the most to the large deviation function at a particular value of ζ are sampled with the highest probability.[264,265] The large deviation function can be estimated from a biased simulation according to

$$\psi(\zeta) = I_\zeta(\mathcal{O}) - I(\mathcal{O}) - \zeta\mathcal{O}/T, \tag{5.8}$$

where $I_\zeta(\mathcal{O})$ is the unnormalized rate function obtained from a simulation with bias ζ. Distributions obtained from multiple biased simulations with different values of ζ are then combined using histogram reweighting techniques, including the weighted histogram analysis method (WHAM),[191] to obtain the large deviation function and unbiased rate function.[266] Although such an approach only determines the large deviation function up to a global constant, the full large deviation function can be uniquely determined since the unbiased trajectory ensemble is normalized, resulting in $\psi(0) = 0$.[264,265] This approach mirrors the use of umbrella sampling simulations[189] to calculate equilibrium free energies and is used below to calculate the large deviation function for the periodic ASEP.

Model: Periodic Asymmetric Simple Exclusion Process

In a simple exclusion process, particles hop stochastically between sites on a lattice. No particles can simultaneously occupy the same lattice site. This hard-core excluded volume constraint mimics short-ranged interactions between particles. Each particle hops clockwise from lattice site l to $l+1$ with rate γ_+ and counterclockwise from lattice site l to $l-1$ with rate γ_-. In ASEPs, $\gamma_+ \neq \gamma_-$. As a result, detailed balance is broken, and the model exhibits a nonequilibrium steady state with a stationary current. On a periodic lattice, the total number of particles, N, is conserved. Figure 5.1A shows a configuration of the periodic ASEP along with the allowed hops and their rates. Because each particle is identical, unique system configurations are determined by the occupancy of each lattice site, $n_l = \{0, 1\}$, where $l = \{1, 2, \ldots, L\}$ indexes each site. When the system configuration is represented in a basis of single site occupancies, $\boldsymbol{n} = [n_1, n_2, \ldots, n_L]$, the transition rate matrix that propagates

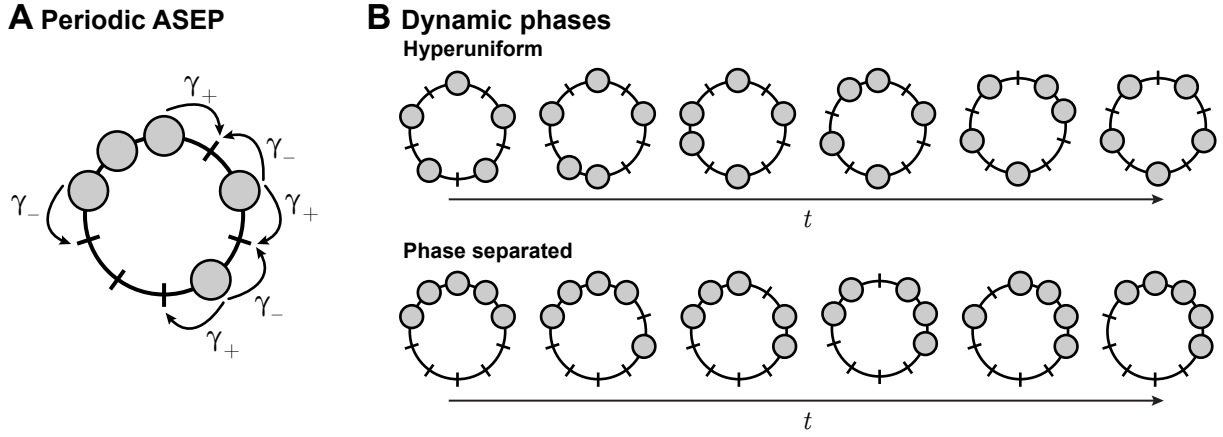


Figure 5.1: The periodic asymmetric simple exclusion process (ASEP) is a nonequilibrium lattice model with a dynamical phase transition. (A) Particles hop between sites on a periodic, 1-dimensional lattice. Clockwise hops occur at rate γ_+, and counterclockwise hops occur at rate γ_-. Particles exclude volume so no two particles can occupy the same lattice site. All allowed hops are indicated with arrows. (B) The periodic ASEP exhbits two dynamical phases: (top) a hyperuniform steady state with high activity and (bottom) a phase separated state with low activity. For each phase, an example trajectory is shown with each configuration separated by a single particle hop.

the system dynamics on a periodic lattice has elements

$$W_{\boldsymbol{n}'\boldsymbol{n}} = -\sum_{\boldsymbol{n}'} w_{\mathrm{ASEP}}(\boldsymbol{n} \to \boldsymbol{n}') + \sum_{\boldsymbol{n}' \neq \boldsymbol{n}} w_{\mathrm{ASEP}}(\boldsymbol{n} \to \boldsymbol{n}') \tag{5.9}$$

where $w_{\mathrm{ASEP}}(\boldsymbol{n} \to \boldsymbol{n}')$ is the transition rate from $\boldsymbol{n}$ to $\boldsymbol{n}'$. The first term in Eq. 5.9 is the total escape rate from configuration $\boldsymbol{n}$, and the second term counts all possible transitions from $\boldsymbol{n}$ to $\boldsymbol{n}'$. The only allowed transitions are ones between configurations separated by a single particle hop to an adjacent empty lattice site. Under periodic boundary conditions, the steady state distribution is uniform, and each configuration has the same probability.

In the following sections, each particle will be treated individually with its configuration x_i given by its position on the lattice, which has sites $x = \{1, 2, \ldots, L\}$. The configuration of all particles will be denoted $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]$. Using this notation, the probability of a trajectory of the periodic ASEP is

$$P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)] = \frac{\rho^{(\mathrm{pbc})}_{\mathrm{ASEP}}(\boldsymbol{x}(0))\theta[\mathscr{X}(T)]}{Z^{(\mathrm{pbc})}_{\mathrm{ASEP}}} \prod_j w_{\mathrm{ASEP}}(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1}))e^{-R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j))\Delta t_j} \tag{5.10a}$$

$$Z^{(\mathrm{pbc})}_{\mathrm{ASEP}} = \sum_{\mathscr{X}(T)} \rho^{(\mathrm{pbc})}_{\mathrm{ASEP}}(\boldsymbol{x}(0))\theta[\mathscr{X}(T)] \prod_j w_{\mathrm{ASEP}}(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1}))e^{-R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j))\Delta t_j}, \tag{5.10b}$$

where $\rho_{\mathrm{ASEP}}^{(\mathrm{pbc})}(\boldsymbol{x}(0))$ is the stationary distribution of initial configurations of the periodic ASEP, which is a uniform distribution, $\theta[\mathscr{X}(T)]$ is a Heaviside function which ensures that configurations that violate the hard-constraint of excluded volume are never sampled during the trajectory, $R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j)) = \sum_{\boldsymbol{x}'} w_{\mathrm{ASEP}}(\boldsymbol{x}(t_j) \to \boldsymbol{x}'(t_{j+1}))$ is the total escape rate from configuration $\boldsymbol{x}$, $\Delta t_j = t_{j+1} - t_j$, and $Z_{\mathrm{ASEP}}^{(\mathrm{pbc})}$ is the path partition function for the periodic ASEP. Trajectories of the periodic ASEP exhibit two characteristic behaviors:[273,274] one is characterized by high dynamical activity and the other by low dynamical activity. The activity, $K[\mathscr{X}(T)]$, is the total number of hops in a trajectory and is calculated according to

$$K[\mathscr{X}(T)] = \sum_j \sum_i^N [\![x_i(t_{j+1}) \neq x_i(t_j)]\!], \tag{5.11}$$

where the Iverson bracket of the statement S, $[\![S]\!] = 1$, is a generalization of the discrete delta function and evaluates to 1 if S is true and 0 otherwise. Trajectories with high activity are termed hyperuniform states since large lengthscale density fluctuations are suppressed and particles are roughly evenly spaced apart.[274,280] Trajectories with low activity are termed phase separated states since macroscopic density inhomogenities arise and particles cluster together.[274,281] Both are schematically illustrated in Figure 5.1B. A dynamical phase transition separates the hyperuniform state from the phase separated state.[273,274] This is evidenced by a singularity in the large deviation function, $\psi(s)$, at $s = 0$, where s is the conjugate field to activity. The hyperuniform state occurs at $s < 0$, and the phase separated state occurs at $s > 0$.

Another dynamical observable that is useful to characterize the steady states of the periodic ASEP is the current.[282] The current, $J[\mathscr{X}(T)]$, is the difference between the number of clockwise hops and counterclockwise hops in a trajectory and is calculated according to

$$J[\mathscr{X}(T)] = \sum_j \sum_i^N [\![x_i(t_{j+1}) = x_i(t_j) + 1]\!] - [\![x_i(t_{j+1}) = x_i(t_j) - 1]\!]. \tag{5.12}$$

The current is determined by the particle's net displacement during the trajectory. We denote the field conjugate to current as λ. Trajectories with large positive current, such as found in the hyperuniform phase, occur at $\lambda \ll 0$, whereas trajectories with negative current, which indicates a change in particle flow from clockwise to counterclockwise, occur at $\lambda \gg 0$.

Transition Path Sampling Algorithms

To calculate the large deviation functions $\psi(s)$ and $\psi(\lambda)$ of the periodic ASEP, we use TPS simulations to sample trajectory ensembles. During a TPS simulation, trajectories are sampled through a Monte Carlo random walk in trajectory space. Each Monte Carlo move generates a new trajectory by altering the present one in a prescribed way. The new

trajectory is then accepted or rejected according to the Metropolis-Rosenbluth-Rosenbluth-Teller-Teller-Hastings (subsequently referred to as Metropolis) criterion to sample the desired distribution. To sample unbiased trajectories of total time T of the periodic ASEP, a new trajectory $\mathscr{X}'(T)$ generated from the previous one $\mathscr{X}(T)$ is accepted with probability

$$\mathrm{acc}[\mathscr{X}(T) \to \mathscr{X}'(T)] = \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)]\mathrm{gen}[\mathscr{X}'(T) \to \mathscr{X}(T)]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]\mathrm{gen}[\mathscr{X}(T) \to \mathscr{X}'(T)]}\right], \tag{5.13}$$

where $\mathrm{gen}[\mathscr{X}(T) \to \mathscr{X}'(T)]$ is the generation probability. The specific form of the generation probability depends on the details of the Monte Carlo move. In the remainder of this section, we review two standard moves and develop two new moves based on Brownian bridges to sample trajectories of the periodic ASEP (Figure 5.2). While acceptance probabilities are derived below for unbiased trajectory ensembles, they can be readily adapted to sample tilted ensembles.

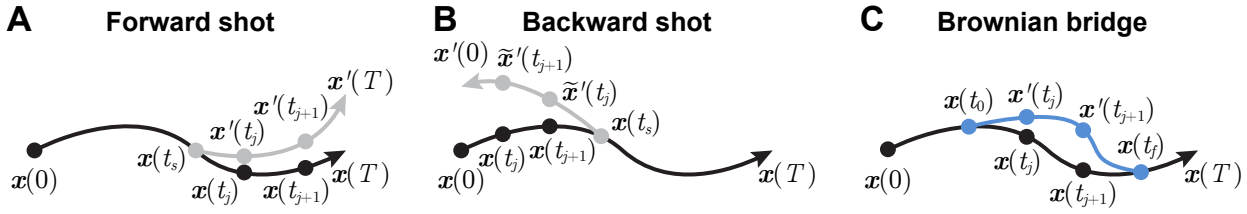


Figure 5.2: TPS moves used to sample trajectories of the periodic ASEP. (A) During a forward shooting move, a new trajectory segment, which is composed of time points $\boldsymbol{x}'(t_s), \ldots, \boldsymbol{x}'(T)$ and colored gray, is obtained by evolving the dynamics forward in time from time t_s. (B) During a backwards shooting move, a new trajectory segment, which is composed of times points $\boldsymbol{x}'(0), \ldots, \boldsymbol{x}'(t_s)$ and colored gray, is obtained by evolving the dynamics backwards in time from time t_s. Time reversed configurations are denoted $\widetilde{\boldsymbol{x}}(t)$. (C) During a Bownian bridge move, a new trajectory segment, which is composed of time points $\boldsymbol{x}'(t_0), \ldots, \boldsymbol{x}'(t_f)$ and colored blue, is obtained by evolving the dynamics according to the master equation for a bridge trajectory constrained to start at configuration $\boldsymbol{x}(t_0)$ and end at $\boldsymbol{x}(t_f)$.

Shooting Moves

Two standard TPS moves for modifying stochastic trajectories involve shooting off a new trajectory segment from a randomly chosen time point, $\boldsymbol{x}(t_s)$, of a previous trajectory.

Forward Shooting Moves. In a forward shooting move, which is illustrated in Figure 5.2A, a new trajectory segment is integrated forward in time, altering times after t_s in the original trajectory. The newly generated trajectory consists of configurations $\mathscr{X}'(T) = [\boldsymbol{x}(0), \ldots, \boldsymbol{x}(t_{s-1}), \boldsymbol{x}(t_s), \boldsymbol{x}'(t_{s+1}), \ldots, \boldsymbol{x}'(T)]$. Because the new trajectory segment

is generated using the propagator for the periodic ASEP given in Eq. 5.9, the generation probability is $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T-t_s)|\boldsymbol{x}_{t_s}]$, and the probability of the new trajectory is $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)] = P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(t_s)]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T-t_s)|\boldsymbol{x}_{t_s}]$. As a result the acceptance probability given in Eq. 5.13 simplifies to

$$\begin{aligned}\mathrm{acc}[\mathscr{X}(T)\to\mathscr{X}'(T)] &= \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T-t_s)|\boldsymbol{x}_{t_s}]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T-t_s)|\boldsymbol{x}_{t_s}]}\right] \\ &= 1. \end{aligned} \tag{5.14}$$

Although unbiased forward shots are always accepted, the acceptance probability of biased shots decreases exponentially with increasing bias value. For this reason, when the bias is large, only the very end of an original trajectory is likely to be modified by forward shots. However, all time points of a trajectory need to be modified for the sampling to be ergodic.

Backward Shooting Moves. Backwards shooting moves are necessary for ergodicity because they alter early time points in the trajectory. In a backward shooting move, which is illustrated in Figure 5.2B, a new segment is integrated backwards in the time from time t_s, altering all earlier times in the original trajectory. The newly generated trajectory consists of configurations $\mathscr{X}'(T) = [\boldsymbol{x}'(0),\ldots,\boldsymbol{x}'(t_{s-1}),\boldsymbol{x}(t_s),\boldsymbol{x}(t_{s+1}),\ldots,\boldsymbol{x}(T)]$. When propagated backwards in time, the clockwise and counterclockwise rates are swapped; the rate for a particle to hop clockwise from site l to $l+1$ backwards in time is $\widetilde{\gamma}_+ = \gamma_-$, and the rate for a particle to hop counterclockwise from site l to $l-1$ backwards in time is $\widetilde{\gamma}_- = \gamma_+$. Then, the time reversed transition rates $\widetilde{w}_{\mathrm{ASEP}}(\widetilde{\boldsymbol{x}}(t_j)\to\widetilde{\boldsymbol{x}}(t_{j+1}))$ between time-reversed configurations $\widetilde{\boldsymbol{x}}(t_j) = \boldsymbol{x}(t_{j+1})$ and $\widetilde{\boldsymbol{x}}(t_{j+1}) = \boldsymbol{x}(t_j)$ are equal to the forward transition rates $w_{\mathrm{ASEP}}(\boldsymbol{x}(t_j)\to\boldsymbol{x}(t_{j+1}))$. Additionally, backwards shooting moves of the periodic ASEP sample initial configurations, $x'(0)$, from a uniform distribution, which is the same stationary distribution sampled when trajectories are propagated forward in time (specifically, $\rho(x'(0)) = \rho(x(0))$). As a result, the generation probability for backwards shooting moves is $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(t_s)|\boldsymbol{x}_{t_s}]$, and the probability of the new trajectory is $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)] = P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)|\boldsymbol{x}_{t_s}]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T-t_s)|\boldsymbol{x}_{\boldsymbol{t_s}}]$. As a result the acceptance probability given in Eq. 5.13 simplifies to

$$\begin{aligned}\mathrm{acc}[\mathscr{X}(T)\to\mathscr{X}'(T)] &= \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T)]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(t_s)|\boldsymbol{x}_{t_s}]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(t_s)|\boldsymbol{x}_{t_s}]}\right] \\ &= 1. \end{aligned} \tag{5.15}$$

As for forward shots, unbiased backwards shots are always accepted but the acceptance probability of biased backwards shots decreases exponentially with increasing bias value. Thus, when the bias is high, backwards shots only modify the very beginning of an original trajectory. Even when both forwards and backwards shooting moves are performed, the middle of the trajectory is seldom altered at high bias.

Brownian Bridge Moves

Brownian bridge moves offer a way to more readily alter the middle of a trajectory, even at high bias. During a Brownian bridge move, which is illustrated in Figure 5.2C, a new trajectory segment between configurations $\boldsymbol{x}(t_0)$ and $\boldsymbol{x}(t_f)$ is generated so that portions of the trajectory away from the initial and final time points are modified. Furthermore, the length of the trajectory that is modified during a Brownian bridge move can be controlled. This is advantageous since longer trajectory segments are accepted with lower probabilities at high bias. In order to alter an entire trajectory with forward and backward shooting moves, shots that are roughly half the total length of the trajectory would need to be accepted with decent probability; however, this is unlikely for long, biased trajectories. Instead, many short bridges, which are each accepted with decent probability, can be performed to alter the entire trajectory, excluding the initial and final time points. Thus, Brownian bridge moves offer two advantages over shooting moves at high bias: (1) The middle of a trajectory can be naturally modified with a bridge. (2) Acceptance probabilities can be tuned by controlling the bridge length. Nevertheless, when using Brownian bridge moves, shooting moves are still required to modify the initial and final time points of a trajectory and to ultimately sample trajectories ergodically.

Implementing Brownian bridge moves requires (1) a simulation algorithm for generating Brownian bridges and (2) an expression for the acceptance probability of a Brownian bridge move. We first present a general algorithm and acceptance criterion for a continuous-time Brownian bridge in discrete space. We then derive expressions specifically for two Brownian bridge moves that we use to sample trajectories of the periodic ASEP.

During a Brownian bridge move, particles are treated as non-interacting. For each particle, a Brownian bridge trajectory is generated independently of all other particles' trajectories. The forward time evolution of a single particle that hops between lattice sites obeys the forward master equation,

$$\begin{aligned}\frac{\partial}{\partial t}p^{(\mathrm{bc})}(x(t)|x_{t_0}) =&\gamma_+ p^{(\mathrm{bc})}(x(t)-1|x_{t_0}) + \gamma_- p^{(\mathrm{bc})}(x(t)-1|x_{t_0}) \\ &- (\gamma_+ + \gamma_-)p^{(\mathrm{bc})}(x(t)|x_{t_0}),\end{aligned} \tag{5.16}$$

where $p^{(\mathrm{bc})}(x(t)|x_{t_0})$ is the probability of the particle to be at position $x = \{1, 2, \ldots, L\}$ on a lattice with boundary conditions labeled "bc" at time t given that it was previously at position x_{t_0} at t_0. In our development of two different bridge moves, we will consider lattices of length L with periodic boundary conditions (labeled "pbc") and lattices of infinite length with free boundary conditions (labeled "fbc"). The corresponding backwards time evolution of a single particle obeys the backward master equation, or adjoint of Eq. 5.16,

$$\begin{aligned}\frac{\partial}{\partial t}q^{(\mathrm{bc})}(x(t)|x_{t_f}) =& -\gamma_+ q^{(\mathrm{bc})}(x(t)+1|x_{t_f}) - \gamma_- q^{(\mathrm{bc})}(x(t)-1|x_{t_f}) \\ &+ (\gamma_+ + \gamma_-)p^{(\mathrm{bc})}(x(t)|x_{t_f}),\end{aligned} \tag{5.17}$$

where $q^{(\mathrm{bc})}(x(t)|x_{t_f})$ is the probability of the particle to be at position x at time t given that it's at position x_{t_f} at the later time t_f. During a bridge, the particle's time evolution from a given initial configuration x_{t_0} is constrained to end at configuration x_{t_f} at time t_f. The probability of the particle to be at position x at time t during a bridge, $\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f})$, is then given in terms of the Green's function solutions to Eq. 5.16 and 5.17 according to Bayes' theorem

$$\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f}) = \frac{p^{(\mathrm{bc})}(x(t)|x_{t_0})q^{(\mathrm{bc})}(x(t)|x_{t_f})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})}. \tag{5.18}$$

It can be shown that $\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f})$ is itself a solution to a master equation as follows

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f}) =& \frac{p^{(\mathrm{bc})}(x(t)|x_{t_0})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})}\frac{\partial}{\partial t}q^{(\mathrm{bc})}(x(t)|x_{t_f}) + \frac{q^{(\mathrm{bc})}(x(t)|x_{t_f})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})}\frac{\partial}{\partial t}p^{(\mathrm{bc})}(x(t)|x_{t_0}) \\
=& -\gamma_+\frac{p^{(\mathrm{bc})}(x(t)|x_{t_0})q^{(\mathrm{bc})}(x(t)+1|x_{t_f})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})} \\
& -\gamma_-\frac{p^{(\mathrm{bc})}(x(t)|x_{t_0})q^{(\mathrm{bc})}(x(t)-1|x_{t_f})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})} \\
& +\gamma_+\frac{q^{(\mathrm{bc})}(x(t)|x_{t_f})p^{(\mathrm{bc})}(x(t)-1|x_{t_0})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})} \\
& +\gamma_-\frac{q^{(\mathrm{bc})}(x(t)|x_{t_f})p^{(\mathrm{bc})}(x(t)-1|x_{t_0})}{p^{(\mathrm{bc})}(x_{t_f}|x_{t_0})} \\
=& -\gamma_+\frac{q^{(\mathrm{bc})}(x(t)+1|x_{t_f})}{q^{(\mathrm{bc})}(x(t)|x_{t_f})}\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f}) \\
& -\gamma_-\frac{q^{(\mathrm{bc})}(x(t)-1|x_{t_f})}{q^{(\mathrm{bc})}(x(t)|x_{t_f})}\mathcal{P}^{(\mathrm{bc})}(x(t)|x_{t_0}, x_{t_f}) \\
& +\gamma_+\frac{q^{(\mathrm{bc})}(x(t)|x_{t_f})}{q^{(\mathrm{bc})}(x(t)-1|x_{t_f})}\mathcal{P}^{(\mathrm{bc})}(x(t)-1|x_{t_0}, x_{t_f}) \\
& +\gamma_-\frac{q^{(\mathrm{bc})}(x(t)|x_{t_f})}{q^{(\mathrm{bc})}(x(t)+1|x_{t_f})}\mathcal{P}^{(\mathrm{bc})}(x(t)+1|x_{t_0}, x_{t_f}).
\end{aligned} \tag{5.19}$$

Eq. 5.19 describes the time evolution of a particle that performs a bridge between configurations x_{t_0} and x_{t_f} by hopping clockwise with rate $\Gamma_+^{(\mathrm{bc})}(x(t))$ and counterclockwise with rate $\Gamma_-^{(\mathrm{bc})}(x(t))$, where

$$\Gamma_+^{(\mathrm{bc})}(x(t)) = \gamma_+\frac{q^{(\mathrm{bc})}(x(t)+1|x_{t_f})}{q^{(\mathrm{bc})}(x(t)|x_{t_f})} \tag{5.20a}$$

$$\Gamma_-^{(\mathrm{bc})}(x(t)) = \gamma_-\frac{q^{(\mathrm{bc})}(x(t)-1|x_{t_f})}{q^{(\mathrm{bc})}(x(t)|x_{t_f})}. \tag{5.20b}$$

Both rates are space- and time-dependent to ensure the bridge trajectory ends at x_{t_f}. When t approaches t_f, the particle preferentially hops in the direction towards x_{t_f}. In contrast, when $t \ll t_f$, the particle can wander broadly and still reach x_{t_f}. As a result, the variance in the particle's position is largest midway through the bridge at time $\frac{t_f - t_0}{2}$. Simulating bridges using these rates requires an analytical expression for $q^{(\mathrm{bc})}(x(t)|x_{t_f})$; however, analytical solutions to the backwards master equation (Eq. 5.17) exist only for specific cases. One such case is a particle undergoing a continuous-time discrete-space random walk. Thus, to generate Brownian bridges, we have chosen to approximate the particles of the periodic ASEP as a collection of non-interacting, random walkers.

We will make use of the relationship between a Brownian bridge of non-interacting particles and continuous-time random walks (CTRW) to derive the acceptance probability of a bridge move. The probability of a bridge of total length $T_\mathrm{B} = t_f - t_0$ from configuration $\boldsymbol{x}_{t_0}$ to $\boldsymbol{x}_{t_f}$, $P^{(\mathrm{bc})}_{\mathrm{BB}}[\mathscr{X}(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})]$, is related to the probability of a trajectory in which all particles perform CTRWs, $P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})]$, according to

$$\begin{aligned} P^{(\mathrm{bc})}_{\mathrm{BB}}[\mathscr{X}(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})] &= P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}] \\ &\propto P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})][\![\boldsymbol{x}(t_0) = \boldsymbol{x}_{t_0}]\!][\![\boldsymbol{x}(t_f) = \boldsymbol{x}_{t_f}]\!], \end{aligned} \tag{5.21}$$

where the discrete delta function is denoted with Iverson brackets, $[\![\cdots]\!]$. The probability of a trajectory in which all particles perform CTRWs is

$$P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})] = \frac{\rho^{(\mathrm{bc})}_{\mathrm{CTRW}}(\boldsymbol{x}(0))}{Z^{(\mathrm{bc})}_{\mathrm{CTRW}}} \prod_j w_{\mathrm{CTRW}}(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1})) e^{-R_{\mathrm{CTRW}}(\boldsymbol{x}(t_j))\Delta t_j} \tag{5.22a}$$

$$Z^{(\mathrm{bc})}_{\mathrm{CTRW}} = \sum_{\mathscr{X}(T_\mathrm{B})} \rho^{(\mathrm{bc})}_{\mathrm{CTRW}}(\boldsymbol{x}(0)) \prod_j w_{\mathrm{CTRW}}(\boldsymbol{x}(t_j) \to \boldsymbol{x}(t_{j+1})) e^{-R_{\mathrm{CTRW}}(\boldsymbol{x}(t_j))\Delta t_j}, \tag{5.22b}$$

where $\rho^{(\mathrm{bc})}_{\mathrm{CTRW}}(\boldsymbol{x}(0))$ is the stationary distribution of initial configurations, $w_{\mathrm{CTRW}}(\boldsymbol{x} \to \boldsymbol{x}')$ is the transition rate from $\boldsymbol{x}$ to $\boldsymbol{x}'$ during the collection of CTRWs, $R_{\mathrm{CTRW}}(\boldsymbol{x})$ is the total escape rate from configuration $\boldsymbol{x}$, and $Z^{(\mathrm{bc})}_{\mathrm{CTRW}}$ is the path partition function for a collection of CTRWs. Eq. 5.21 is the generation probability for a Brownian bridge move. In general, the acceptance probability for a Brownian bridge move is

$$\begin{aligned} \mathrm{acc}[\mathscr{X}(T_\mathrm{B}) \to \mathscr{X}'(T_\mathrm{B})] &= \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T_\mathrm{B})] P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})][\![\boldsymbol{x}(t_0) = \boldsymbol{x}_{t_0}]\!][\![\boldsymbol{x}(t_f) = \boldsymbol{x}_{t_f}]\!]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T_\mathrm{B})] P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}'(T_\mathrm{B})][\![\boldsymbol{x}'(t_0) = \boldsymbol{x}_{t_0}]\!][\![\boldsymbol{x}'(t_f) = \boldsymbol{x}_{t_f}]\!]}\right] \\ &= \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T_\mathrm{B})] P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T_\mathrm{B})] P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}'(T_\mathrm{B})]}\right]. \end{aligned} \tag{5.23}$$

Using Eqs. 5.20 and 5.23, we can now formulate a general simulation algorithm for Brownian bridge moves that sample trajectories of the periodic ASEP. Pseudocode for a Brownian bridge moves is provided in Algorithm 1. A bridge is first generated through a kinetic Monte

Algorithm 1: Brownian bridge move

/* Generate new trajectory segment $\mathscr{X}'(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$ with a Brownian bridge */

Select random $t_0 \in [0, T]$

$t_f \leftarrow t_0 + T_\mathrm{B}$

Initialize $t \leftarrow t_0$, $\boldsymbol{x}'(t) \leftarrow \boldsymbol{x}_{t_0}$

while $t < t_f$ **do**

  Evaluate the $2N$ rates $\Gamma^{(\mathrm{bc})}_{\{+/-\}}(x'_{\{i\}}(t))$ for each potential single particle hop

  Evaluate $R_\mathrm{CTRW}(\boldsymbol{x}'(t)) = \sum_i^N \Gamma^{(\mathrm{bc})}_+(x'_i(t)) + \Gamma^{(\mathrm{bc})}_-(x'_i(t))$

  Sample $r_1, r_2 \sim \mathcal{U}(0,1)$

  $\Delta t \leftarrow -\ln(r_1)/R_\mathrm{CTRW}(\boldsymbol{x}'(t))$

  if $\Delta t > 0.01/R_\mathrm{CTRW}(\boldsymbol{x}'(t))$ **then**

    $\Delta t \leftarrow 0.01/R_\mathrm{CTRW}(\boldsymbol{x}'(t))$

    $\boldsymbol{x}'(t + \Delta t) \leftarrow \boldsymbol{x}'(t)$

  else

    $t \leftarrow t + \Delta t$

    Obtain $\boldsymbol{x}'(t + \Delta t)$ by updating $\boldsymbol{x}'(t)$ according to the single particle hop j which first satisfies $\sum_{j'=1}^{j} \Gamma^{(\mathrm{bc})}_{j'}(x'_{j'}(t)) < r_2 R_\mathrm{CTRW}(\boldsymbol{x}'(t))$

  end

end

/* Accept $\mathscr{X}'(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$ according to Metropolis criterion */

Evaluate $\alpha = \frac{P^{(\mathrm{pbc})}_\mathrm{ASEP}[\mathscr{X}'(T_\mathrm{B})] P^{(\mathrm{bc})}_\mathrm{CTRW}[\mathscr{X}(T_\mathrm{B})]}{P^{(\mathrm{pbc})}_\mathrm{ASEP}[\mathscr{X}(T_\mathrm{B})] P^{(\mathrm{bc})}_\mathrm{CTRW}[\mathscr{X}'(T_\mathrm{B})]}$

Sample $r_3 \sim \mathcal{U}(0,1)$

if $r_3 \leq \alpha$ **then**

  $\mathscr{X}(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}) \leftarrow \mathscr{X}'(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$

else

  $\mathscr{X}(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}) \leftarrow \mathscr{X}(T_\mathrm{B}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$

end

Carlo (KMC) simulation. During a conventional KMC simulation using the Gillespie algorithm, time is continuously advanced as particles hop between sites with specified rates. The time between hops is determined by sampling from the waiting time distribution. Because the hopping rates required to generate a bridge (Eq. 5.20) are time-dependent, the waiting time distribution also varies in time. During a bridge, we approximately sample from the exact waiting time distribution by only advancing time in intervals short enough that the exact waiting time distribution is nearly constant. Specifically, from $\boldsymbol{x}(t)$, time is advanced by no more than $\frac{0.01}{R_\mathrm{CTRW}(\boldsymbol{x}(t))}$, and only hops that would occur within a time of $\frac{0.01}{R_\mathrm{CTRW}(\boldsymbol{x}(t))}$ are performed. Further reducing the time interval where hops can occur would better approximate the exact dynamics; however, we have found a maximal time interval of $\frac{0.01}{R_\mathrm{CTRW}(\boldsymbol{x}(t))}$ to

be sufficient. Once a new trajectory segment, $\mathscr{X}'(T_{\mathrm{B}}; \boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$, is generated with a bridge, it is accepted with the probability given in Eq. 5.23 using the Metropolis algorithm.

Having developed a general algorithm, we now present two types of Brownian bridge moves, which each utilize different expressions for the hopping rates $\Gamma^{(\mathrm{bc})}_{+/-}$ and for the acceptance probability. The first move generates bridges on a periodic lattice. All such bridges end at the same configuration, $\boldsymbol{x}_{t_f}$, on the periodic lattice. Thus, we call these moves "fixed-end Brownian bridge moves." During these moves, the particles' net displacements typically do not change since they all start and end at the same configuration. As a result, the current does not change either. Only when these bridges are long enough for particles to make multiple laps around the lattice does their net displacement, and thus the current, change. We additionally developed another bridge move that alters the final configuration by exploiting the fact that all particles are identical and by making use of a mapping between a lattice with periodic boundaries and free boundaries. We call these moves "variable-end Brownian bridge moves." By slightly altering the final configuration, the particles' net displacement, and thus, current, also changes.

