Impact of increased membrane realism on conformational sampling of proteins

Austin T. Weigle[‡], *Matthew Carr*[a], *Diwakar Shukla*[*†,§,‖,#,⊥,¶,□,×]

[†]Department of Chemical & Biomolecular Engineering, [‡]Department of Chemistry, [§]Center for Biophysics and Computational Biology, [‖]Beckman Institute for Advanced Science and Technology, [#]Cancer Center at Illinois, [⊥]Center for Digital Agriculture, [¶]Department of Plant Biology, [□]National Center for Supercomputing Applications, and [×]NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States
[a]Independent Researcher. Email: matt.carr@me.com
Corresponding author: diwakar@illinois.edu

KEYWORDS. molecular dynamics, membrane complexity, membrane asymmetry, plant plasma membrane, membrane transporter function, Markov state models.

ABSTRACT

The realism and accuracy of lipid bilayer simulations through molecular dynamics (MD) is heavily dependent on the lipid composition. While the field is pushing towards implementing more heterogeneous and realistic membrane compositions, a lack of high-resolution lipidomic data prevents some membrane protein systems from being modeled with the highest level of realism. Given the additional diversity of real-world cellular membranes and protein-lipid interactions, it is still not fully understood how altering membrane complexity affects modeled membrane protein function or if it matters over long timescale simulations. This is especially true for organisms whose membrane environments have little to no computational study, such as the plant plasma membrane. Tackling these issues in tandem, a generalized, realistic, and asymmetric plant plasma with more than 10 different lipid species membrane is constructed herein. Classical MD simulations of pure membrane constructs were performed to evaluate how altering the compositional complexity of the membrane impacted the plant membrane properties. The *apo* form of a plant sugar transporter, OsSWEET2b, was inserted into membrane models where lipid diversity was calculated in either a size-dependent or -independent manner. An adaptive sampling simulation regime validated by Markov-state models was performed to capture the gating dynamics of OsSWEET2b in each of these membrane constructs. In comparison to previous OsSWEET2b simulations performed in a pure POPC bilayer, we confirm that simulations performed within a native-like membrane composition alter the stabilization of *apo* OsSWEET2b conformational states by ~1 kcal/mol. The free energy barriers of intermediate conformational states decrease when realistic membrane complexity is simplified, albeit roughly within sampling error, suggesting that protein-specific responses to membranes differ due to altered packing caused by compositional fluctuations. This work serves as a case study where a more realistic bilayer composition makes unbiased conformational sampling easier to achieve than with simplified bilayers.

INTRODUCTION

Biological membranes of natural origin are nonhomogeneous, multicomponent and asymmetric with a large diversity of lipid chemical structures. Resulting protein-lipid interactions are as diverse as the membrane bilayer composition itself and deserve further study. Although recent realistic membrane studies have emerged, the modeling of pure complex membranes and their interactions with membrane proteins is still in a characterization phase. Molecular dynamics (MD) simulations account for the positions of all atoms of a given system at any moment in time, providing an exceptional level of atomistic, label-free resolution for studying membrane properties and membrane protein function. As the popularity of MD has grown in parallel to the advent of more efficient computing capabilities, modern practices concerning the simulation of membrane systems have also advanced.[1,2] Specifically, increasing attention has been paid to appropriate modeling of membrane realism in the hopes of better capitulating native biological processes.[3] How "realistic" a membrane's simulated behavior can be is influenced by the selection of forcefields, system setup tools, and the constituent lipid diversity representing the modeled cellular environment. System preparation protocols and forcefield development behind the assembly and parameterization of membrane-containing MD systems have continuously improved over the last decade.[4,5] The refinement of these necessary components of membrane system setup have resulted in increased predictive power; for instance, coarse-grained models are widely accepted for accurately predicting lipid-protein interactions.[3,6,7] Still, uncertainty surrounds how realistic any

membrane composition should be for a given MD application, where the benefit of "better" lipid representation is not fully understood.

Simulation of realistic membranes is not yet a standard modeling convention for a number of reasons. Firstly, high-resolution lipidomic data is not available for all cell types of all organisms to reveal accurate chemical diversity nor leaflet asymmetry. This is mainly due to a lack of interest in resolving the complete "lipidomic picture" for a diversity of organisms' native membrane environments or because spectrometry technology was previously not advanced enough to fully resolve the identity of both lipid acyl chains when first studied. Given crystal structure availability of biodiverse origins without corresponding membrane characterizations available, MD practitioners resort to incorporating simplified membranes into their models. Up until the past decade, this practice of using simplified models has resulted in regularly unchallenged "default" modeling practices. For instance, single component mammalian membranes are commonly represented as POPC, and bacterial membranes are represented by POPE/POPG. Meanwhile, realistic model bilayers include at least six different lipid species.[3]

Secondly, the modeling of an asymmetric membrane is computationally difficult, due to both leaflets being constrained by the same box volume and pressure coupling. Compounded by a lack of data validating explicit leaflet-specific ratios of lipid species, determining how different lipid types should be allotted between the upper and lower leaflets often results from tedious trial and error. Some methods have been developed to make the construction of asymmetric membranes become more systematic,[8–12] but these approaches do not completely eradicate the lateral pressure differences inherent to leaflet asymmetry.[13] Additionally, the majority of realistic membranes vary the number of different lipid species between two to six different lipids,[3] minus a few exceptions.[14–19] Again, the question of how complex a realistic membrane should be remains unanswered. Taken together, the inherent challenges and increased computational costs involved in validating realistic membrane modeling explain why relatively few computational studies have incorporated such complex bilayers.

Computationally, it is now understood that protein-lipid interactions are recruited and protein-specific.[20] Experimentally, it has been seen that some protein conformations are inaccessible when reconstituted within a given membrane composition, a finding further consolidated by the ability of detergents to facilitate or deter membrane protein crystal structures from certain conformational states.[21] Performing long timescale simulations studies with realistic membranes can serve as a starting point for improved understanding. Such simulations present an opportunity to expand upon current computational knowledge of protein-lipid interactions and enhance insights to lipid-dependent membrane protein conformational ensembles.

Herein, a realistic plant plasma membrane construct is chosen as a model system to observe the effects of variable membrane complexity on pure membrane properties and membrane protein function. Currently there exists no standard protocol for conducting MD simulations on membrane systems modeling a plant cell; unlike for mammalian systems, whose bilayers are generally assumed to be POPC-rich,[22] plant membrane systems lack extensive study or even default compositions when simulated via MD. Some plant membrane compositions exist, though have mostly been used for a specific application, such as the thylakoid membrane for photosynthesis studies, and are not asymmetric.[23–27] While plant species present biological differences, an averaged, generalizable model would facilitate membrane MD study in plants. Furthermore, by establishing a membrane composition with valid asymmetry whose physical properties are robust against size-dependent changes, the effects of membrane realism on molecular simulation timescales can be evaluated. For comparing the effects of membrane composition on protein

function, two formulas for plant plasma membrane construction were then implemented: (1) a composition which was purely based off experimental ratios scaling with respect to the membrane size (called "Maximum Complexity"), and (2) a size-independent composition where only the top ten most abundant lipids species – those which had a literature calculated relative abundance greater than 1% - were retained (termed "Top10"). Depending on the system size, our all-atom plant plasma membranes are represented by the asymmetrical arrangement of 9, 10, 15, or 16 different lipid species. Such chemical diversity is rarely seen in all-atom MD approaches to membrane modeling. Model bilayers containing 10 or more different lipid species exhibited more or less equivalent physical properties.

To this end, the *apo* form of OsSWEET2b, a rice sugar transporter, was inserted into each membrane type. The complete transport cycle of *apo* and *holo* variants of OsSWEET2b had previously been elucidated in a pure POPC membrane,[28] allowing for a direct comparison to evaluate the differences brought by introduction of a native membrane composition. Between the two membrane constructs, OsSWEET2b *apo* transport cycles were capitulated across an aggregate of 61 and 91 μs using Maximum Complexity and Top10 membranes, respectively. Markov state models (MSMs) were used to remove sampling bias and interpret aggregate data from many short simulations into an ensemble view for characterizing longer timescale behaviors (i.e., OsSWEET2b gating dynamics). Thus, the use of MSMs helps in better estimating the free energy barriers between conformationally distinct gating states. While the establishment of a realistic plant plasma membrane is intended for enabling future general simulation applications in plant-based systems, our results concerning transporter dynamics are intended to serve as a reference for MD practitioners considering how complex to make their own membrane models when studying membrane protein function. Ultimately, we advocate for constructing the most realistic membranes when lipidomic data is not limiting. Even in cases where simplification of membrane complexity is unavoidable, we present readers with discussion of some practical considerations when modeling membrane protein function.

METHODS

Curation of literature for plant plasma membrane construction. An extensive literature search was performed to determine the lipid composition of a realistic plant plasma membrane representing a generalized plant cell. Curated data including relative percentages of sterol and lipid headgroup composition were averaged regardless of plant species and tissues sampled. Fatty acid chain distributions for each headgroup were estimated only from sources where mass spectrometry information was provided on both the *sn-1* and *sn-2* acyl chains.[29–43] To make this membrane as realistic as possible, we wanted to account for leaflet asymmetry. In terms of asymmetric lipid distribution across different leaflets within the plant plasma membrane, literature is limited; specifically, it has been reported that the distributions of lipid species across the plasma membrane upper:lower leaflets are 35:65 and 70:30 for phospholipids and sterols, respectively.[44] However, to the best of our knowledge, no such data is provided on the specific distributions of phospholipid headgroups and acyl chain types across the plant plasma membrane leaflets. We proceeded under the assumption that the headgroup and acyl chain asymmetries seen across the outer and inner membranes of chloroplasts and mitochondria would be similar to that for the plant plasma membrane. From our literature search,[45–58] we compiled the data of lipid species distributions across the outer and inner membranes of chloroplasts and plant mitochondria to determine relative

ratios for lipid species with different acyl chain lengths and degrees of unsaturation within each headgroup type.

Membrane and membrane protein system assembly. Pure membrane constructs were made with the designated compositions in the CHARMM-GUI Membrane Builder webtool.[10] Pure membrane constructs were sized at 64, 128, 256, and 512 total lipids. These system sizes were chosen to reflect the necessary membrane sizes needed for common MD applications (e.g., pure membrane property calculations, small-molecule membrane interaction studies, membrane protein structure-function capture). The largest system size which could be made using CHARMM-GUI was 512 lipids; the larger sterol content resulted in lipid tail-ring penetration issues which could not be overcome when using CHARMM-GUI at larger sizes. For OsSWEET2b simulations, select structures representing previously published inward-facing (IF), outward-facing (OF), occluded (OC), and hourglass (HG) conformations were used for system construction.[28] One of each of these four protein structures was then submitted to the Membrane Builder webtool for membrane building using the "Maximum Complexity" and "Top10" compositions at a membrane size of 256 total lipids. Addition of all water molecules for each system was then repacked using PACKMOL 18.169.[59]

System parameterization. CHARMM36 protein and lipid forcefields were applied to each system in the psfgen VMD-plugin,[60] where protein protonation states were determined using the PDB2PQR (PROPKA) server.[61] CHARMM-formatted topology and structure files for pure membrane systems over 100,000 atoms in size were generated using the TopoTools plugin to output final structure files in a format other than the pdb format, as pdb files cannot be written to number more than 99,999 atoms.[62] The default TIPS3P water model for CHARMM was used for all simulations. Potassium and chloride ions were added using the autoionize package to neutralize the systems. All systems prepared through VMD were then converted from CHARMM to AMBER format using the CHAMBER within the ParmEd toolkit to enable parallel computing capabilities in AMBER.[63] During conversion, all membrane protein systems were subject to hydrogen mass repartitioning (HMR) in order to allow faster timesteps for simulation, which would help with capturing slower kinetic processes within single trajectories.[64] Pure membrane simulations were not performed using HMR for simplifying characterization work.[65]

Simulation details – Pure membrane simulations. All simulations were performed by using classical molecular dynamics (MD) in the AMBER18 software package.[66] For pure membrane simulations, minimization was performed for 50000 cycles, where steepest descent was used for the first 5000 and then conjugate gradient for the remaining 45000 cycles. No positional restraints were employed. Each system was then heated for 2 ns from 0 to 10 K in an NVT ensemble. Each system was then heated at 10K in an NPT ensemble for 2 ns. NPT heating was then increased to 300K for 2 ns. All prior heating stages were run using the `sander` CPU simulation code. In order to ease box dimension changes prior to using the `pmemd.cuda` GPU simulation code, a 5 ns NPT hold was performed at 300K using the `pmemd` CPU simulation code. Following the 5ns hold, all systems then underwent equilibration and production runs. Pure membrane simulations used semi-isotropic pressure scaling and were run with a 2fs timestep. A Berendsen thermostat and barostat were used throughout heating steps, where the pressure was maintained at 1 bar.[67] A Langevin thermostat and Monte Carlo barostat were then implemented for temperature and pressure maintenance in all equilibration runs, respectively.[68,69] A Langevin collision frequency of 2 ps^{-1}

was used, while default settings for the Monte Carlo barostat were kept. Equilibration and production runs proceeded using the `pmemd.cuda` GPU simulation code except for the 64-sized lipid membranes for reasons discussed in the main text, which were run using `sander`. The SHAKE algorithm was applied to all stages of simulation initialization except for minimization,[70] while the Particle Mesh Ewald method was used for treating long-range electrostatics at a 12 Å cutoff in correspondence with the CHARMM36 lipid forcefields.[71,72] Pure membrane simulations were equilibrated for 1 μs and in triplicate, where the frame save rate was set to every 50000 frames. Both Maximum Complexity and Top10 models appeared to equilibrate within 400 ns.