Fixed-End Brownian Bridge Moves.

Hopping Rates for Generating a Bridge on a Periodic Lattice. To obtain expressions for $\Gamma^{(\mathrm{pbc})}_{+/-}$ that yield bridges on a periodic lattice, we derive the Green's function solutions $p^{(\mathrm{pbc})}(x(t)|x_{t_0})$ and $q^{(\mathrm{pbc})}(x(t)|x_{t_f})$. In doing so, we also obtain an expression for the probability of a particle's position during a bridge according to Eq. 5.18.

The forward master equation (Eq. 5.16) for a single particle performing an asymmetric CTRW on a lattice with periodic boundary conditions can be solved using a generating function. Commensurate with periodic boundary conditions, the generating function $G(k,t)$ is a discrete Fourier transform of $p^{(\mathrm{pbc})}(x(t))$,

$$G(k,t) = \frac{1}{L}\sum_{x=0}^{L-1} z_k^x p^{(\mathrm{pbc})}(x(t)), \tag{5.24}$$

where $z_k = \exp(2\pi i k/L)$ and $p^{(\mathrm{pbc})}(x(t))$ is the probability of the particle being at position x at time t on a periodic lattice. The time derivative of $G(k,t)$ is

$$\begin{aligned}
\frac{\partial}{\partial t}G(k,t) =& \frac{1}{L}\sum_{x=0}^{L-1} z_k^x \frac{\partial}{\partial t} p^{(\mathrm{pbc})}(x(t)) \\
=& \gamma_+ \frac{1}{L}\sum_{x=0}^{L-1} z_k^{x+1} p^{(\mathrm{pbc})}(x(t)) + \gamma_- \frac{1}{L}\sum_{x=0}^{L-1} z_k^{x-1} p^{(\mathrm{pbc})}(x(t)) \\
& - (\gamma_+ + \gamma_-)\frac{1}{L}\sum_{x=0}^{L-1} z_k^x p^{(\mathrm{pbc})}(x(t)) \\
=& \left[\gamma_+(z_k - 1) + \gamma_-(z_k^{-1} - 1)\right] G(k,t),
\end{aligned} \tag{5.25}$$

where the forward master equation for $p^{(\mathrm{pbc})}(x(t))$ was substituted for $\frac{\partial}{\partial t}p^{(\mathrm{pbc})}(x(t))$. Integrating $\frac{\partial}{\partial t}G(k,t)$ yields

$$G(k,t) = G(k_0)e^{\lambda_k t}, \tag{5.26}$$

where $G(k_0)$ describes the initial condition and

$$\lambda_k \;=\; \gamma_+(z_k - 1) + \gamma_-(z_k^{-1} - 1).$$

Without loss of generality, we set the time of the initial configuration to $t = 0$. In the case of a Brownian bridge, the initial configuration is fixed at x_{t_0} and $p^{(\mathrm{pbc})}(x(t_0)) = [\![x(t_0) = x_{t_0}]\!]$. Thus, from Eq. 5.24, $G(k_0) = z_k^{x_{t_0}}/L$. The Green's function solution for $p^{(\mathrm{pbc})}(x(t))$ is then obtained by an inverse Fourier transform of $G(k,t)$,

$$p^{(\mathrm{pbc})}(x(t)|x_{t_0}) = \frac{1}{L}\sum_{k=0}^{L-1} e^{\lambda_k t} z_k^{x_{t_0}-x(t)}. \tag{5.27}$$

The Green's function solution to the backwards master equation (Eq. 5.17) for a particle performing an asymmetric CTRW on a periodic lattice is derived analogously to the solution to the forward master equation. Using the change of variables $\tau = t_f - t$, the generating function,

$$F(k,\tau) = \frac{1}{L}\sum_{x=0}^{L-1} z_k^x q^{(\mathrm{pbc})}(x(t_f - \tau)), \tag{5.28}$$

is a discrete Fourier transform of $q^{(\mathrm{pbc})}(x(t))$. The time derivative of $F(k,\tau)$ is

$$\begin{aligned}
\frac{\partial}{\partial \tau}F(k,\tau) =& \frac{\partial t}{\partial \tau}\frac{\partial}{\partial t}\left[\frac{1}{L}\sum_{x=0}^{L-1} z_k^x q^{(\mathrm{pbc})}(x(t))\right] \\
=& \gamma_+\frac{1}{L}\sum_{x=0}^{L-1} z_k^{x-1} q^{(\mathrm{pbc})}(x(t_f-\tau)) + \gamma_-\frac{1}{L}\sum_{x=0}^{L-1} z_k^{x+1} q^{(\mathrm{pbc})}(x(t_f-\tau)) \\
& - (\gamma_+ + \gamma_-)\frac{1}{L}\sum_{x=0}^{L-1} z_k^x q^{(\mathrm{pbc})}(x(t_f-\tau)) \\
=& \left[\gamma_+(z_k^{-1} - 1) + \gamma_-(z_k - 1)\right] F(k,\tau),
\end{aligned} \tag{5.29}$$

where the backwards master equation for $q^{(\mathrm{pbc})}(x(t))$ was substituted for $\frac{\partial}{\partial t}q^{(\mathrm{pbc})}(x(t))$. Integrating $\frac{\partial}{\partial \tau}F(k,\tau)$ yields

$$F(k,\tau) = F(k_0)e^{\lambda_k^\dagger \tau}, \tag{5.30}$$

where $F(k_0)$ describes the initial condition at $\tau = 0$, which is also the configuration at the final time $t = t_f$, and

$$\lambda_k^\dagger = \gamma_+(z_k^{-1} - 1) + \gamma_-(z_k - 1). \tag{5.31}$$

In the case of a Brownian bridge, the final configuration is fixed at x_{t_f} and $q^{(\mathrm{pbc})}(x(t_f)) = [\![x(t_f) = x_{t_f}]\!]$. Thus, from Eq. 5.28, $F(k_0) = z_k^{x_{t_f}}/L$. The Green's function solution for $q^{(\mathrm{pbc})}(x(t))$ is then obtained by an inverse Fourier transform of $F(k,\tau)$,

$$q^{(\mathrm{pbc})}(x(t)|x_{t_f}) = \sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t)}. \tag{5.32}$$

Bridges between configurations x_{t_0} and x_{t_f} on a periodic lattice are then generated using the hopping rates given by substituting Eq. 5.32 into Eq. 5.20,

$$\Gamma_+^{(\mathrm{pbc})}(x(t)) = \gamma_+ \frac{\sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t) - 1}}{\sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t)}} \tag{5.33a}$$

$$\Gamma_-^{(\mathrm{pbc})}(x(t)) = \gamma_- \frac{\sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t) + 1}}{\sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t)}}. \tag{5.33b}$$

From Eq. 5.18, bridges generated on a periodic lattice using these hopping rates sample the probability distribution

$$\mathcal{P}^{(\mathrm{pbc})}(x(t)|x_{t_0}, x_{t_f}) = \frac{\left(\sum_{k=0}^{L-1} e^{\lambda_k t} z_k^{x_{t_0} - x(t)}\right)\left(\sum_{k=0}^{L-1} e^{\lambda_k^\dagger (t_f - t)} z_k^{x_{t_f} - x(t)}\right)}{L \sum_{k=0}^{L-1} e^{\lambda_k^\dagger t_f} z_k^{x_{t_f} - x_{t_0}}}. \tag{5.34}$$

As shown in Figure 5.3, bridges generated using the hopping rates in Eq. 5.33 accurately sample the exact distribution $\mathcal{P}^{(\mathrm{pbc})}(x(t)|x_{t_0}, x_{t_f})$.

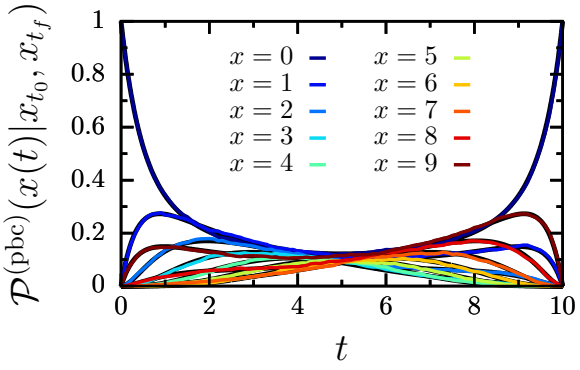


Figure 5.3: Brownian bridges on a periodic lattice are accurately sampled with KMC simulations using the hopping rates $\boldsymbol{\Gamma}^{(\mathrm{pbc})}_{\{+/-\}}(\boldsymbol{x(t)})$. The probability of a particle being at position x at time t during a Brownian bridge between $x_{t_0=0} = 0$ and $x_{t_f=10} = 0$ is plotted. Results from simulations are plotted in color, and analytical results (Eq. 5.34) are plotted in black. The particle hops between sites with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ on a lattice of $L = 10$ sites.

Acceptance Probability of fixed-end Brownian Bridge Moves. Once a new trajectory segment is generated through a bridge with the hopping rates in Eq. 5.33, it is accepted with probability

$$\mathrm{acc}[\mathscr{X}(T_\mathrm{B}) \to \mathscr{X}'(T_\mathrm{B})] = \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T_\mathrm{B})]P^{(\mathrm{pbc})}_{\mathrm{CTRW}}[\mathscr{X}(T_\mathrm{B})]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T_\mathrm{B})]P^{(\mathrm{pbc})}_{\mathrm{CTRW}}[\mathscr{X}'(T_\mathrm{B})]}\right]. \tag{5.35}$$

The acceptance probability is further simplified using the ratio of the probability of trajectory of the periodic ASEP (Eq. 5.10) to the probability of a trajectory of a collection of asymmetric CTRWs (Eq. 5.21),

$$\begin{aligned}\frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]}{P^{(\mathrm{pbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)]} =& \frac{\rho^{(\mathrm{pbc})}_{\mathrm{ASEP}}(\boldsymbol{x}(0))\theta[\mathscr{X}(T)]Z^{(\mathrm{pbc})}_{\mathrm{CTRW}}}{\rho^{(\mathrm{pbc})}_{\mathrm{CTRW}}(\boldsymbol{x}(0))Z^{(\mathrm{pbc})}_{\mathrm{ASEP}}} \\ &\times \exp\left[\sum_j (R_{\mathrm{CTRW}}(\boldsymbol{x}(t_j)) - R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j)))\Delta t_j\right].\end{aligned} \tag{5.36}$$

Thus, fixed-end Brownian bridge moves are accepted with probability

$$\mathrm{acc}[\mathscr{X}(T_\mathrm{B}) \to \mathscr{X}'(T_\mathrm{B})] = \min\left[1, \frac{\exp\left[-\sum_j R_{\mathrm{ASEP}}(\boldsymbol{x}'(t_j))\Delta t'_j\right]\theta[\mathscr{X}'(T_\mathrm{B})]}{\exp\left[-\sum_j R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j))\Delta t_j\right]\theta[\mathscr{X}(T_\mathrm{B})]}\right]. \tag{5.37}$$

This acceptance probability ensures that only bridges that respect constraints of volume exclusion are accepted. Because particles do not interact during a bridge, they can occupy the same lattice site. Due to the terms $R_{\mathrm{ASEP}}(\boldsymbol{x})$ and $\theta[\mathscr{X}(T_\mathrm{B})]$, only bridges that respect the constraints of volume exclusion are accepted. Figure 5.4 demonstrates that fixed-end bridge moves accurately sample trajectories of particles that exclude volume: During fixed-end bridge trajectories of a periodic ASEP with two particles, each particle occupies a given lattice site with the same probability as it does during a trajectory generated to obey the constraints of volume exclusion using conventional KMC simulations. Because the acceptance probability contains a term that scales exponentially with bridge length, $\exp\left[-\sum_j R_{\mathrm{ASEP}}(\boldsymbol{x}(t_j))\Delta t_j\right]$, bridges are increasingly rejected as their length is increased. As shown in Figure 5.5, the acceptance probability of unbiased bridge moves decreases as T_B increases. Good acceptance probabilities (greater than roughly 30%) are nevertheless obtained for unbiased bridges of lengths $T_\mathrm{B} \leq 5$. Even though the acceptance probability will decrease with an applied bias, the bridge length can be aptly chosen to obtain a desired acceptance probability. Thus, unlike shooting moves, bridges can be used to sample new segments in the middle of a trajectory even at high bias.

Variable-End Brownian Bridge Moves. We also developed a move that samples bridges with variable final configurations. To do so, we exploit the fact that all particles are identical.

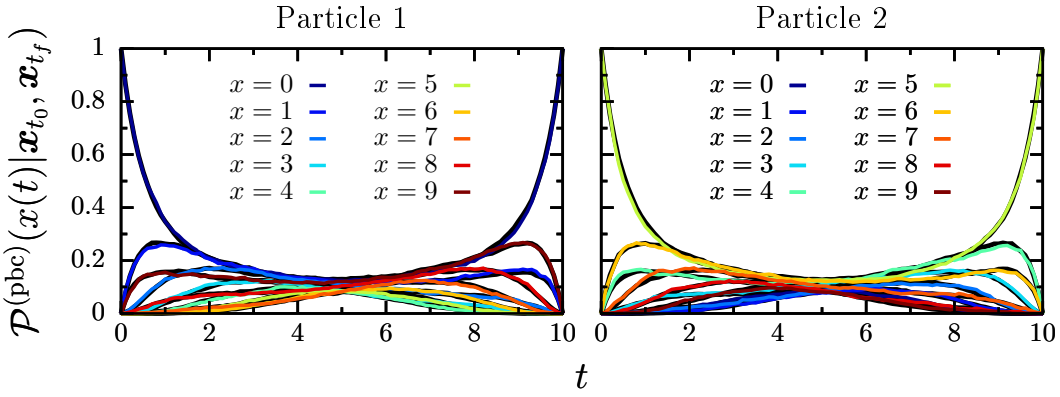


Figure 5.4: Fixed-end Brownian bridge moves accurately sample trajectories of particles that exclude volume. The probability of each particle being at position x at time t during a Brownian bridge between $\boldsymbol{x}_{t_0=0} = [0, 5]$ and $\boldsymbol{x}_{t_f=10} = [0, 5]$ is plotted. Results from fixed-end Brownian bridge moves are plotted in color, and results from conventional KMC simulations that were generated to obey the constraints of excluded volume are plotted in black. All KMC simulations were started from $\boldsymbol{x}_{t_0=0} = [0, 5]$ and those that did not end at $\boldsymbol{x}_{t_f=10} = [0, 5]$ were discarded to calculate $\mathcal{P}^{(\mathrm{pbc})}(x(t)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$. The two particles hop between sites with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ on a lattice with $L = 10$ sites.

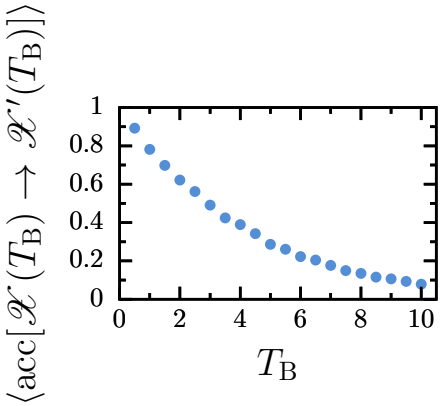


Figure 5.5: Acceptance of fixed-end Brownian bridge moves decreases with bridge length. Average acceptance probability is plotted versus bridge length, T_B. Results are shown for periodic ASEP trajectories of length $T = 10$ and of $N = 2$ particles that hop with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ between sites on a lattice with $L = 10$ sites.

This allows us to change the final configuration during a bridge by permuting particle indices such that same lattice sites are occupied in both the new and old final configurations, but by different particles. In order to generate variable-end bridge, we additionally use a mapping between a periodic lattice and infinite, linear lattice. On a periodic lattice, a particle that starts at position $\bar{x}_{t_0}$ may make any number of laps around the lattice during a trajectory and end at position $\bar{x}_{t_f}$. Positions on a periodic lattice, which we denote $\bar{x}$, map to positions

on a lattice with free boundaries, which we denote as x, according to

$$x_i = \bar{x}_i + n_i L \tag{5.38a}$$

$$\bar{x}_i = x_i - L \mod (x_i, L), \tag{5.38b}$$

where n is the signed number of laps that the particle makes around the lattice. In all previous equations, x denotes the positions commensurate with the specified lattice. When the particle's net displacement is clockwise, n is positive, whereas, when the particle's net displacement is counterclockwise, n is negative. When a particle's trajectory is mapped to a lattice with free boundaries, the particle would end at distinct positions x_{t_f} based on the number of laps it made around the lattice. An example of a new final configuration, $\boldsymbol{x}'_{t_f}$, obtained by permuting particle indices in the original final configuration, $\boldsymbol{x}_{t_f}$, is shown in Figure 5.6. On the lattice with free boundaries, it is evident that the net displacement of each particle in the new trajectory segment differs from its displacement in the old trajectory, thus changing the current. To generate bridges that end at a new final configuration, we first map configurations on a periodic lattice to those on a linear lattice with free boundary conditions. Next, to construct a $\boldsymbol{x}'_{t_f}$, each particle's index i in the final configuration is permuted by $\Delta i = \pm 1$ such that $x'_i(t_f) = x_{i+\Delta i}(t_f)$ while maintaining the overall order of particles. Each particle then performs an asymmetric CTRW from x_{t_0} to x'_{t_f} on the linear lattice with free boundary conditions. Configurations sampled by the bridge are subsequently mapped back to the periodic lattice. Finally, the new trajectory segment is accepted or rejected according to a Metropolis criterion that accounts for the fact that the bridge was generated for non-interacting particles on a lattice with free boundary conditions.

Hopping Rates for Generating a Bridge on an Infinite, Linear Lattice. To obtain expressions for $\Gamma^{(\mathrm{fbc})}_{+/-}$ that yield bridges on a linear lattice with free boundary conditions, we derive the Green's function solutions $p^{(\mathrm{fbc})}(x(t)|x_{t_0})$ and $q^{(\mathrm{fbc})}(x(t)|x_{t_f})$. In doing so, we also obtain an expression for the probability of a particle's position during the bridge according to Eq. 5.18.

The forward master equation (Eq. 5.16) for a single particle performing an asymmetric CTRW on a lattice with free boundary conditions can be solved by setting

$$p^{(\mathrm{fbc})}(x(t)|x_{t_0}) = \left(\frac{\gamma_+}{\gamma_-}\right)^{x(t)/2} \pi_p(x(t)|x_{t_0}) e^{-(\gamma_+ + \gamma_- - 2\sqrt{\gamma_+\gamma_-})t}, \tag{5.39}$$

where $\pi_p(x(t)|x_{t_0})$ is the probability of observing the particle at position x at time t during a symmetric CTRW with hopping rate $\gamma_{\mathrm{eff}} = \sqrt{\gamma_+\gamma_-}$ given that it started at configuration x_{t_0}. Without loss of generality, we set the initial time to $t = 0$. $\pi_p(x(t)|x_{t_0})$ evolves in time according to the forward master equation

$$\frac{\partial}{\partial t}\pi_p(x(t)|x_{t_0}) = \gamma_{\mathrm{eff}}\pi_p(x(t)-1|x_{t_0}) + \gamma_{\mathrm{eff}}\pi_p(x(t)+1|x_{t_0}) - 2\gamma_{\mathrm{eff}}\pi_p(x(t)|x_{t_0}). \tag{5.40}$$

The Green's function solution for this master equation is

$$\pi_p(x(t)|x_{t_0}) \;=\; I_{x(t)-x_{t_0}}(2\gamma_{\mathrm{eff}}t)e^{-2\gamma_{\mathrm{eff}}t}, \tag{5.41}$$

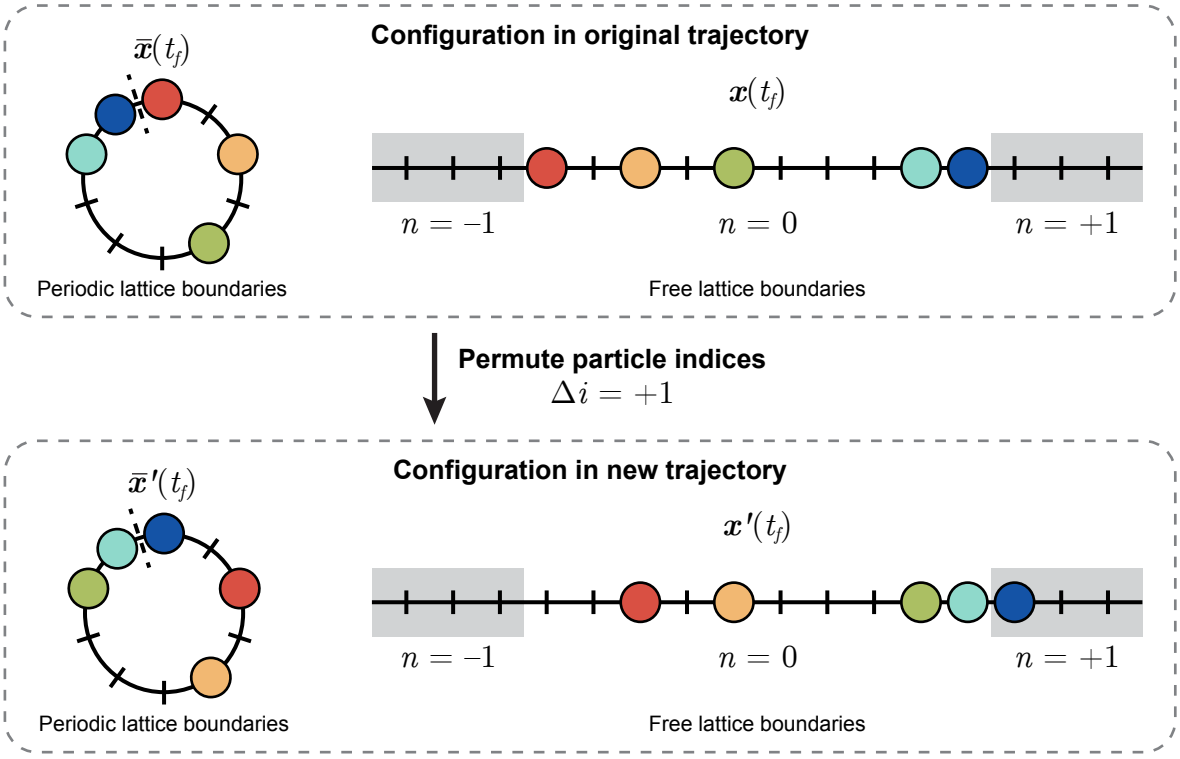


Figure 5.6: Variable-end Brownian bridge moves change the current by altering the final configuration of the bridge. The positions of each particle, which are illustrated in different colors, in the final configuration of the original trajectory $\bar{\boldsymbol{x}}(t_f)$ (top box) are altered to construct the final configuration in the new trajectory $\bar{\boldsymbol{x}}'(t_f)$ (bottom box). The dashed line indicates an arbitrarily chosen origin of the periodic lattice; the closest site clockwise from the origin is position $\bar{x} = 0, n = 0$, and the closest site counterclockwise from the origin is position $\bar{x} = 9, n = 0$. Particle indices i are permuted such that $x'_i(t_f) = x_{i+\Delta i}(t_f)$. On a lattice with periodic boundaries (left), uncolored $\bar{\boldsymbol{x}}(t_f)$ and $\bar{\boldsymbol{x}}'(t_f)$ are identical, but they are distinguishable on a lattice with free boundaries. Because the particle order is maintained when their indices are permuted, the net displacement of all particles is increased when their indices are permuted by $\Delta i = +1$, and particles appear on another periodic image of the lattice, which are highlighted in gray. Each periodic image of the lattice is labeled by how many images it is away from the original periodic lattice, n.

where $I_\nu(y)$ is the modified Bessel function of the first kind, since

$$\begin{aligned}\frac{\partial}{\partial t}\left[I_{x(t)-x_{t_0}}(2\gamma_{\text{eff}}t)e^{-2\gamma_{\text{eff}}t}\right] =& \gamma_{\text{eff}}\left(I_{x(t)-1-x_{t_0}}(2\gamma_{\text{eff}}t) + I_{x(t)+1-x_{t_0}}(2\gamma_{\text{eff}}t)\right)e^{-2\gamma_{\text{eff}}t} \\ &- 2\gamma_{\text{eff}}I_{x(t)-x_{t_0}}(2\gamma_{\text{eff}}t)e^{-2\gamma_{\text{eff}}t} \\ =& \gamma_{\text{eff}}\left(\pi_p(x(t)-1|x_{t_0}) + \pi_p(x(t)+1|x_{t_0})\right) - 2\gamma_{\text{eff}}\pi_p(x(t)|x_{t_0}) \\ =& \frac{\partial}{\partial t}\pi_p(x(t)|x_{t_0}). \end{aligned} \tag{5.42}$$

Substituting Eq. 5.41 into Eq. 5.39 yields the Green's function solution for $p^{(\text{fbc})}(x(t)|x_{t_0})$,

$$p^{(\text{fbc})}(x(t)|x_{t_0}) = \left(\frac{\gamma_+}{\gamma_-}\right)^{x(t)/2} I_{x(t)-x_{t_0}}(2\gamma_{\text{eff}}t)e^{-(\gamma_++\gamma_-)t}. \tag{5.43}$$

The Green's function solution to the backwards master equation (Eq. 5.17) for a particle performing an asymmetric CTRW on a lattice with free boundary conditions is derived analogously to the solution to the forward master equation. The solution can be obtained by setting

$$q^{(\text{fbc})}(x(t)|x_{t_f}) = \left(\frac{\gamma_+}{\gamma_-}\right)^{\frac{x_{t_f}-x(t)}{2}} \pi_q(x(t)|x_{t_f})e^{-(\gamma_++\gamma_--2\gamma_{\text{eff}})(t_f-t)}, \tag{5.44}$$

where $\pi_q(x(t)|x_{t_f})$ is the probability of observing the particle at position x at time t during a symmetric CTRW with hopping rate $\gamma_{\text{eff}} = \sqrt{\gamma_+\gamma_-}$ given that it ends at configuration x_{t_f}. $\pi_q(x(t)|x_{t_f})$ evolves in time according to the backwards master equation

$$\frac{\partial}{\partial t}\pi_q(x(t)|x_{t_f}) = -\gamma_{\text{eff}}\pi_q(x(t)+1|x_{t_f}) - \gamma_{\text{eff}}\pi_q(x(t)-1|x_{t_f}) + 2\gamma_{\text{eff}}\pi_q(x(t)|x_{t_f}). \tag{5.45}$$

The Green's function solution for this master equation is

$$\pi_q(x(t)|x_{t_f}) \;=\; I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))e^{-2\gamma_{\text{eff}}(t_f-t)}, \tag{5.46}$$

since

$$\begin{aligned}
&\frac{\partial}{\partial t}\left[I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))e^{-2\gamma_{\text{eff}}(t_f-t)}\right] \\
=&\gamma_{\text{eff}}\left(I_{x_{t_f}-x(t)-1}(2\gamma_{\text{eff}}(t_f-t)) + I_{x_{t_f}-x(t)+1}(2\gamma_{\text{eff}}(t_f-t))\right)e^{-2\gamma_{\text{eff}}(t_f-t)} \\
&- 2\gamma_{\text{eff}}I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))e^{-2\gamma_{\text{eff}}(t_f-t)} \\
=&\gamma_{\text{eff}}\left(\pi_q(x(t)+1|x_{t_f}) + \pi_q(x(t)-1|x_{t_f})\right) - 2\gamma_{\text{eff}}\pi_q(x(t)|x_{t_f}) \\
=&\frac{\partial}{\partial t}\pi_q(x(t)|x_{t_f}).
\end{aligned} \tag{5.47}$$

Substituting Eq. 5.46 into Eq. 5.44 yields the Green's function solution for $q^{(\text{fbc})}(x(t)|x_{t_f})$,

$$q^{(\text{fbc})}(x(t)|x_{t_f}) = \left(\frac{\gamma_+}{\gamma_-}\right)^{\frac{x_{t_f}-x(t)}{2}} I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))e^{-(\gamma_++\gamma_-)(t_f-t)}. \tag{5.48}$$

Bridges between configurations x_{t_0} and x_{t_f} on a lattice with free boundaries are then generated using the hopping rates given by substituting Eq. 5.48 into Eq. 5.20,

$$\Gamma_+^{(\text{fbc})}(x(t)) = \gamma_{\text{eff}}\frac{I_{x_{t_f}-x(t)-1}(2\gamma_{\text{eff}}(t_f-t))}{I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))} \tag{5.49a}$$

$$\Gamma_-^{(\text{fbc})}(x(t)) = \gamma_{\text{eff}}\frac{I_{x_{t_f}-x(t)+1}(2\gamma_{\text{eff}}(t_f-t))}{I_{x_{t_f}-x(t)}(2\gamma_{\text{eff}}(t_f-t))}. \tag{5.49b}$$

From Eq. 5.18, the bridges generated with these hopping rates sample the probability distribution

$$\mathcal{P}^{(\mathrm{fbc})}(x(t)|x_{t_0}, x_{t_f}) = \frac{I_{x(t)-x_{t_0}}(2\gamma_{\mathrm{eff}}t) I_{x_{t_f}-x(t)}(2\gamma_{\mathrm{eff}}(t_f - t))}{I_{x_{t_f}-x_{t_0}}(2\gamma_{\mathrm{eff}}t_f)}. \tag{5.50}$$

As shown in Figure 5.7, bridges generated using the hopping rates in Eq. 5.49 accurately sample the exact distribution $\mathcal{P}^{(\mathrm{fbc})}(x(t)|x_{t_0}, x_{t_f})$.

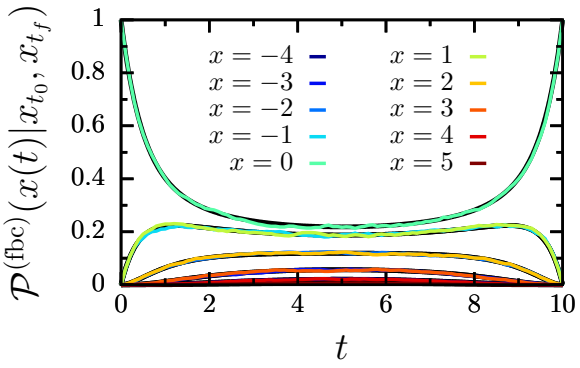


Figure 5.7: Brownian bridges on a linear lattice with free boundaries are accurately sampled with KMC simulations using the hopping rates $\boldsymbol{\Gamma}^{(\mathbf{fbc})}_{\{+/-\}}(\boldsymbol{x}(\boldsymbol{t}))$. The probability of a particle being at position x at time t during a Brownian bridge between $x_{t_0=0} = 0$ and $x_{t_f=10} = 0$ is plotted. Results from simulations are plotted in color, and analytical results (Eq. 5.50) are plotted in black. The particle hops between sites with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$.