Simulation details – Membrane protein simulations. Membrane protein simulations were minimized in a similar fashion to the pure membrane simulations, although positional restraints at strengths of 5 kcal/mol were applied to all backbone atoms. Because of HMR, all membrane protein simulations were run with a 4fs timestep. Identical heating protocols were used for membrane protein simulations except for the following changes. Anisotropic pressure scaling was implemented throughout and a Berendsen thermostat and barostat were used for equilibration and production runs. Despite their widespread accessibility in MD software, it has not been recommended to use the Berendsen thermostat or barostat for production runs unless comparing results to already published work which had used them.[73] Accordingly, we used a Berendsen thermostat and barostat to compare the effects of increased membrane complexity on the gating dynamics of OsSWEET2b.[28] All atom restraints were removed from membrane protein simulations during the 5 ns holding period. OsSWEET2b simulations were equilibrated for 50 ns. Based on expected timescales, OsSWEET2b production runs ran for 25 μs, where only coordinates and not velocities were read from restart files. Velocities are purposefully reinitialized in adaptive sampling in order to achieve greater randomized seeding, and in turn, more efficient sampling over the desired conformational space. Frame recording rates were set for every 50000 steps.

Adaptive sampling protocols. Pure membrane simulations were run using continuous MD while membrane protein simulations were run using adaptive sampling-based MD. In adaptive sampling regimes, any state recorded along a trajectory can be used as the starting point to seed a new trajectory.[74–77] A k-means clustering of trajectory data based on some reaction coordinate allows for the potential states used for seeding new trajectories to be representative of the entire round of data collected.[78] For production trajectories run from each starting OsSWEET2b conformer within each membrane construct, the top 50 structures from the least populated clusters were used to reseed new adaptive rounds of simulation. For *apo* simulations, distances between gating residues (intracellular gating residues Phe42 and Phe164; extracellular gating residues Arg69 and Asp189) were used as the reaction coordinates for clustering. Systematic adaptive sampling was performed until the gating landscape for OsSWEET2b gating dynamics was filled or completed. All distance measurements between atoms were conducted using MDTraj 1.9.3.[79] Clustered states and trajectories were visualized using VMD 1.9.3, while aggregate data collected were visualized as projected landscapes using Matplotlib 3.2.0.[80]

Pure membrane simulation analyses – validation of physicality among constructs. The physical correctness of asymmetric membrane simulations can be assessed in two different ways, where either approach attempts to evaluate to what extent lateral pressure distributions differ between leaflets. If the difference in lateral pressure experienced by either leaflet is too great, then the pressure experienced by one leaflet will dominate the dynamics experienced by the bilayer as a

whole. One method to clarifying satisfactory lateral pressure differences between leaflets is to compare the area-per-lipid (APL) mismatch between the two leaflets of the asymmetric construct.[11] As long as APL mismatch is between 5 and 10%, the impacts of asymmetry on bilayer properties are expected to be mild, where the best-case scenario is to have an APL mismatch between 0 and 5%. Another approach is to use the upper and lower leaflet compositions of the desired asymmetric leaflet membrane to build two symmetric membrane constructs.[22] So long as the membrane-spanning box dimensions of the symmetric systems - in this case, the XY dimensions – demonstrate a similar box area difference from the asymmetric simulations, the defined leaflet asymmetry can be deemed acceptable. Symmetric membranes were built for each membrane size tested, prepared as described above, and with production runs totaling for 50 ns. The box dimensions from these trial simulations were then compared with the first 50 ns of asymmetric membrane simulations.

Pure membrane simulation analyses – bulk and headgroup-specific physical property measurements. Pure membrane simulations were analyzed for bulk physical properties using two software: Membrainy and SuAVE.[81,82] The following physical properties were calculated using these software: area-per-lipid (APL), entropy, gel fluidity, membrane thickness, membrane curvature order parameter, membrane curvature angle, lipid S_{CD} order parameters, headgroup axial tilt angle, and headgroup lateral tilt angle. Electron density calculations were performed using the CPPTRAJ module found in AMBER.[83] Membrainy input files were generated by converting the AMBER-formatted trajectory and topology files to GROMACS formats using GROMACS 2020.[84] SuAVE input coordinates were generated from trajectories using VMD and MDTraj.

Membrainy software updates. Membrainy was updated to allow new lipids to be parameterized via a user-created external lipid library (i.e., accept custom lipid forcefields).[81] The leaflet detection method was also modified for improved efficiency and to allow for a vector-based solution to better account for leaflet undulation when discriminating between leaflets for calculations. In this leaflet-identification approach, the vectors are created by using the phosphorous atom and the last carbon atom of each lipid tail. For lipids with two tails, the average vector orientation was taken. Sterols are treated as having only one tail. Vectors that point up indicate an individual lipid getting placed in the upper leaflet, with vectors pointing down indicating lower leaflet placement. An updated version, Membrainy 2020, capable of performing all calculations done in this manuscript is available online (http://www.membrainy.net/) and is compatible with Gromacs-2020.[84]

OsSWEET2b Markov state model (MSM) construction and validation All stages of MSM construction were performed using PyEMMA software.[85] Feature selection for OsSWEET2b MSM construction followed the same protocol as employed in previous studies on membrane proteins.[28,86–88] Maintaining the same choice of input features enables a direct comparison to the previous OsSWEET2b work done in a pure POPC bilayer. An iterative process of hyperparameter selection for optimization of the maximum VAMP score was performed.[89] The number of clusters tested for data clustering varied from 100 to 1000 clusters, increasing by increments of 50. The number of time-Independent Component (tIC) were varied from 1 to 9. The time-independent component analysis (tICA) lag time varied from 5, 6, 7, 10 and 12 ns. The optimal combination of hyperparameters for OsSWEET2b Scale MSM construction was 9 tICs, 1000 clusters, and a tICA lag time of 12 ns. For constructing MSM of OsSWEET2b in the Top10 membrane model, the

optimal combination of hyperparameters was 9 tICs, 650 clusters, and a tICA lag time of 7 ns. Provided the nine residue pair distances used for featurization, the slowest timescale process for OsSWEET2b dynamics in either membrane converged at an MSM lag time of 9 ns. Accordingly, final MSMs were constructed using the optimal hyperparameters with an MSM lag time of 9 ns.

Adaptive sampling error analysis. The data ensemble containing nine residue pair distances used for MSM featurization were replicated into ten slices. Each of these slices were comprised of a random selection of 80% of the overall data. MSMs were constructed off each of these data subsets using the same combination of hyperparameters as the entire dataset. With the gating dynamics represented as a 2D histogram plotting extracellular versus intracellular gating distances, the average error for each bin among each of the data slices was calculated as previously shown.[28]
OsSWEET2b annular shell analyses. The linear interaction energy between OsSWEET2b residues and lipid species was calculated in CPPTRAJ. Snapshots were rendered using Chimera 1.14.[90]

RESULTS AND DISCUSSION

Establishing a realistic plant plasma membrane composition
Based on the literature survey, the plant plasma membrane presents an ~48:52 split between sterols and phospholipids (PL) across the entire bilayer. However, the upper leaflet has a sterol:PL split of 65:35.[44] Meanwhile, the lower leaflet exhibits a sterol:PL split of 30:70,[44] which is consistent with the "positive-inside" rule, as more electropositive phospholipid headgroups are present along the intracellular leaflet to counter cytosolic electronegativity (Figure 1).[91–93] In order of abundance, the plant plasma membrane is comprised of SITO (~32%), STIG (~16%), PLPE (~13%), PLPC (~12%), DLiPC and DLiPE (~5%), LLPC and PLPG (~4%), LLPE (~3%), POPC (~1%), and DSPC, DOPC, POPE, SLPE, POPG and DPPG (< 1%). While PLs represent the aggregate bulk of the plant plasma membrane, the most abundant individual lipid species are sterols, with SITO and STIG at ~32% and ~16%, respectively (Figure 2).

Such sterol abundance has been seen in other membrane compositions.[94] The most abundant individual PLs are PLPE (~14%) and PLPC (~12%) (Figure 2). Still, PC headgroup lipids account for the majority of PLs, followed by PE and PG. Regardless of leaflet, PLs predominantly exhibit 16:0/18:2 unsaturation (~60%), followed by 18:2/18:2 (~18%). Taken together, ~91% of PLs exhibit 18:2 unsaturation along one of their acyl chains. This is a far departure from simple mammalian membrane models which are often simulated using just POPC (16:0/18:1), serving as a reminder that "default" membrane practices for one organism type cannot be assumed to be appropriate for modeling another organism. For comparison, POPC represents ~1% of the plant plasma membrane composition, although the type of 18:1 unsaturation present in POPC-like lipids accounts for ~5% of all lipid unsaturation (Figure 1). The exact ratios for Maximum Complexity lipid bilayers for each size constructed (64, 128, 256, and 512 total lipids) are supplied in Tables S1-S5.

Reducing the Maximum Complexity membranes to that of a Top10 composition, where any lipid species present under 1% relative abundance is removed from consideration, causes the relative abundances of the remaining lipid species (SITO, STIG, PLPE, PLPC, DLiPC, DLiPE, LLPC, PLPG, LLPE, and POPC) to fluctuate slightly. Removal of just five lipid species with < 1% relative

abundance causes the remaining ten species to be represented in greater abundance, as the MD system membrane size is the same but now with fewer types of lipids. Accordingly, this exacerbates leaflet-specific asymmetries as lipid counts become redistributed (i.e., the presence of slightly less sterols in the upper leaflet, or slightly more PLs in the lower leaflet, than would typically occur in nature). For the sterols, the disparity in relative abundance between the two leaflets decreases, where more sterols are redistributed to the lower leaflet. The reallocation of trace lipid counts to other lipids within their same headgroups results in relative increases for all PLs except for LLPE and POPC. PLPE content is increased by ~3%, making it now the second most abundant lipid species above STIG (Figure S1).

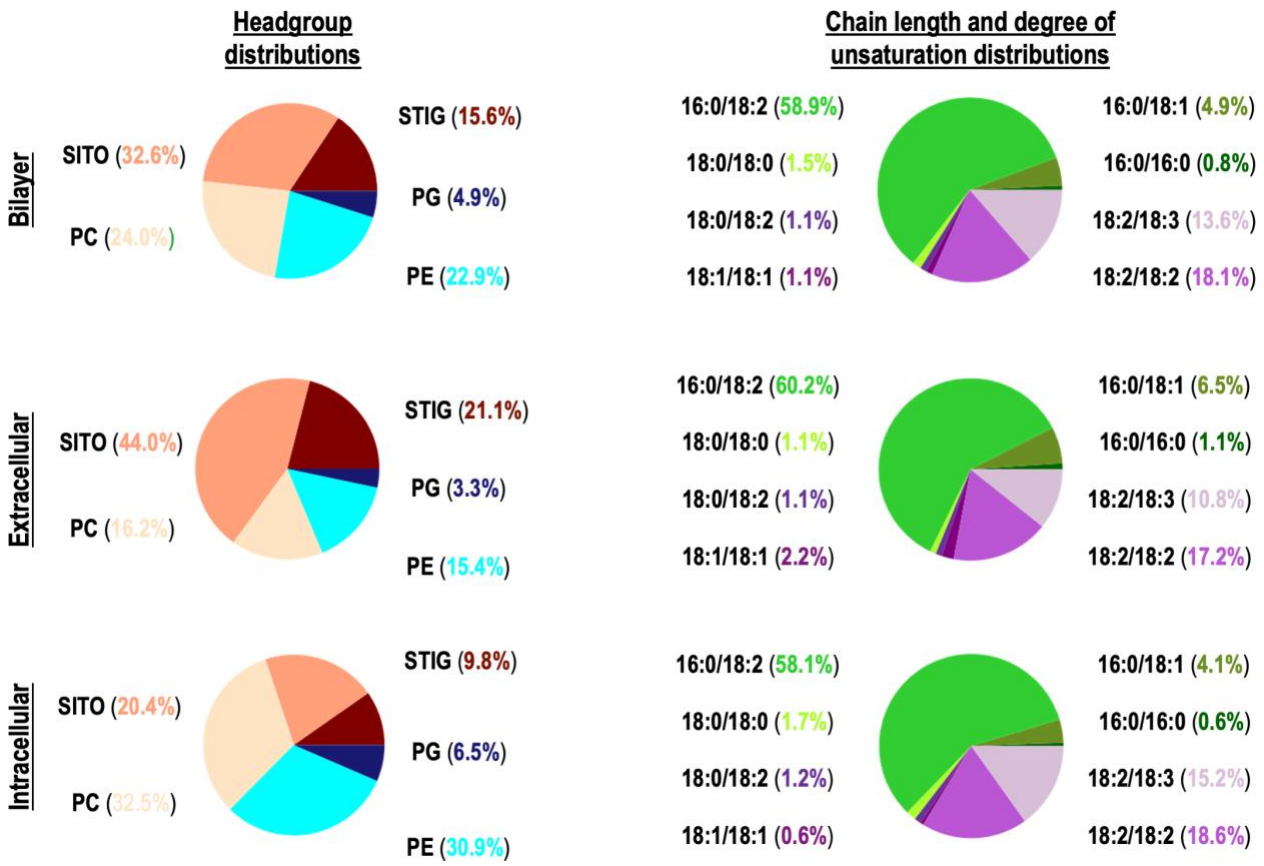

Figure 1. Distribution of different lipid headgroups and degrees of tail unsaturation in model membranes constructed following a Maximum Complexity recipe.

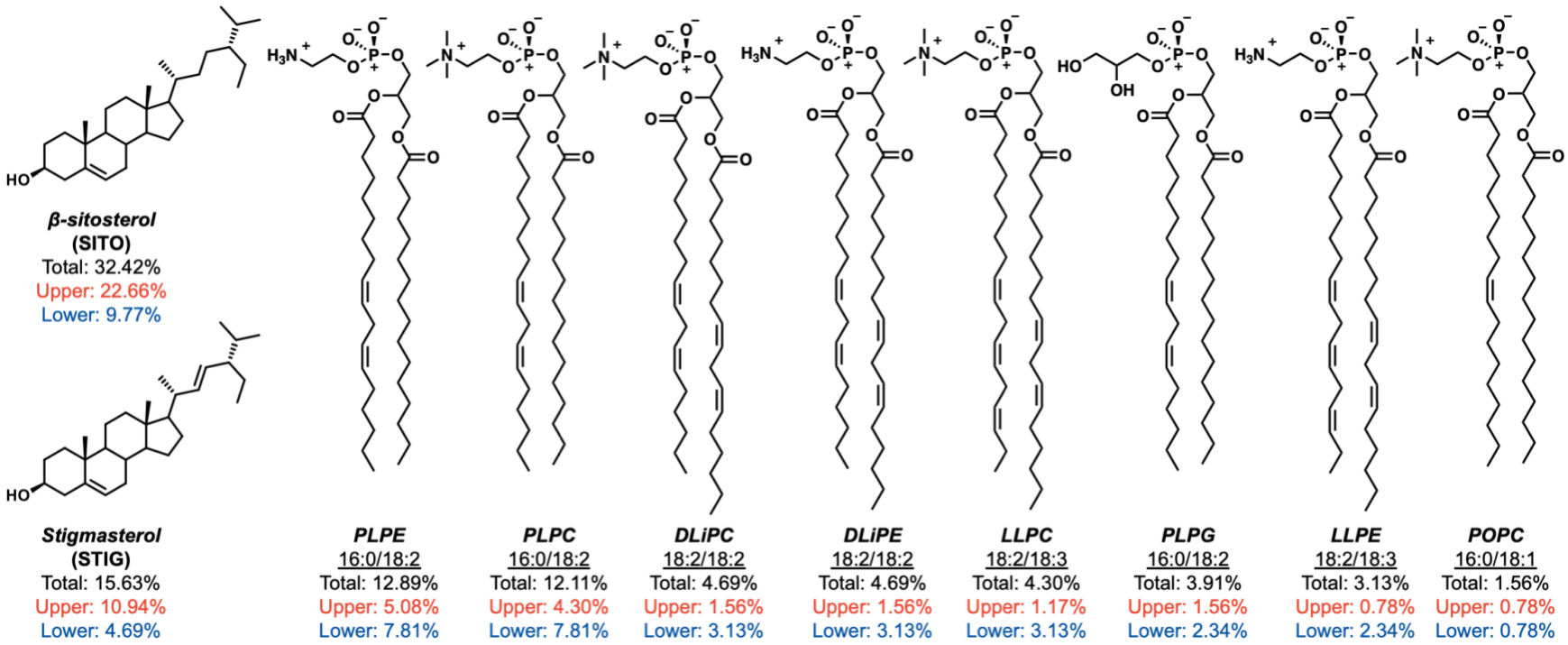

Figure 2. Leaflet-specific percentage of different lipids present in a Maximum Complexity membrane construct. Ratios presented are based off those seen in the 256-lipid construct.

Pressure-related analyses indicate membrane physicality

To evaluate the generalizability of this informatically-driven Maximum Complexity plant plasma membrane composition and its simplified Top10 version, common membrane physical properties (box area, area per lipid, entropy, gel fluidity, thickness, curvature, acyl chain order parameters, headgroup orientation, electron density) were calculated.

Asymmetric membranes inherently experience differing lateral pressure profiles due to the different number of lipids present in each leaflet. Numerous developments by the field have been made to mitigate the damage to membrane physicality as a result of inherent differences between the leaflets of an asymmetric membrane. The generation of optimized and adjusted APL profiles for a given lipid species and zeroed lateral tension settings are some ways to help.[8–12] Ultimately, asymmetric membrane physicality (i.e., how distinct lateral pressure profiles are for each leaflet) needs to be verified. Short simulations using symmetric bilayers for each leaflet of each membrane construct can be used to retrieve an MD box size specific to that leaflet pairing. For example, the upper leaflet from a Maximum Complexity construct was used to create a symmetric bilayer. This is critical for validating asymmetric bilayers, as the individual leaflets must have stable individual pressure constraints. The box size for a membrane system is a valuable metric for evaluating membrane construct stability. Given that the membrane bilayer spans the entire MD box in the X and Y dimensions, box size fluctuations give an estimate as to how the bilayer responds to the barostat. So long as the calculated box sizes between the symmetric leaflets are not too different from each other and the final asymmetric construct, then it can be viewed as reasonable for the specified leaflet pairing to comprise a single bilayer.[22] Another validation approach is to compare the percentage difference in APL between the two leaflets of an asymmetric bilayer.[11] Similar to box size, APL serves as another measure to evaluate how the bilayer is responding to pressure. When APL converges, a membrane is often said to have equilibrated. According to this approach developed by the Klauda and Im groups in 2015, the ideal APL mismatch between leaflets of an asymmetric membrane should be between 0 and 5%.[11] Yet membranes demonstrating 5 to 10% mismatch only demonstrate mild effects from asymmetry, the extents of which are less noticeable in membranes with more unsaturated tails.

A general rule of thumb is to assume that membranes containing more than one type of lipid species requires 100 ns of equilibration for each of the different lipid species present.[95] From APL analyses it appears that each of the membrane constructs equilibrated after 400 ns of simulation, although the APL for each 64-sized bilayer leaflet appear to be distinct from the other membranes' (Figure 3). APL mismatch analysis shows the 128- and 256-sized Maximum Complexity membrane constructs with 6.06% mismatch between leaflets (Table 1). Meanwhile, the Maximum Complexity constructs containing 64 and 512 lipids have 11.76% and 7.52% leaflet APL mismatch, respectively (Table 1). When comparing the symmetric constructs for each asymmetric leaflet to the asymmetric bilayer, the leaflet area differences are all below 5%, except for the 64-lipid lower leaflet construct (Table 2).

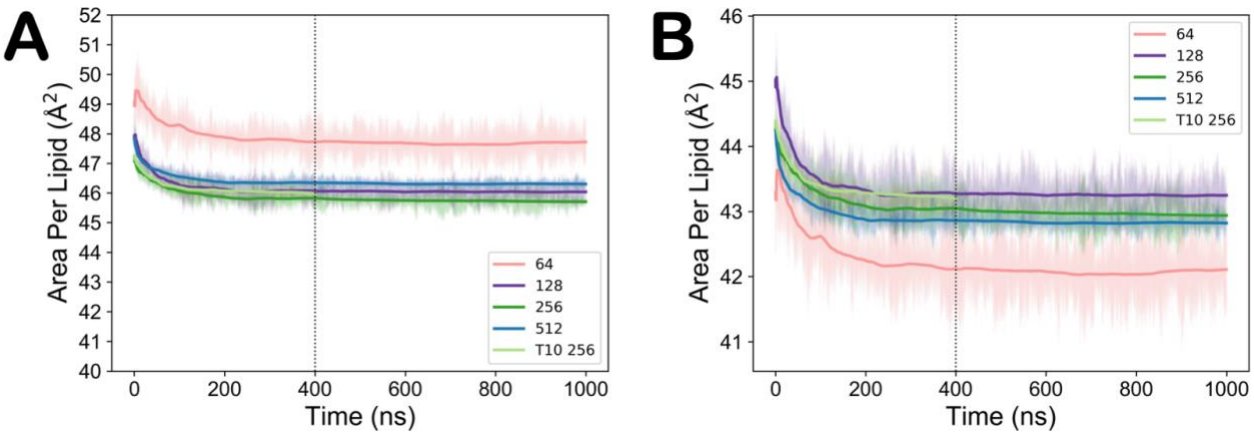

Figure 3. Area per lipid (APL) values for the (A) lower and (B) upper leaflet of different sized membrane constructs.

Table 1. APL and APL % mismatch between leaflets of membrane constructs

Systems	AVG APL Lower Leaflet	AVG APL Upper Leaflet	AVG APL % mismatch	SEM APL % mismatch
Maximum Complexity 64	47.72 ± 0.26	42.11 ± 0.23	11.76	6.84E-14
Maximum Complexity 128	46.04 ± 0.14	43.25 ± 0.13	6.06	4.23E-14
Maximum Complexity 256	45.71 ± 0.17	42.92 ± 0.16	6.06	1.29E-13
Top10 256	45.98 ± 0.03	42.82 ± 0.03	6.06	4.62E-14
Maximum Complexity 512	46.30 ± 0.04	42.82 ± 0.04	7.52	2.92E-13

Table 2. Leaflet-specific pressure tests. Comparing box area difference between symmetric leaflet constructs and final asymmetric bilayers.