Acceptance Probability of variable-end Brownian Bridge Moves. Once a new trajectory segment is generated through a bridge with the hopping rates in Eq. 5.49, it is accepted with probability

$$\mathrm{acc}[\mathscr{X}(T_{\mathrm{B}}) \to \mathscr{X}'(T_{\mathrm{B}})] = \min\left[1, \frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}'(T_{\mathrm{B}})] P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T_{\mathrm{B}})|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}]}{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T_{\mathrm{B}})] P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}'(T_{\mathrm{B}})|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}]}\right]. \tag{5.51}$$

For any given model (ASEP or CTRW), the probability of a trajectory that starts at $\bar{\boldsymbol{x}}_{t_0} = \boldsymbol{x}_{t_0}$ and ends at $\bar{\boldsymbol{x}}_{t_f}$ on a lattice with periodic boundaries, $P^{(\mathrm{pbc})}_{\mathrm{model}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \bar{\boldsymbol{x}}_{t_f}]$, is related to the probability on a lattice with free boundaries, $P^{(\mathrm{fbc})}_{\mathrm{model}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \bar{\boldsymbol{x}}_{t_f}]$, according to

$$P^{(\mathrm{pbc})}_{\mathrm{model}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \bar{\boldsymbol{x}}_{t_f}] = \frac{1}{Z^{(\mathrm{pbc})}_{\mathrm{model}}(\bar{\boldsymbol{x}}_{t_f}|\boldsymbol{x}_{t_0})} \sum_n P^{(\mathrm{fbc})}_{\mathrm{model}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}] \prod_i^N [\![x_i(t_f) = \bar{x}_{t_f,i} + n_i L]\!] \tag{5.52a}$$

$$Z^{(\mathrm{pbc})}_{\mathrm{model}}(\bar{\boldsymbol{x}}_{t_f}|\boldsymbol{x}_{t_0}) = \sum_{\mathscr{X}(t_f)} \sum_n P^{(\mathrm{fbc})}_{\mathrm{model}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}] \prod_i^N [\![x_i(t_f) = \bar{x}_{t_f,i} + n_i L]\!], \tag{5.52b}$$

where $Z^{(\mathrm{pbc})}_{\mathrm{model}}(\bar{\boldsymbol{x}}_{t_f}|\boldsymbol{x}_{t_0})$ is the path partition function for trajectories on a periodic lattice that start at $\boldsymbol{x}_{t_0}$ and end at $\bar{\boldsymbol{x}}_{t_f}$. Thus, trajectories on a periodic lattice can be sampled by generating Brownian bridges on a lattice with free boundaries as long as different values of n are also sampled. The fact that each particle performs an asymmetric CTRW independently of all other particles allows us to simplify the path partition function for a collection of CTRWs,

$$\begin{aligned} Z^{(\mathrm{bc})}_{\mathrm{CTRW}}(\boldsymbol{x}_{t_f}|\boldsymbol{x}_{t_0}) &= \prod_i^N \sum_{\mathscr{X}_i(T)} P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}_i(T)|x_{t_0,i}][\![x_i(t_f) = x_{t_f,i}]\!] \\ &= \prod_i^N p^{(\mathrm{bc})}(x_{t_f,i}|x_{t_0,i}). \end{aligned} \tag{5.53}$$

Here we have identified $Z^{(\mathrm{bc})}_{\mathrm{CTRW}}(\boldsymbol{x}_{t_f}|\boldsymbol{x}_{t_0})$ as the average of $[\![x_i(t_f) = x_{t_f,i}]\!]$, which is the Green's function $p^{(\mathrm{bc})}(x(t)|x_{t_0})$ evaluated at time $t = t_f$. Thus, the probability of a trajectory that is a collection of CTRWs that start and end at the same configuration is

$$P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}] = \left(\prod_i^N p^{(\mathrm{bc})}(x_{t_f,i}|x_{t_0,i})\right)^{-1} P^{(\mathrm{bc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}] \prod_i^N [\![x_i(t_f) = x_{t_f,i}]\!]. \tag{5.54}$$

Using the above relationships, we simplify the ratio $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]/P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}]$ in the acceptance probability in Eq. 5.51 as follows: Using Eqs. 5.36 and 5.52, the probability of a periodic ASEP trajectory can be written in terms of the probability of a collection of independent asymmetric CTRWs on a lattice with free boundaries. Using Eq. 5.54, the probability of a collection of independent asymmetric CTRWs that start from $\boldsymbol{x}_{t_0}$ is then

written in terms of bridge trajectories that not only start at $\boldsymbol{x}_{t_0}$ but also end at $\boldsymbol{x}_{t_f}$.

$$\begin{aligned}
P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)] \propto & P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \bar{\boldsymbol{x}}_{t_f}] \\
\propto & \sum_n P^{(\mathrm{fbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}] \prod_i^N [\![x_i(t_f) = \bar{x}_{t_f,i} + n_i L]\!] \\
\propto & \sum_n P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}] \prod_i^N [\![x_i(t_f) = \bar{x}_{t_f,i} + n_i L]\!] \\
& \times \exp\left[\sum_j \Delta R(\bar{\boldsymbol{x}}(t_j))\Delta t_j\right] \theta[\mathscr{X}(T)] \\
= & \sum_n P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \bar{\boldsymbol{x}}_{t_f} + \boldsymbol{n}L] \prod_i^N p^{(\mathrm{fbc})}(\bar{x}_{t_f,i} + n_i L|x_{t_0,i}) \\
& \times \exp\left[\sum_j \Delta R(\bar{\boldsymbol{x}}(t_j))\Delta t_j\right] \theta[\mathscr{X}(T)] \\
= & P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}] \prod_i^N p^{(\mathrm{fbc})}(x_{t_f,i}|x_{t_0,i}) [\![\bar{x}_i(t_f) = \bar{x}_{t_f,i}]\!] \\
& \times \exp\left[\sum_j \Delta R(\bar{\boldsymbol{x}}(t_j))\Delta t_j\right] \theta[\mathscr{X}(T)],
\end{aligned} \tag{5.55}$$

where $\Delta R(\bar{\boldsymbol{x}}(t_j)) = R_{\mathrm{CTRW}}(\bar{\boldsymbol{x}}(t_j)) - R_{\mathrm{ASEP}}(\bar{\boldsymbol{x}}(t_j))$. The final simplification was made using the mapping between positions on a periodic lattice and positions on a lattice with free boundaries (Eq. 5.38). Using this expression for $P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]$,

$$\begin{aligned}
\frac{P^{(\mathrm{pbc})}_{\mathrm{ASEP}}[\mathscr{X}(T)]}{P^{(\mathrm{fbc})}_{\mathrm{CTRW}}[\mathscr{X}(T)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f}]} \propto & \prod_i^N p^{(\mathrm{fbc})}(x_{t_f,i}|x_{t_0,i}) [\![\bar{x}_i(t_f) = \bar{x}_{t_f,i}]\!] \\
& \times \exp\left[\sum_j \Delta R(\bar{\boldsymbol{x}}(t_j))\Delta t_j\right] \theta[\mathscr{X}(T)].
\end{aligned} \tag{5.56}$$

Thus, variable-end Brownian bridge moves are accepted with probability

$$\begin{aligned}
&\mathrm{acc}[\mathscr{X}(T_{\mathrm{B}}) \to \mathscr{X}'(T_{\mathrm{B}})] = \\
&\quad \min\left[1, \frac{\prod_i^N p^{(\mathrm{fbc})}(x'_{t_f,i}|x_{t_0,i}) \exp\left[-\sum_j R_{\mathrm{ASEP}}(\bar{\boldsymbol{x}}'(t_j))\Delta t'_j\right] \theta[\mathscr{X}'(T_{\mathrm{B}})]}{\prod_i^N p^{(\mathrm{fbc})}(x_{t_f,i}|x_{t_0,i}) \exp\left[-\sum_j R_{\mathrm{ASEP}}(\bar{\boldsymbol{x}}(t_j))\Delta t_j\right] \theta[\mathscr{X}(T_{\mathrm{B}})]}\right].
\end{aligned} \tag{5.57}$$

The acceptance probability includes the same terms in the acceptance probability of fixed-end bridge moves (Eq. 5.37) to ensure that only bridges that respect constraints of volume

exclusion are accepted. Figure 5.8 demonstrates that variable-end bridge moves accurately sample trajectories of particles that exclude volume and can swap final positions: During variable-end bridge trajectories of a periodic ASEP with two particles, each particle occupies a given lattice site with the same probability as it does during a trajectory generated to obey the constraints of volume exclusion using conventional KMC simulations. The additional

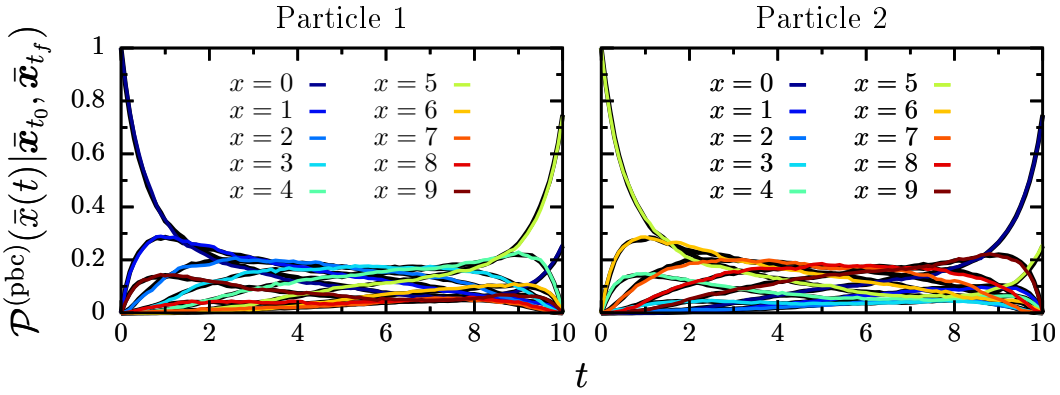


Figure 5.8: Variable-end Brownian bridge moves accurately sample trajectories of particles that exclude volume. The probability of each particle being at position x at time t during a Brownian bridge between $\boldsymbol{x}_{t_0=0} = [0, 5]$ and $\boldsymbol{x}_{t_f=10} = \{[0, 5], [5, 0]\}$ is plotted. Results from variable-end Brownian bridge moves in which $\Delta i = \{-1, 0, +1\}$ are plotted in color, and results from conventional KMC simulations that were generated to obey the constraints of excluded volume are plotted in black. All KMC simulations were started from $\boldsymbol{x}_{t_0=0} = [0, 5]$ and those that did not end at $\boldsymbol{x}_{t_f=10} = \{[0, 5], [5, 0]\}$ were discarded to calculate $\mathcal{P}^{(\mathrm{pbc})}(x(t)|\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_f})$. The two particles hop between sites with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ on a lattice with $L = 10$ sites.

term in the acceptance probability, $\frac{\prod_i^N p^{(\mathrm{fbc})}(x'_{t_f,i}|x_{t_0,i})}{\prod_i^N p^{(\mathrm{fbc})}(x_{t_f,i}|x_{t_0,i})}$, accounts for differences between the new and old final configurations. Due to this term, variable-end bridge moves are accepted with lower probability than fixed-end bridges when $\Delta i \neq 0$ (Figure 5.9). Non-negligible acceptance probabilities for $\Delta i = \pm 1$ and $N = 2$ are nevertheless obtained for unbiased bridges of lengths $T_{\mathrm{B}} \geq 2$. When $\Delta i = \pm 1$ and $N = 2$, longer bridges are required for a particle to complete an extra lap around the lattice. As N increases, though, particles do not have to hop as far to reach their new final position when $\Delta i = \pm 1$ (see Figure 5.6 for example), such that shorter variable-end bridges may be accepted with higher probability. Thus, these bridge moves can be used to efficiently sample new segments in the middle of a trajectory with different currents.

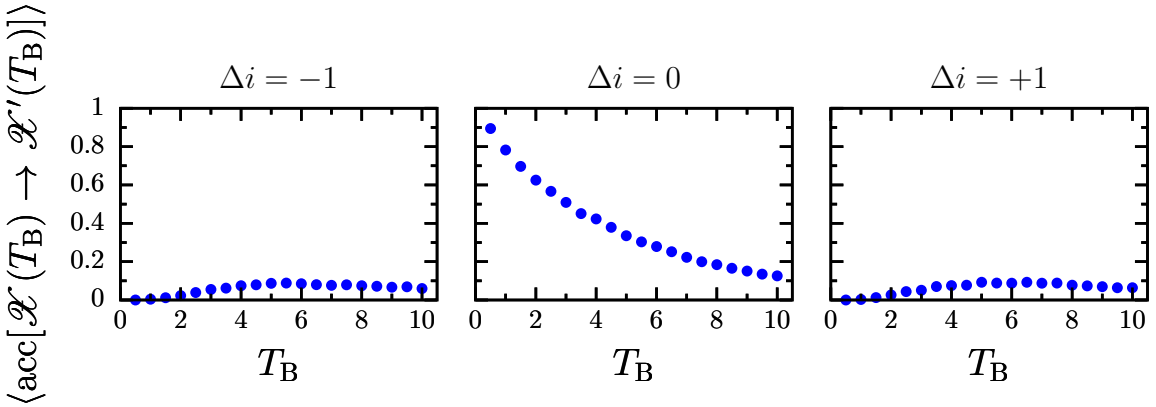


Figure 5.9: Acceptance of variable-end Brownian bridge moves depends on bridge length and Δi. Average acceptance probabilities for particle permutations of $\Delta i = \{-1, 0, +1\}$ are plotted versus bridge length, T_B. Results are shown for periodic ASEP trajectories of length $T = 10$ and of $N = 2$ particles that hop with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ between sites on a lattice with $L = 10$ sites.

Brownian Bridge Moves Sample Biased Trajectories of the Periodic ASEP More Efficiently than Shooting Moves

As an initial test of our new moves, we evaluate the efficiency of Brownian bridge moves versus shooting moves at sampling trajectories of a periodic ASEP with $N = 2$ particles, which each hop with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ between sites on a lattice with $L = 10$ sites. All other simulation parameters are given in Appendix D. Specifically, we evaluate the moves' abilities to sample rare values of the activity (Eq. 5.11) and current (Eq. 5.12) over an observation time of $T = 10$. To enhance the sampling of rare values of the activity and current we tilt the trajectory ensemble by applying a bias s on the activity or a bias λ on the current (Eq. 5.7). This amounts to including the factor $\exp[-sK[\mathscr{X}'(T)] + sK[\mathscr{X}(T)]]$ or $\exp[-\lambda J[\mathscr{X}'(T)] + \lambda J[\mathscr{X}(T)]]$ in the acceptance probabilities given in the previous section (Eqs. 5.14, 5.15, 5.37, and 5.57) to bias the activity, K, and current, J, respectively.

Moves must be accepted with reasonable probability (roughly 30% is often optimal) at high bias, to efficiently sample trajectories with rare activity or current. As shown in Figure 5.10, black points, more than half of all shooting moves are accepted when the magnitude of the bias is less than 0.25, but very few moves are accepted at high bias. At high bias, trajectory segments generated by shooting moves are seldom accepted, making shooting moves inefficient at sampling rare values of the activity or current. Shooting moves of shorter lengths (roughly lengths $T_\mathrm{B} < 1.5$) are accepted with reasonable probabilities at high bias; however, they only alter the very beginning or end of the trajectory and leave the middle unchanged. In contrast, fixed-end Brownian bridge moves are accepted frequently at high bias when the bridge length, T_B, is chosen appropriately, and they alter the middle of the trajectory. Figure 5.11 shows the acceptance probability of biased fixed-end bridges with

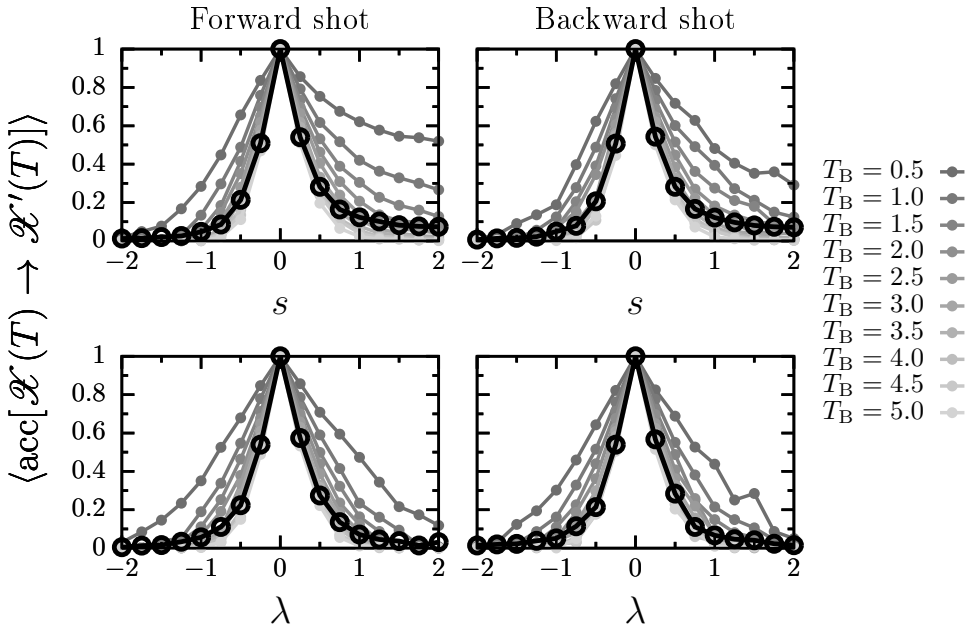


Figure 5.10: Shooting moves are seldom accepted at high bias. Average acceptance probabilities of (left) forward and (right) backward shooting moves are plotted versus bias value. Results for biasing (top) the activity by s and (bottom) the current by λ are shown for periodic ASEP trajectories of length $T = 10$ and of $N = 2$ particles that hop with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ between sites on a lattice with $L = 10$ sites. Average acceptances of shooting moves of any length are shown in black, and acceptances of shooting moves of length T_B are shown in shades of gray.

$0.5 \leq T_\mathrm{B} \leq 5$. As $|s|$ increases, the acceptance probability of fixed-end bridge decreases in a fashion similar to shooting moves. The acceptance probability decays asymmetrically from $s = 0$ due to an underlying asymmetry in the large deviation function $\psi(s)$, which further hinders sampling for $s < 0$ compared to $s > 0$. For all values of s, shorter fixed-end bridges are accepted with higher probability. When the current is biased, the acceptance probability of longer fixed-end bridges (roughly $T_\mathrm{B} > 3.5$) also decreases in a fashion similar to shooting moves. The acceptance probability decays roughly symmetrically from $\lambda = 0$ since the large deviation function $\psi(\lambda)$ is symmetric. For all values of λ, shorter fixed-end bridges are accepted with higher probability. The acceptance probability of short fixed-end bridges (roughly $T_\mathrm{B} < 1.5$) does not vary significantly with λ. When fixed-end bridges are short, particles are unlikely to make additional laps around the periodic lattice during the bridge such that the particles' net displacements do not change. As a result, the current in the new trajectory segment is equal to the current in the old segment, and the term in the acceptance probability that accounts for the bias has no influence since $\exp[-\lambda J[\mathscr{X}'(T)] + \lambda J[\mathscr{X}(T)]] = 1$. Thus, short fixed-end bridges do not efficiently sample different values of the current. However, when fixed-end bridges are long enough for particles to make multiple laps around

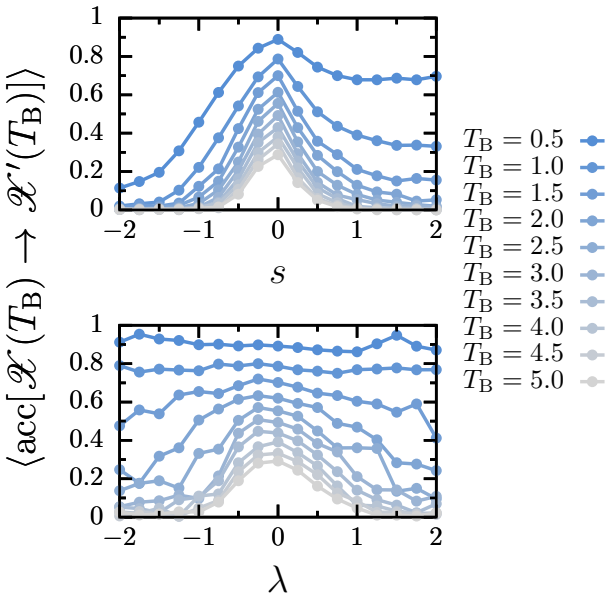


Figure 5.11: Short fixed-end Brownian bridge moves are accepted at high bias. Average acceptance probabilities of fixed-end Brownian bridge of different bridge lengths, T_B, are plotted versus bias value. Results for biasing (top) the activity by s and (bottom) the current by λ are shown for periodic ASEP trajectories of length $T = 10$ and of $N = 2$ particles that hop with rates $\gamma_+ = 1$ and γ_- between sites on a lattice with $L = 10$ sites.

the periodic lattice, the current in the new segment differs from the current in the old segment, and the acceptance probability varies with λ. Importantly, fixed-end bridge moves are accepted at large $|s|$ and $|\lambda|$ when T_B is chosen appropriately. Thus, unlike shooting moves, fixed-end bridge moves can be used to efficiently sample rare values of the activity or current.

Variable-end Brownian bridge moves allow the current to be altered in a controlled way by swapping particle positions in the final configuration of the bridge. When the final configuration is not changed, which amounts to permuting particle indices by $\Delta i = 0$, the acceptance probability of variable-end bridges biased by s (Figure 5.12) is similar to the acceptance probability of fixed-end bridges (Figure 5.11). When $\Delta i = 0$, the acceptance probability of variable-end bridges of any length does not vary with λ since the current does not change during the bridge. Although variable-end bridges with $\Delta i = \pm 1$ are accepted with much lower probability, their acceptance probability is non-negligible for biases of roughly $-1 \leq s \leq 0$ and $-0.5 \leq \lambda \leq 0.5$. At low bias, longer moves are accepted with higher probability than shorter ones because longer bridges provide the particles with more time to make the necessary number of hops to reach their new final configuration. When the indices of the two particles are permuted, one particle must make at least one full additional lap

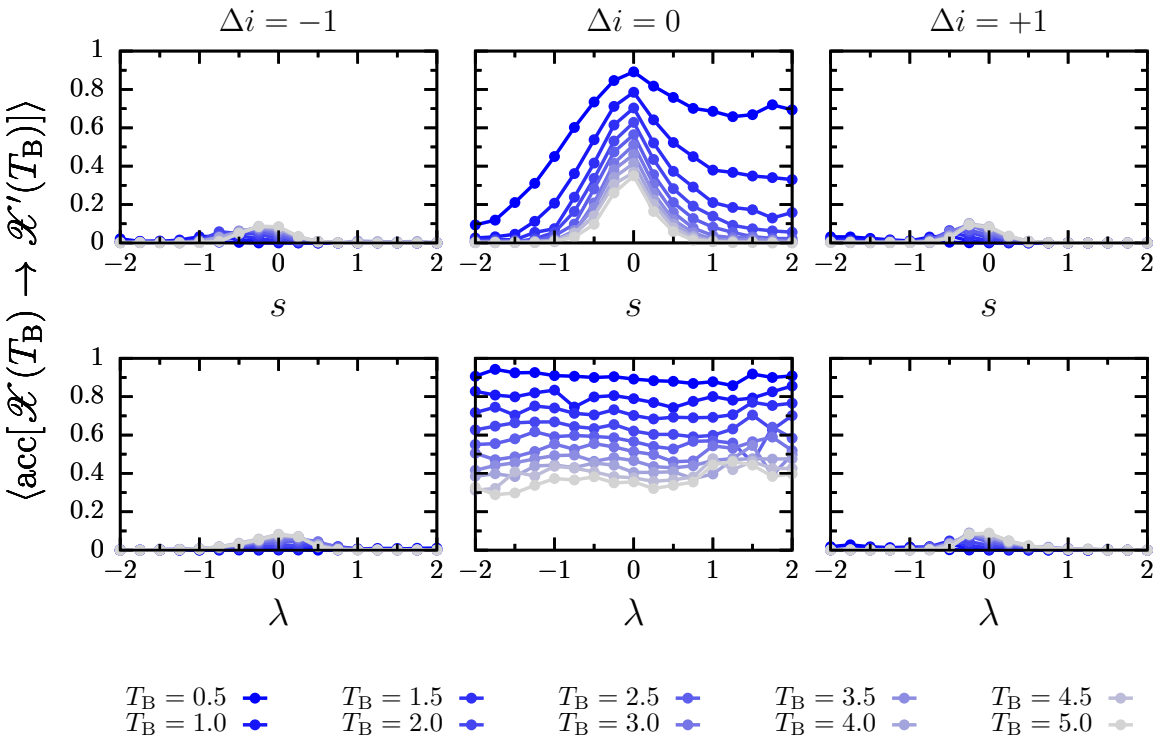


Figure 5.12: Variable-end Brownian bridge moves are accepted at small bias. Average acceptance probabilities of variable-end Brownian bridge with particle permutations of $\Delta i = \{-1, 0, +1\}$ and different bridge lengths, T_B, are plotted versus bias value. Results for biasing (top) the activity by s and (bottom) the current by λ are shown for periodic ASEP trajectories of length $T = 10$ and of $N = 2$ particles that hop with rates $\gamma_+ = 1$ and γ_- between sites on a lattice with $L = 10$ sites.

around the periodic lattice in the new trajectory compared to the old one. In an unbiased trajectory of the periodic ASEP with our chosen parameters, this takes roughly a total time of $T_\mathrm{B} = 5$ and is less likely to occur as T_B decreases. For $\lambda > 0$, which biases trajectories towards negative currents, variable-end bridges with particle indices permuted by $\Delta i = -1$ are accepted more since they can create negative currents. In contrast, for $\lambda < 0$, which biases trajectories towards positive currents, bridge with particle indices permuted by $\Delta i = +1$ are accepted more since they can create more positive currents. Since variable-end bridges with $\Delta i = \pm 1$ change the current, which is correlated with the activity, judicious use of variable-end bridge moves with $\Delta i = \pm 1$ in addition to fixed-end bridge moves can further enhance the sampling of rare values of the activity or current.

The efficiency of TPS simulations depends not only on moves being accepted with sufficient probabilities, but more importantly by how readily those moves generate new, independent trajectories. If accepted moves only make small changes to an existing trajectory,

trajectories remain highly correlated. If accepted moves instead make large changes, trajectories decorrelate quickly, fewer moves are required to fully sample the trajectory ensemble, and overall computational efficiency of the simulation increases. To quantify how quickly trajectories decorrelate during biased simulations, we compute correlations functions of either the activity or current as a function of the number of moves, τ, according to

$$C_\zeta(\tau) = \frac{\langle \delta\mathcal{O}[\mathscr{X}_0(T)]\delta\mathcal{O}[\mathscr{X}_\tau(T)]\rangle_\zeta}{\langle \delta\mathcal{O}[\mathscr{X}(T)]^2\rangle_\zeta}, \tag{5.58}$$

where $\langle\cdots\rangle_\zeta$ indices an ensemble average at a bias value of ζ, $\delta\mathcal{O}[\mathscr{X}(T)] = \mathcal{O}[\mathscr{X}(T)] - \langle\mathcal{O}\rangle_\zeta$ is the variation in the observable biased by ζ, and each trajectory is indexed by the number of moves performed. Correlation functions from biased simulations using (1) only shooting moves, (2) fixed-end bridges and shooting moves, and (3) variable-end bridges with $\Delta i = \pm 1$, fixed-end bridges, and shooting moves are shown in Figure 5.13. For all combinations of moves, trajectories decorrelate slower as the magnitude of the bias is increased since fewer moves are accepted at high bias (Figures 5.10, 5.11, and 5.12). When the bias on the activity is large and negative, using fixed-end bridges in addition to shooting moves leads to more rapid decorrelation of the activity. Because the activity and current are correlated, additionally using variable-end bridges with particle permutations of $\Delta i = \pm 1$ further enhances sampling of trajectories with large activity typical of $s \ll 0$. When the bias on the current is large, using fixed-end bridges in addition to shooting moves can lead to more rapid decorrelation of the current since the bridges can modify the middle of the trajectory. However, because short fixed-end bridges ($T_\mathrm{B} \leq 3$) were used in the simulations, they do not readily change the current and in fact slow decorrelation for some values of λ. Additionally using variable-end bridges to change the current in a controlled way leads to significantly more rapid decorrelation. Thus, Brownian bridge moves, and especially variable-end bridges that alter the current, improve the efficiency of simulations at high bias.

Finally, we examined whether Brownian bridge moves enable accurate sampling of exceptionally rare values of the activity and current, which are otherwise very difficult to sample with shooting moves alone. Sampling such trajectories is necessary to accurately calculate the large deviation functions $\psi(s)$ and $\psi(\lambda)$ over a wide range of values from TPS simulations. Rate functions and corresponding large deviation functions were calculated from biased simulations using WHAM. Simulations were performed using (1) only shooting moves, (2) fixed-end bridges and shooting moves, and (3) variable-end bridges with $\Delta i = \pm 1$, fixed-end bridges, and shooting moves. Rate functions and large deviation functions calculated from each of these simulations are shown in Figure 5.14. For comparison, the exact rate functions and large deviation functions obtained through matrix diagonalization are also plotted. Trajectories with exceptionally large values of the activity and current are sampled when Brownian bridge moves are used. This allows the rate functions $I(K/T)$ and $I(J/T)$ to be accurately estimated over a wider range than possible using shooting moves alone. Simulations using fixed-end bridges and simulations using fixed-end and variable-end bridges both yield accurate estimates of $I(K/T)$. Furthermore, the large deviation function $\psi(s)$ calculated from both simulations that use bridges agrees with the exact result, with only

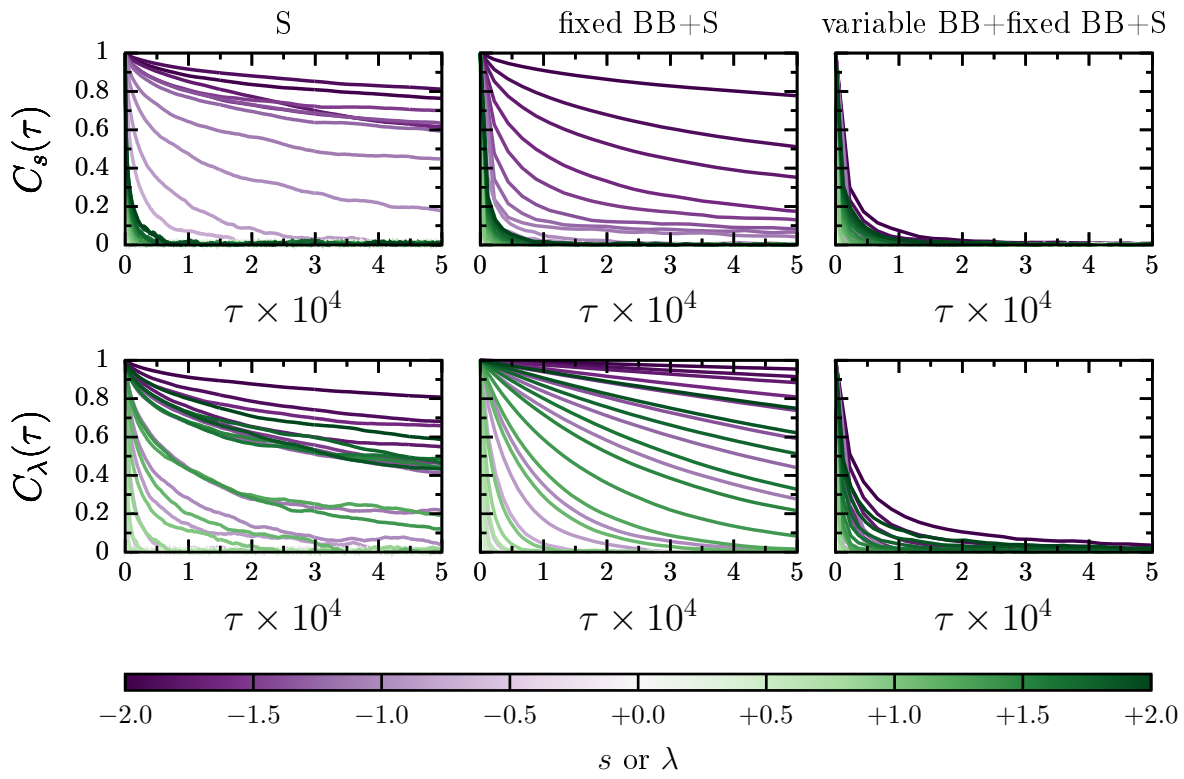


Figure 5.13: Brownian bridge moves decorrelate trajectories more efficiently than shooting moves at high bias. (Top) Correlation function of the activity, $C_s(\tau)$ as a function of moves, τ, during simulations biased by s. (Bottom) Correlation function of the current, $C_\lambda(\tau)$ as a function of moves during simulations biased by λ. Results are shown for TPS simulations using (left) only forward and backward shooting moves ("S"), (middle) a majority of fixed-end Brownian bridges in addition to shooting moves ("fixed BB+S"), and (right) an equal amount of fixed-end bridge moves and variable-end bridge moves that permute particles by $\Delta i = \pm 1$ in addition to shooting moves ("variable BB+fixed BB+S").

minor deviations at large positive values of s. In contrast, $\psi(s)$ calculated from simulations that only use shooting moves significantly deviates from the exact result at large negative values of s because trajectories with large activity, which contribute most to $\psi(s)$ at large negative s are not sampled. Simulations using bridges also yield reasonably accurate estimates of $I(J/T)$. The large deviation function $\psi(\lambda)$ calculated from both simulations that use bridges reasonably agrees with the exact result; better agreement is obtained at large negative λ when only fixed-end bridges are performed whereas better agreement is obtained at large positive λ when variable-end bridges are also used. By further tuning the relative number of fixed-end and variable-end bridges and their respective bridge lengths, accurate results for both large deviation functions over the range $-2 \leq \lambda \leq 2$ ought to be obtainable from simulations. In contrast, $\psi(s)$ calculated from simulations that only use shooting moves

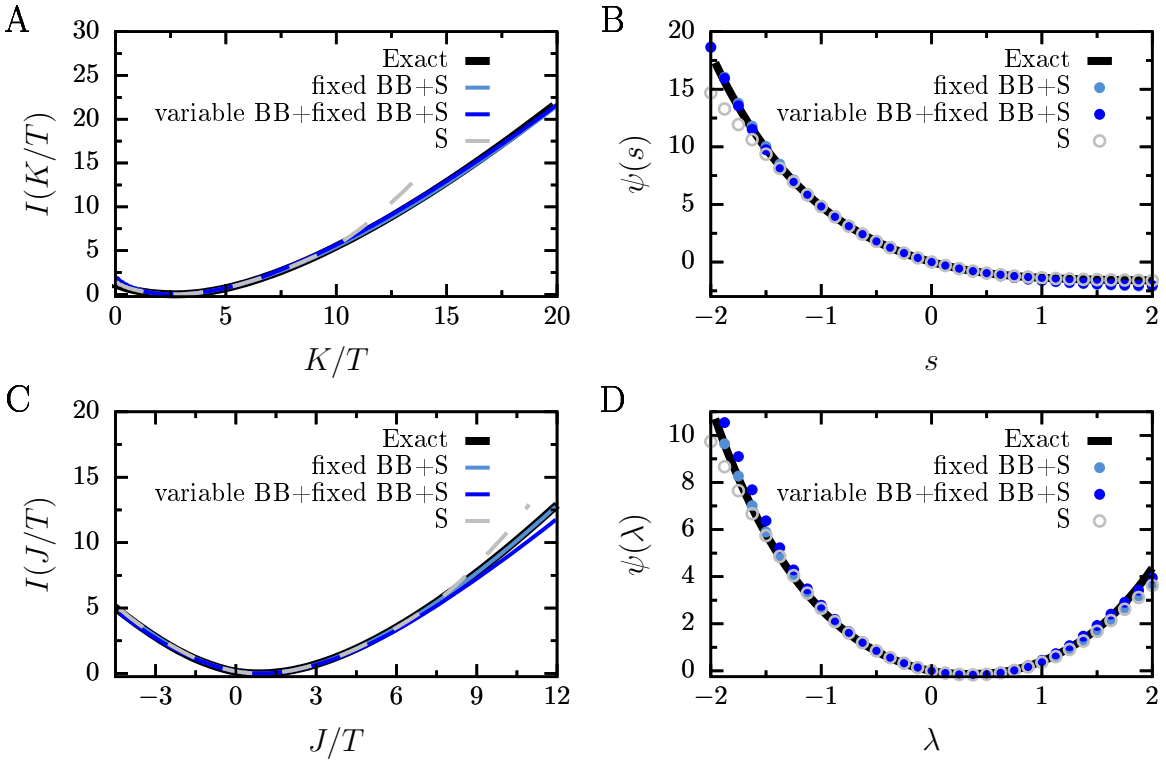


Figure 5.14: Brownian bridge moves sample rare trajectories of the periodic ASEP more readily than shooting moves. (A and C) Rate functions and (B and D) large deviation functions for the (A and B) activity and (C and D) current of a periodic ASEP with $N = 2$ particles that hop with rates $\gamma_+ = 1$ and $\gamma_- = 1/2$ between sites on a lattice with $L = 10$ sites over an observation time of $T = 10$. Results are shown for TPS simulations using fixed-end Brownian bridge moves and shooting moves ("fixed BB+S"), an equal amount of fixed-end bridge moves and variable-end bridge moves that permute particles by $\Delta i = \pm 1$ in addition to shooting moves ("variable BB+fixed BB+S"), and only shooting moves ("S"). For reference, exact results from matrix diagonalization are shown.

deviates more substantially from the exact result at both large negative and large positive values of λ. Thus, due to their ability to achieve reasonable acceptance probabilities at high bias and to quickly decorrelate trajectories, Brownian bridge moves enable large deviation functions of periodic ASEPs to be accurately calculated over a wide range of bias values from simulations.