Systems	Symmetric System Data	
	Lower Leaflet Area Difference (%)	Upper Leaflet Area Difference (%)
Maximum Complexity 64	5.86	2.59
Maximum Complexity 128	5.03	4.7
Maximum Complexity 256	4.52	4.14
Top10 256	3.28	0.71
Maximum Complexity 512	4.53	3.17

The leaflet area deviations from the asymmetric bilayer are most similar between symmetric Maximum Complexity constructs containing 128 and 256 lipids, which is in agreement with the APL analyses. Thus, when generating plant plasma membranes for simulation, asymmetric bilayers between 128 and 256 lipids in size would be expected to be least affected by lateral pressure differences caused by asymmetry, followed by the 512-sized membrane. The APL mismatch for the 64-sized bilayer suggests that lateral pressure differences experienced by either leaflet would compromise simulation integrity, likely leading to unrealistic or unphysical

representation of the membrane. No 512-sized Top10 construct which arrived at the equilibration stage of simulation made from the given lipid could pass either of these pressure tests because of leaflet-specific asymmetry being too different. Thus, the asymmetry recipe applied to this simplified Top10 construct was only valid for a 256-sized bilayer.

Physical property analyses demonstrate robust compositional recipes

The remaining bulk property analyses included here are APL, entropy, gel fluidity, thickness, and curvature. Mixing entropy indicates the probability that a lipid of one type will have a chemically different lipid as its neighbor. The higher the lipid type mixing entropy, the more mixed the membrane.[96] For each of the leaflets, entropy analyses follow a size-dependent trend. Given that compositional complexity scales with membrane size when following the Maximum Complexity recipe, it is expected that larger membranes would have higher entropy or mixing probabilities (Figure 4). However, the 256-sized Top10 construct violates this relationship, where its upper leaflet has a higher entropy than the 512-sized membrane. Overall upper leaflet entropy values appear to be more similar between each of the membrane constructs, suggesting that their shared sterol-rich profiles may underlie similarities in lipid organization. Like upper leaflet entropy, 256-sized Top10 construct bilayer entropy matches that for the 512-sized Maximum Complexity construct.

Gel fluidity measures the linearity of the lipid tails, which can provide insights into the bulk ordering of the membrane.[81] Here, all membrane constructs demonstrate the same bilayer properties, but the fluidity of each of the leaflets differs. For either leaflet, the 64-sized Maximum Complexity and 256-sized Top10 constructs appear to have distinct distributions of gel fluidity percentage from the remaining construct sizes, where the Top10 construct's lower leaflet is less and its upper leaflet is more fluid than the Maximum Complexity constructs. Membrane thickness calculations follow similar, albeit less pronounced per leaflet, trends as gel fluidity (Figure 4). Membrane curvature calculations lastly show that regardless of size, all Maximum Complexity membranes demonstrate the same curvature distribution, peaking at a curvature angle of 18º (indicative of a planar surface).[82] Meanwhile, the Top10 membrane construct differs, where the lower leaflet is slightly more planar and the upper leaflet more curved (Figure 4).

Headgroup-specific properties measured include acyl chain order parameters, headgroup orientation, and electron density. Surprisingly, when comparing properties between lipid species which are shared between each of the membrane constructs, all of these results are nearly equivalent or within error (Figures S2-S13). It appears that headgroup-specific properties are not as influenced by the relative system complexity nor size. On the other hand, bulk membrane properties differ only slightly between a few membrane constructs. The 256-sized Top10 construct appears to have altered distributions from the 256-sized Maximum Complexity membrane likely because of composition-dependent effects. Again, the major differences seen in the 256-sized Top10 construct were the relative redistribution of sterols to the lower leaflet and the amplification of 18:2 lipid species across the bilayer. It is possible that slight alteration of relative lipid abundances could impact the initial packing schemes within CHARMM-GUI. As for the 64-sized Maximum Complexity construct, differences in bulk properties can be accredited to a number of factors. Pressure tests (Tables 1 and 2) indicate that the asymmetry seen in such a small membrane may not be physically valid. This discrepancy may be exacerbated by an analytical effect from the MD engine. The box-size for 64-sized Maximum Complexity constructs approached the minimum nonbonded cutoff specifications required to perform simulations using Amber software. Thus, 64-sized membranes could only be simulated using the more sensitive CPU-based code, `sander`,

which may amplify any size-dependent effects when compared to the systems run using the GPU-based code `pmemd.cuda`.[66] Aside from the abovementioned constructs, Maximum Complexity constructs sized 128 to 512 lipids appear robust enough to be used for any plant membrane MD application and demonstrate equivalent properties.

Effect of simplified complexity on membrane transporter dynamics

OsSWEET2b, a vacuolar sugar transporter, was selected for modeling the effects of reducing membrane complexity on membrane protein function. Although OsSWEET2b is not found in the plant plasma membrane, it is suitable to study for two demonstrative purposes. Firstly, the *apo* transport cycle for this protein has previously been elucidated in a symmetric and homogenous POPC membrane.[28] Simulation of OsSWEET2b enables direct comparison of the influence of using a complex membrane construct versus an overly simplified one. Secondly, detailed lipidomic data providing both acyl chain identities for constituent lipids of a tonoplast (vacuolar) membrane does not exist. Indeed, this lack of high-resolution lipidomic data concerning relative lipid species abundances is a fundamental driver causing MD practitioners to resort to very simple membranes. While exact structure elucidation is missing, plant tonoplast membrane composition has been measured at low resolution with the measurements of headgroup and single acyl chain abundance.[43,50,53,97–107] Overall, the plant tonoplast offers similar trends in relative headgroup and individual acyl chain representation to suggest that the same lipids would be represented as shown in these plasma membrane constructs. Thus, based on rationale and experimental evidence suggesting enough similarity between tonoplast and plasma membrane compositions, we proceed with simulating the transport cycles of OsSWEET2b.

To evaluate the effect of protein-lipid interactions, it is advisable to generate multiple versions of the same membrane composition with different initial packings.[108] By using multiple membrane starting structures to generate different membranes with the same overall composition, a greater diversity of potential protein-lipid interactions can be observed, thereby providing greater perspective into the effects of a complex membrane composition.[108] Such a project design also makes the capture of OsSWEET2b gating dynamics more efficient as it enables greater initial starting structures for seeding adaptive sampling-based trajectories.[74,109]

OsSWEET2b gating dynamics cover roughly the same conformational spread of intracellular and extracellular gating distances when simulated with the either the Maximum Complexity or Top10 membrane constructs. Both landscapes present reasonable sampling of extracellular gating distances ranging between ~4-20 Å and intracellular gating distance ranges of ~5-25 Å (Figure 5). Small deviations between the energetic accessibility of intermediate states of each landscape exist, although the majority of them fall within the 0.2 kcal/mol sampling error (Figure S14). For example, the extent of stabilization for IF states differs between the constructs, where the Top10 construct demonstrates an energetic cost of ~0.6-0.8 kcal/mol versus ~1.2-1.4 kcal/mol seen in the Maximum Complexity construct (state 2 of each landscape), which falls within error (Figure 5). Energetic costs underlying some transitions exceed sampling error. The difference in transition energy from Maximum Complexity states 2→4 and the corresponding Top10 path between states 2→3 exceeds the error limit, exhibiting free energy costs of 2.8 ± 0.2 and 3.5 ± 0.3 kcal/mol, respectively (Figure 5). A similar analysis on the most probable path between Maximum Complexity states 3→5 and the corresponding Top10 states show respective free energy barriers of 1.8 ± 0.2 versus 2.4 ± 0.3 kcal/mol.

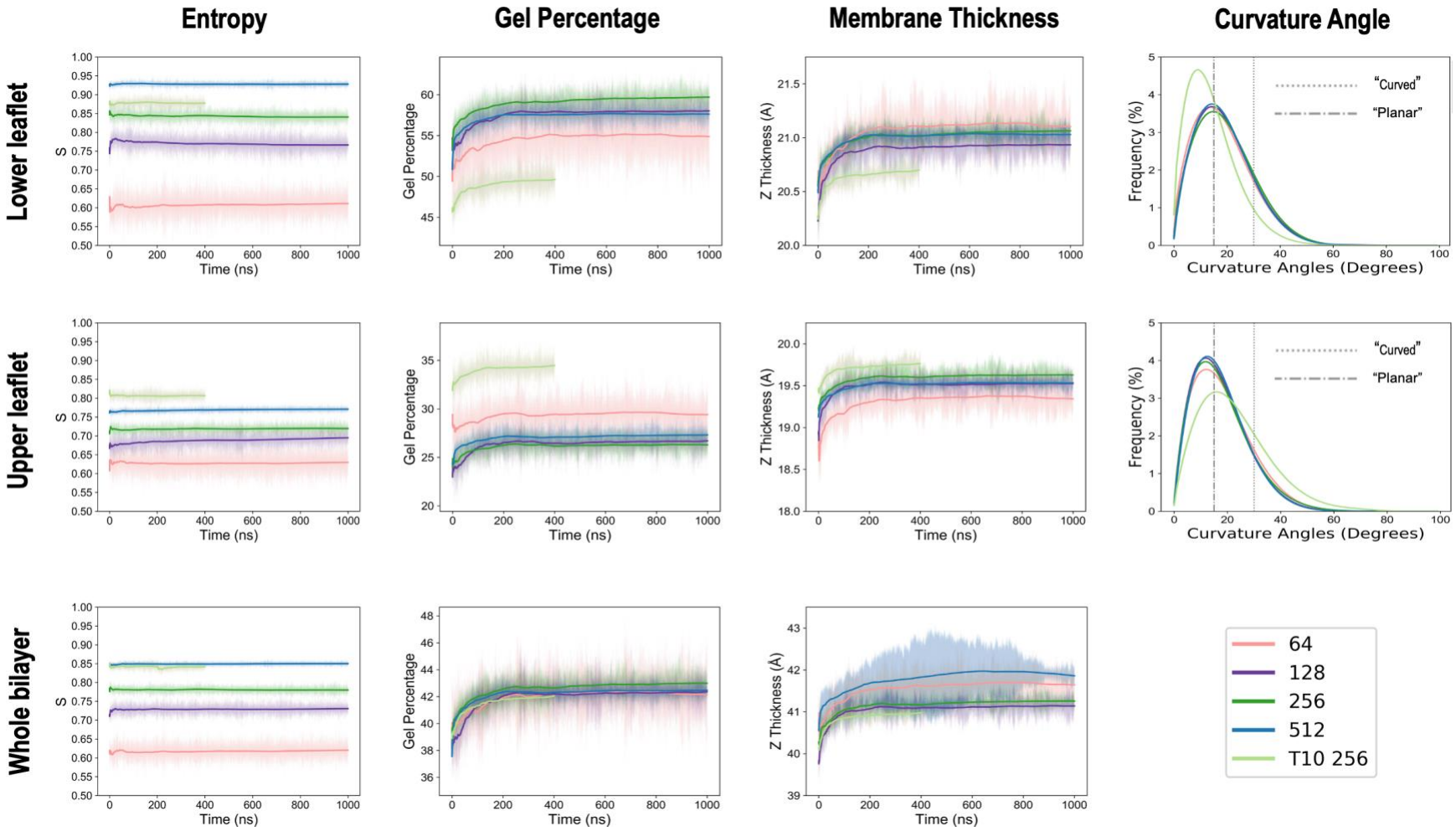

Figure 4. Membrane construct physical property analyses demonstrate robustness of compositional recipe.

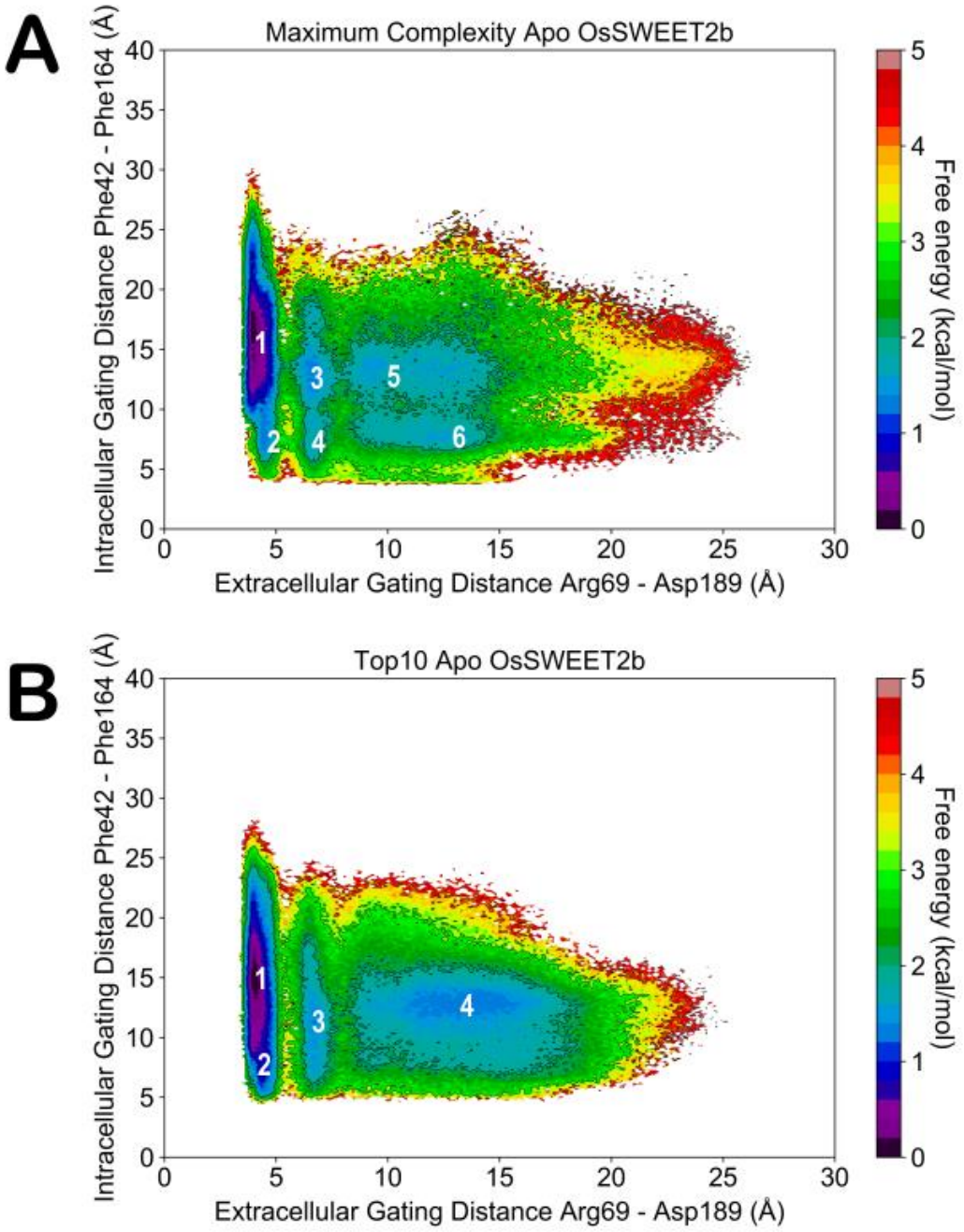

Figure 5. MSM-weighted OsSWEET2b gating dynamics. (A) apo OsSWEET2b simulations done in the 256 Maximum Complexity membrane construct. (B) apo OsSWEET2b simulations done in the 256 Top10 Complexity membrane construct. Key conformational states are numerically labeled.

Aside from more stringent quantitative analyses, visual inspection of the gating landscape shows how the Maximum Complexity membrane construct is able to better resolve microstates seen with respect to IF, extended IF-OC, and extended OF-OC gating states (Figure 5).

OsSWEET2b Top10 simulations indicate more coarse and less comparable lumping of OF states, where sampling of a more OF and less OC state is less likely when compared to the Maximum Complexity gating landscape. This coarser representation of Top10 microstates also reflects the slightly higher barriers between intermediate OC states and fully IF or OF states. When compared to the Maximum Complexity construct, the simplification of lipid composition provided by the Top10 construct does provide a smoothening effect to the one-dimensional projections of either the intracellular and extracellular gating distances, particularly when it comes to resolving IF-OC and OF-OC states (Figure S15). The thermodynamic differences between Maximum Complexity and Top10 conformational landscapes are small (~ 1 kcal/mol) but these differences based on lipid composition could lead to significant changes in kinetics and impact conformational sampling.

Exploring specific protein-lipid interactions with varying membrane complexity

Regardless of which realistic membrane composition was used, *apo* simulations performed using a native versus a purely POPC composition demonstrate overall lower energetic costs for transport (Figure S16). In fact, the *apo* gating dynamics landscape for simulations done in a realistic membrane composition more closely resemble the previously published *holo* gating dynamics performed in a purely POPC membrane composition.[28] Between the ~1 kcal/mol lower barrier for all transitions and the protein spending less time in higher energy intermediate states, the extent of stabilization provided by a native membrane environment mimics the energetic stability provided by substrate introduction, suggesting that specific interactions between OsSWEET2b and annular shell lipids could be a contributing factor.

An annular shell analysis, which differentiates the lipids in the immediate vicinity of the protein compared to the bulk, was performed on OsSWEET2b states from a higher energy transitionary region containing extracellular gating distances between 5.25-5.75 Å and intracellular gating distances between 13.25-13.75 Å from simulations involving either membrane construct. This region covers a transition between an IF and extended IF-OC conformational state and is one where minor discrepancy in free energy differences between each of the membranes exists. Each of the states were visualized to see which lipids presented themselves within 7 Å of the OsSWEET2b transmembrane-spanning surface. The annular shell compositions provided by the four different system setups for each membrane construct mainly differ in leaflet-dependent manners. Upper-leaflet sterol and PE content, as well as 16:0, 18:1, and 18:2 unsaturation lipid tails differ between annular shells formed in Maximum Complexity and Top10 constructs (Figure S17). Meanwhile the lower leaflet annular shell composition is more or less equivalent minus the presence of 18:0 and 18:1 unsaturated lipid tails (Figure S17). Analyzing snapshots revealed the presence of a putative lipid binding site with interaction energies largely dominated by Tyr168. Defining a "full binder" as a lipid species whose scaffold fits entirely into the cleft formed by residues Ile17, Phe18, Leu24, Val27, Thr28, Tyr168, Leu171, and Leu175, Maximum Complexity membrane setups never pose a sterol as the "full binder" while Top10 membrane setups can (Table S6**,** Figure S18). Top10 membrane setups can additionally pose sterols as partial binders which are able to interact with Tyr168; this was not seen in any of the Maximum Complexity setups, each of which had more stable interaction energies with a PL by at most ~1.8 kcal/mol (Table S6). It is possible that even slight modification in relative lipid abundance results in altered packing of specific lipid species. For example, redistribution of sterols to the lower leaflet in Top10 system setups could alter the ordering of PLs with large polar headgroups, like PC.[110–113] Small modifications in bulk lipid ordering could present changes to which conformations of what lipid species would be compatible for annular shell constitution. It could be that altered contacts, such

as those seen in this putative lipid binding site, found in different membrane constructs are a result of tiny changes in membrane complexity and lipid species representation. Given some combination of specific membrane composition and protein, such small changes could manifest into energetic differences between the accessibility of some conformational transition paths over others. Again, the specificity of protein-lipid interactions and dependence of initial or achieved annular packings given a bilayer composition offer an element of kinetic control in making conformational sampling feasible.

CONCLUSIONS

This work establishes the composition of a realistic model plant plasma membrane and utilizes this construct as a model system for examining the effects of realistic membrane composition on long timescale membrane protein dynamics. Hurdles to simulating realistic membranes include lack of lipidomic data and trial-and-error workflows to appropriately account for leaflet asymmetry. After establishing the robust properties of the Maximum Complexity membrane composition, we show that even the removal of trace lipid species can alter the free energy barriers of intermediate conformational states for a given membrane protein. Simulations performed herein using native lipid bilayers provide easier access to intermediate states and may give faster kinetics than a nonnative lipid bilayer. The use of an appropriate membrane appears to become crucial when sampling for rare events and optimizing computational efficiency for membrane protein simulations. More importantly, a more native lipid bilayer produces different results than overly simplified homogeneous lipid bilayers, providing greater confidence and accuracy in the resolution of the membrane protein conformations sampled.

Either for simplicity's sake or out of necessity, MD practitioners are too often without proper lipidomic data (i.e., mass spectrometry results showing relative abundances for both lipid acyl chains in a leaflet-specific manner) when assembling membrane lipid bilayers for their own studies, ultimately resorting to "default" membrane compositions. Simplistic lipid bilayers for one cell type may be unacceptable for another organism; only 1% of plant plasma membranes are comprised of POPC, which may not even be represented in membrane systems of smaller scale. True, membrane proteins can still function so long as the surrounding lipid environment provides headgroups and acyl chains which are reminiscent of the native membrane composition,[114,115] but function in a nonnative setting may not be optimal. From the perspective of simulation, protein insertion into an incorrect membrane could result in misrepresentation of state probability, or the generation of a different landscape altogether. In this sense, the ability of an incorrect membrane environment to eschew an energy landscape is similar to that imposed by mutation.

With the increasing availability of computational power and the demand for larger, more accurately represented lipid bilayers, traditional simplistic approaches will no doubt be pushed aside for more complex systems, allowing true representations of realistic membrane-protein dynamics. OsSWEET2b simulations performed in a realistic plant membrane mirrored the dynamics seen in *holo* simulations performed using purely POPC membranes, emphasizing how lipid species can behave as ligands themselves (Figure S16). General protein-lipid binding motifs have been shown,[116,117] but the lack of consistent realistic membrane modeling when it comes to membrane protein simulations currently leaves this knowledge unexploited. Introducing a large diversity of lipid species into a finite-sized MD membrane system also complicates the annular shell composition for a given protein. While additional membrane simulation approaches have been suggested for modeling the effect of protein-lipid interactions,[108] we recommend that attention be paid to packing procedures chosen for complex membrane assembly.[118] Replicating

complex membrane assembly is one way to better account for how lipid diversity interfaces with membrane protein function via varied annular shell composition. Small differences in annular shell composition between the two membrane constructs accompanied OsSWEET2b gating dynamics which could not equally access the same intermediate states. It is possible that small changes in very complex membrane compositions could result in altered lipid organization, which may directly impact modeled membrane protein function and dynamics. For instance, hydrophobic mismatch – when the thicknesses of the transmembrane spanning region of a membrane protein and the surrounding lipid bilayer are incompatible – can result in energetic penalties which either restrict or promote certain protein conformational changes.[114,119,120] Thus even small manipulations in lipid bilayer composition could directly impact membrane protein function by affecting the probability for the optimal annular shell composition to be assembled for a given protein conformation during bilayer construction. This may explain why explain why some intermediate states are not as accessible for OsSWEET2b while embedded in the Top10 versus the Maximum Complexity construct. Furthermore, optimal protein-lipid interactions can be completely lost when using traditional simplistic lipid bilayers (e.g., just POPC) where hydrophobic mismatch penalties could be maximized.

A caveat to these findings is that comparative membrane modeling was only performed on one protein, OsSWEET2b. Here we show that removal of minor lipid species with relative abundances < 1% do indeed alter the accessibility of some intermediate conformational states when using a realistic membrane. Still, more studies on other membrane proteins are required to obtain a generalizable understanding of how non-optimal versus realistic membrane compositions affect the thermodynamics and kinetics of protein dynamics. Current understanding of how to universally approach and contextualize protein-lipid interactions in MD simulations is as unrefined as the present implementation of realistic membranes in biomolecular simulations. The study of complex membrane composition and membrane protein function often appear as mutually exclusive works. Increased prevalence of realistic membranes in membrane protein function studies will ease these two simulation communities to become more conjunctive. As the modeling of more realistic membranes for protein function simulations becomes more standard, deeper simulation analyses will become more routine and facilitate improved translation of *in silico* findings to *in vitro* or *vivo*. With the intent of being used as a modeling practices resource, the results of this work will aid in the encouragement of realistic membrane simulations and increased research output in plant-based membrane MD applications.

ASSOCIATED CONTENT

Supporting Information

Simulation results: membrane compositions for specified construct sizes, order parameters, headgroup orientations, electron density, MSM sampling error, 1-dimensional gating landscapes, comparative landscape with previously published work on OsSWEET2b embedded in a POPC membrane, linear interaction energies with residues in putative lipid binding site, annular shell snapshots.