Conclusions

Much like equilibrium free energies, large deviation functions characterize the likelihood of nonequilibrium steady states and fluctuations thereabout. However, calculating large

deviation functions from simulations is often quite challenging. To overcome some of the limitations of current simulation methods, we developed new path sampling moves that extended the capabilities of transition path sampling simulations to efficiently calculate large deviation functions of periodic ASEPs. Specifically, we developed two new moves based on Brownian bridges that allow segments in the middle of a trajectory to be modified in a controlled manner. Unlike shooting moves, Brownian bridge moves include adjustable parameters that enable them to achieve good acceptance probabilities even at high bias. When optimally tuned, Brownian bridge moves enhance the sampling of periodic ASEP trajectories that exhibit rare values of activity and current. Thus, large deviation functions are more accurately calculated from transition path sampling simulations that use Brownian bridge moves compared to those that only use shooting moves.

Our initial tests demonstrated that Brownian bridge moves efficiently sample rare values of the activity and current for a periodic ASEP with $N = 2$ particles over an observation time of $T = 10$. Because a long time limit is not quite reached by $T = 10$, a more accurate estimate of the large deviation function may require sampling longer trajectories. We expect Brownian bridge moves to be especially valuable to sample longer trajectories since they modify the middle of trajectories, which comprise the bulk of longer trajectories, but shooting moves only modify the ends at high bias. Because the dynamical phase transition between a hyperuniform state and phase separated state becomes more pronounced at high packing fractions, a thorough investigation of this phase transition may require sampling periodic ASEPs with a larger number of particles. Variable-end bridges may be accepted with higher probability as N increases since alterations to the final configuration of the bridge are less substantial than at small N. However, both fixed-end and variable-end bridges may be less efficient at sampling systems with increased N since particles are treated as non-interacting during a Brownian bridge. During a bridge, they are more likely to occupy the same lattice site when N is large and, thus, more likely to be rejected since they do not respect the constraints of volume exclusion. A potential way to overcome this limitation is to develop a simulation algorithm that generates bridges of interacting particles. This approach would require a solution to the backwards Kolmogorov equation, which generally does not exist for interacting systems. Recently, however, solutions were obtained for continuous-space Brownian bridges during which particles never cross paths.[283] Thus, it may be possible to develop a simulation algorithm that generates discrete-space Brownian bridges of excluded-volume particles and further increases the efficiency of Brownian bridge moves.

While we have focused on developing Brownian bridge moves to sample trajectories of periodic ASEPs, Brownian bridge moves may enhance the sampling of other systems. For example, a continuous-space formulation[284] can be used to study analogues of ASEPs whose configuration spaces are continuous in addition to more molecular systems. Furthermore, Brownian bridge moves may also be useful for sampling equilibrium transitions and especially rare events characterized by long transitions, many metastable intermediates, or rugged transition landscapes. Shooting moves are ineffective at sampling such rare events since they generate trajectories that unproductively wander among metastable states and that cannot surmount barriers separating different transition paths.[135] In contrast, Brownian

bridge moves modify the middle of a trajectory in a controlled way such that new trajectory segments do not get stuck in metastable states. Furthermore, the acceptance probability of biased bridges can be tuned so that enhanced sampling techniques can be used to help overcome barriers in the transition landscape. Thus, Brownian bridge moves may greatly expand the capabilities of transition path sampling simulations to study the dynamics of both equilibrium and nonequilibrium systems.

References

1. Harayama, T.; Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 281–296.
2. van Meer, G.; Voelker, D. R.; Feigenson, G. W. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124.
3. Hishikawa, D.; Hashidate, T.; Shimizu, T.; Shindou, H. Diversity and function of membrane glycerophospholipids generated by the remodeling pathway in mammalian cells. *J. Lipid Res.* **2014**, *55*, 799–807.
4. Gault, C. R.; Obeid, L. M.; Hannun, Y. A. An overview of sphingolipid metabolism: From synthesis to breakdown. *Adv. Exp. Med. Biol.* **2010**, *688*, 1–23.
5. Griffiths, W. J.; Abdel-Khalik, J.; Yutuc, E.; Morgan, A. H.; Gilmore, I.; Hearn, T.; Wang, Y. Cholesterolomics: An update. *Anal. Biochem.* **2017**, *524*, 56–67.
6. Miranda, W. E.; Guo, J.; Mesa-Galloso, H.; Corradi, V.; Lees-Miller, J. P.; Tieleman, D. P.; Duff, H. J.; Noskov, S. Y. Lipid regulation of hERG1 channel function. *Nat. Commun.* **2021**, *12*, 1409.
7. Weinberg, Z. Y.; Puthenveedu, M. A. Regulation of G protein-coupled receptor signaling by plasma membrane organization and endocytosis. *Traffic* **2019**, *20*, 121–129.
8. Jones, A. J. Y.; Gabriel, F.; Tandale, A.; Nietlispach, D. Structure and dynamics of gpcrs in lipid membranes: Physical principles and experimental approaches. *Molecules* **2020**, *25*, 4729.
9. Yogurtcu, O. N.; Johnson, M. E. Cytosolic proteins can exploit membrane localization to trigger functional assembly. *PLoS Comput. Biol.* **2018**, *14*, e1006031.
10. Thiam, A. R.; Farese Jr, R. V.; Walther, T. C. The biophysics and cell biology of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 775–786.
11. Wymann, M. P.; Schneiter, R. Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 162–176.
12. Hammond, G. R. V.; Burke, J. E. Novel roles of phosphoinositides in signaling, lipid transport, and disease. *Curr. Opin. Cell Biol.* **2020**, *63*, 57–67.
13. Raghu, P.; Joseph, A.; Krishnan, H.; Singh, P.; Saha, S. Phosphoinositides: Regulators of nervous system function in health and disease. *Front. Mol. Neurosci.* **2019**, *12*, 208.
14. Bunney, T. D.; Katan, M. Phosphoinositide signalling in cancer: Beyond PI3K and

PTEN. *Nat. Rev. Cancer* **2010**, *10*, 342–352.

15. Hannun, Y. A.; Obeid, L. M. Sphingolipids and their metabolism in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 175–191.
16. Maceyka, M.; Spiegel, S. Sphingolipid metabolites in inflammatory disease. *Nature* **2014**, *510*, 58–67.
17. Ogretmen, B. Sphingolipid metabolism in cancer signalling and therapy. *Nat. Rev. Cancer* **2018**, *18*, 33–50.
18. Sunshine, H.; Iruela-Arispe, M. L. Membrane lipids and cell signaling. *Curr. Opin. Lipidol.* **2017**, *28*.
19. Dennis, E. A.; Norris, P. C. Eicosanoid storm in infection and inflammation. *Nat Rev Immunol* **2015**, *15*, 511–523.
20. Manning, B. D.; Toker, A. AKT/PKB signaling: Navigating the network. *Cell* **2017**, *169*, 381–405.
21. Lamour, N. F.; Stahelin, R. V.; Wijesinghe, D. S.; Maceyka, M.; Wang, E.; Allegood, J. C.; Merrill, A. H.; Cho, W.; Chalfant, C. E. Ceramide kinase uses ceramide provided by ceramide transport protein: Localization to organelles of eicosanoid synthesis. *J. Lipid Res.* **2007**, *48*, 1293–1304.
22. Lamour, N. F.; Subramanian, P.; Wijesinghe, D. S.; Stahelin, R. V.; Bonventre, J. V.; Chalfant, C. E. Ceramide-1-phosphate is required for the translocation of group IVA cytosolic phospholipase a_2 and prostaglandin synthesis. *J. Biol. Chem.* **2009**, *284*, 26897–26907.
23. Gable, K.; Gupta, S. D.; Han, G.; Niranjanakumari, S.; Harmon, J. M.; Dunn, T. M. A disease-causing mutation in the active site of serine palmitoyltransferase causes catalytic promiscuity. *J. Biol. Chem.* **2010**, *285*, 22846–22852.
24. Duan, J.; Merrill, A. H. 1-deoxysphingolipids encountered exogenously and made de novo: Dangerous mysteries inside an enigma. *J. Biol. Chem.* **2015**, *290*, 15380–15389.
25. Guri, Y.; Colombi, M.; Dazert, E.; Hindupur, S. K.; Roszik, J.; Moes, S.; Jenoe, P.; Heim, M. H.; Riezman, I.; Riezman, H.; Hall, M. N. mTORC2 promotes tumorigenesis via lipid synthesis. *Cancer Cell* **2017**, *32*, 807–823.e12.
26. Perrotti, F.; Rosa, C.; Cicalini, I.; Sacchetta, P.; Del Boccio, P.; Genovesi, D.; Pieragostino, D. Advances in lipidomics for cancer biomarkers discovery. *Int. J. Mol. Sci.* **2016**, *17*, 1992.
27. Wigger, L. et al. Plasma dihydroceramides are diabetes susceptibility biomarker candidates in mice and humans. *Cell Rep.* **2017**, *18*, 2269–2279.
28. Ollero, M.; Astarita, G.; Guerrera, I. C.; Sermet-Gaudelus, I.; Trudel, S.; Piomelli, D.; Edelman, A. Plasma lipidomics reveals potential prognostic signatures within a cohort of cystic fibrosis patients. *J. Lipid Res.* **2011**, *52*, 1011–1022.
29. Signorelli, P.; Pivari, F.; Barcella, M.; Merelli, I.; Zulueta, A.; Dei Cas, M.; Rosso, L.;

Ghidoni, R.; Caretti, A.; Paroni, R.; Mingione, A. Myriocin modulates the altered lipid metabolism and storage in cystic fibrosis. *Cell. Signal.* **2021**, *81*, 109928.

30. Schaefer, E. J.; Bongard, V.; Beiser, A. S.; Lamon-Fava, S.; Robins, S. J.; Au, R.; Tucker, K. L.; Kyle, D. J.; Wilson, P. W. F.; Wolf, P. A. Plasma phosphatidylcholine docosahexaenoic acid content and risk of dementia and alzheimer disease: The framingham heart study. *Arch. Neurol.* **2006**, *63*, 1545–1550.
31. Zhang, X.; Liu, W.; Zan, J.; Wu, C.; Tan, W. Untargeted lipidomics reveals progression of early alzheimer's disease in APP/PS1 transgenic mice. *Sci. Rep.* **2020**, *10*, 14509.
32. Shamim, A.; Mahmood, T.; Ahsan, F.; Kumar, A.; Bagga, P. Lipids: An insight into the neurodegenerative disorders. *Clin. Nutr. Exp.* **2018**, *20*, 1–19.
33. Hannich, J. T.; Umebayashi, K.; Riezman, H. Distribution and functions of sterols and sphingolipids. *CSH Perspect. Biol.* **2011**, *3*.
34. Yamashita, A.; Hayashi, Y.; Nemoto-Sasaki, Y.; Ito, M.; Oka, S.; Tanikawa, T.; Waku, K.; Sugiura, T. Acyltransferases and transacylases that determine the fatty acid composition of glycerolipids and the metabolism of bioactive lipid mediators in mammalian cells and model organisms. *Prog. Lipid Res.* **2014**, *53*, 18–81.
35. Grosch, S.; Schiffmann, S.; Geisslinger, G. Chain length-specific properties of ceramides. *Prog. Lipid Res.* **2012**, *51*, 50–62.
36. Harayama, T.; Eto, M.; Shindou, H.; Kita, Y.; Otsubo, E.; Hishikawa, D.; Ishii, S.; Sakimura, K.; Mishina, M.; Shimizu, T. Lysophospholipid acyltransferases mediate phosphatidylcholine diversification to achieve the physical properties required in vivo. *Cell Metab.* **2014**, *20*, 295–305.
37. Frechin, M.; Stoeger, T.; Daetwyler, S.; Gehin, C.; Battich, N.; Damm, E. M.; Stergiou, L.; Riezman, H.; Pelkmans, L. Cell-intrinsic adaptation of lipid composition to local crowding drives social behaviour. *Nature* **2015**, *523*, 88–91.
38. Antonny, B.; Vanni, S.; Shindou, H.; Ferreira, T. From zero to six double bonds: Phospholipid unsaturation and organelle function. *Trends Cell Biol.* **2015**, *25*, 427–436.
39. Lorent, J. H.; Levental, K. R.; Ganesan, L.; Rivera-Longsworth, G.; Sezgin, E.; Doktorova, M.; Lyman, E.; Levental, I. Plasma membranes are asymmetric in lipid unsaturation, packing and protein shape. *Nat. Chem. Biol.* **2020**, *16*, 644–652.
40. Sezgin, E.; Levental, I.; Mayor, S.; Eggeling, C. The mystery of membrane organization: Composition, regulation and roles of lipid rafts. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 361–374.
41. Russo, D.; Capolupo, L.; Loomba, J. S.; Sticco, L.; D'Angelo, G. Glycosphingolipid metabolism in cell fate specification. *J. Cell Sci.* **2018**, *131*.
42. Denz, M.; Chiantia, S.; Herrmann, A.; Mueller, P.; Korte, T.; Schwarzer, R. Cell cycle dependent changes in the plasma membrane organization of mammalian cells. *Biochm. Biophys. Acta Biomembr.* **2017**, *1859*, 350–359.
43. Lagny, T. J.; Bassereau, P. Bioinspired membrane-based systems for a physical approach

of cell organization and dynamics: Usefulness and limitations. *Interface Focus* **2015**, *5*.

44. Deleu, M.; Crowet, J. M.; Nasir, M. N.; Lins, L. Complementary biophysical tools to investigate lipid specificity in the interaction between bioactive molecules and the plasma membrane: A review. *Biochm. Biophys. Acta Biomembr.* **2014**, *1838*, 3171–3190.
45. Poyry, S.; Vattulainen, I. Role of charged lipids in membrane structures: Insight given by simulations. *Biochm. Biophys. Acta Biomembr.* **2016**, *1858*, 2322–2333.
46. Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingolfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational modeling of realistic cell membranes. *Chem. Rev.* **2019**, *119*, 6184–6226.
47. Bigay, J.; Antonny, B. Curvature, lipid packing, and electrostatics of membrane organelles: Defining cellular territories in determining specificity. *Dev. Cell* **2012**, *23*, 886–895.
48. Magalhaes, M. A. O.; Glogauer, M. Pivotal advance: Phospholipids determine net membrane surface charge resulting in differential localization of active rac1 and rac2. *J. Leukoc. Biol.* **2010**, *87*, 545–555.
49. Yeung, T.; Terebiznik, M.; Yu, L. M.; Silvius, J.; Abidi, W. M.; Philips, M.; Levine, T.; Kapus, A.; Grinstein, S. Receptor activation alters inner surface potential during phagocytosis. *Science* **2006**, *313*, 347–351.
50. Papayannopoulos, V.; Co, C.; Prehoda, K. E.; Snapper, S.; Taunton, J.; Lim, W. A. A polybasic motif allows N-WASP to act as a sensor of PIP2 density. *Mol. Cell* **2005**, *17*, 181–191.
51. Bretscher, M. S.; Munro, S. Cholesterol and the golgi apparatus. *Science* **1993**, *261*, 1280.
52. Sharpe, H. J.; Stevens, T. J.; Munro, S. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* **2010**, *142*, 158–169.
53. Kandasamy, S. K.; Larson, R. G. Molecular dynamics simulations of model transmembrane peptides in lipid bilayers: A systematic investigation of hydrophobic mismatch. *Biophys. J.* **2006**, *90*, 2326–2343.
54. Kim, T.; Im, W. Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation. *Biophys. J.* **2010**, *99*, 175–183.
55. Cosson, P.; Perrin, J.; Bonifacino, J. S. Anchors aweigh: Protein localization and transport mediated by transmembrane domains. *Trends Cell Biol.* **2013**, *23*, 511–517.
56. Welch, L. G.; Munro, S. A tale of short tails, through thick and thin: Investigating the sorting mechanisms of golgi enzymes. *FEBS Lett.* **2019**, *593*, 2452–2465.
57. Hanulová, M.; Weiss, M. Protein sorting and membrane-mediated interactions. *Biophys. Rev.* **2012**, *4*, 117–124.

58. Kaiser, H.-J.; Orłowski, A.; Róg, T.; Nyholm, T. K. M.; Chai, W.; Feizi, T.; Lingwood, D.; Vattulainen, I.; Simons, K. Lateral sorting in model membranes by cholesterol-mediated hydrophobic matching. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16628.
59. Milovanovic, D.; Honigmann, A.; Koike, S.; Göttfert, F.; Pähler, G.; Junius, M.; Müllar, S.; Diederichsen, U.; Janshoff, A.; Grubmüller, H.; Risselada, H. J.; Eggeling, C.; Hell, S. W.; van den Bogaart, G.; Jahn, R. Hydrophobic mismatch sorts snare proteins into distinct membrane domains. *Nat. Commun.* **2015**, *6*, 5984.
60. Vanni, S.; Hirose, H.; Barelli, H.; Antonny, B.; Gautier, R. A sub-nanometre view of how membrane curvature and composition modulate lipid packing and protein recruitment. *Nat. Commun.* **2014**, *5*, 4916.
61. Gautier, R.; Bacle, A.; Tiberti, M. L.; Fuchs, P. F.; Vanni, S.; Antonny, B. Packmem: A versatile tool to compute and visualize interfacial packing defects in lipid bilayers. *Biophys. J.* **2018**, *115*, 436–444.
62. Vamparys, L.; Gautier, R.; Vanni, S.; Bennett, W. F. D.; Tieleman, D. P.; Antonny, B.; Etchebest, C.; Fuchs, P. F. J. Conical lipids in flat bilayers induce packing defects similar to that induced by positive curvature. *Biophys. J.* **2013**, *104*, 585–593.
63. Deguil, J.; Pineau, L.; Snyder, E. C. R.; Dupont, S.; Beney, L.; Gil, A.; Frapper, G.; Ferreira, T. Modulation of lipid-induced ER stress by fatty acid shape. *Traffic* **2011**, *12*, 349–362.
64. Monje-Galvan, V.; Klauda, J. B. Peripheral membrane proteins: Tying the knot between experiment and computation. *Biochm. Biophys. Acta Biomembr.* **2016**, *1858*, 1584–1593.
65. Wildermuth, K. D.; Monje-Galvan, V.; Warburton, L. M.; Klauda, J. B. Effect of membrane lipid packing on stable binding of the ALPS peptide. *J. Chem. Theory Comput.* **2019**, *15*, 1418–1429.
66. Mabrey, S.; Sturtevant, J. M. Investigation of phase-transitions of lipids and lipid mixtures by high sensitivity differential scanning calorimetry. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 3862–3866.
67. Melchior, D. L.; Steim, J. M. Thermotropic transitions in biomembranes. *Annu. Rev. Biophys. Bioeng.* **1976**, *5*, 205–238.
68. Hjort Ipsen, J.; Karlström, G.; Mourtisen, O. G.; Wennerström, H.; Zuckermann, M. J. Phase equilibria in the phosphatidylcholine-cholesterol system. *Biochm. Biophys. Acta Biomembr.* **1987**, *905*, 162–172.
69. Rosetti, C. M.; Mangiarotti, A.; Wilke, N. Sizes of lipid domains: What do we know from artificial lipid membranes? What are the possible shared features with membrane rafts in cells? *Biochm. Biophys. Acta Biomembr.* **2017**, *1859*, 789–802.
70. Kaiser, H.-J.; Lingwood, D.; Levental, I.; Sampaio, J. L.; Kalvodova, L.; Rajendran, L.; Simons, K. Order of lipid phases in model and plasma membranes. *Proc. Natl. Acad.*

Sci. U. S. A. **2009**, *106*, 16645.

71. Stone, M. B.; Shelby, S. A.; Nunez, M. F.; Wisser, K.; Veatch, S. L. Protein sorting by lipid phase-like domains supports emergent signaling function in B lymphocyte plasma membranes. *Elife* **2017**, *6*.
72. Zick, M.; Stroupe, C.; Orr, A.; Douville, D.; Wickner, W. T. Membranes linked by trans-snare complexes require lipids prone to non-bilayer structure for progression to fusion. *Elife* **2014**, *3*.
73. Irie, A.; Yamamoto, K.; Miki, Y.; Murakami, M. Phosphatidylethanolamine dynamics are required for osteoclast fusion. *Sci. Rep.* **2017**, *7*.
74. Shibata, Y.; Hu, J. J.; Kozlov, M. M.; Rapoport, T. A. Mechanisms shaping the membranes of cellular organelles. *Annu. Rev. Cell. Dev. Biol.* **2009**, *25*, 329–354.
75. de Mendoza, D.; Pilon, M. Control of membrane lipid homeostasis by lipid-bilayer associated sensors: A mechanism conserved from bacteria to humans. *Prog. Lipid Res.* **2019**, *76*, 100996.
76. Ernst, R.; Ballweg, S.; Levental, I. Cellular mechanisms of physicochemical membrane homeostasis. *Curr. Opin. Cell Biol.* **2018**, *53*, 44–51.
77. Holthuis, J. C. M.; Menon, A. K. Lipid landscapes and pipelines in membrane homeostasis. *Nature* **2014**, *510*, 48–57.
78. Ernst, R.; Ejsing, C. S.; Antonny, B. Homeoviscous adaptation and the regulation of membrane lipids. *J. Mol. Biol.* **2016**, *428*, 4776–4791.
79. Hazel, J. R. Thermal adaptation in biological-membranes: Is homeoviscous adaptation the explanation? *Annu. Rev. Physiol.* **1995**, *57*, 19–42.
80. Levental, K. R.; Malmberg, E.; Symons, J. L.; Fan, Y.-Y.; Chapkin, R. S.; Ernst, R.; Levental, I. Lipidomic and biophysical homeostasis of mammalian membranes counteracts dietary lipid perturbations to maintain cellular fitness. *Nat. Commun.* **2020**, *11*, 1339.
81. Loewen, C. J. R.; Gaspar, M. L.; Jesch, S. A.; Delon, C.; Ktistakis, N. T.; Henry, S. A.; Levine, T. P. Phospholipid metabolism regulated by a transcription factor sensing phosphatidic acid. *Science* **2004**, *304*, 1644–1647.
82. Halbleib, K.; Pesek, K.; Covino, R.; Hofbauer, H. F.; Wunnicke, D.; Hanelt, I.; Hummer, G.; Ernst, R. Activation of the unfolded protein response by lipid bilayer stress. *Mol. Cell* **2017**, *67*, 673–684.
83. Mesmin, B.; Bigay, J.; Polidori, J.; Jamecna, D.; Lacas-Gervais, S.; Antonny, B. Sterol transfer, PI4P consumption, and control of membrane lipid order by endogenous OSBP. *EMBO J.* **2017**, *36*, 3156–3174.
84. Drin, G. In *Annual review of biochemistry, vol 83*; Kornberg, R. D., Ed.; Annual review of biochemistry; 2014; Vol. 83; pp 51–77.
85. Lev, S. Non-vesicular lipid transport by lipid-transfer proteins and beyond. *Nat. Rev.*

Mol. Cell Biol. **2010**, *11*, 739–750.

86. van Meer, G. Lipid traffic in animal cells. *Annu. Rev. Cell Biol.* **1989**, *5*, 247–275.
87. Wong, L. H.; Gatta, A. T.; Levine, T. P. Lipid transfer proteins: The lipid commute via shuttles, bridges and tubes. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 85–101.
88. Jackson, C. L.; Walch, L.; Verbavatz, J. M. Lipids and their trafficking: An integral part of cellular organization. *Dev. Cell* **2016**, *39*, 139–153.
89. Prinz, W. A.; Toulmay, A.; Balla, T. The functional universe of membrane contact sites. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 7–24.
90. Cockcroft, S.; Raghu, P. Phospholipid transport protein function at organelle contact sites. *Curr. Opin. Cell Biol.* **2018**, *53*, 52–60.
91. Wong, L. H.; Levine, T. P. Lipid transfer proteins do their thing anchored at membrane contact sites . . . But what is their thing? *Biochem. Soc. Trans.* **2016**, *44*, 517–527.
92. Aniansson, E. A. G.; Wall, S. N.; Almgren, M.; Hoffmann, H.; Kielmann, I.; Ulbricht, W.; Zana, R.; Lang, J.; Tondre, C. Theory of kinetics of micellar equilibria and quantitative interpretation of chemical relaxation studies of micellar solutions of ionic surfactants. *J. Phys. Chem.* **1976**, *80*, 905–922.
93. McLean, L. R.; Phillips, M. C. Mechanism of cholesterol and phosphatidylcholine exchange or transfer between unilamellar vesicles. *Biochemistry* **1981**, *20*, 2893–2900.
94. Nichols, J. W.; Pagano, R. E. Kinetics of soluble lipid monomer diffusion between vesicles. *Biochemistry* **1981**, *20*, 2783–2789.
95. Nichols, J. W.; Pagano, R. E. Use of resonance energy-transfer to study the kinetics of amphiphile transfer between vesicles. *Biochemistry* **1982**, *21*, 1720–1726.
96. Massey, J. B.; Gotto, A. M.; Pownall, H. J. Kinetics and mechanism of the spontaneous transfer of fluorescent phosphatidylcholines between apolipoprotein-phospholipid recombinants. *Biochemistry* **1982**, *21*, 3630–3636.
97. Pownall, H. J.; Hickson, D. L.; Smith, L. C. Transport of biological lipophiles: Effect of lipophile structure. *J. Am. Chem. Soc.* **1983**, *105*, 2440–2445.
98. McLean, L. R.; Phillips, M. C. Kinetics of phosphatidylcholine and lysophosphatidylcholine exchange between unilamellar vesicles. *Biochemistry* **1984**, *23*, 4624–4630.
99. Nichols, J. W. Thermodynamics and kinetics of phospholipid monomer vesicle interaction. *Biochemistry* **1985**, *24*, 6390–6398.
100. Ferrell, J. E.; Lee, K. J.; Huestis, W. H. Lipid transfer between phosphatidylcholine vesicles and human erythrocytes: Exponential decrease in rate with increasing acyl chain length. *Biochemistry* **1985**, *24*, 2857–2864.
101. Bayerl, T. M.; Schmidt, C. F.; Sackmann, E. Kinetics of symmetric and asymmetric phospholipid transfer between small sonicated vesicles studied by high-sensivity differential scanning calorimetry, NMR, electron-microscopy, and dynamic light-scattering. *Biochemistry* **1988**, *27*, 6078–6085.

102. Elvington, S. M.; Nichols, J. W. Spontaneous, intervesicular transfer rates of fluorescent, acyl chain-labeled phosphatidylcholine analogs. *Biochm. Biophys. Acta Biomembr.* **2007**, *1768*, 502–508.
103. Richens, J. L.; Tyler, A. I. I.; Barriga, H. M. G.; Bramble, J. P.; Law, R. V.; Brooks, N. J.; Seddon, J. M.; Ces, O.; O'Shea, P. Spontaneous charged lipid transfer between lipid vesicles. *Sci. Rep.* **2017**, *7*, 12606.
104. Wong, L. H.; Copic, A.; Levine, T. P. Advances on the transfer of lipids by lipid transfer proteins. *Trends Biochem. Sci.* **2017**, *42*, 516–530.
105. Chiapparino, A.; Maeda, K.; Turei, D.; Saez-Rodriguez, J.; Gavin, A.-C. The orchestra of lipid-transfer proteins at the crossroads between metabolism and signaling. *Prog. Lipid Res.* **2016**, *61*, 30–39.
106. Kumagai, K.; Yasuda, S.; Okemoto, K.; Nishijima, M.; Kobayashi, S.; Hanada, K. CERT mediates intermembrane transfer of various molecular species of ceramides. *J. Biol. Chem.* **2005**, *280*, 6488–6495.
107. Helmkamp, G. M. Phospholipid transfer proteins: Mechanism of action. *J. Bioenerg. Biomembr.* **1986**, *18*, 71–91.
108. Gadella, T. W. J.; Wirtz, K. W. A. Phospholipid-binding and transfer by the nonspecific lipid-transfer protein (sterol carrier protein-2): A kinetic-model. *Eur. J. Biochem.* **1994**, *220*, 1019–1028.
109. Lipp, N.-F.; Gautier, R.; Magdeleine, M.; Renard, M.; Albanèse, V.; Čopič, A.; Drin, G. An electrostatic switching mechanism to control the lipid transfer activity of Osh6p. *Nat. Commun.* **2019**, *10*, 3926.
110. Rogaski, B.; Klauda, J. B. Membrane-binding mechanism of a peripheral membrane protein through microsecond molecular dynamics simulations. *J. Mol. Biol.* **2012**, *423*, 847–861.
111. Sugiki, T.; Takeuchi, K.; Yamaji, T.; Takano, T.; Tokunaga, Y.; Kumagai, K.; Hanada, K.; Takahashi, H.; Shimada, I. Structural basis for the golgi association by the pleckstrin homology domain of the ceramide trafficking protein (CERT). *J. Biol. Chem.* **2012**, *287*, 33706–33718.
112. Kawano, M.; Kumagai, K.; Nishijima, M.; Hanada, K. Efficient trafficking of ceramide from the endoplasmic reticulum to the golgi apparatus requires a VAMP-associated protein-interacting FFAT motif of CERT. *J. Biol. Chem.* **2006**, *281*, 30279–30288.
113. Im, Y. J.; Raychaudhuri, S.; Prinz, W. A.; Hurley, J. H. Structural mechanism for sterol sensing and transport by OSBP-related proteins. *Nature* **2005**, *437*, 154–158.
114. Watanabe, Y.; Tamura, Y.; Kawano, S.; Endo, T. Structural and mechanistic insights into phospholipid transfer by Ups1–Mdm35 in mitochondria. *Nat. Commun.* **2015**, *6*, 7922.
115. Xiao, T.; Lu, Y.; Fan, J.-S.; Yang, D. Ligand entry into fatty acid binding protein via local unfolding instead of gap widening. *Biophys. J.* **2020**, *118*, 396–402.