Additional files including raw data, analysis and plotting code, as well as downloadable membrane structures and trajectories, are available free of charge at:
https://uofi.box.com/s/6s4jhq5j5pifxofr6ptdvtaply40sslo

AUTHOR INFORMATION

Corresponding Author

* University of Illinois at Ubana-Champaign, 203 Roger Adams Laboratory MC 712, 600 S Mathews Ave, Urbana, IL 61801. Phone: 217-300-0021. Email: diwakar@illinois.edu

Author Contributions

ATW and DS conceived the project. DS supervised the project. ATW gathered literature for membrane design and performed simulations. MC customized Membrainy code for complex membrane analyses. ATW and DS analyzed the data. ATW wrote the manuscript with input from DS and MC. All authors have given approval to the final version of the manuscript.

Funding Sources

DS acknowledges support from the Foundation for Food and Agriculture Research via the New Innovator Award and NSF Early CAREER Award (NSF-MCB-1845606).

ACKNOWLEDGMENT

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) the State of Illinois, and as of December, 2019, the National Geospatial-Intelligence Agency. Blue Waters is a joint effort of the University of Illinois and its National Center for Supercomputing Allocations. We thank Balaji Selvam for providing MSM-weighted OsSWEET2b starting structures and landscapes from a previous work for more amenable comparisons. We also thank Matthew Chan and Soumajit Dutta for useful input during the preparation of this manuscript.

ABBREVIATIONS

MSM, Markov state models; MD, molecular dynamics; PDB, protein data bank; GPU, graphical processing unit; CPU, central processing unit; HMR, hydrogen mass repartitioning; APL, area-per-lipid; tIC, time-independent component; tICA, time-independent component analysis; PL, phospholipid; SITO, β-sitosterol; STIG, stigmasterol; PLPE, 1-palmitoyl-2-linoleoyl-*sn*-glycero-3-phosphoethanolamine; PLPC, 1-palmitoyl-2-linoleoyl-*sn*-glycero-3-phosphocholine; DLiPC, 1,2-dilinoleoyl-*sn*-glycero-3-phosphocholine; DLiPE, 1,2-dilinoleoyl-*sn*-glycero-3-phosphoethanolamine; LLPC, 1-linoleoyl-2-linolenoyl-*sn*-glycero-3-phosphocholine; PLPG, 1-palmitoyl-2-linoleoyl-*sn*-glycero-3-phosphatidylglycerol; LLPE, 1-linoleoyl-2-linolenoyl-*sn*-glycero-3-phosphoethanolamine; POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; DSPC, 1,2-distearoyl-*sn*-glycero-3-phosphocholine; DOPC, 1,2-dioleoyl-*sn*-glycero-3-

phosphocholine; POPE, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine; POPG, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylglycerol; DPPG, 1,2-dipalmitoyl-*sn*-glycero-3-phosphatidylglycerol; SLPE, 1-stearoyl-2-linoleoyl-*sn*-glycero-3-phosphoethanolamine; HG, hourglass; IF, inward-open; OC, occluded; OF, outward-open

REFERENCES

(1) Chavent, M.; Duncan, A. L.; Sansom, M. S. Molecular Dynamics Simulations of Membrane Proteins and Their Interactions: From Nanoscale to Mesoscale. *Curr. Opin. Struct. Biol.* **2016**, *40*, 8–16. https://doi.org/10.1016/j.sbi.2016.06.007.

(2) Friedman, R.; Khalid, S.; Aponte-Santamaría, C.; Arutyunova, E.; Becker, M.; Boyd, K. J.; Christensen, M.; Coimbra, J. T. S.; Concilio, S.; Daday, C.; van Eerden, F. J.; Fernandes, P. A.; Gräter, F.; Hakobyan, D.; Heuer, A.; Karathanou, K.; Keller, F.; Lemieux, M. J.; Marrink, S. J.; May, E. R.; Mazumdar, A.; Naftalin, R.; Pickholz, M.; Piotto, S.; Pohl, P.; Quinn, P.; Ramos, M. J.; Schiøtt, B.; Sengupta, D.; Sessa, L.; Vanni, S.; Zeppelin, T.; Zoni, V.; Bondar, A.-N.; Domene, C. Understanding Conformational Dynamics of Complex Lipid Mixtures Relevant to Biology. *J. Membr. Biol.* **2018**, *251* (5), 609–631. https://doi.org/10.1007/s00232-018-0050-y.

(3) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119* (9), 6184–6226. https://doi.org/10.1021/acs.chemrev.8b00460.

(4) Loschwitz, J.; Olubiyi, O. O.; Hub, J. S.; Strodel, B.; Poojari, C. S. Computer Simulations of Protein–Membrane Systems. *Prog. Mol. Biol. Transl. Sci.* **2020**, *170*, 273–403. https://doi.org/10.1016/bs.pmbts.2020.01.001.

(5) Javanainen, M.; Martinez-Seara, H. Efficient Preparation and Analysis of Membrane and Membrane Protein Systems. *Biochim. Biophys. Acta BBA - Biomembr.* **2016**, *1858* (10), 2468–2482. https://doi.org/10.1016/j.bbamem.2016.02.036.

(6) Muller, M. P.; Jiang, T.; Sun, C.; Lihan, M.; 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* (9), 6086–6161. https://doi.org/10.1021/acs.chemrev.8b00608.

(7) Corradi, V.; Sejdiu, B. I.; Mesa-Galloso, H.; Abdizadeh, H.; Noskov, S. Yu.; Marrink, S. J.; Tieleman, D. P. Emerging Diversity in Lipid–Protein Interactions. *Chem. Rev.* **2019**, *119* (9), 5775–5848. https://doi.org/10.1021/acs.chemrev.8b00451.

(8) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865. https://doi.org/10.1002/jcc.20945.

(9) Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S.-J.; Patel, D. S.; Beaven, A. H.; Lee, K. I.; Rui, H.; Park, S.; Lee, H. S.; Roux, B.; MacKerell, A. D.; Klauda, J. B.; Qi, Y.; Im, W. CHARMM-GUI 10 Years for Biomolecular Modeling and Simulation. *J. Comput. Chem.* **2017**, *38* (15), 1114–1124. https://doi.org/10.1002/jcc.24660.

(10) 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* (27), 1997–2004. https://doi.org/10.1002/jcc.23702.

(11) Park, S.; Beaven, A. H.; Klauda, J. B.; Im, W. How Tolerant Are Membrane Simulations with Mismatch in Area per Lipid between Leaflets? *J. Chem. Theory Comput.* **2015**, *11* (7), 3466–3477. https://doi.org/10.1021/acs.jctc.5b00232.
(12) Esteban-Martín, S.; Risselada, H. J.; Salgado, J.; Marrink, S. J. Stability of Asymmetric Lipid Bilayers Assessed by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2009**, *131* (42), 15194–15202. https://doi.org/10.1021/ja904450t.
(13) Doktorova, M.; Weinstein, H. Accurate in Silico Modeling of Asymmetric Bilayers Based on Biophysical Principles. *Biophys. J.* **2018**, *115* (9), 1638–1643. https://doi.org/10.1016/j.bpj.2018.09.008.
(14) Ingólfsson, H. I.; Melo, M. N.; van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; de Vries, A. H.; Tieleman, D. P.; Marrink, S. J. Lipid Organization of the Plasma Membrane. *J. Am. Chem. Soc.* **2014**, *136* (41), 14554–14559. https://doi.org/10.1021/ja507832e.
(15) Ingólfsson, H. I.; Bhatia, H.; Zeppelin, T.; Bennett, W. F. D.; Carpenter, K. A.; Hsu, P.-C.; Dharuman, G.; Bremer, P.-T.; Schiøtt, B.; Lightstone, F. C.; Carpenter, T. S. Capturing Biologically Complex Tissue-Specific Membranes at Different Levels of Compositional Complexity. *J. Phys. Chem. B* **2020**, *124* (36), 7819–7829. https://doi.org/10.1021/acs.jpcb.0c03368.
(16) Flinner, N.; Schleiff, E. Dynamics of the Glycophorin A Dimer in Membranes of Native-like Composition Uncovered by Coarse-Grained Molecular Dynamics Simulations. *PLOS ONE* **2015**, *10* (7), e0133999. https://doi.org/10.1371/journal.pone.0133999.
(17) Ingólfsson, H. I.; Carpenter, T. S.; Bhatia, H.; Bremer, P.-T.; Marrink, S. J.; Lightstone, F. C. Computational Lipidomics of the Neuronal Plasma Membrane. *Biophys. J.* **2017**, *113* (10), 2271–2280. https://doi.org/10.1016/j.bpj.2017.10.017.
(18) Andoh, Y.; Okazaki, S.; Ueoka, R. Molecular Dynamics Study of Lipid Bilayers Modeling the Plasma Membranes of Normal Murine Thymocytes and Leukemic GRSL Cells. *Biochim. Biophys. Acta BBA - Biomembr.* **2013**, *1828* (4), 1259–1270. https://doi.org/10.1016/j.bbamem.2013.01.005.
(19) Kadri, L.; Ferru-Clément, R.; Bacle, A.; Payet, L.-A.; Cantereau, A.; Hélye, R.; Becq, F.; Jayle, C.; Vandebrouck, C.; Ferreira, T. Modulation of Cellular Membrane Properties as a Potential Therapeutic Strategy to Counter Lipointoxication in Obstructive Pulmonary Diseases. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **2018**, *1864* (9, Part B), 3069–3084. https://doi.org/10.1016/j.bbadis.2018.06.021.
(20) Corradi, V.; Mendez-Villuendas, E.; Ingólfsson, H. I.; Gu, R.-X.; Siuda, I.; Melo, M. N.; Moussatova, A.; DeGagné, L. J.; Sejdiu, B. I.; Singh, G.; Wassenaar, T. A.; Delgado Magnero, K.; Marrink, S. J.; Tieleman, D. P. Lipid–Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent. Sci.* **2018**, *4* (6), 709–717. https://doi.org/10.1021/acscentsci.8b00143.
(21) Lasitza-Male, T.; Bartels, K.; Jungwirth, J.; Wiggers, F.; Rosenblum, G.; Hofmann, H.; Löw, C. Membrane Chemistry Tunes the Structure of a Peptide Transporter. *Angew. Chem. Int. Ed.* **2020**, *59* (43), 19121–19128. https://doi.org/10.1002/anie.202008226.
(22) Shahane, G.; Ding, W.; Palaiokostas, M.; Orsi, M. Physical Properties of Model Biological Lipid Bilayers: Insights from All-Atom Molecular Dynamics Simulations. *J. Mol. Model.* **2019**, *25* (3), 76. https://doi.org/10.1007/s00894-019-3964-0.