116. Miliara, X.; Tatsuta, T.; Berry, J. L.; Rouse, S. L.; Solak, K.; Chorev, D. S.; Wu, D.; Robinson, C. V.; Matthews, S.; Langer, T. Structural determinants of lipid specificity within Ups/PRELI lipid transfer proteins. *Nat. Commun.* **2019**, *10*.

117. Owens, T. W.; Taylor, R. J.; Pahil, K. S.; Bertani, B. R.; Ruiz, N.; Kruse, A. C.; Kahne, D. Structural basis of unidirectional export of lipopolysaccharide to the cell surface. *Nature* **2019**, *567*, 550–553.

118. Li, Y.; Orlando, B. J.; Liao, M. Structural basis of lipopolysaccharide extraction by the $LptB_2FGC$ complex. *Nature* **2019**, *567*, 486–490.

119. Enkavi, G.; Javanainen, M.; Kulig, W.; Rog, T.; Vattulainen, I. Multiscale simulations of biological membranes: The challenge to understand biological phenomena in a living substance. *Chem. Rev.* **2019**, *119*, 5607–5774.

120. Ingolfsson, H. I.; Arnarez, C.; Periole, X.; Marrink, S. J. Computational 'microscopy' of cellular membranes. *J. Cell Sci.* **2016**, *129*, 257–268.

121. Knecht, V.; Marrink, S. J. Molecular dynamics simulations of lipid vesicle fusion in atomic detail. *Biophys. J.* **2007**, *92*, 4254–4261.

122. Kasson, P. M.; Lindahl, E.; Pande, V. S. Atomic-resolution simulations predict a transition state for vesicle fusion defined by contact of a few lipid tails. *PLoS Comput. Biol.* **2010**, *6*, e1000829.

123. Larsson, P.; Kasson, P. M. Lipid tail protrusion in simulations predicts fusogenic activity of influenza fusion peptide mutants and conformational models. *PLoS Comput. Biol.* **2013**, *9*, e1002950.

124. Mondal, J. A brief appraisal of computational modeling of antimicrobial peptides' activity. *Drug Dev. Res.* **2019**, *80*, 28–32.

125. Reid, L. M.; Verma, C. S.; Essex, J. W. The role of molecular simulations in understanding the mechanisms of cell-penetrating peptides. *Drug Discov. Today* **2019**, *24*, 1821–1835.

126. Muller, M. P.; Jiang, T.; Sun, C.; Lihan, M. Y.; Pant, S.; Mahinthichaichan, P.; Trifan, A.; Tajkhorshid, E. Characterization of lipid-protein interactions and lipid-mediated modulation of membrane protein function through molecular simulation. *Chem. Rev.* **2019**, *119*, 6086–6161.

127. Chavent, M.; Duncan, A. L.; Sansom, M. S. P. Molecular dynamics simulations of membrane proteins and their interactions: From nanoscale to mesoscale. *Curr. Opin. Struct. Biol.* **2016**, *40*, 8–16.

128. Marrink, S. J.; Tieleman, D. P. Perspective on the martini model. *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.

129. Cooke, I. R.; Deserno, M. Solvent-free model for self-assembling fluid bilayer membranes: Stabilization of the fluid phase based on broad attractive tail potentials. *J. Chem. Phys.* **2005**, *123*, 224710.

130. Li, H. L.; Gorfe, A. A. Membrane remodeling by surface-bound protein aggregates:

Insights from coarse-grained molecular dynamics simulation. *J. Phys. Chem. Lett.* **2014**, *5*, 1457–1462.

131. Pezeshkian, W.; Konig, M.; Wassenaar, T. A.; Marrink, S. J. Backmapping triangulated surfaces to coarse-grained membrane models. *Nat. Commun.* **2020**, *11*.
132. Pezeshkian, W.; Marrink, S. J. Simulating realistic membrane shapes. *Curr. Opin. Cell Biol.* **2021**, *71*, 103–111.
133. Argudo, D.; Bethel, N. P.; Marcoline, F. V.; Grabe, M. Continuum descriptions of membranes and their interaction with proteins: Towards chemically accurate models. *Biochm. Biophys. Acta Biomembr.* **2016**, *1858*, 1619–1634.
134. Pabis, A.; Rawle, R. J.; Kasson, P. M. Influenza hemagglutinin drives viral entry via two sequential intramembrane mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 7200.
135. Bolhuis, P. G.; Swenson, D. W. H. Transition path sampling as markov chain monte carlo of trajectories: Recent algorithms, software, applications, and future outlook. *Adv. Theory Simul.* **2021**, 2000237.
136. Dellago, C.; Bolhuis, P. G.; Geissler, P. L. In *Advances in chemical physics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons: New York, 2002; Vol. 123; pp 1–78.
137. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.
138. Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. On the transition coordinate for protein folding. *J. Chem. Phys.* **1998**, *108*, 334–350.
139. Zuckerman, D. M. *Statistical physics of biomolecules: An introduction*; CRC Press: Boca Raton, FL, 2010.
140. Peters, B.; Trout, B. L. Obtaining reaction coordinates by likelihood maximization. *J. Chem. Phys.* **2006**, *125*, 054108.
141. Best, R. B.; Hummer, G. Reaction coordinates and rates from transition paths. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6732–6737.
142. Klosek, M. M.; Matkowsky, B. J.; Schuss, Z. The kramers problem in the turnover regime: The role of the stochastic separatrix. *Ber. Bunsenges. Phys. Chem.* **1991**, *95*, 331–337.
143. Dzierlenga, M. W.; Varga, M. J.; Schwartz, S. D. In *Methods in enzymology*; Voth, G. A., Ed.; Academic Press, 2016; Vol. 578; pp 21–43.
144. Basner, J. E.; Schwartz, S. D. How enzyme dynamics helps catalyze a reaction in atomic detail: A transition path sampling study. *J. Am. Chem. Soc.* **2005**, *127*, 13822–13831.
145. Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. The mechanism of cellulose hydrolysis by a two-step, retaining cellobiohydrolase elucidated by structural and

transition path sampling studies. *J. Am. Chem. Soc.* **2014**, *136*, 321–329.

146. Bolhuis, P. G. Transition-path sampling of β-hairpin folding. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12129.
147. Juraszek, J.; Bolhuis, P. G. Sampling the multiple folding mechanisms of trp-cage in explicit solvent. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15859.
148. Vreede, J.; Juraszek, J.; Bolhuis, P. G. Predicting the reaction coordinates of millisecond light-induced conformational changes in photoactive yellow protein. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2397.
149. Hagan, M. F.; Dinner, A. R.; Chandler, D.; Chakraborty, A. K. Atomistic understanding of kinetic pathways for single base-pair binding and unbinding in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13922.
150. Radhakrishnan, R.; Schlick, T. Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β's closing. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5970.
151. Riccardi, E.; van Mastbergen, E. C.; Navarre, W. W.; Vreede, J. Predicting the mechanism and rate of h-ns binding to at-rich DNA. *PLoS Comput. Biol.* **2019**, *15*, e1006845.
152. Okazaki, K.-i.; Wohlert, D.; Warnau, J.; Jung, H.; Yildiz, O.; Kuhlbrandt, W.; Hummer, G. Mechanism of the electroneutral sodium/proton antiporter panhap from transition-path shooting. *Nat. Commun.* **2019**, *10*, 1742.
153. Mori, T.; Miyashita, N.; Im, W.; Feig, M.; Sugita, Y. Molecular dynamics simulations of biological membranes and membrane proteins using enhanced conformational sampling algorithms. *Biochm. Biophys. Acta Biomembr.* **2016**, *1858*, 1635–1651.
154. Harpole, T. J.; Delemotte, L. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. *Biochm. Biophys. Acta Biomembr.* **2018**, *1860*, 909–926.
155. Girard, M.; Bereau, T. Regulating lipid composition rationalizes acyl tail saturation homeostasis in ectotherms. *Biophys. J.* **2020**, *119*, 892–899.
156. Tieleman, D. P.; Marrink, S. J. Lipids out of equilibrium: Energetics of desorption and pore mediated flip-flop. *J. Am. Chem. Soc.* **2006**, *128*, 12462–12467.
157. Sapay, N.; Bennett, W. F. D.; Tieleman, D. P. Thermodynamics of flip-flop and desorption for a systematic series of phosphatidylcholine lipids. *Soft Matter* **2009**, *5*, 3295–3302.
158. Grafmuller, A.; Lipowsky, R.; Knecht, V. Effect of tension and curvature on the chemical potential of lipids in lipid aggregates. *Phys. Chem. Chem. Phys.* **2013**, *15*, 876–881.
159. Sapay, N.; Bennett, W. F. D.; Tieleman, D. P. Molecular simulations of lipid flip-flop in the presence of model transmembrane helices. *Biochemistry* **2010**, *49*, 7665–7673.
160. Bennett, W. F. D.; MacCallum, J. L.; Tieleman, D. P. Thermodynamic analysis of the effect of cholesterol on dipalmitoylphosphatidylcholine lipid membranes. *J. Am. Chem.*

Soc. **2009**, *131*, 1972–1978.

161. Xia, Y.; Li, M.; Charubin, K.; Liu, Y.; Heberle, F. A.; Katsaras, J.; Jing, B. X.; Zhu, Y. X.; Nieh, M. P. Effects of nanoparticle morphology and acyl chain length on spontaneous lipid transfer rates. *Langmuir* **2015**, *31*, 12920–12928.
162. De Cuyper, M.; Joniau, M. Spontaneous intervesicular transfer of anionic phospholipids differing in the nature of their polar headgroup. *Biochim. Biophys. Acta* **1985**, *814*, 374–380.
163. Homan, R.; Pownall, H. J. Transbilayer diffusion of phospholipids: Dependence on headgroup structure and acyl chain-length. *Biochim. Biophys. Acta* **1988**, *938*, 155–166.
164. Jencks, W. P. *Catalysis in chemistry and enzymology*; McGraw-Hill: New York, 1969.
165. Vermaas, J. V.; Tajkhorshid, E. A microscopic view of phospholipid insertion into biological membranes. *J. Phys. Chem. B* **2014**, *118*, 1754–1764.
166. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.
167. Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
168. Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144–2155.
169. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
170. Parrinello, M.; Rahman, A. Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
171. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
172. Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. The “hot-solvent/cold-solute" problem revisited. *J. Chem. Theory Comput.* **2008**, *4*, 1293–306.
173. Hockney, W. R. The potential calculation and some applications. *Methods Comput. Phys.* **1970**, *9*, 136–211.
174. de Jong, D. H.; Baoukina, S.; Ingolfsson, H. I.; Marrink, S. J. MARTINI straight: Boosting performance using a shorter cutoff and gpus. *Comput. Phys. Commun.* **2016**, *199*, 1–7.
175. Páll, S.; Hess, B. A flexible algorithm for calculating pair interactions on simd architectures. *Comput. Phys. Commun.* **2013**, *184*, 2641–2650.

176. Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.

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

178. MacKerell, A. D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

179. Pluhackova, K.; Kirsch, S. A.; Han, J.; Sun, L. P.; Jiang, Z. Y.; Unruh, T.; Bockmann, R. A. A critical comparison of biomembrane force fields: Structure and dynamics of model DMPC, POPC, and POPE bilayers. *J. Phys. Chem. B* **2016**, *120*, 3888–3903.

180. Piggot, T. J.; Pineiro, A.; Khalid, S. Molecular dynamics simulations of phosphatidylcholine membranes: A comparative force field study. *J. Chem. Theory Comput.* **2012**, *8*, 4593–4609.

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

182. Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **2014**, *35*, 1997–2004.

183. Nosé, S. A molecular-dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

184. Hoover, W. G. Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.

185. Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. Lincs: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

186. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

187. Lin, D.; Grossfield, A. Thermodynamics of micelle formation and membrane fusion modulate antimicrobial lipopeptide activity. *Biophys. J.* **2015**, *109*, 750–759.

188. Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.

189. Torrie, G. M.; Valleau, J. P. Non-physical sampling distributions in monte-carlo free-energy estimation - umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

190. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

191. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The

weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

192. Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284–304.
193. Smirnova, Y. G.; Marrink, S. J.; Lipowsky, R.; Knecht, V. Solvent-exposed tails as prestalk transition states for membrane fusion at low hydration. *J. Am. Chem. Soc.* **2010**, *132*, 6710–6718.
194. Mirjanian, D.; Dickey, A. N.; Hoh, J. H.; Woolf, T. B.; Stevens, M. J. Splaying of aliphatic tails plays a central role in barrier crossing during liposome fusion. *J. Phys. Chem. B* **2010**, *114*, 11061–11068.
195. Stevens, M. J.; Hoh, J. H.; Woolf, T. B. Insights into the molecular mechanism of membrane fusion from simulation: Evidence for the association of splayed tails. *Phys. Rev. Lett.* **2003**, *91*, 188102.
196. Ohta-Iino, S.; Pasenkiewicz-Gierula, M.; Takaoka, Y.; Miyagawa, H.; Kitamura, K.; Kusumi, A. Fast lipid disorientation at the onset of membrane fusion revealed by molecular dynamics simulations. *Biophys. J.* **2001**, *81*, 217–224.
197. Holopainen, J. M.; Lehtonen, J. Y.; Kinnunen, P. K. Evidence for the extended phospholipid conformation in membrane fusion and hemifusion. *Biophys. J.* **1999**, *76*, 2111–2120.
198. Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **2005**, *437*, 640–647.
199. Peters, B. Reaction coordinates and mechanistic hypothesis tests. *Annu. Rev. Phys. Chem.* **2016**, *67*, 669–690.
200. Peters, B.; Beckham, G. T.; Trout, B. L. Extensions to the likelihood maximization approach for finding reaction coordinates. *J. Chem. Phys.* **2007**, *127*, 034109.
201. Nakano, M.; Fukuda, M.; Kudo, T.; Endo, H.; Handa, T. Determination of interbilayer and transbilayer lipid transfers by time-resolved small-angle neutron scattering. *Phys. Rev. Lett.* **2007**, *98*, 238101.
202. Jones, J. D.; Thompson, T. E. Mechanism of spontaneous, concentration-dependent phospholipid transfer between bilayers. *Biochemistry* **1990**, *29*, 1593–1600.
203. Drazenovic, J.; Ahmed, S.; Tuzinkiewicz, N. M.; Wunder, S. L. Lipid exchange and transfer on nanoparticle supported lipid bilayers: Effect of defects, ionic strength, and size. *Langmuir* **2015**, *31*, 721–731.
204. Kramer, A.; Pickard, F. C.; Huang, J.; Venable, R. M.; Simmonett, A. C.; Reith, D.; Kirschner, K. N.; Pastor, R. W.; Brooks, B. R. Interactions of water and alkanes: Modifying additive force fields to account for polarization effects. *J. Chem. Theory Comput.* **2019**, *15*, 3854–3867.
205. Paloncyova, M.; Fabre, G.; DeVane, R. H.; Trouillas, P.; Berka, K.; Otyepka, M. Benchmarking of force fields for molecule-membrane interactions. *J. Chem. Theory Comput.*

2014, *10*, 4143–4151.

206. Jämbeck, J. P. M.; Lyubartsev, A. P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B* **2012**, *116*, 3164–3179.
207. Jämbeck, J. P. M.; Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **2012**, *8*, 2938–2948.
208. Gurtovenko, A. A.; Vattulainen, I. Effect of NaCl and KCl on phosphatidylcholine and phosphatidylethanolamine lipid membranes: Insight from atomic-scale simulations for understanding salt-induced effects in the plasma membrane. *J. Phys. Chem. B* **2008**, *112*, 1953–1962.
209. Bockmann, R. A.; Hac, A.; Heimburg, T.; Grubmuller, H. Effect of sodium chloride on a lipid bilayer. *Biophys. J.* **2003**, *85*, 1647–1655.
210. Oncins, G.; Garcia-Manyes, S.; Sanz, F. Study of frictional properties of a phospholipid bilayer in a liquid environment with lateral force microscopy as a function of NaCl concentration. *Langmuir* **2005**, *21*, 7373–7379.
211. Cheng, Z. L.; Luisi, P. L. Coexistence and mutual competition of vesicles with different size distributions. *J. Phys. Chem. B* **2003**, *107*, 10940–10945.
212. Luisi, P. L. Are micelles and vesicles chemical equilibrium systems? *J. Chem. Educ.* **2001**, *78*, 380–384.
213. Cho, N. J.; Hwang, L. Y.; Solandt, J. J. R.; Frank, C. W. Comparison of extruded and sonicated vesicles for planar bilayer self-assembly. *Materials* **2013**, *6*, 3294–3308.
214. Israelachvili, J. N. *Intermolecular and surface forces*; Academic: New York, 1985.
215. François-Martin, C.; Rothman, J. E.; Pincet, F. Low energy cost for optimal speed and control of membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 1238–1241.
216. Nguyen, M. H. L.; DiPasquale, M.; Rickeard, B. W.; Stanley, C. B.; Kelley, E. G.; Marquardt, D. Methanol accelerates DMPC flip-flop and transfer: A SANS study on lipid dynamics. *Biophys. J.* **2019**, *116*, 755–759.
217. Huston, K. J.; Kiemen, A.; Larson, R. G. Search for the source of an apparent interfacial resistance to mass transfer of cnem surfactants to the water/oil interface. *Langmuir* **2019**, *35*, 2898–2908.
218. Lin, D.; Grossfield, A. Thermodynamics of antimicrobial lipopeptide binding to membranes: Origins of affinity and selectivity. *Biophys. J.* **2014**, *107*, 1862–1872.
219. Bhaskara, R. M.; Linker, S. M.; Vögele, M.; Köfinger, J.; Hummer, G. Carbon nanotubes mediate fusion of lipid vesicles. *ACS Nano* **2017**, *11*, 1273–1280.
220. Pownall, H. J.; Bick, D. L. M.; Massey, J. B. Spontaneous phospholipid transfer: Development of a quantitative model. *Biochemistry* **1991**, *30*, 5696–5700.
221. Silvius, J. R.; Leventis, R. Spontaneous interbilayer transfer of phospholipids: Depen-

dence of acyl chain composition. *Biochemistry* **1993**, *32*, 13318–13326.

222. Wimley, W. C.; Thompson, T. E. Exchange and flip-flop of dimyristoylphosphatidylcholine in liquid-crystalline, gel, and 2-component, 2-phase large unilamellar vesicles. *Biochemistry* **1990**, *29*, 1296–1303.
223. Rogers, J. R.; Geissler, P. L. Breakage of hydrophobic contacts limits the rate of passive lipid exchange between membranes. *J. Phys. Chem. B* **2020**, *124*, 5884–5898.
224. Khakbaz, P.; Klauda, J. B. Investigation of phase transitions of saturated phosphocholine lipid bilayers via molecular dynamics simulations. *Biochm. Biophys. Acta Biomembr.* **2018**, *1860*, 1489–1501.
225. Domański, J.; Stansfeld, P. J.; Sansom, M. S. P.; Beckstein, O. Lipidbook: A public repository for force-field parameters used in membrane simulations. *J. Membr. Biol.* **2010**, *236*, 255–258.
226. Harris, C. R. et al. Array programming with numpy. *Nature* **2020**, *585*, 357–362.
227. Seu, K. J.; Cambrea, L. R.; Everly, R. M.; Hovis, J. S. Influence of lipid chemistry on membrane fluidity: Tail and headgroup interactions. *Biophys. J.* **2006**, *91*, 3727–3735.
228. Dickey, A.; Faller, R. Examining the contributions of lipid shape and headgroup charge on bilayer behavior. *Biophys. J.* **2008**, *95*, 2636–2646.
229. Lee, D. E.; Lew, M. G.; Woodbury, D. J. Vesicle fusion to planar membranes is enhanced by cholesterol and low temperature. *Chem. Phys. Lipids* **2013**, *166*, 45–54.
230. Liao, Z.; Cimakasky, L. M.; Hampton, R.; Nguyen, D. H.; Hildreth, J. E. K. Lipid rafts and HIV pathogenesis: Host membrane cholesterol is required for infection by HIV type 1. *AIDS Res. Hum. Retrovir.* **2001**, *17*, 1009–1019.
231. Klug, Y. A.; Rotem, E.; Schwarzer, R.; Shai, Y. Mapping out the intricate relationship of the HIV envelope protein and the membrane environment. *Biochm. Biophys. Acta Biomembr.* **2017**, *1859*, 550–560.
232. Takeda, M.; Leser, G. P.; Russell, C. J.; Lamb, R. A. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14610–14617.
233. Bavari, S.; Bosio, C. M.; Wiegand, E.; Ruthel, G.; Will, A. B.; Geisbert, T. W.; Hevey, M.; Schmaljohn, C.; Schmaljohn, A.; Aman, M. J. Lipid raft microdomains: A gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J. Exp. Med.* **2002**, *195*, 593–602.
234. Freitas, M. S.; Follmer, C.; Costa, L. T.; Vilani, C.; Bianconi, M. L.; Achete, C. A.; Silva, J. L. Measuring the strength of interaction between the Ebola fusion peptide and lipid rafts: Implications for membrane fusion and virus infection. *PLoS One* **2011**, *6*, e15756–e15756.
235. Fecchi, K.; Anticoli, S.; Peruzzu, D.; Iessi, E.; Gagliardi, M. C.; Matarrese, P.; Ruggieri, A. Coronavirus interplay with lipid rafts and autophagy unveils promising therapeutic targets. *Front. Microbiol.* **2020**, *11*, 1821.

236. Patel, A. J.; Garde, S. Efficient method to characterize the context-dependent hydrophobicity of proteins. *J. Phys. Chem. B* **2014**, *118*, 1564–1573.

237. Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. Sitting at the edge: How biomolecules use hydrophobicity to tune their interactions and function. *J. Phys. Chem. B* **2012**, *116*, 2498–2503.

238. Rego, N. B.; Xi, E.; Patel, A. J. Protein hydration waters are susceptible to unfavorable perturbations. *J. Am. Chem. Soc.* **2019**, *141*, 2080–2086.

239. Rego, N. B.; Xi, E.; Patel, A. J. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2018234118.

240. Brown, M. F. Soft matter in lipid–protein interactions. *Annu. Rev. Biophys.* **2017**, *46*, 379–410.

241. Brouillette, C. G.; Anantharamaiah, G. M.; Engler, J. A.; Borhani, D. W. Structural models of human apolipoprotein A-I: A critical analysis and review. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2001**, *1531*, 4–46.

242. Lentz, B. R.; Barenholz, Y.; Thompson, T. E. Fluorescence depolarization studies of phase transitions and fluidity in phospholipid bilayers. 1. Single component phosphatidylcholine liposomes. *Biochemistry* **1976**, *15*, 4521–4528.

243. Yesylevskyy, S. O.; Rivel, T.; Ramseyer, C. The influence of curvature on the properties of the plasma membrane. Insights from atomistic molecular dynamics simulations. *Sci. Rep.* **2017**, *7*, 16078.

244. Liu, C.; Elvati, P.; Majumder, S.; Wang, Y.; Liu, A. P.; Violi, A. Predicting the time of entry of nanoparticles in lipid membranes. *ACS Nano* **2019**, *13*, 10221–10232.

245. Iaea, D. B.; Dikiy, I.; Kiburu, I.; Eliezer, D.; Maxfield, F. R. STARD4 membrane interactions and sterol binding. *Biochemistry* **2015**, *54*, 4623–4636.

246. Olson, D. K.; Frohlich, F.; Farese, R. V.; Walther, T. C. Taming the sphinx: Mechanisms of cellular sphingolipid homeostasis. *Biochim. Biophys. Acta* **2016**, *1861*, 784–792.

247. Simanshu, D. K.; Kamlekar, R. K.; Wijesinghe, D. S.; Zou, X.; Zhai, X.; Mishra, S. K.; Molotkovsky, J. G.; Malinina, L.; Hinchcliffe, E. H.; Chalfant, C. E.; Brown, R. E.; Patel, D. J. Non-vesicular trafficking by a ceramide-1-phosphate transfer protein regulates eicosanoids. *Nature* **2013**, *500*, 463–467.

248. Mishra, S. K.; Gao, Y. G.; Deng, Y. B.; Chalfant, C. E.; Hinchcliffe, E. H.; Brown, R. E. CPTP: A sphingolipid transfer protein that regulates autophagy and inflammasome activation. *Autophagy* **2018**, *14*, 862–879.

249. Rogers, J. R.; Espinoza Garcia, G.; Geissler, P. L. Membrane hydrophobicity determines the activation free energy of passive lipid transport. **2021**, bioRxiv doi: 10.1101/2021.03.17.435885.

250. Zhai, X. H.; Gao, Y. G.; Mishra, S. K.; Simanshu, D. K.; Boldyrev, I. A.; Benson, L. M.; Bergen, H. R.; Malinina, L.; Mundy, J.; Molotkovsky, J. G.; Patel, D. J.; Brown, R. E.

Phosphatidylserine stimulates ceramide-1-phosphate (C1P) intermembrane transfer by C1P transfer proteins. *J. Biol. Chem.* **2017**, *292*, 2531–2541.

251. Venable, R. M.; Sodt, A. J.; Rogaski, B.; Rui, H.; Hatcher, E.; MacKerell, A. D.; Pastor, R. W.; Klauda, J. B. CHARMM all-atom additive force field for sphingomyelin: Elucidation of hydrogen bonding and of positive curvature. *Biophys. J.* **2014**, *107*, 134–145.
252. Khan, H. M.; Souza, P. C. T.; Thallmair, S.; Barnoud, J.; de Vries, A. H.; Marrink, S. J.; Reuter, N. Capturing choline-aromatics cation-pi interactions in the MARTINI force field. *J. Chem. Theory Comput.* **2020**, *16*, 2550–2560.
253. Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*.
254. Periole, X.; Cavalli, M.; Marrink, S. J.; Ceruso, M. A. Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics, and intermolecular recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.
255. Webb, B.; Sali, A. Comparative protein structure modeling using modeller. *Current Protocols in Bioinformatics* **2016**, *54*, 5.6.1–5.6.37.
256. Fiser, A.; Do, R. K. G.; Sali, A. Modeling of loops in protein structures. *Protein Sci.* **2000**, *9*, 1753–1773.
257. de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013**, *9*, 687–697.
258. Wassenaar, T. A.; Pluhackova, K.; Bockmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. *J. Chem. Theory Comput.* **2014**, *10*, 676–690.
259. Schmidtke, P.; Bidon-Chanal, A.; Luque, F. J.; Barril, X. Mdpocket: Open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics (Oxford, England)* **2011**, *27*, 3276–3285.
260. Best, R. B.; Hummer, G.; Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17874.
261. Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B* **2006**, *110*, 3533–3539.
262. Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*.
263. Tiwary, P.; Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **2015**, *119*, 736–742.
264. Touchette, H. The large deviation approach to statistical mechanics. *Phys. Rep.* **2009**, *478*, 1–69.

265. Touchette, H. A basic introduction to large deviations: Theory, applications, simulations. **2011**, arXiv:1106.4146.

266. Ray, U.; Chan, G. K. L.; Limmer, D. T. Importance sampling large deviations in nonequilibrium steady states. I. *J. Chem. Phys.* **2018**, *148*.

267. Hurtado, P. I.; Garrido, P. L. Large fluctuations of the macroscopic current in diffusive systems: A numerical test of the additivity principle. *Phys. Rev. E* **2010**, *81*.

268. Speck, T.; Garrahan, J. P. Space-time phase transitions in driven kinetically constrained lattice models. *Eur. Phys. J. B* **2011**, *79*, 1–6.

269. Hurtado, P. I.; Espigares, C. P.; del Pozo, J. J.; Garrido, P. L. Thermodynamics of currents in nonequilibrium diffusive systems: Theory and simulation. *J. Stat. Phys.* **2014**, *154*, 214–264.

270. Dellago, C.; Bolhuis, P. G.; Chandler, D. Efficient transition path sampling: Application to lennard-jones cluster rearrangements. *J. Chem. Phys.* **1998**, *108*, 9236–9245.

271. Chou, T.; Mallick, K.; Zia, R. K. P. Non-equilibrium statistical mechanics: From a paradigmatic model to biological transport. *Rep. Prog. Phys.* **2011**, *74*.

272. Mallick, K. Some exact results for the exclusion process. *J. Stat. Mech.: Theory Exp.* **2011**, P01024.

273. Jack, R. L.; Thompson, I. R.; Sollich, P. Hyperuniformity and phase separation in biased ensembles of trajectories for diffusive systems. *Phys. Rev. Lett.* **2015**, *114*.

274. Dolezal, J.; Jack, R. L. Large deviations and optimal control forces for hard particles in one dimension. *J. Stat. Mech.: Theory Exp.* **2019**, 123208.

275. MacDonald, C. T.; Gibbs, J. H.; Pipkin, A. C. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers* **1968**, *6*, 1–25.

276. MacDonald, C. T.; Gibbs, J. H. Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers* **1969**, *7*, 707–725.

277. Kolomeisky, A. B. Channel-facilitated molecular transport across membranes: Attraction, repulsion, and asymmetry. *Phys. Rev. Lett.* **2007**, *98*.

278. Van Kampen, N. G. *Stochastic processes in physics and chemistry*; North-holland personal library; Elsevier Science, 1992.

279. Chetrite, R.; Touchette, H. Nonequilibrium markov processes conditioned on large deviations. *Ann. Henri Poincare* **2015**, *16*, 2005–2057.

280. Popkov, V.; Schutz, G. M.; Simon, D. ASEP on a ring conditioned on enhanced flux. *J. Stat. Mech.: Theory Exp.* **2010**, P10007.

281. Lecomte, V.; Garrahan, J. P.; van Wijland, F. Inactive dynamical phase of a symmetric exclusion process on a ring. *J. Phys. A* **2012**, *45*, 175001.

282. Nyawo, P. T.; Touchette, H. Large deviations of the current for driven periodic diffusions. *Phys. Rev. E* **2016**, *94*.

283. Grela, J.; Majumdar, S. N.; Schehr, G. Non-intersecting brownian bridges in the flat-

to-flat geometry. **2021**, arXiv:2103.02545.

284. Majumdar, S. N.; Orland, H. Effective langevin equations for constrained stochastic processes. *J. Stat. Mech.: Theory Exp.* **2015**, P06039.

Appendix A

Appendix for "Breakage of Hydrophobic Contacts Limits the Rate of Passive Lipid Exchange between Membranes"

Additional Order Parameters Used to Characterize Transition Paths

In addition to the order parameters described in Chapter 1, d_{lip}, d_{sn1}, d_{sn2}, $\min(d_{\mathrm{CC}})$, and n_{CC}, we evaluated additional order parameters for their potential as reaction coordinates. They are as follows:

$n_{\mathrm{w,X}}^{(\mathrm{lip})}$ The number of water molecules within the first solvation shell of bead (atom) X of the tagged MARTINI (CHARMM36) lipid. Water beads (water molecules with oxygen atoms) within a distance $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{X})$ of X were counted. $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{X})$ is the location of the first minimum of the radial distribution function between water beads (oxygens) and X calculated from a simulation of a single lipid in solution.

$n_{\mathrm{w,X}}^{(\mathrm{mem})}$ The number of water molecules within the first solvation shell of bead (atom) X of MARTINI (CHARMM36) lipids in the membrane that are nearby the tagged lipid. Only lipids in the membrane whose distance to the tagged lipid in the plane parallel to the bilayer surface (xy plane) was within 8 Å were considered. Water beads (water molecules with oxygen atoms) within a distance $d_{\mathrm{solv}}^{(\mathrm{mem})}(\mathrm{X})$ of X were counted. $d_{\mathrm{solv}}^{(\mathrm{mem})}(\mathrm{X})$ is the location of the first minimum of the radial distribution function between water beads (oxygens) and X calculated from a simulation of an unperturbed lipid bilayer. The value of $d_{\mathrm{solv}}^{(\mathrm{mem})}(\mathrm{X})$ for beads C1A and C1B (hydrophobic carbons closest to the glycerol group) were used for all hydrophobic carbons in the lipid tails.

$n_{\mathrm{X,Y}}^{(\mathrm{cyl})}$ The number of X within a cylinder extending in z from Y to the midplane of the bilayer. X is either water beads (oxygens) with label X = w, all lipid beads (atoms) with label X = lip, hydrophilic lipid beads (atoms) with label X = philic, hydrophobic lipid beads (atoms) with label X = phobic, or all beads (atoms) with label X = all. Y is either the center-of-mass (COM) of the tagged lipid with label Y = lip, the terminal carbon of the sn1 tail of the tagged lipid with label Y = sn1, or the terminal carbon of the sn2 tail of the tagged lipid with label Y = sn2. For Y = lip, the radius of the cylinder is 8 Å; for Y = sn1, the radius is $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{C3B})$ for MARTINI and $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{C314})$ for CHARMM36; for Y = sn2 the radius is $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{C3A})$ for MARTINI and $d_{\mathrm{solv}}^{(\mathrm{lip})}(\mathrm{C214})$ for CHARMM36.

$\rho_{\mathrm{X,Y}}^{(\mathrm{cyl})}$ The number density of X within a cylinder extending in z from Y to the midplane of the bilayer, where X = w, lip, philic, phobic, or all and Y = lip, sn1, or sn2. The radius of the cylinder depends on Y just as for $n_{\mathrm{X,Y}}^{(\mathrm{cyl})}$.

$\langle \delta z_{\mathrm{phos}} \rangle_{\mathrm{mem}}$ Average height fluctuation of lipids in the membrane that are within 8 Å of the tagged lipid in the xy plane. $\delta z_{\mathrm{phos}} = z_{\mathrm{phos}} - \langle z_{\mathrm{phos}} \rangle$ is the deviation of the z position of the phosphate group of a lipid nearby the tagged lipid from the average z position of the phosphate group of all lipids in the leaflet.

$\langle n_{\mathrm{ET}} \rangle_{\mathrm{mem}}$ Average number of exposed lipid tails flipped out of the bilayer. An exposed tail has at least one hydrophobic carbon bead (atom) with a z position greater than $\langle z_{\mathrm{phos}} \rangle$. Only lipids nearby the tagged lipid, which are within 8 Å of the tagged lipid in the xy plane, were considered.

$\min(d_{\mathrm{HT,Y}})$ The minimum distance between head groups of lipids in the membrane and the terminal carbon of one tail of the tagged lipid in the xy plane. Specifically, $d_{\mathrm{HT,Y}}$ is measured from bead NC3 (atom N) to either terminal carbon of the tagged lipid's tails, labeled Y = sn1 or sn2, for MARTINI (CHARMM36).

$A_{\text{defect,Y}}$	Area of packing defects in the membrane near one tail of the tagged lipid, labeled Y = sn1 or sn2. Defects that expose hydrophobic portions of the membrane were identified with PackMem[60,61] on a 1 Å × 1 Å grid. The tagged lipid was excluded in the analysis to identify defects. A defect is considered near Y if any grid point that makes up the defect is within 1 Å of Y in the xy plane.
E_{int}	Interaction energy between the tagged lipid and the rest of the system.
$\min(d_{\text{CC,Y}})$	Hydrophobic contacts between one tail of the tagged lipid and membrane. $d_{\text{CC,Y}}$ is the distance between a hydrophobic carbon in the sn1 tail, labeled Y = sn1, or sn2 tail, labeled Y = sn2, of the tagged lipid and of the closest leaflet.
$n_{\text{CC,Y}}$	Number of close hydrophobic contacts between one tail of the tagged lipid, labeled Y = sn1 or sn2, and closest leaflet. Any pair of hydrophobic carbons with a distance $d_{\text{CC}} \leq 14$ Å for MARTINI lipids or with $d_{\text{CC}} \leq 10$ Å for CHARMM36 lipids were counted as close contacts.

Stockholm Lipids (Slipids) Simulations

Molecular Dynamics Simulations

To simulate a bilayer of 128 Slipids[206,207] DMPC lipids for use in subsequent umbrella sampling (US) simulations, we used a protocol similar to that used for CHARMM36 lipids. First, the bilayer was energy minimized prior to undergoing a two-stage equilibration at 320 K and 1 bar. The first 250 ps equilibration utilized the Berendsen barostat[169] for semi-isotropic pressure coupling with a coupling time constant of 2 ps and isothermal compressibility of 4.5×10^{-5} bar^{-1}, and the second 250 ps equilibration utilized the Parinello-Rahman barostat[170] with a coupling time constant of 10 ps. Then, a 50 ns production run was performed to allow the bilayer to fully equilibrate (Figure A.1). The lipids and solvent were coupled to separate Nosé-Hoover thermostats[183,184] using a coupling time constant of 1 ps to maintain the temperature. Dynamics were evolved according to the leapfrog algorithm[173] using a 2 fs time step. All bonds were constrained using the LINCS algorithm.[185] Lennard-Jones forces were smoothly switched off between 1.4 and 1.5 nm. Coulomb interactions were truncated at 1.5 nm, and long-ranged Coulomb interactions were calculated using Particle Mesh Ewald (PME) summation[186] with a Fourier spacing of 0.12 nm and an interpolation order of 4. A long-range analytic dispersion correction was applied to both energy and pressure to account for the truncation of Lennard-Jones interactions. Neighbor lists were constructed with the Verlet algorithm.[175]

Next, a system with a tagged lipid in solution was built for use in US simulations. This system was energy minimized and equilibrated using the protocol described for the Slipids bilayer system with the addition of harmonic restraints on the z coordinates of all heavy atoms of the tagged lipid.

Free Energy Calculations

To obtain the 2D free energy surface $\Delta F(\min(d_{\mathrm{CC}}), n_{\mathrm{CC}})$ for the Slipids system using US simulations, the same number of windows and the same harmonic biases on $\min(d_{\mathrm{CC}})$ and n_{CC} were used as for the CHARMM36 system. To generate initial configurations for each window, a 20 ns steered molecular dynamics (SMD) simulation was performed. During the SMD simulation, the system was pulled along the minimum free energy path determined for CHARMM36 (Figure A.9) using a harmonic bias on $\min(d_{\mathrm{CC}})$ with a spring constant of 5,000 kJ/mol/nm^2 and on n_{CC} with a spring constant of 0.005 kJ/mol/contact^2. Each US window was initialized with a configuration obtained from the SMD simulation that has a value of ($\min(d_{\mathrm{CC}})$, n_{CC}) close to the center of the window's bias. Each window was simulated for 24 ns. The first 4 ns of data from these windows was discarded to account for equilibration. Finally, data from all windows was combined with the weighted histogram analysis method (WHAM)[191] to obtain a free energy surface as a function of $\min(d_{\mathrm{CC}})$ and n_{CC}. Error bars were calculated as the standard error of free energy surfaces estimated from five independent 5 ns blocks.

Maximum Likelihood Approach to Identify a Reaction Coordinate

The maximum likelihood approach[140,200] identifies a reaction coordinate r_c, which is a linear combination of order parameters q_i, that best represents the outcomes of individual committor calculations when inserted into a model for the committor. The model for the committor is

$$p_{\mathrm{B}}(r_c(\boldsymbol{q})) = \frac{1 + \tanh(r_c(\boldsymbol{q}))}{2}. \tag{A.1}$$

Specifically, the form of the reaction coordinate is

$$r_c(\boldsymbol{q}) = \alpha_0 + \sum_{i=1}^{M} \alpha_i q_i, \tag{A.2}$$

where M is the number of order parameters used to construct r_c. All order parameters q_i are scaled to the range [0,1], and their units are absorbed into the coefficients α_i. The coefficient α_0 is included as an offset so that $r_c = 0$ for transition states. The coefficients α_i

are determined by maximizing the likelihood

$$\mathcal{L} = \prod_{k=1}^{N_\mathrm{B}} p_\mathrm{B}(r_c(\boldsymbol{q})) \prod_{k=1}^{N_\mathrm{A}} \left[1 - p_\mathrm{B}(r_c(\boldsymbol{q}))\right], \tag{A.3}$$

where N_B is the number of trials used in the committor calculations that end in state B and N_A is the number of trials that end in state A. As M is increased, the number of fitting parameters α_i increases, resulting in an improved reaction coordinate and increased likelihood at the cost of physical insight. A Bayesian information criterion can be used to determine when including an additional order parameter results in a significantly improved reaction coordinate. For $N = N_\mathrm{A} + N_\mathrm{B}$ realizations of the likelihood, a reaction coordinate composed of $M + 1$ order parameters provides a significant improvement over a reaction coordinate composed of M order parameters if $\ln \mathcal{L}$ increases by more than the Bayesian information criterion, $\frac{1}{2} \ln N$.

We used the maximum likelihood approach to construct $r_c^{(M=48)}$ as a linear combination of up to 48 order parameters, excluding $\min(d_\mathrm{CC})$ and n_CC, for the MARTINI model. Additionally, we used the maximum likelihood approach to construct the reaction coordinate for lipid (L) transport via solvent (xS), r_LxS, from order parameters including $\min(d_\mathrm{CC})$ and n_CC for both MARTINI and CHARMM36 models. r_LxS constructed as a linear combination of $\min(d_\mathrm{CC})$ and n_CC was used to identify dividing surfaces between states A and B where $r_\mathrm{LxS} = 0$. For MARTINI, the likelihood for each tested reaction coordinate was evaluated on $N = 4,569,150$ commitor trials found along transition paths, corresponding to a Bayesian information criterion of $\frac{1}{2} \ln N = 7.667$. For CHARMM36, the likelihood for each tested reaction coordinate was evaluated based on $N = 2,760$ commitor trials found along transition paths, corresponding to a Bayesian information criterion of $\frac{1}{2} \ln N = 3.961$.

Additional Figures

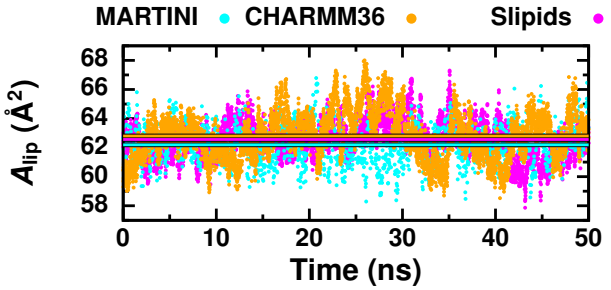


Figure A.1: Equilibration of initial lipid bilayer. Area per lipid, A_{lip}, during MD simulations of a lipid bilayer composed of 128 lipids. The straight lines that are outlined in black indicate the average area per lipid during the MARTINI simulation (cyan), CHARMM36 simulation (orage), and Slipids simulation (magenta).

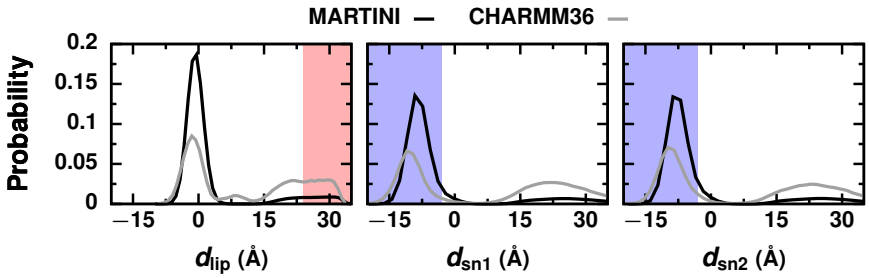


Figure A.2: Determination of stable state definitions. Probability distributions of d_{lip}, d_{sn1}, and d_{sn2} from MARTINI and CHARMM36 spontaneous insertion trajectories. State A configurations, in which the tagged lipid is fully solvated, are located in the red region. State B configurations, in which the tagged lipid is in the bilayer, are located in the blue region. The small peak around $d_{\mathrm{lip}} = 20$ Å arises from configurations in which the tagged lipid is absorbed to the bilayer. The peak in the distribution from CHARMM36 simulations around $d_{\mathrm{lip}} = 8$ Å arises from splayed lipid configurations.

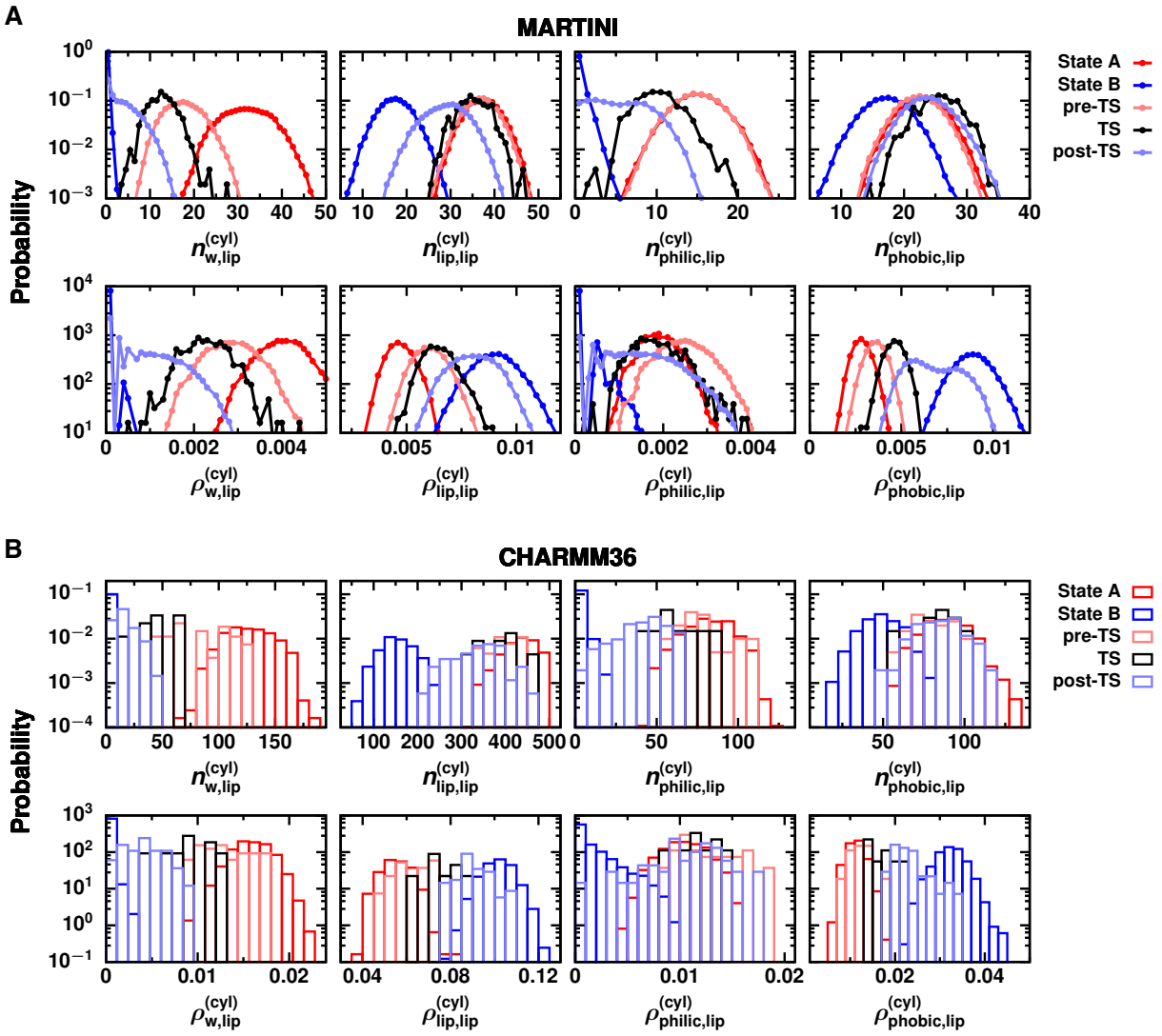


Figure A.3: Cavities around the tagged lipid are not observed during insertion. The number of water molecules and hydrophilic lipid fragments below the tagged lipid decreases during insertion while the density of hydrophobic lipid fragments increases. Probability distributions of the number, $n^{(\mathrm{cyl})}_{\mathrm{X,lip}}$, and number density, $\rho^{(\mathrm{cyl})}_{\mathrm{X,lip}}$, of water and different types of molecular fragments between the tagged lipid and midplane of the bilayer for (A) MARTINI and (B) CHARMM36 simulations. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_{\mathrm{B}} = 0$, transition states (TS) identified by $0.45 \leq p_{\mathrm{B}} \leq 0.55$ for MARTINI and by $0.4 \leq p_{\mathrm{B}} \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_{\mathrm{B}} = 1$.

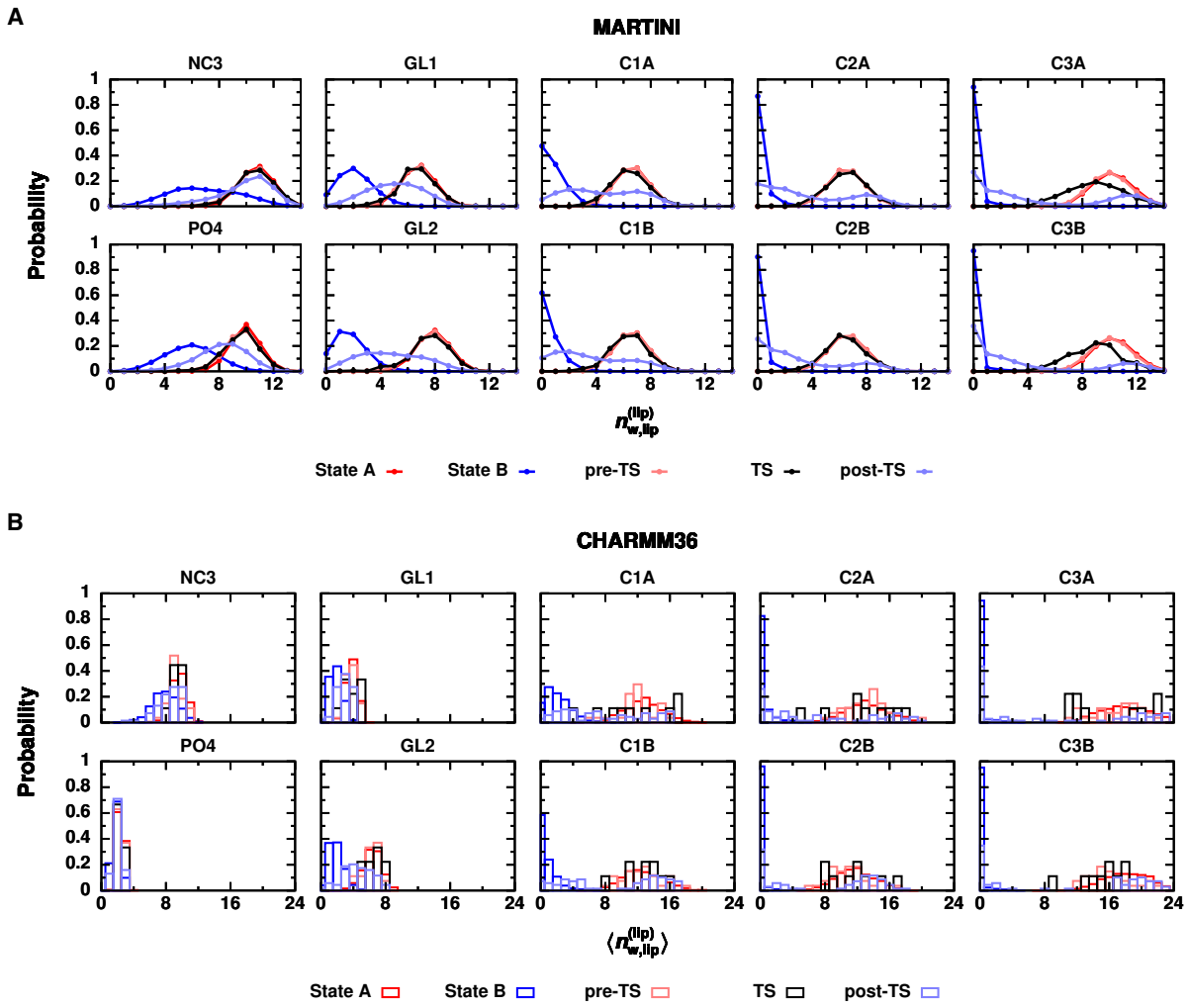


Figure A.4: Water molecules are shed from the tagged lipid's first solvation shell during insertion. Probability distributions of the number of water molecules within the first solvation shell of the tagged lipid, $n_{\mathrm{w,X}}^{(\mathrm{lip})}$. (A) For MARTINI, the number of water beads around each bead of the tagged lipid is shown. (B) For CHARMM36, $n_{\mathrm{w,X}}^{(\mathrm{lip})}$ is averaged over the atoms that map to a given MARTINI bead: Atoms N, C13, C14, C15, C12, and C11 are mapped to bead NC3; atoms P, O13, O14, O11, and O12 are mapped to bead PO4; atoms C1, C2, O21, C21, O22, and C22 are mapped to bead GL1; atoms C3, O31, C31, O32, and C32 are mapped to bead GL2; atoms C23, C24, C25, and C26 are mapped to bead C1A; atoms C27, C28, C29, and C210 are mapped to bead C2A; atoms C211, C212, C213, and C214 are mapped to bead C3A; atoms C33, C34, C35, and C36 are mapped to bead C1B; atoms C37, C38, C39, and C310 are mapped to bead C2B; atoms C311, C312, C313, and C314 are mapped to bead C3B. Beads C1B, C2B, and C3B comprise the sn1 tail, and beads C1A, C2A, and C3A comprise the sn2 tail. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_{\mathrm{B}} = 0$, transition states (TS) identified by $0.45 \leq p_{\mathrm{B}} \leq 0.55$ for MARTINI and by $0.4 \leq p_{\mathrm{B}} \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_{\mathrm{B}} = 1$.

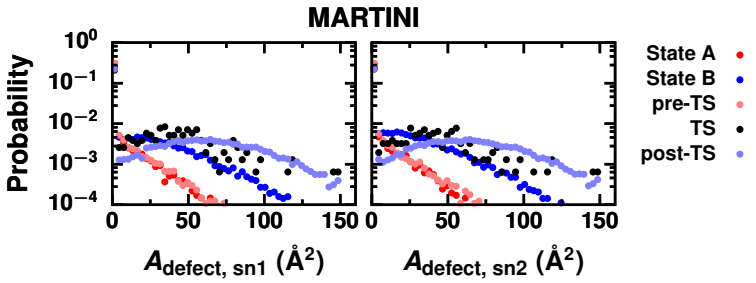


Figure A.5: Membrane packing defects that expose hydrophobic lipid fragments are increased at the transition state and when the lipid is committed to entering the membrane. Probability distributions of the area of defects located nearby either tail of the tagged lipid in the xy plane, $A_{\mathrm{defect,Y}}$. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_{\mathrm{B}} = 0$, transition states (TS) identified by $0.45 \leq p_{\mathrm{B}} \leq 0.55$ for MARTINI and by $0.4 \leq p_{\mathrm{B}} \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_{\mathrm{B}} = 1$. Becuase the tagged lipid is excluded from the analysis of defects, larger defects are observed in state B than in state A; these defects correspond to the tagged lipid.

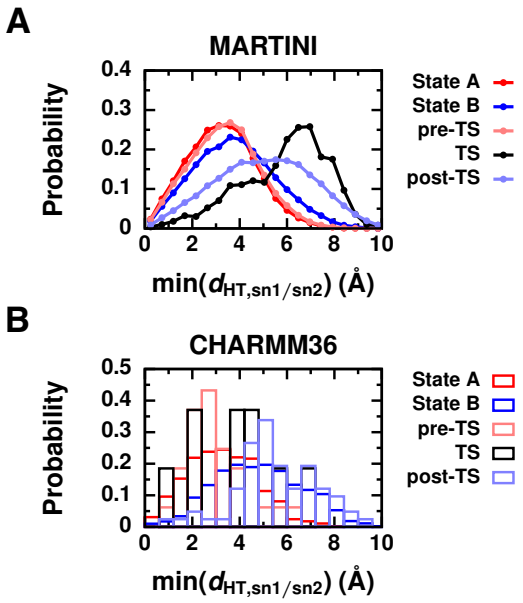


Figure A.6: The distance between the tagged lipid's tails and hydrophilic head groups of membrane lipids is increased at the transition state. Probability distributions of the minimum distance between the head groups of the membrane lipids and either tail of the tagged lipid measured in the xy plane, $\min(d_{\mathrm{HT,sn1/sn2}})$, from (A) MARTINI and (B) CHARMM36 simulations. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_{\mathrm{B}} = 0$, transition states (TS) identified by $0.45 \leq p_{\mathrm{B}} \leq 0.55$ for MARTINI and by $0.4 \leq p_{\mathrm{B}} \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_{\mathrm{B}} = 1$.

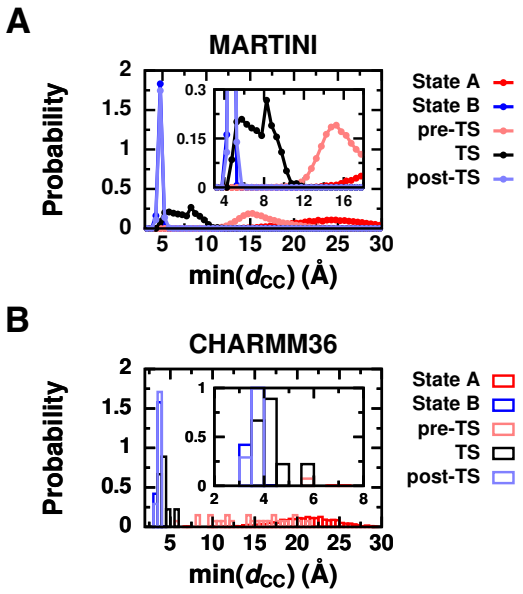


Figure A.7: The minimum distance between hydrophobic lipid and membrane carbons may be the reaction coordinate. Probability distributions of the min(d_{CC}) from (A) MARTINI and (B) CHARMM36 simulations. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_B = 0$, transition states (TS) identified by $0.45 \leq p_B \leq 0.55$ for MARTINI and by $0.4 \leq p_B \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_B = 1$. Magnified views of the transition state distributions are shown in the insets.

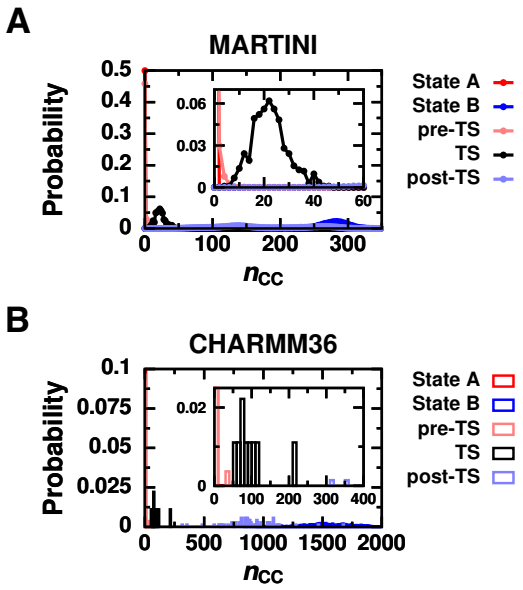


Figure A.8: The number of close hydrophobic lipid–membrane contacts may be the reaction coordinate. Probability distributions of n_{CC} from (A) MARTINI and (B) CHARMM36 simulations. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_B = 0$, transition states (TS) identified by $0.45 \leq p_B \leq 0.55$ for MARTINI and by $0.4 \leq p_B \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_B = 1$. Magnified views of the transition state distributions are shown in the insets.

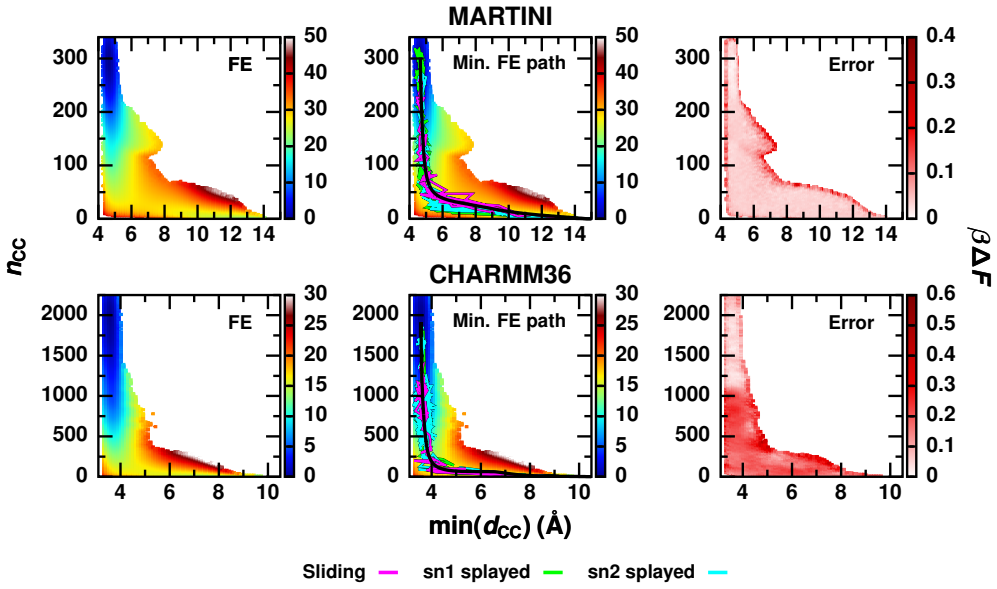


Figure A.9: Transition paths follow the minimum free energy path in the space of min(d_{CC}) and n_{CC}. (Left panel) Free energy surfaces as a function of min(d_{CC}) and n_{CC} obtained from umbrella sampling simulations. (Middle panel) Example transitions paths for each lipid insertion pathway are plotted along with the minimum free energy path (black line). (Right panel) Block standard error of the free energy surface calculated from five simulation blocks.

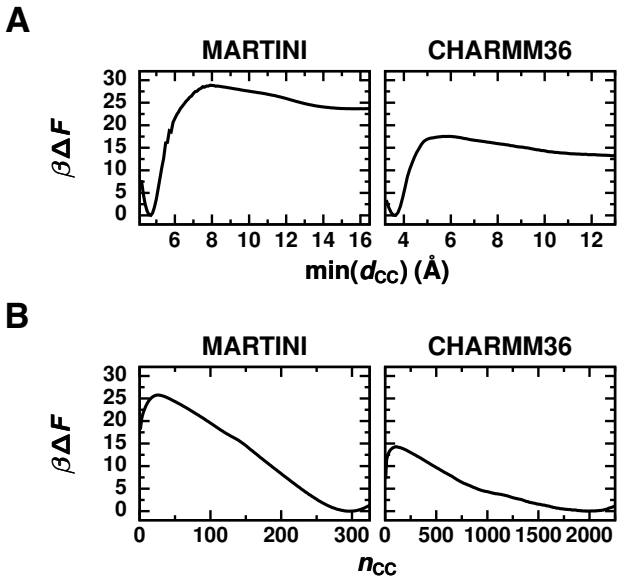


Figure A.10: A free energy barrier for lipid insertion exists due to hydrophobic lipid–membrane contact formation. 1D free energy profiles of (A) $\min(d_{\mathrm{CC}})$ and (B) n_{CC} obtained from 2D free energy surfaces shown in Figure SA.9 by numerically integrating out the other degree of freedom.