(23) Zhuang, X.; Ou, A.; Klauda, J. B. Simulations of Simple Linoleic Acid-Containing Lipid Membranes and Models for the Soybean Plasma Membranes. *J. Chem. Phys.* **2017**, *146* (21), 215103. https://doi.org/10.1063/1.4983655.
(24) van Eerden, F. J.; de Jong, D. H.; de Vries, A. H.; Wassenaar, T. A.; Marrink, S. J. Characterization of Thylakoid Lipid Membranes from Cyanobacteria and Higher Plants by Molecular Dynamics Simulations. *Biochim. Biophys. Acta BBA - Biomembr.* **2015**, *1848* (6), 1319–1330. https://doi.org/10.1016/j.bbamem.2015.02.025.
(25) Ogata, K.; Yuki, T.; Hatakeyama, M.; Uchida, W.; Nakamura, S. All-Atom Molecular Dynamics Simulation of Photosystem II Embedded in Thylakoid Membrane. *J. Am. Chem. Soc.* **2013**, *135* (42), 15670–15673. https://doi.org/10.1021/ja404317d.
(26) van Eerden, F. J.; van den Berg, T.; Frederix, P. W. J. M.; de Jong, D. H.; Periole, X.; Marrink, S. J. Molecular Dynamics of Photosystem II Embedded in the Thylakoid Membrane. *J. Phys. Chem. B* **2017**, *121* (15), 3237–3249. https://doi.org/10.1021/acs.jpcb.6b06865.
(27) Van Eerden, F. J.; Melo, M. N.; Frederix, P. W. J. M.; Marrink, S. J. Prediction of Thylakoid Lipid Binding Sites on Photosystem II. *Biophys. J.* **2017**, *113* (12), 2669–2681. https://doi.org/10.1016/j.bpj.2017.09.039.
(28) Selvam, B.; Yu, Y.-C.; Chen, L.-Q.; Shukla, D. Molecular Basis of the Glucose Transport Mechanism in Plants. *ACS Cent. Sci.* **2019**, *5* (6), 1085–1096. https://doi.org/10.1021/acscentsci.9b00252.
(29) Bohn, M.; Lüthje, S.; Sperling, P.; Heinz, E.; Dörffling, K. Plasma Membrane Lipid Alterations Induced by Cold Acclimation and Abscisic Acid Treatment of Winter Wheat Seedlings Differing in Frost Resistance. *J. Plant Physiol.* **2007**, *164* (2), 146–156. https://doi.org/10.1016/j.jplph.2005.12.008.
(30) Borner, G. H. H.; Sherrier, D. J.; Weimar, T.; Michaelson, L. V.; Hawkins, N. D.; MacAskill, A.; Napier, J. A.; Beale, M. H.; Lilley, K. S.; Dupree, P. Analysis of Detergent-Resistant Membranes in Arabidopsis. Evidence for Plasma Membrane Lipid Rafts. *Plant Physiol.* **2005**, *137* (1), 104–116. https://doi.org/10.1104/pp.104.053041.
(31) Brown, D. J.; DuPont, F. M. Lipid Composition of Plasma Membranes and Endomembranes Prepared from Roots of Barley (Hordeum Vulgare L.): Effects of Salt. *Plant Physiol.* **1989**, *90* (3), 955–961. https://doi.org/10.1104/pp.90.3.955.
(32) Furt, F.; Simon-Plas, F.; Mongrand, S. Lipids of the Plant Plasma Membrane. In *The Plant Plasma Membrane*; Murphy, A. S., Schulz, B., Peer, W., Eds.; Plant Cell Monographs; Springer: Berlin, Heidelberg, 2011; pp 3–30. https://doi.org/10.1007/978-3-642-13431-9_1.
(33) Grandmougin, A.; Bouvier-Navé, P.; Ullmann, P.; Benveniste, P.; Hartmann, M.-A. Cyclopropyl Sterol and Phospholipid Composition of Membrane Fractions from Maize Roots Treated with Fenpropimorph. *Plant Physiol.* **1989**, *90* (2), 591–597. https://doi.org/10.1104/pp.90.2.591.
(34) Grison, M. S.; Brocard, L.; Fouillen, L.; Nicolas, W.; Wewer, V.; Dörmann, P.; Nacir, H.; Benitez-Alfonso, Y.; Claverol, S.; Germain, V.; Boutté, Y.; Mongrand, S.; Bayer, E. M. Specific Membrane Lipid Composition Is Important for Plasmodesmata Function in Arabidopsis. *Plant Cell* **2015**, *27* (4), 1228–1250. https://doi.org/10.1105/tpc.114.135731.
(35) Lynch, D. V.; Steponkus, P. L. Plasma Membrane Lipid Alterations Associated with Cold Acclimation of Winter Rye Seedlings (Secale Cereale L. Cv Puma). *Plant Physiol.* **1987**, *83* (4), 761–767.
(36) Minami, A.; Fujiwara, M.; Furuto, A.; Fukao, Y.; Yamashita, T.; Kamo, M.; Kawamura, Y.; Uemura, M. Alterations in Detergent-Resistant Plasma Membrane Microdomains in

Arabidopsis Thaliana during Cold Acclimation. *Plant Cell Physiol.* **2009**, *50* (2), 341–359. https://doi.org/10.1093/pcp/pcn202.
(37) Mongrand, S.; Morel, J.; Laroche, J.; Claverol, S.; Carde, J.-P.; Hartmann, M.-A.; Bonneu, M.; Simon-Plas, F.; Lessire, R.; Bessoule, J.-J. Lipid Rafts in Higher Plant Cells Purification and Characterization of Triton X-100-Insoluble Microdomains from Tobacco Plasma Membrane. *J. Biol. Chem.* **2004**, *279* (35), 36277–36286. https://doi.org/10.1074/jbc.M403440200.
(38) Norberg, P.; Liljenberg, C. Lipids of Plasma Membranes Prepared from Oat Root Cells: Effects of Induced Water-Deficit Tolerance. *Plant Physiol.* **1991**, *96* (4), 1136–1141. https://doi.org/10.1104/pp.96.4.1136.
(39) Palta, J. P.; Whitaker, B. D.; Weiss, L. S. Plasma Membrane Lipids Associated with Genetic Variability in Freezing Tolerance and Cold Acclimation of Solanum Species. *Plant Physiol.* **1993**, *103* (3), 793–803. https://doi.org/10.1104/pp.103.3.793.
(40) Uemura, M.; Steponkus, P. L. A Contrast of the Plasma Membrane Lipid Composition of Oat and Rye Leaves in Relation to Freezing Tolerance. *Plant Physiol.* **1994**, *104* (2), 479–496. https://doi.org/10.1104/pp.104.2.479.
(41) Uemura, M.; Joseph, R. A.; Steponkus, P. L. Cold Acclimation of Arabidopsis Thaliana (Effect on Plasma Membrane Lipid Composition and Freeze-Induced Lesions). *Plant Physiol.* **1995**, *109* (1), 15–30. https://doi.org/10.1104/pp.109.1.15.
(42) Wright, L. C.; McMurchie, E. J.; Pomeroy, M. K.; Raison, J. K. Thermal Behavior and Lipid Composition of Cauliflower Plasma Membranes in Relation to ATPase Activity and Chilling Sensitivity. *Plant Physiol.* **1982**, *69* (6), 1356–1360.
(43) Yoshida, S.; Uemura, M. Lipid Composition of Plasma Membranes and Tonoplasts Isolated from Etiolated Seedlings of Mung Bean (Vigna Radiata L.). *Plant Physiol.* **1986**, *82* (3), 807–812. https://doi.org/10.1104/pp.82.3.807.
(44) Tjellström, H.; Hellgren, L. I.; Wieslander, Å.; Sandelius, A. S. Lipid Asymmetry in Plant Plasma Membranes: Phosphate Deficiency-Induced Phospholipid Replacement Is Restricted to the Cytosolic Leaflet. *FASEB J.* **2010**, *24* (4), 1128–1138. https://doi.org/10.1096/fj.09-139410.
(45) Freeman Allen, C.; Good, P.; Davis, H. F.; Chisum, P.; Fowler, S. D. Methodology for the Separation of Plant Lipids and Application to Spinach Leaf and Chloroplast Lamellae. *J. Am. Oil Chem. Soc.* **1966**, *43* (4), 223–231. https://doi.org/10.1007/BF02641091.
(46) Bahl, J.; Francke, B.; Monéger, R. Lipid Composition of Envelopes, Prolamellar Bodies and Other Plastid Membranes in Etiolated, Green and Greening Wheat Leaves. *Planta* **1976**, *129* (3), 193–201. https://doi.org/10.1007/BF00398257.
(47) Bligny, R.; Douce, R. A Precise Localization of Cardiolipin in Plant Cells. *Biochim. Biophys. Acta BBA - Lipids Lipid Metab.* **1980**, *617* (2), 254–263. https://doi.org/10.1016/0005-2760(80)90168-X.
(48) Block, M. A.; Dorne, A. J.; Joyard, J.; Douce, R. Preparation and Characterization of Membrane Fractions Enriched in Outer and Inner Envelope Membranes from Spinach Chloroplasts. II. Biochemical Characterization. *J. Biol. Chem.* **1983**, *258* (21), 13281–13286.
(49) Douce, R.; Holtz, R. B.; Benson, A. A. Isolation and Properties of the Envelope of Spinach Chloroplasts. *J. Biol. Chem.* **1973**, *248* (20), 7215–7222.
(50) Harwood, J. L. Plant Acyl Lipids: Structure, Distribution, and Analysis. In *Lipids: Structure and Function*; Stumpf, P. K., Ed.; Academic Press, 1980; pp 1–55. https://doi.org/10.1016/B978-0-12-675404-9.50007-2.

(51) Hashimoto, H.; Murakami, S. Dual Character of Lipid Composition of the Envelope Membrane of Spinach Chloroplasts. *Plant Cell Physiol.* **1975**, *16* (5), 895–902. https://doi.org/10.1093/oxfordjournals.pcp.a075212.
(52) Mackender, R. O.; Leech, R. M. The Galactolipid, Phospholipid, and Fatty Acid Composition of the Chloroplast Envelope Membranes of Vicia Faba. L. *Plant Physiol.* **1974**, *53* (3), 496–502. https://doi.org/10.1104/pp.53.3.496.
(53) Mannella, C. A.; Bonner, W. D. Biochemical Characteristics of the Outer Membranes of Plant Mitochondria. *Biochim. Biophys. Acta BBA - Biomembr.* **1975**, *413* (2), 213–225. https://doi.org/10.1016/0005-2736(75)90105-4.
(54) Mazliak, P. Glyco- and Phospholipids of Biomembranes in Higher Plants. In *Lipids and Lipid Polymers in Higher Plants*; Tevini, M., Lichtenthaler, H. K., Eds.; Springer: Berlin, Heidelberg, 1977; pp 48–74. https://doi.org/10.1007/978-3-642-66632-2_3.
(55) McCarty, R. E.; Douce, R.; Benson, A. A. The Acyl Lipids of Highly Purified Plant Mitochondria. *Biochim. Biophys. Acta BBA - Lipids Lipid Metab.* **1973**, *316* (2), 266–270. https://doi.org/10.1016/0005-2760(73)90019-2.
(56) Meunier, D.; Mazliak, P. Difference in lipid composition between two membranes of mitochondria of the potato. *Acad Sci Paris C R Ser D* **1972**.
(57) Moreau, F.; Dupont, J.; Lance, C. Phospholipid and Fatty Acid Composition of Outer and Inner Membranes of Plant Mitochondria. *Biochim. Biophys. Acta BBA - Biomembr.* **1974**, *345* (3), 294–304. https://doi.org/10.1016/0005-2736(74)90191-6.
(58) Poincelot, R. P. Isolation and Lipid Composition of Spinach Chloroplast Envelope Membranes. *Arch. Biochem. Biophys.* **1973**, *159* (1), 134–142. https://doi.org/10.1016/0003-9861(73)90437-2.
(59) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30* (13), 2157–2164. https://doi.org/10.1002/jcc.21224.
(60) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
(61) Rostkowski, M.; Olsson, M. H. M.; Søndergaard, C. R.; Jensen, J. H. Graphical Analysis of PH-Dependent Properties of Proteins Predicted Using PROPKA. *BMC Struct. Biol.* **2011**, *11*, 6. https://doi.org/10.1186/1472-6807-11-6.
(62) Kohlmeyer, A. *TopoTools*; http://doi.org/10.5281/zenodo.50249, 2016.
(63) Crowley, M. F.; Williamson, M. J.; Walker, R. C. CHAMBER: Comprehensive Support for CHARMM Force Fields within the AMBER Software. *Int. J. Quantum Chem.* **2009**, *109* (15), 3767–3772. https://doi.org/10.1002/qua.22372.
(64) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864–1874. https://doi.org/10.1021/ct5010406.
(65) Balusek, C.; Hwang, H.; Lau, C. H.; Lundquist, K.; Hazel, A.; Pavlova, A.; Lynch, D. L.; Reggio, P. H.; Wang, Y.; Gumbart, J. C. Accelerating Membrane Simulations with Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2019**, *15* (8), 4673–4686. https://doi.org/10.1021/acs.jctc.9b00160.
(66) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. I.; Cruziero, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T.-S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.;

Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER*; University of California, San Francisco: San Francisco, CA, 2018.

(67) 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* (8), 3684–3690. https://doi.org/10.1063/1.448118.

(68) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N'-Methylamide. *Biopolymers* **1992**, *32* (5), 523–535. https://doi.org/10.1002/bip.360320508.