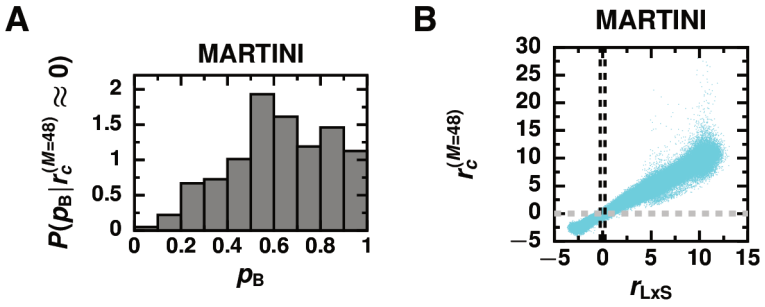


Figure A.11: Reaction coordinate, $r_c^{(M=48)}$, composed of $M = 48$ order parameters, excluding $\min(d_{\mathrm{CC}})$ and n_{CC}, with coefficients determined using a maximum likelihood approach (Table SA.2) accurately distinguishes transition states. (A) Histogram of committor values, p_{B}, for MARTINI configurations found along transition paths with $-0.25 \leq r_c^{(M=48)} \leq 0.25$. (B) Scatter plot of the reaction coordinate, $r_{\mathrm{LxS}} = \alpha_1 \min(d_{\mathrm{CC}}) + \alpha_2 n_{\mathrm{CC}} + \alpha_0$ with coefficients determined using the maximum likelihood approach (Table SA.3) versus $r_c^{(M=48)}$ constructed from $M = 48$ other order parameters for all MARTINI configurations found along transition paths. The black dashed lines outline the values of r_{LxS} characteristic of transition states. The gray dashed lines outline the values of $r_c^{(M=48)}$ characteristic of transition states.

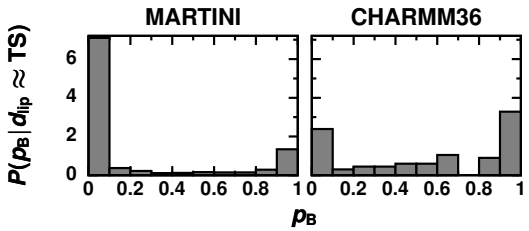


Figure A.12: The COM displacement is not the reaction coordinate for lipid exchange. Histograms of committor values, p_B, for configurations found along transition paths and that have values of d_{lip} typical of transition states (TS). Specifically, MARTINI configurations with $14 \leq d_{lip} \leq 19$ Å and CHARMM36 configurations with $10 \leq d_{lip} \leq 19$ Å are included.

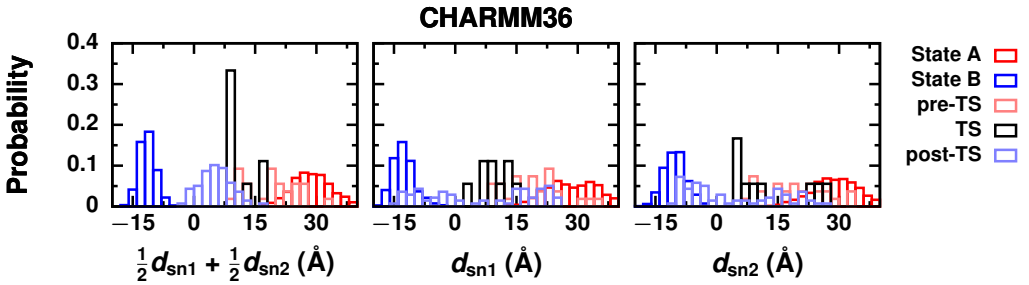


Figure A.13: Displacement-based order parameters do not reliably identify transition states found along transition paths during CHARMM36 simulations when the transition paths are not separated based on pathway. Probability distributions of $\frac{1}{2}d_{sn1} + \frac{1}{2}d_{sn2}$, d_{sn1}, and d_{sn2} from transition paths obtained from CHARMM36 simulations. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_B = 0$, transition states (TS) identified by $0.45 \leq p_B \leq 0.55$ for MARTINI and by $0.4 \leq p_B \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_B = 1$.

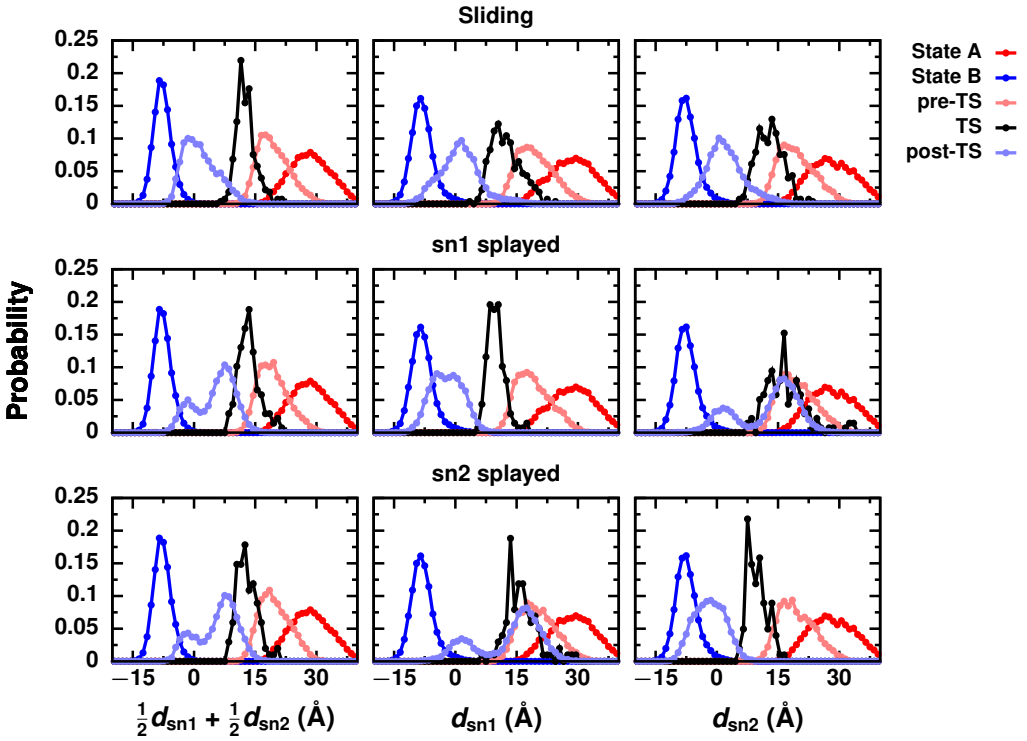


Figure A.14: Displacement-based order parameters focused on the two lipid tails may be pathway specific reaction coordinates. Probability distributions of $\frac{1}{2}d_{\mathrm{sn1}} + \frac{1}{2}d_{\mathrm{sn2}}$, d_{sn1}, and d_{sn2} from MARTINI simulations that follow each lipid insertion pathway. Distributions are plotted for state A, for state B, and for three ensembles of configurations drawn from transition paths: pre-transition state (pre-TS) configurations identified by $p_{\mathrm{B}} = 0$, transition states (TS) identified by $0.45 \leq p_{\mathrm{B}} \leq 0.55$ for MARTINI and by $0.4 \leq p_{\mathrm{B}} \leq 0.6$ for CHARMM36, and post-transition state (post-TS) configurations identified by $p_{\mathrm{B}} = 1$.

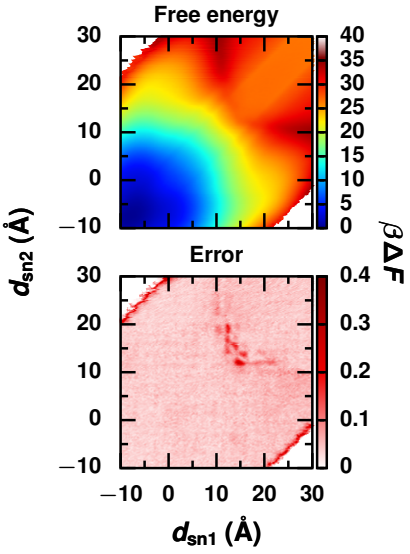


Figure A.15: Error in the free energy surface as a function of the two lipid tail distances is comparable to the barrier height. (Top panel) Free energy surface as a function of d_{sn1} and d_{sn2} obtained from umbrella sampling simulations using the MARTINI force field. (Bottom panel) Block standard error of the free energy surface calculated from five 200 ns blocks.

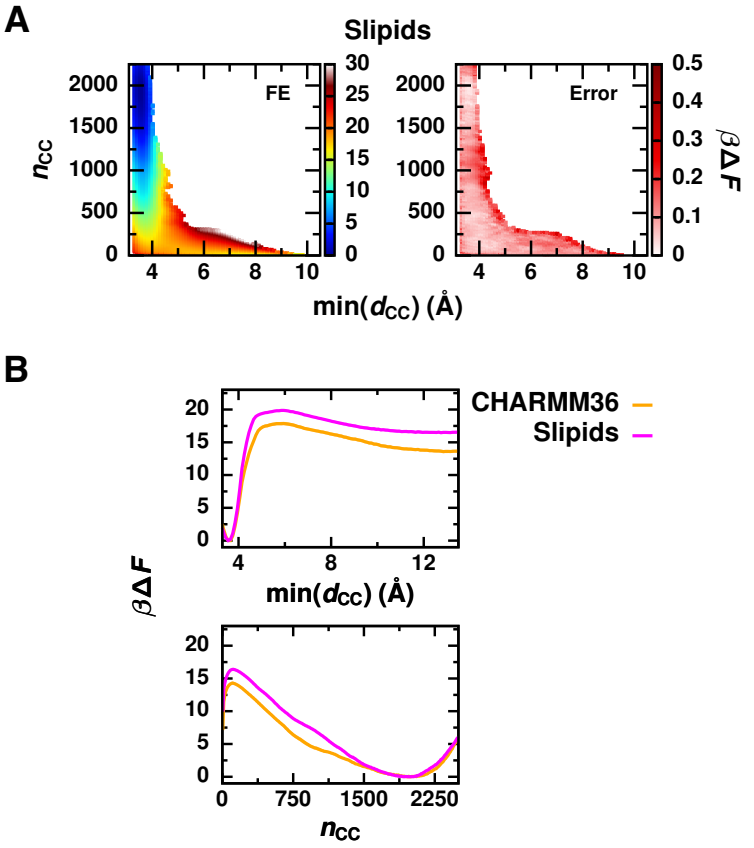


Figure A.16: The height of the free energy barrier for hydrophobic-contact formation/breakage for the Slipids force field is similar but slighly higher than the barrier height for CHARMM36. (A) (Left) Free energy surfaces as a function of min(d_{CC}) and n_{CC} obtained from umbrella sampling simulations using the Slipids force field. (Right) Block standard error of the free energy surface calculated from five simulation blocks. (B) 1D free energy profiles of min(d_{CC}) and n_{CC} for Slipids compared to CHARMM36.

Additional Tables

Table A.1: Frequency of each insertion pathway.

Pathway	MARTINI	CHARMM36
Sliding	51%	20%
sn1 splayed	27%	40%
sn2 splayed	22%	40%

Table A.2: Construction of a reaction coordinate, $r_c^{(M=48)}$, excluding min(d_{CC}) and n_{CC} using a maximum likelihood approach. The best reaction coordinates for the MARTINI model with $M = 1, 2$, and 48 are reported. To determine the best reaction coordinate with $M = 2$, only linear combinations with d_{lip}, d_{sn1}, or d_{sn2} as one of the order parameters were considered. To determine the best reaction coordinate with $M = 3$, only linear combinations with either (1) d_{lip} and $\rho^{(\text{cyl})}_{\text{philic,lip}}$ or (2) d_{sn1} and d_{sn2} were considered. All reported reaction coordinates are constructed from order parameters scaled to the range [0,1]. Scaled order parameters are denoted with a tilde to distinguish them from the unscaled values.

M	$r_c^{(M=48)}$	ln $\mathcal{L}$
1	$-12.635\tilde{d}_{\text{lip}} + 8.060$	−937984
2	$-13.583\tilde{d}_{\text{lip}} - 5.663\tilde{\rho}^{(\text{cyl})}_{\text{philic,lip}} + 11.296$	−681129
3	$-11.889\tilde{d}_{\text{lip}} - 4.411\tilde{\rho}^{(\text{cyl})}_{\text{philic,lip}} + 18.892\tilde{\rho}^{(\text{cyl})}_{\text{phobic,sn1}} + 7.680$	−596942
48	$-3.919\tilde{d}_{\text{lip}} - 2.455\tilde{d}_{\text{sn1}} - 1.529\tilde{d}_{\text{sn2}} - 0.116\tilde{n}^{(\text{lip})}_{\text{w,NC3}} - 0.054\tilde{n}^{(\text{lip})}_{\text{w,PO4}} - 0.363\tilde{n}^{(\text{lip})}_{\text{w,GL1}} - 0.436\tilde{n}^{(\text{lip})}_{\text{w,GL2}} + 0.051\tilde{n}^{(\text{lip})}_{\text{w,C1A}} + 0.278\tilde{n}^{(\text{lip})}_{\text{w,C2A}} + 0.231\tilde{n}^{(\text{lip})}_{\text{w,C3A}} - 0.028\tilde{n}^{(\text{lip})}_{\text{w,C1B}} + 0.208\tilde{n}^{(\text{lip})}_{\text{w,C2B}} + 0.099\tilde{n}_{\text{w,C3B}} + 0.174\tilde{n}^{(\text{mem})}_{\text{w,NC3}} - 0.256\tilde{n}^{(\text{mem})}_{\text{w,PO4}} - 0.90\tilde{n}^{(\text{mem})}_{\text{w,GL1}} - 0.401\tilde{n}^{(\text{mem})}_{\text{w,GL2}} - 0.001\tilde{n}^{(\text{mem})}_{\text{w,C1A}} + 0.188\tilde{n}^{(\text{mem})}_{\text{w,C2A}} + 0.007\tilde{n}^{(\text{mem})}_{\text{w,C3A}} - 0.023\tilde{n}^{(\text{mem})}_{\text{w,C1B}} + 0.043\tilde{n}^{(\text{mem})}_{\text{w,C2B}} + 0.048\tilde{n}^{(\text{mem})}_{\text{w,C3B}} + 1.447\langle\delta\tilde{z}_{\text{phos}}\rangle_{\text{mem}} + 0.813\langle\tilde{n}_{\text{ET}}\rangle_{\text{mem}} + 0.169\tilde{A}_{\text{defect,sn1}} - 0.013\tilde{A}_{\text{defect,sn2}} + 0.359\min(\tilde{d}_{\text{HT,sn1}}) + 0.255\min(\tilde{d}_{\text{HT,sn2}}) - 0.572\tilde{\rho}^{\text{cyl})}_{\text{w,lip}} - 5.839\tilde{\rho}^{(\text{cyl})}_{\text{w,sn1}} - 5.215\tilde{\rho}^{(\text{cyl})}_{\text{w,sn2}} - 0.521\tilde{\rho}^{(\text{cyl})}_{\text{lip,lip}} - 0.865\tilde{\rho}^{(\text{cyl})}_{\text{philic,lip}} + 1.499\tilde{\rho}^{(\text{cyl})}_{\text{phobic,lip}} + 4.337\tilde{\rho}^{(\text{cyl})}_{\text{lip,sn1}} + 2.885\tilde{\rho}^{(\text{cyl})}_{\text{lip,sn2}} - 1.264\tilde{\rho}^{(\text{cyl})}_{\text{philic,sn1}} - 1.504\tilde{\rho}^{(\text{cyl})}_{\text{philic,sn2}} + 17.591\tilde{\rho}^{(\text{cyl})}_{\text{phobic,sn1}} + 18.334\tilde{\rho}^{(\text{cyl})}_{\text{phobic,sn2}} + 7.904\tilde{n}^{(\text{cyl})}_{\text{w,sn1}} + 7.161\tilde{n}^{(\text{cyl})}_{\text{w,sn2}} - 2.827\tilde{n}^{(\text{cyl})}_{\text{phobic,sn1}} - 4.252\tilde{n}^{(\text{cyl})}_{\text{phobic,sn2}} - 0.442\tilde{n}^{(\text{cyl})}_{\text{all,sn1}} + 0.443\tilde{n}^{(\text{cyl})}_{\text{all,sn2}} - 1.184\tilde{E}_{\text{int}} + 5.291$	−415253

Table A.3: Construction of a reaction coordinate that can include n_{CC} and $\min(d_{CC})$ using a maximum likelihood approach demonstrates that a linear combination of n_{CC} and $\min(d_{CC})$ is the ideal reaction coordinate. To construct r_c, 54 and 32 order parameters were considered for the MARTINI and CHARMM36 models, respectively. The best reaction coordinates with $M = 1$ and 2 are reported. To determine the best reaction coordinate with $M = 2$, only linear combinations with $\min(d_{CC})$ or n_{CC} (and tail specific variants) as one of the order parameters were considered. All reported reaction coordinates are constructed from order parameters scaled to the range [0,1]. Scaled order parameters are denoted with a tilde to distinguish them from the unscaled values.

M	r_c	$\ln \mathcal{L}$
	MARTINI	
1	$24.917\tilde{n}_{CC} - 1.922$	-419147
2	$-3.136 \min(\tilde{d}_{CC}) + 13.643\tilde{n}_{CC} - 0.516$	-399237
	CHARMM36	
1	$18.583\tilde{n}_{CC} - 1.217$	-584
2	$-3.982 \min(\tilde{d}_{CC}) + 8.102\tilde{n}_{CC} - 0.151$	-515

Appendix B

Appendix for "Membrane Hydrophobicity Determines the Activation Free Energy of Passive Lipid Transport"

Analysis of Membrane Properties

Membrane Thickness. The average membrane thickness was calculated as the average distance in z between the phosphorous atoms of the top and bottom leaflets.

Density Profiles. The density of various groups along the membrane normal relative to the center of mass of the membrane at $z = 0$ were calculated with the GROMACS tool `gmx density`. The groups are: (1) polar, which includes all atoms present in the system that are not hydrophobic carbons; (2) solvent, which includes all water molecules and any sodium ions present; (3) membrane, which includes all lipid atoms; (4) polar membrane, which includes all lipid atoms that are not hydrophobic carbons; and (5) hydrophobic, which includes all hydrophobic carbons.

Acyl Chain Order Parameters. Deuterium order parameters for carbons in the lipid tails, except the ester and terminal carbons, were calculated with the GROMACS tool `gmx order`. The order parameters of carbons in double bonds were calculated following the instructions provided by Pluhackova et al.[179]

Radial Distribution Functions. To characterize the intermolecular structure of the membrane interface, we calculated radial distribution functions between the oxygen atoms of water molecules and either the phosphorous atoms or carbonyl oxygen atoms of the lipids using the GROMACS tool `gmx rdf`. Additionally, 2D radial distribution functions in the

xy-plane were calculated between the phosphorous atoms of lipids in each leaflet individually and then averaged to obtain one 2D radial distribution function for each membrane.

Additional Figures

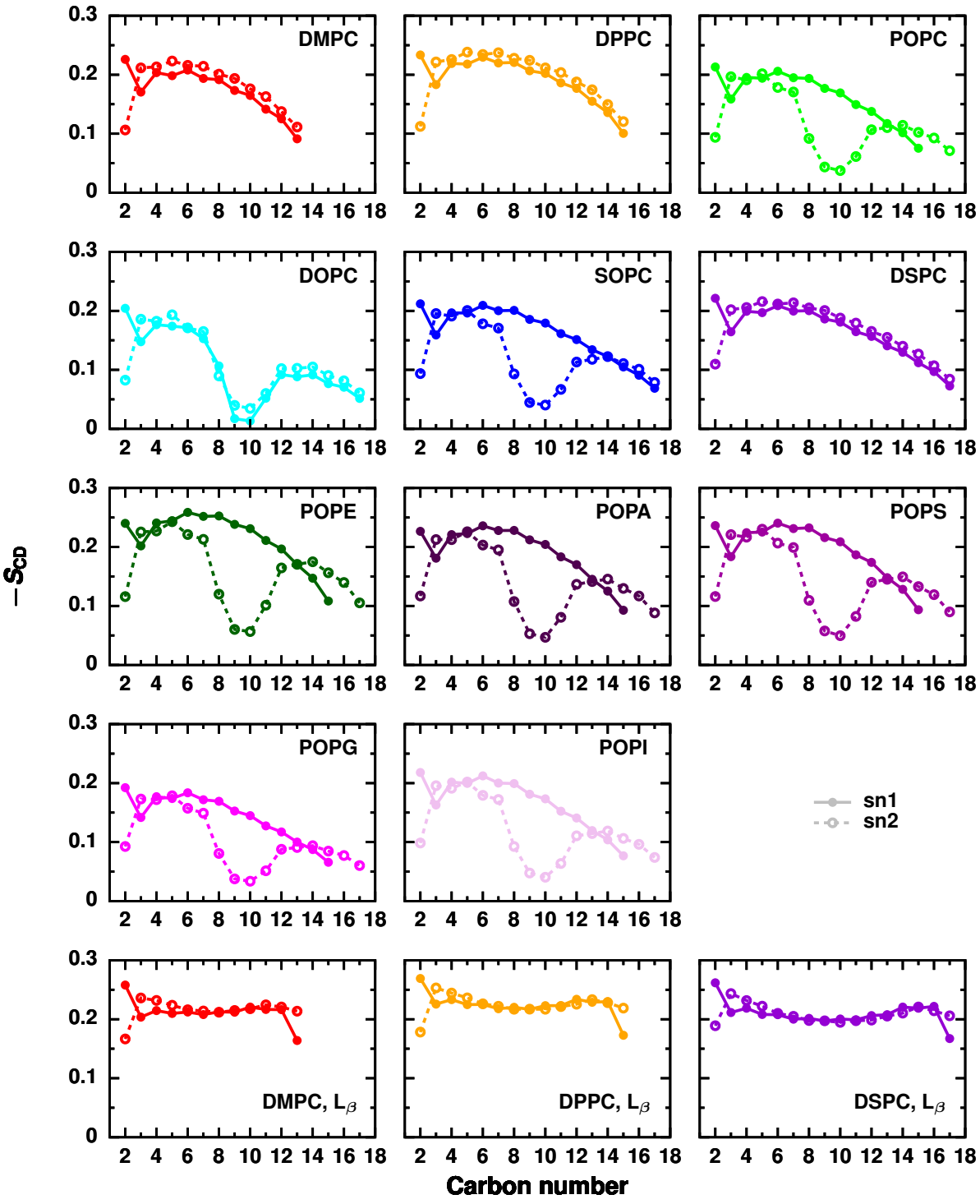


Figure B.1: Lipid chemistry and membrane phase influence the membrane's order. Deuterium order parameters of the carbons in both acyl chains of the lipids in each membrane.

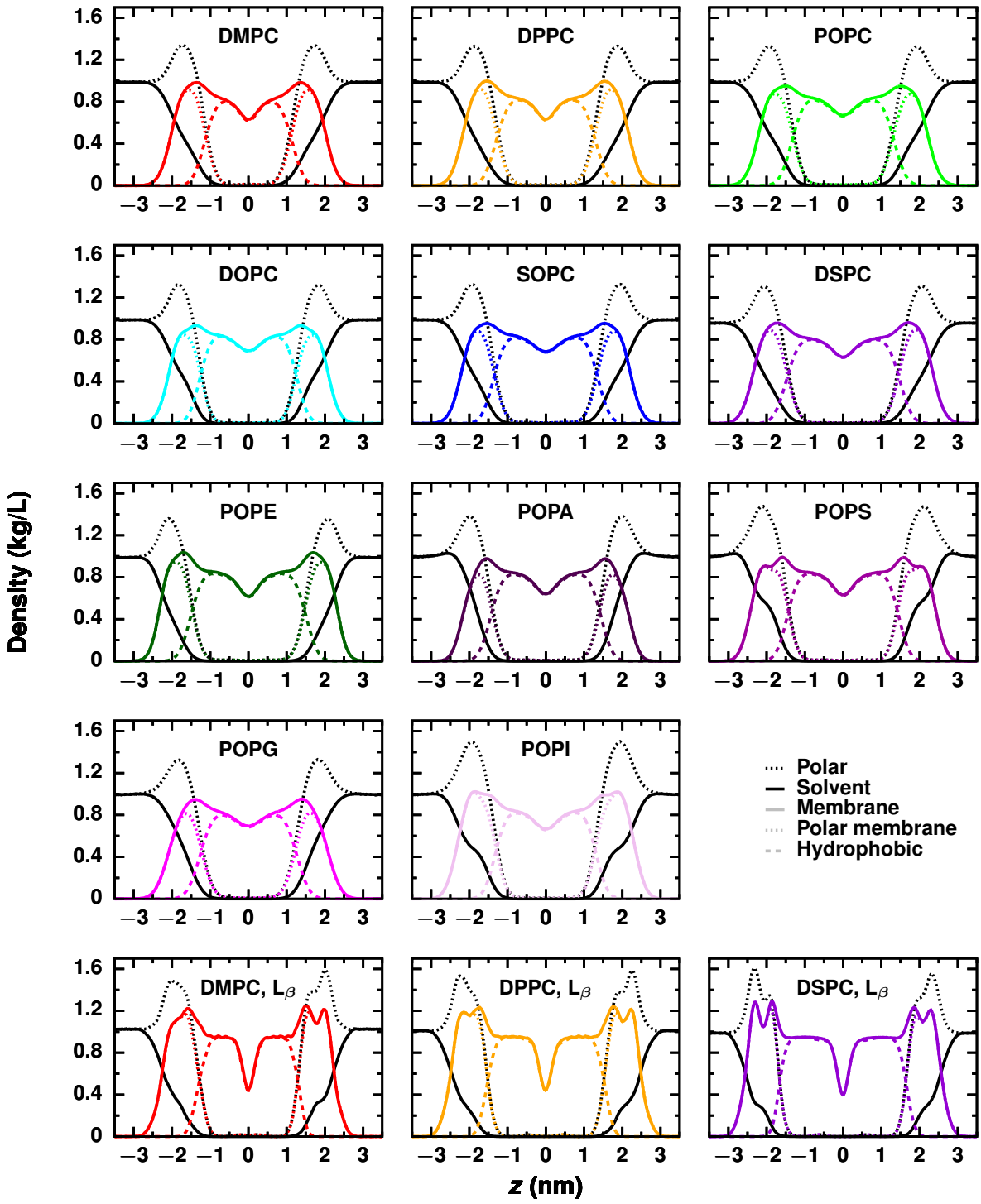


Figure B.2: System composition along the axis normal to the membranes varies with lipid chemistry and membrane phase. Density profiles along the membrane normal, z, for each membrane. The system is grouped into (1) polar, which includes all atoms that are not hydrophobic carbons; (2) solvent; (3) membrane; (4) polar membrane, which includes all lipid atoms that are not hydrophobic carbons; and (5) hydrophobic, which includes all hydrophobic carbons.

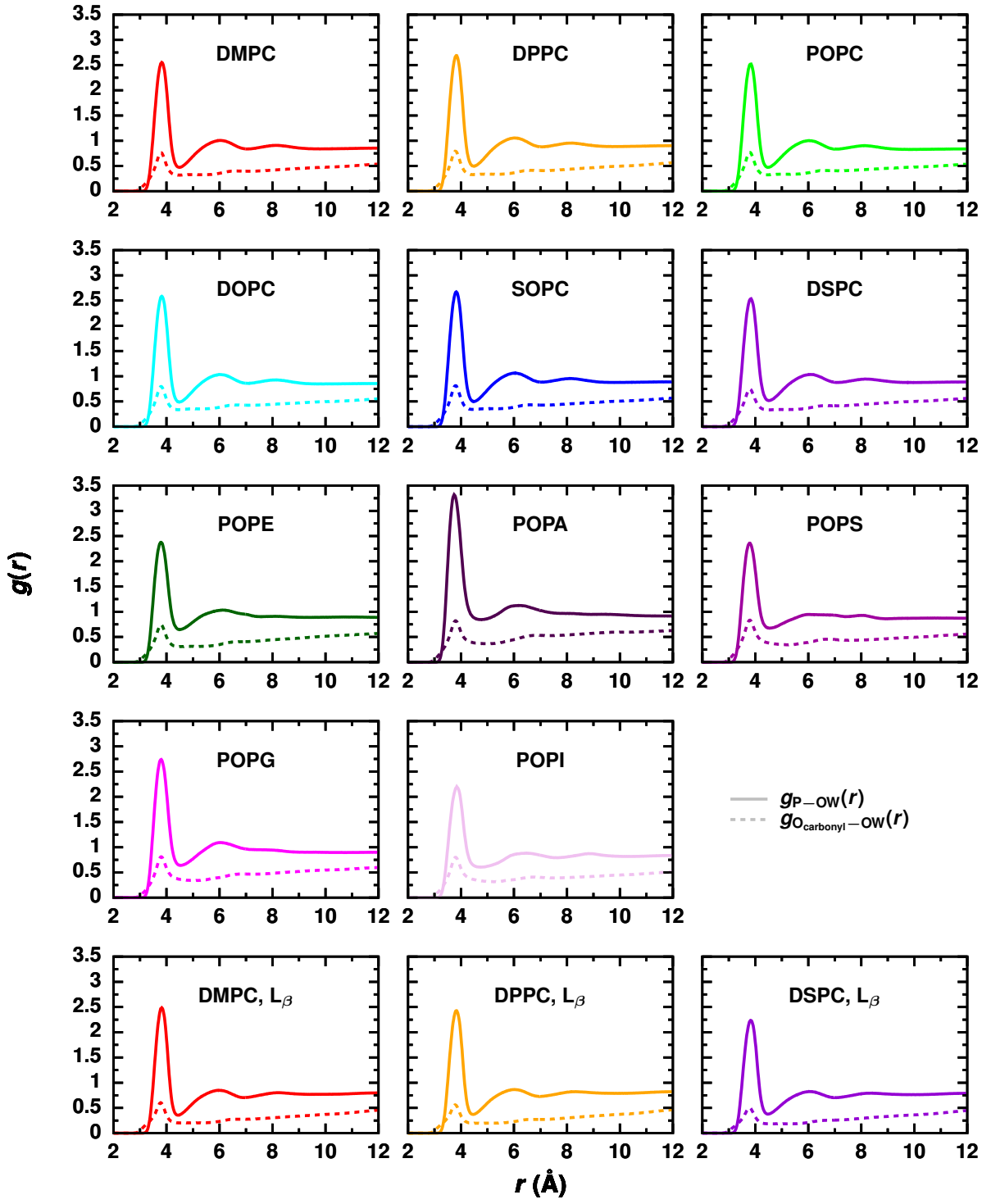


Figure B.3: Water's structure at the membrane surface depends on the lipid headgroup. Radial distribution functions, $g(r)$, between water oxygen atoms (OW) and either the phosphorous atoms (P) or carbonyl oxygen atoms ($O_{carbonyl}$) of the lipids for each membrane.

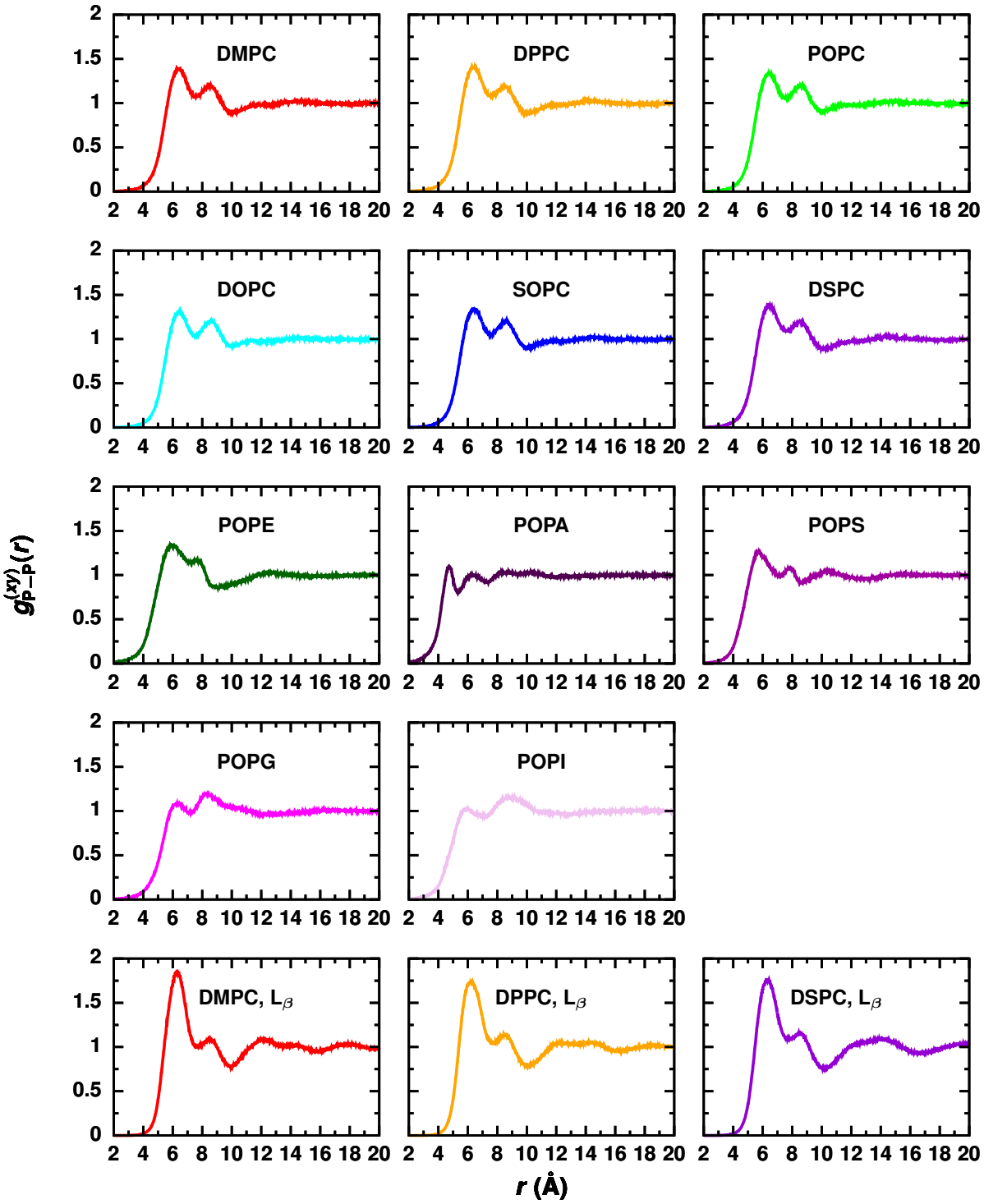


Figure B.4: Membrane surface structure, specifically the arrangement of phosphorous atoms, depends on the lipid headgroup. In L_β phase membranes, phosphorous atoms have an increased density of nearest neighbor phosphorous atoms compared to L_α phase membranes. 2D Radial distribution function, $g^{(xy)}_{\mathrm{P-P}}(r)$, between phosphorous atoms of lipids in the same leaflet for each membrane.