(69) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm. *Chem. Phys. Lett.* **2004**, *384* (4–6), 288–294. https://doi.org/10.1016/j.cplett.2003.12.039.

(70) Krautler, V.; van Gunsteren, W. F.; Hunenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22*, 501–508.

(71) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.

(72) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–2145. https://doi.org/10.1002/jcc.23354.

(73) Braun, E.; Moosavi, S. M.; Smit, B. Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited. *J. Chem. Theory Comput.* **2018**, *14* (10), 5262–5272. https://doi.org/10.1021/acs.jctc.8b00446.

(74) Bowman, G. R.; Ensign, D. L.; Pande, V. S. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. *J. Chem. Theory Comput.* **2010**, *6* (3), 787–794. https://doi.org/10.1021/ct900620b.

(75) Shukla, D.; Hernández, C. X.; Weber, J. K.; Pande, V. S. Markov State Models Provide Insights into Dynamic Modulation of Protein Function. *Acc. Chem. Res.* **2015**, *48* (2), 414–422. https://doi.org/10.1021/ar5002999.

(76) Shamsi, Z.; Cheng, K. J.; Shukla, D. Reinforcement Learning Based Adaptive Sampling: REAPing Rewards by Exploring Protein Conformational Landscapes. *J. Phys. Chem. B* **2018**, *122* (35), 8386–8395.

(77) Shamsi, Z.; Moffett, A. S.; Shukla, D. Enhanced Unbiased Sampling of Protein Dynamics Using Evolutionary Coupling Information. *Sci. Rep.* **2017**, *7* (1), 12700.

(78) Harrigan, M. P.; Sultan, M. M.; Hernández, C. X.; Husic, B. E.; Eastman, P.; Schwantes, C. R.; Beauchamp, K. A.; McGibbon, R. T.; Pande, V. S. MSMBuilder: Statistical Models for Biomolecular Dynamics. *Biophys. J.* **2017**, *112* (1), 10–15. https://doi.org/10.1016/j.bpj.2016.10.042.

(79) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015.

(80) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95. https://doi.org/10.1109/MCSE.2007.55.
(81) Carr, M.; MacPhee, C. E. Membrainy: A "Smart", Unified Membrane Analysis Tool. *Source Code Biol. Med.* **2015**, *10*, 3. https://doi.org/10.1186/s13029-015-0033-7.
(82) Santos, D. E. S.; Pontes, F. J. S.; Lins, R. D.; Coutinho, K.; Soares, T. A. SuAVE: A Tool for Analyzing Curvature-Dependent Properties in Chemical Interfaces. *J. Chem. Inf. Model.* **2020**, *60* (2), 473–484. https://doi.org/10.1021/acs.jcim.9b00569.
(83) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.
(84) 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. https://doi.org/10.1016/j.softx.2015.06.001.
(85) Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Pérez-Hernández, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J.-H.; Noé, F. PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **2015**, *11* (11), 5525–5542. https://doi.org/10.1021/acs.jctc.5b00743.
(86) Cheng, K. J.; Selvam, B.; Chen, L.-Q.; Shukla, D. Distinct Substrate Transport Mechanism Identified in Homologous Sugar Transporters. *J. Phys. Chem. B* **2019**, *123* (40), 8411–8418. https://doi.org/10.1021/acs.jpcb.9b08257.
(87) Selvam, B.; Mittal, S.; Shukla, D. Free Energy Landscape of the Complete Transport Cycle in a Key Bacterial Transporter. *ACS Cent. Sci.* **2018**, *4* (9), 1146–1154.
(88) Selvam, B.; Shamsi, Z.; Shukla, D. Universality of the Sodium Ion Binding Mechanism in Class A G-Protein-Coupled Receptors. *Angew. Chem. Int. Ed Engl.* **2018**, *57* (12), 3048–3053.
(89) Wu, H.; Noé, F. Variational Approach for Learning Markov Processes from Time Series Data. *J. Nonlinear Sci.* **2020**, *30* (1), 23–66. https://doi.org/10.1007/s00332-019-09567-y.
(90) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/jcc.20084.
(91) von Heijne, G. Mitochondrial Targeting Sequences May Form Amphiphilic Helices. *EMBO J.* **1986**, *5* (6), 1335–1342. https://doi.org/10.1002/j.1460-2075.1986.tb04364.x.
(92) Heijne, G. von; Gavel, Y. Topogenic Signals in Integral Membrane Proteins. *Eur. J. Biochem.* **1988**, *174* (4), 671–678. https://doi.org/10.1111/j.1432-1033.1988.tb14150.x.
(93) van Klompenburg, W.; Nilsson, I.; von Heijne, G.; de Kruijff, B. Anionic Phospholipids Are Determinants of Membrane Protein Topology. *EMBO J.* **1997**, *16* (14), 4261–4266. https://doi.org/10.1093/emboj/16.14.4261.
(94) 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* (6), 644–652. https://doi.org/10.1038/s41589-020-0529-6.
(95) Braun, E.; Gilmer, J.; Mayes, H. B.; Mobley, D. L.; Monroe, J. I.; Prasad, S.; Zuckerman, D. M. Best Practices for Foundations in Molecular Simulations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2019**, *1* (1). https://doi.org/10.33011/livecoms.1.1.5957.

(96) Brandani, G. B.; Schor, M.; Macphee, C. E.; Grubmüller, H.; Zachariae, U.; Marenduzzo, D. Quantifying Disorder through Conditional Entropy: An Application to Fluid Mixing. *PloS One* **2013**, *8* (6), e65617. https://doi.org/10.1371/journal.pone.0065617.
(97) Benabdelkader, A.; Mazliak, P. Echanges de Lipides Entre Mitochondries, Microsomes et Surnageant Cytoplasmique de Cellules de Pomme de Terre Ou de Chou-Fleur. *Eur. J. Biochem.* **1970**, *15* (2), 250–262. https://doi.org/10.1111/j.1432-1033.1970.tb01002.x.
(98) Douce, R. Site of Biosynthesis of Galactolipids in Spinach Chloroplasts. *Science* **1974**, *183* (4127), 852–853. https://doi.org/10.1126/science.183.4127.852.
(99) Meunier, D.; Mazliak, P. Comptes rendus hebdomadaires des séances de l'Académie des sciences. Série D, Sciences naturelles https://gallica.bnf.fr/ark:/12148/bpt6k57786873 (accessed Jan 4, 2021).
(100) Moreau, F.; Dupont, J.; Lance, C. Phospholipid and Fatty Acid Composition of Outer and Inner Membranes of Plant Mitochondria. *Biochim. Biophys. Acta BBA - Biomembr.* **1974**, *345* (3), 294–304. https://doi.org/10.1016/0005-2736(74)90191-6.
(101) Philipp, E. I.; Franke, W. W.; Keenan, T. W.; Stadler, J.; Jarasch, E. D. Characterization of Nuclear Membranes and Endoplasmic Reticulum Isolated from Plant Tissue. *J. Cell Biol.* **1976**, *68* (1), 11–29. https://doi.org/10.1083/jcb.68.1.11.
(102) Brandt, R. D.; Benveniste, P. Isolation and Identification of Sterols from Subcellular Fractions of Bean Leaves (Phaseolus Vulgaris). *Biochim. Biophys. Acta BBA - Biomembr.* **1972**, *282*, 85–92. https://doi.org/10.1016/0005-2736(72)90313-6.
(103) Brown, D. J.; Dupont, F. M. Lipid Composition of Plasma Membranes and Endomembranes Prepared from Roots of Barley (Hordeum Vulgare L.): Effects of Salt. *Plant Physiol.* **1989**, *90* (3), 955–961. https://doi.org/10.1104/pp.90.3.955.
(104) Dörmann, P. *Membrane Lipids*; Blackwell, 2020; pp 123–161. https://doi.org/10.1201/9780367813123-5.
(105) Zhang, C.; Hicks, G. R.; Raikhel, N. V. Molecular Composition of Plant Vacuoles: Important but Less Understood Regulations and Roles of Tonoplast Lipids. *Plants* **2015**, *4* (2), 320–333. https://doi.org/10.3390/plants4020320.
(106) Zhou, Y.; Pan, X.; Qu, H.; Underhill, S. J. R. Tonoplast Lipid Composition and Proton Pump of Pineapple Fruit during Low-Temperature Storage and Blackheart Development. *J. Membr. Biol.* **2014**, *247* (5), 429–439. https://doi.org/10.1007/s00232-014-9650-3.
(107) Tavernier, E.; Lê Quôc, D.; Lê Quôc, K. Lipid Composition of the Vacuolar Membrane of Acer Pseudoplatanus Cultured Cells. *Biochim. Biophys. Acta* **1993**, *1167* (3), 242–247. https://doi.org/10.1016/0005-2760(93)90225-x.
(108) Corey, R. A.; Stansfeld, P. J.; Sansom, M. S. P. The Energetics of Protein–Lipid Interactions as Viewed by Molecular Simulations. *Biochem. Soc. Trans.* **2019**, *48* (1), 25–37. https://doi.org/10.1042/BST20190149.
(109) Betz, R. M.; Dror, R. O. How Effectively Can Adaptive Sampling Methods Capture Spontaneous Ligand Binding? *J. Chem. Theory Comput.* **2019**, *15* (3), 2053–2063. https://doi.org/10.1021/acs.jctc.8b00913.
(110) Dai, J.; Alwarawrah, M.; Huang, J. Instability of Cholesterol Clusters in Lipid Bilayers and the Cholesterol's Umbrella Effect. *J. Phys. Chem. B* **2010**, *114* (2), 840. https://doi.org/10.1021/jp909061h.
(111) Huang, J.; Feigenson, G. W. A Microscopic Interaction Model of Maximum Solubility of Cholesterol in Lipid Bilayers. *Biophys. J.* **1999**, *76* (4), 2142–2157. https://doi.org/10.1016/S0006-3495(99)77369-8.

(112) Parker, A.; Miles, K.; Cheng, K. H.; Huang, J. Lateral Distribution of Cholesterol in Dioleoylphosphatidylcholine Lipid Bilayers: Cholesterol-Phospholipid Interactions at High Cholesterol Limit. *Biophys. J.* **2004**, *86* (3), 1532–1544. https://doi.org/10.1016/S0006-3495(04)74221-6.
(113) Mesmin, B.; Maxfield, F. R. Intracellular Sterol Dynamics. *Biochim. Biophys. Acta* **2009**, *1791* (7), 636–645. https://doi.org/10.1016/j.bbalip.2009.03.002.
(114) Andersen, O. S.; Koeppe, R. E. Bilayer Thickness and Membrane Protein Function: An Energetic Perspective. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36* (1), 107–130. https://doi.org/10.1146/annurev.biophys.36.040306.132643.
(115) Lee, A. G. How Lipids and Proteins Interact in a Membrane: A Molecular Approach. *Mol. Biosyst.* **2005**, *1* (3), 203–212. https://doi.org/10.1039/B504527D.
(116) Hunte, C. Specific Protein-Lipid Interactions in Membrane Proteins. *Biochem. Soc. Trans.* **2005**, *33* (Pt 5), 938–942. https://doi.org/10.1042/BST20050938.
(117) Palsdottir, H.; Hunte, C. Lipids in Membrane Protein Structures. *Biochim. Biophys. Acta BBA - Biomembr.* **2004**, *1666* (1), 2–18. https://doi.org/10.1016/j.bbamem.2004.06.012.
(118) Sommer, B. Membrane Packing Problems: A Short Review on Computational Membrane Modeling Methods and Tools. *Comput. Struct. Biotechnol. J.* **2013**, *5* (6), e201302014. https://doi.org/10.5936/csbj.201302014.
(119) Ernst, A. M.; Contreras, F. X.; Brügger, B.; Wieland, F. Determinants of Specificity at the Protein-Lipid Interface in Membranes. *FEBS Lett.* **2010**, *584* (9), 1713–1720. https://doi.org/10.1016/j.febslet.2009.12.060.
(120) Harayama, T.; Riezman, H. Understanding the Diversity of Membrane Lipid Composition. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (5), 281–296. https://doi.org/10.1038/nrm.2017.138.