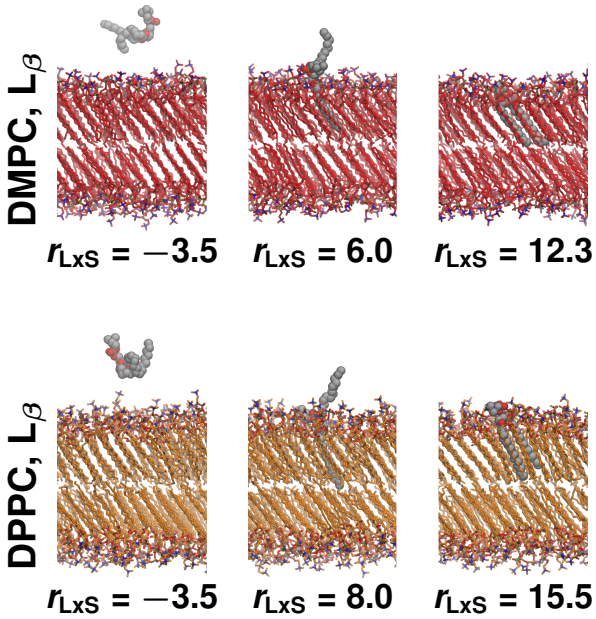


Figure B.5: Splayed lipid configurations are locally thermodynamically stable in L_β phase membranes. Configurations of L_β phase (top) DMPC and (bottom) DPPC that are representative of all minimum in $\Delta F(r_{\mathrm{LxS}})$: the lipid in solution, splayed lipid intermdiate, and lipid in the membrane.

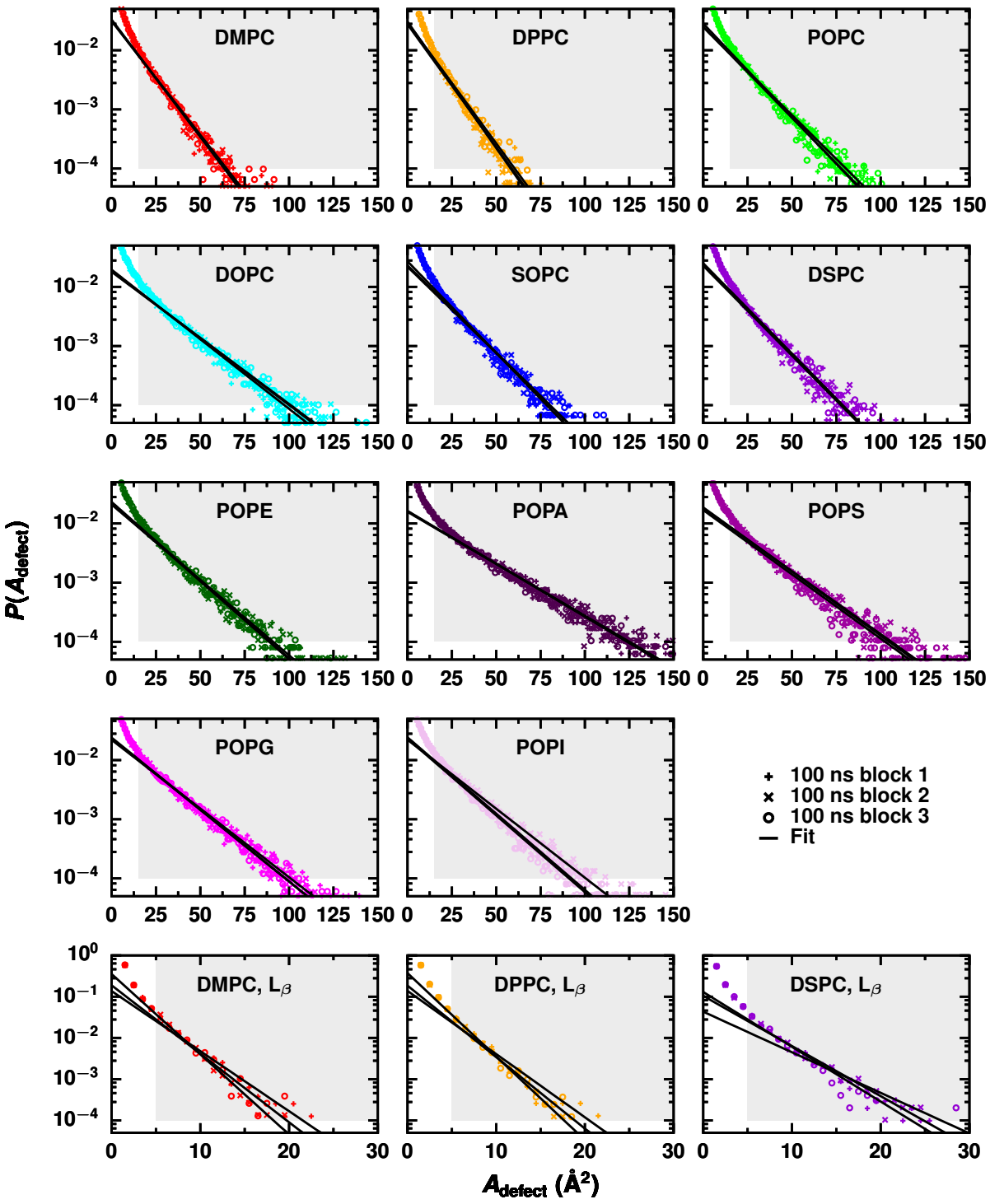


Figure B.6: The size of packing defects that expose the membrane's hydrophobic core to solvent vary with lipid chemistry and membrane phase. Distributions of the area of packing defects for each membrane. For each membrane, three independent fits were performed to distributions constructed from three 100ns simulation blocks. The region used to fit to an exponential distribution is highlighted in gray.

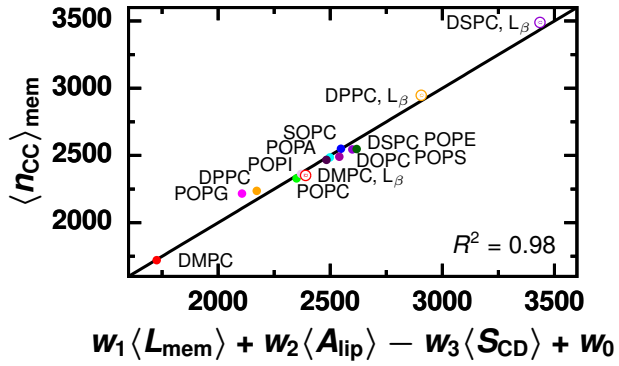


Figure B.7: $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$ varies with membrane thickness, lipid packing, and membrane order. A linear combination of a membrane's average thickness, $\langle L_{\mathrm{mem}} \rangle$, average area per lipid, $\langle A_{\mathrm{lip}} \rangle$, and average deuterium order parameter of all carbons in the lipid tails, $-\langle S_{\mathrm{CD}} \rangle$, with coefficients $w_1 = 126$ Å^{-1}, $w_2 = -40$ Å^{-2}, $w_3 = 9964$, and $w_0 = 1584$ accurately models $\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$. Lipids in thicker, more tightly packed (smaller area per lipid), and more ordered (larger $-\langle S_{\mathrm{CD}} \rangle$) membranes form more hydrophobic contacts.

Additional Tables

Table B.1: Umbrella sampling parameters.

Membrane	Spring constant (kJ/mol)	Location of first center	Location of last center	Number of windows
DMPC	20	−5.4468	13	52
DPPC	40	−6	14	54
POPC	40	−6	14	54
DOPC	40	−6	19.2828	68
SOPC	40	−6	19.2828	68
DSPC	40	−6	19.2828	68
POPE	40	−6	14	54
POPA	40	−6	14	54
POPS	40	−6	14	54
POPG	40	−6	14	54
POPI	40	−6	14	54
DMPC, L_β	40	−6	15.1358	67*
DPPC, L_β	40	−6	17.7735	67†
DSPC, L_β	40	−6	19.2828	68

* Count includes 10 windows centered at −0.9048, −0.5273, −0.1499, 1.7372, 2.8695, 7.3985, 7.7760, 8.1534, and 8.9082, which were added to decrease the spacing between some original windows and improve sampling.

† Count includes 3 windows centered at 9.2830, 9.6604, and 9.5660, which were added to improve sampling around a local maximum.

Table B.2: Average membrane physical properties.

Membrane	$\langle A_{\mathrm{lip}} \rangle$ (Å^2)	Av. Thickness (nm)	$\langle n_{\mathrm{CC}} \rangle_{\mathrm{mem}}$	π_{defect} (Å^2)	$\langle E_{\mathrm{mem-solv}} \rangle$ (kcal/mol)
DMPC	62.88 ± 0.06	3.523 ± 0.003	1720 ± 2	11.10 ± 0.08	-11361 ± 9
DPPC	60.9 ± 0.2	3.95 ± 0.01	2237 ± 6	10.4 ± 0.1	-9414 ± 16
POPC	66.77 ± 0.01	3.834 ± 0.001	2329 ± 1	14.1 ± 0.2	-11717 ± 1
DOPC	70.14 ± 0.07	3.815 ± 0.003	2485 ± 2	19.0 ± 0.2	-11993 ± 12
SOPC	66.60 ± 0.05	4.002 ± 0.003	2550 ± 2	14.5 ± 0.2	-11688 ± 7
DSPC	65.04 ± 0.02	4.401 ± 0.001	2544 ± 1	14.25 ± 0.02	-10934 ± 2
POPE	58.98 ± 0.06	4.148 ± 0.004	2548 ± 2	16.6 ± 0.2	-11168 ± 18
POPA	62.0 ± 0.1	3.982 ± 0.005	2467 ± 3	24.5 ± 0.1	-14616 ± 5
POPS	61.62 ± 0.06	4.042 ± 0.004	2491 ± 3	20.1 ± 0.1	-19022 ± 53
POPG	70.70 ± 0.07	3.630 ± 0.003	2216 ± 2	18.2 ± 0.2	-13889 ± 3
POPI	65.9 ± 0.2	3.848 ± 0.009	2363 ± 5	17.4 ± 0.5	-15181 ± 25
DMPC, L_β	49.3 ± 0.1	3.92 ± 0.05	2351 ± 5	2.6 ± 0.2	-9986 ± 52
DPPC, L_β	48.75 ± 0.02	4.376 ± 0.002	2947 ± 2	2.5 ± 0.2	-9416 ± 16
DSPC, L_β	50.08 ± 0.07	4.718 ± 0.003	3489 ± 8	3.8 ± 0.3	-9037 ± 11

Table B.3: Free energy barriers for lipid desorption (ΔF_{des}), lipid insertion (ΔF_{ins}), and the creation of splayed lipid configurations ($\Delta F_{\mathrm{splay}}$).

Membrane	ΔF_{des} (kcal/mol)	ΔF_{ins} (kcal/mol)	$\Delta F_{\mathrm{splay}}$ (kcal/mol)
DMPC	10.4 ± 0.1	2.7 ± 0.2	—
DPPC	12.1 ± 0.2	3.2 ± 0.4	—
POPC	13.8 ± 0.4	2.8 ± 0.6	—
DOPC	15.0 ± 0.4	2.5 ± 0.5	—
SOPC	14.4 ± 0.2	3.0 ± 0.3	—
DSPC	15.4 ± 0.3	3.1 ± 0.4	—
POPE	17.3 ± 0.1	1.9 ± 0.1	—
POPA	16.6 ± 0.2	3.3 ± 0.3	—
POPS	16.9 ± 0.1	4.6 ± 0.2	—
POPG	12.6 ± 0.1	2.6 ± 0.2	—
POPI	15.1 ± 0.3	3.6 ± 0.3	—
DMPC, L_β	13.7 ± 0.4	7.0 ± 0.9	9.4 ± 0.4
DPPC, L_β	16.3 ± 0.5	8.0 ± 0.8	11.4 ± 0.5
DSPC, L_β	20.7 ± 0.4	5.0 ± 0.6	14.1 ± 0.4

Appendix C

Appendix for "Ceramide-1-Phosphate Transfer Protein Enhances Lipid Transport by Disrupting Hydrophobic Lipid–Membrane Contacts"

Additional Figures

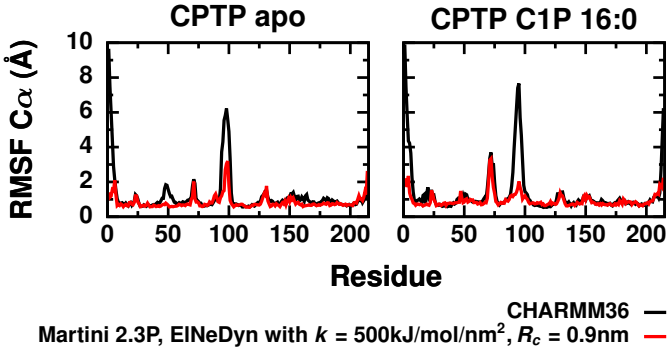


Figure C.1: Coarse-grained representations of CPTP reproduce its internal dynamics observed in all-atom simulations. Root-mean-square fluctuation (RMSF) of Cα atoms (backbone beads) of CPTP in its apo and C1P-bound forms during solution all-atom (coarse-grained) simulations.

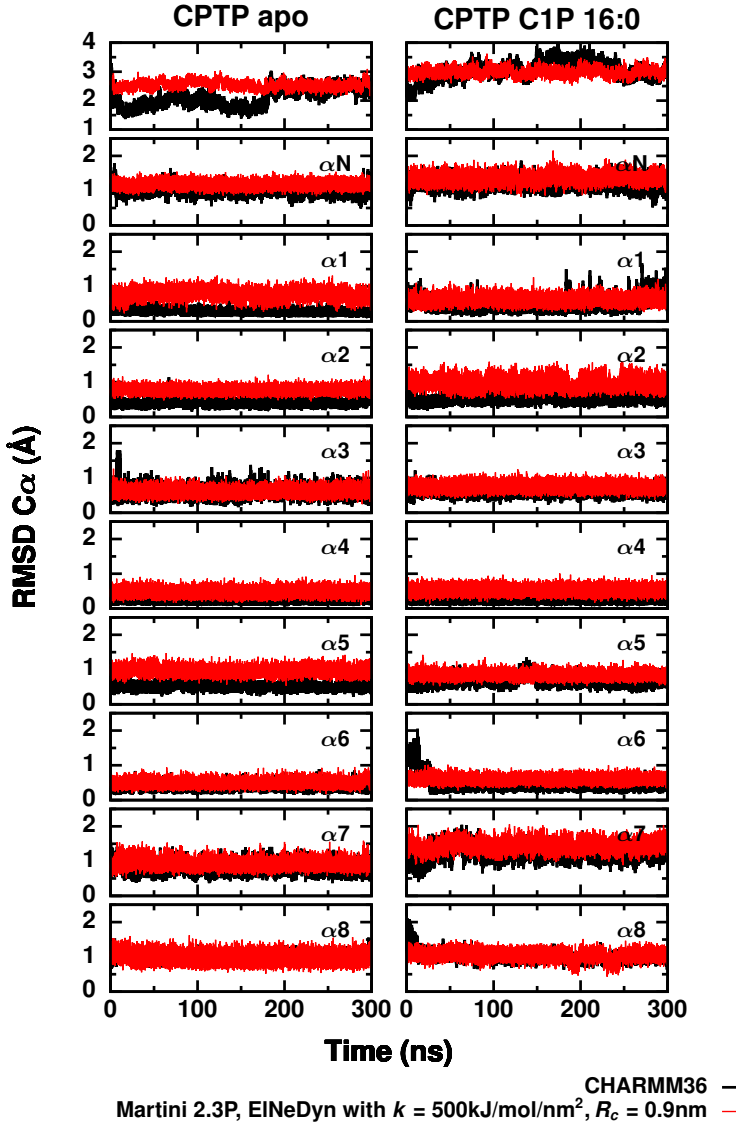
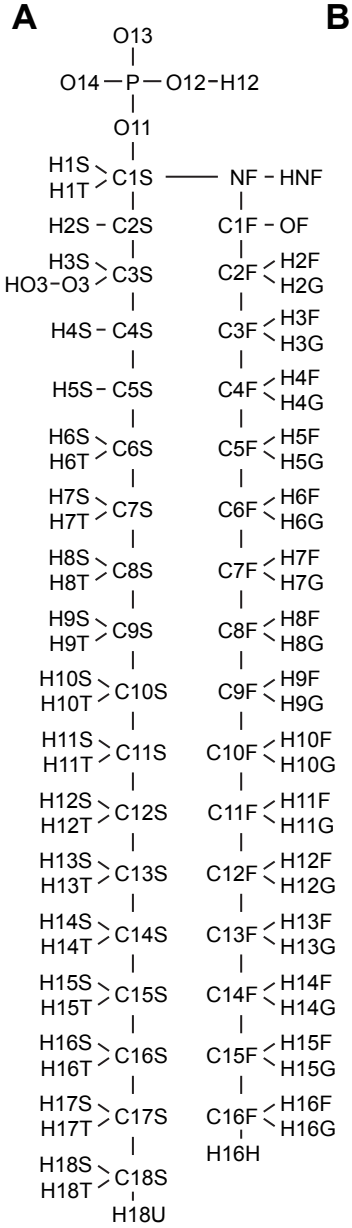


Figure C.2: Coarse-grained representations of CPTP match its structure in all-atom simulations. Root-mean-square deviation (RMSD) of Cα atoms (backbone beads) between CPTP in its apo and C1P-bound forms during solution all-atom (coarse-grained) simulations and the cystral structure of CPTP in PDB 4K85. In the top row, the RMSD for residues $8-214$ is plotted. All other rows show the RMSD for individual helices.

A

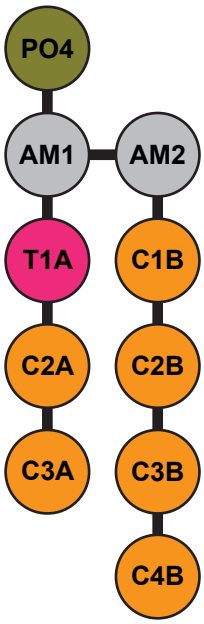


B

```
[ DPCP ]
[ atoms]
     P    PL    1.500  0
   O13   O2L   -0.820  1
   O14   O2L   -0.820  2
   O12   OHL   -0.680  3
   H12   HOL    0.340  4
   O11  OSLP   -0.620  5
   C1S   CTL2  -0.080  6
   H1S   HAL2   0.090  7
   H1T   HAL2   0.090  8
    NF   NHL   -0.700  9
   HNF     H    0.350 10
   C2S   CTL1   0.300 11
   H2S   HAL1   0.050 12
   C3S   CTO1   0.500 13
   H3S   HAL1  -0.110 14
    O3   OHL   -0.690 15
   HO3   HOL    0.300 16
   C4S   CEL3  -0.150 17
   H4S   HEL1   0.150 18
   C5S   CEL3  -0.150 19
   H5S   HEL1   0.150 20
   C6S   CTL2  -0.180 21
   H6S   HAL2   0.090 22
   H6T   HAL2   0.090 23
   C7S   CTL2  -0.180 24
   H7S   HAL2   0.090 25
   H7T   HAL2   0.090 26
   C8S   CTL2  -0.180 27
   H8S   HAL2   0.090 28
   H8T   HAL2   0.090 29
   C9S   CTL2  -0.180 30
   H9S   HAL2   0.090 31
   H9T   HAL2   0.090 32
  C10S   CTL2  -0.180 33
  H10S   HAL2   0.090 34
  H10T   HAL2   0.090 35
  C11S   CTL2  -0.180 36
  H11S   HAL2   0.090 37
  H11T   HAL2   0.090 38
  C12S   CTL2  -0.180 39
  H12S   HAL2   0.090 40
  H12T   HAL2   0.090 41
  C13S   CTL2  -0.180 42
  H13S   HAL2   0.090 43
  H13T   HAL2   0.090 44
  C14S   CTL2  -0.180 45
  H14S   HAL2   0.090 46
  H14T   HAL2   0.090 47
  C15S   CTL2  -0.180 48
  H15S   HAL2   0.090 49
  H15T   HAL2   0.090 50
  C16S   CTL2  -0.180 51
  H16S   HAL2   0.090 52
  H16T   HAL2   0.090 53
  C17S   CTL2  -0.180 54
  H17S   HAL2   0.090 55
  H17T   HAL2   0.090 56
  C18S   CTL3  -0.270 57
  H18S   HAL3   0.090 58
  H18T   HAL3   0.090 59
  H18U   HAL3   0.090 60
   C1F     C    0.550 61
    OF     O   -0.600 62
   C2F   CTL2  -0.070 63
   H2F   HAL2   0.060 64
   H2G   HAL2   0.060 65
   C3F   CTL2  -0.180 66
   H3F   HAL2   0.090 67
   H3G   HAL2   0.090 68
   C4F   CTL2  -0.180 69
   H4F   HAL2   0.090 70
   H4G   HAL2   0.090 71
   C5F   CTL2  -0.180 72
   H5F   HAL2   0.090 73
   H5G   HAL2   0.090 74
   C6F   CTL2  -0.180 75
   H6F   HAL2   0.090 76
   H6G   HAL2   0.090 77
   C7F   CTL2  -0.180 78
   H7F   HAL2   0.090 79
   H7G   HAL2   0.090 80
   C8F   CTL2  -0.180 81
   H8F   HAL2   0.090 82
   H8G   HAL2   0.090 83
   C9F   CTL2  -0.180 84
   H9F   HAL2   0.090 85
   H9G   HAL2   0.090 86
  C10F   CTL2  -0.180 87
  H10F   HAL2   0.09
  H10G   HAL2   0.09
  C11F   CTL2  -0.18
  H11F   HAL2   0.09
  H11G   HAL2   0.09
  C12F   CTL2  -0.18
  H12F   HAL2   0.09
  H12G   HAL2   0.09
  C13F   CTL2  -0.18
  H13F   HAL2   0.09
  H13G   HAL2   0.09
  C14F   CTL2  -0.18
  H14F   HAL2   0.09
  H14G   HAL2   0.09
  C15F   CTL2  -0.18
  H15F   HAL2   0.090 103
  H15G   HAL2   0.090 104
  C16F   CTL3  -0.270 105
  H16F   HAL3   0.090 106
  H16G   HAL3   0.090 107
  H16H   HAL3   0.090 108
[ bonds ]
   C18S   H18S
   C18S   H18T
   C18S   H18U
   C18S   C17S
   C17S   H17S
   C17S   H17T
   C17S   C16S
   C16S   H16S
   C16S   H16T
   C16S   C15S
   C15S   H15S
   C15S   H15T
   C15S   C14S
   C14S   H14S
   C14S   H14T
   C14S   C13S
   C13S   H13S
   C13S   H13T
   C13S   C12S
   C12S   H12S
   C12S   H12T
   C12S   C11S
   C11S   H11S
   C11S   H11T
   C11S   C10S
   C10S   H10S
   C10S   H10T
   C10S    C9S
    C9S    H9S
    C9S    H9T
    C9S    C8S
    C8S    H8S
    C8S    H8T
    C8S    C7S
    C7S    H7S
    C7S    H7T
    C7S    C6S
    C6S    H6S
    C6S    H6T
    C6S    C5S
    C5S    H5S
    C4S    H4S
    C3S    H3S
    C3S     O3
     O3    HO3
    C3S    C2S
    C2S    H2S
    C2S    C1S
    C1S    H1S
    C1S    H1T
    C1S    O11
   O11      P
   O12      P
   O13      P
   O14      P
   O12    H12
   C2S     NF
    NF    HNF
    NF    C1F
   C1F    C2F
   C2F    H2F
   C2F    H2G
   C2F    C3F
   C3F    H3F
   C3F    H3G
   C3F    C4F
   C4F    H4F
   C4F    H4G
   C4F    C5F
   C5F    H5F
   C5F    H5G
   C5F    C6F
   C6F    H6F
   C6F    H6G
   C6F    C7F
   C7F    H7F
   C7F    H7G
   C7F    C8F
   C8F    H8F
   C8F    H8G
   C8F    C9F
   C9F    H9F
   C9F    H9G
   C9F   C10F
  C10F   H10F
  C10F   H10G
  C10F   C11F
  C11F   H11F
  C11F   H11G
  C11F   C12F
  C12F   H12F
  C12F   H12G
  C12F   C13F
  C13F   H13F
  C13F   H13G
  C13F   C14F
  C14F   H14F
  C14F   H14G
  C14F   C15F
  C15F   H15F
  C15F   H15G
  C15F   C16F
  C16F   H16F
  C16F   H16G
  C16F   H16H
   C4S    C5S
   C1F     OF
[ impropers ]
   C1F    C2F    NF     OF
    NF    C1F   C2S    HNF
```

Figure C.3: CHARMM36 topology of 16:0-C1P. (A) Atoms and their connectivity (B) Parameters in GROMACS format.

A



B

```
[moleculetype]
; molname       nrexcl
  DPCP          1

[atoms]
; id    type    resnr   residu  atom    cgnr    charge
   1    Qa       1      DPCP    PO4      1      -1.0
   2    P1       1      DPCP    AM1      2      0
   3    P5       1      DPCP    AM2      3      0
   4    C3       1      DPCP    T1A      4      0
   5    C1       1      DPCP    C2A      5      0
   6    C1       1      DPCP    C3A      6      0
   7    C1       1      DPCP    C1B      7      0
   8    C1       1      DPCP    C2B      8      0
   9    C1       1      DPCP    C3B      9      0
  10    C1       1      DPCP    C4B     10      0

[bonds]
;  i  j          funct   length  force.c.
   1  2          1       0.47    1250
   2  3          1       0.37    1250
   2  4          1       0.47    1250
   4  5          1       0.47    1250
   5  6          1       0.47    1250
   3  7          1       0.47    1250
   7  8          1       0.47    1250
   8  9          1       0.47    1250
   9 10          1       0.47    1250

[angles]
;  i  j  k       funct   angle   force.c.
   1  2  3       2       120.0   25.0
   1  2  4       2       180.0   25.0
   2  4  5       2       180.0   45.0
   4  5  6       2       180.0   25.0
   3  7  8       2       180.0   25.0
   7  8  9       2       180.0   25.0
   8  9 10       2       180.0   25.0
```

Figure C.4: MARTINI topology of 16:0-C1P. (A) Illustration of beads and their connectivity (B) Parameters in GROMACS format.

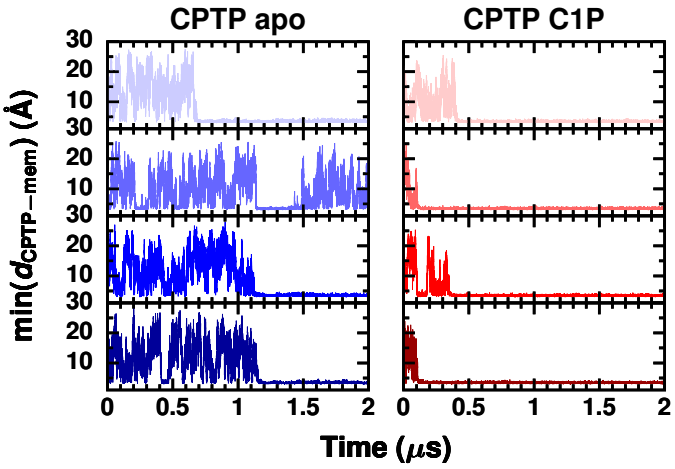


Figure C.5: Membrane binding occurs in all but one coarse-grained simulation. Minimum distance between CPTP and the membrane during coarse-grained simulations of membrane binding. Four simulation runs were performed for each form of CPTP. Within 2 μs, CPTP stably bound the membrane in all but one simulation of the apo form (right column, plot second from the top).

Appendix D

Appendix for "Development of Brownian Bridge Based Path Sampling Methods for a Nonequilibrium Transport Model"

Simulation parameters

To determine how the acceptance probabilities of shooting, fixed-end bridge, and variable-end bridge moves depend on the bias value, TPS simulations were performed for bias values ranging from -2 to 2 in increments of 0.25. For bridge moves, the acceptance probability was also obtained for bridge lengths ranging from $T_{\mathrm{B}} = 0.5$ to 5.0. The acceptance probabilities of forward and backward shooting moves of any length were each calculated from a TPS simulation in which 10^4 forward shooting moves and 10^4 backward shooting moves were performed. The acceptance probabilities of forward and backward shooting moves of a given length T_{B} were each calculated from a TPS simulation in which 10^4 forward shooting moves of length T_{B}, 10^4 forward shooting moves of any length, 10^4 backward shooting moves of length T_{B}, and 10^4 backward shooting moves of any length were performed. The acceptance probability of fixed-end bridge moves of a given length T_{B} was calculated from a TPS simulation in which 10^4 fixed-end bridge moves, 10^4 forward shooting moves, and 10^4 backward shooting moves were performed. The acceptance probability of variable-end bridge moves of a given length T_{B} was calculated from a TPS simulation in which 10^4 variable-end bridge moves, 10^4 forward shooting moves, and 10^4 backward shooting moves were performed.

To calculate large deviation functions, biased TPS simulations of the periodic ASEP were performed for bias values ranging from -2 to 2 in increments of 0.125. During TPS simulations using only forward and backward shooting moves (labeled "S"), a total of 6×10^7 moves were performed. During TPS simulations using fixed-end Brownian bridges in addition to shooting moves (labeled "fixed BB+S"), or both fixed-end and variable-end Brownian bridges in addition to shooting moves (labeled "variable BB+fixed BB+S"), a total of 10^7 sweeps were performed. A sweep consisted of $\lceil \frac{T}{T_B} \rceil$ bridge moves, one forward shooting

move, and one backward shooting move. During variable BB+fixed BB+S simulations, half of the bridges performed were variable-end bridges and the other half were fixed-end bridges. Different bridge lengths were used for different bias values, and these are tabulated in Table D.1.

Table D.1: Bridge lengths used during biased TPS simulations of the periodic ASEP.

Bias value (s or λ)	T_{B}
−2.000	0.5
−1.875	0.5
−1.750	0.5
−1.500	0.5
−1.375	0.5
−1.250	0.5
−1.125	0.5
−1.000	1.0
−0.875	1.0
−0.750	1.0
−0.625	2.0
−0.500	2.0
−0.375	2.0
−0.250	3.0
−0.125	3.0
0.000	3.0
0.125	3.0
0.250	3.0
0.375	2.0
0.500	2.0
0.625	2.0
0.750	2.0
0.875	2.5
1.000	2.5
1.125	1.0
1.250	1.0
1.375	1.0
1.500	1.0
1.625	1.0
1.750	1.0
1.875	1.0
2.000	1.0