Symmetric and Asymmetric Models for the *Arabidopsis thaliana* Plasma Membrane: A Simulation Study

Yalun Yu[1] and Jeffery B. Klauda[1,2,*]

[1]Biophysics Graduate Program, [2]Department of Chemical and Biomolecular Engineering,
University of Maryland,
College Park, Maryland 20742, USA

*Corresponding Author: jbklauda@umd.edu

Abstract

Arabidopsis thaliana is an important model organism which has attracted many biologists. While most research efforts have been on studying the genetics and proteins of this organism, a systematic study of its lipidomics is lacking. Here, we present a novel, asymmetric model of its cell membrane with its lipid composition consisting of five glycerophoslipids, two glycolipids and sitosterol determined from multiple independent experiments. A typical lipid type in plant membranes is Glycosyl Inositol Phosphoryl Ceramide (GIPC), which accounts for about 10 % of the total lipids in the outer leaflet in our model. Two symmetric models representing the inner and outer leaflets of the membrane were built and simulated until equilibrium was reached and then combined to form the asymmetric model. Our results indicate that the outer leaflet is more rigid and tightly packed compared to the inner leaflet. Pressure profiles for the two leaflets are overall similar though the outer leaflet exhibits larger oscillations. A special focus on lipid organization is discussed and the interplay between glycolipids and sitosterols is found to be important. The current model provides a baseline for future modeling of similar membranes and can be used to study partioning of small molecules in the membrane or further developed to study the interaction between plant membrane proteins and lipids.

1. Introduction

Arabidopsis thaliana, a small flowering plant, was widely used as a model organism in the past half-century to study plant biology and beyond, and the pioneers to study it can be traced back to the 1930s.[1] During the 1970s, suspicion about using it as a model organism was raised because it seemed to resist most initial attempts to proliferate and regenerate in culture and its small-sized chromosomes.[1] However, with the emergence of efficient tissue culture methods[2-4] and alternative protocols,[5] and the realization in the field of the advantages of its small-sized chromosomes,[6] it quickly regained the focus of biologists. Taking gene sequencing as an example, the latest annotation of the *Arabidopsis* genome identified 2655 protein-coding and 5178 non-protein coding genes.[7] Moreover, comprehensive studies have been carried out by research groups from all over the world and results have been assembled into books[8] and a free-access database.[9] Studies using *Arabidopsis* have also played a leading role in basic biological discoveries,[7, 10] such as proteins identified as human orthologs in the innate immune system,[11] the role of auxin on the ubiquitin pathway conserved among eukaryotes,[12] and a light signaling component COP1 whose mammalian orthologs has a role in tumorigenesis.[13]

While most molecular biologists studying *Arabidopsis* have focused on its nucleic acids (chromosomes) and proteins, its membrane composition and chemical physics have not been explored thoroughly. In fact, the plasma membrane (PM), in particular, plays a critical role in plant physiological processes including growth and development, ion and metabolite transport, perception of environmental changes, and disease resistance.[14] The PM is the cellular interface composed of a lipid bilayer and proteins that separates the cytoplasm from the surrounding environment and regulates the exchange of molecules and information between the cell and enviroment. For many years, it has been considered as a two-dimensional fluid consisting of homogenously distributed lipids and proteins,[15] but now it is clear that they are more mosaic than fluid and can form distinct membrane microdomains of various sizes and mobilities.[16, 17] When it comes to plants, their PM is composed of three major lipid types, which are sterols (both free and conjugated), glycerolipids and sphingolipids.[18] Sterols are essential for plant development, particularly embryogenesis, cell elongation, and vascular differentiation.[19] They can also act as structural lipids and induce the liquid-ordered phase in membrane.[18] The plant PM also has a high sphingolipid composition. In tobacco leaves, for example, study has shown that glycosylinositolphosphoceramides (GIPCs) may represent up to 30-40 mol % of the PM.[20]

Sphingolipids are engaged in lipid raft formation in conjunction with sterols.[21] These rafts are sometimes referred to as detergent insoluble domains (DIM) by biologist and have putative functions in plant cell signaling in cooperation with their associated proteins.[21-23] Glycerolipids, primarily phospholipids, are another major lipid type in plant PM. Historically, they are classified by abundance into structural lipids and signaling lipids, though more recent studies have shown that some "structural lipids" are also involved in signal-transducing.[18] The dominating phospholipids in plant PM are phosphatidylcholine (PC) and phosphatidylethanolamine (PE), which represent up to 68–80% of structural phospholipids. The remainder consists of phosphatidylglycerol (PG), phosphatidylinositol (PI), phosphatidylserine (PS), and phosphatidic acid (PA).[18]

Besides their interesting physical chemistry properties, lipids are top players in protein functions.[24] Experimentally, the interactions between lipids and proteins can be studied at many different levels, and methods can be classified into two groups: *in vivo* and *in vitro*.[25] While being successful in many cases, *in vitro* methods suffers from the fact that they are usually conducted under artificial conditions such as non-physiological concentrations,[26] hence the physiological relevance of the interactions needs to be investigated through orthogonal and more physiological methods.[25] *In vivo* methods, on the other hand, can reduce such artifacts greatly. They usually involve the perturbation of a membrane component or a protein segment followed by the measurement of the effects of these perturbations based on phenotypic readouts.[27-33] However, due to the complex nature of many physiological pathways, interactions cannot always be measured directly and metabolized products of the mutated lipids may disturb the measurements. To compensate for these shortcomings, molecular dynamics (MD) simulation can be used. Not only can MD simulation model the lipid diversity,[34-36] but it also provides an atomistic view of the lipid-lipid and lipid-protein interactions at timescales relevant to biology.[37-39] Apart from these benefits, asymmetric membranes can be modeled in MD simulations,[40, 41] though they are not commonly seen in *in vitro* experiments. In terms of plants, the PM is asymmetric with the outer membrane containing a fair portion of sphingolipids, while the inner membrane consists nearly exclusively of sterols and phospholipids.[42]

In this paper, an asymmetric membrane model is built to represent the PM of *Arabidopsis thaliana*. The model uses five different phospholipids (see next section for details) which are representative of the major types found in the PM of *Arabidopsis*. Two sphingolipids are considered and only used in the apoplastic (outer) leaflet of the membrane. Sitosterol is used as the only sterol in the model and sterol derivatives are not considered due to their minor amounts.[43] Since the component area of these lipids are not known beforehand, two symmetric models representing the cytosolic (inner) and outer leaflets are built and simulated until reaching equilibrium before they are merged to form the asymmetric model. The analysis of the simulations covers overall structure of the membrane (surface area, density profile), mechanical properties of the membrane (area compressibility and pressure profile), lipid-specific structure (deuterium order parameter, lipid tilt), lipid-lipid interaction (hydrogen bonding), and pattern of lipid organization within a membrane (clustering, interleaflet coupling). Hidden Markov Modeling (HMM) was utilized to further study lipid correlation and compositional phase separation (potential raft formation).

One purpose of this study is to provide readily useful models of plant membrane for future study (for example, auxin partitioning). Although the membranes are modest in size, they can be easily expanded to simulate large bilayers with proteins or other biologically relevant molecules. By way of outline, Section 2 describes the protocols to build and simulate the model membranes and the methods to analyze the simulation data. Section 3 shows a wide range of physical and chemical properties of the model membranes. A specific focus of Section 3 is the lateral arrangement of lipids within the membrane, which is investigated by multiple approaches. Section 4 presents the Discussion and Conclusions.

2. Methods

2.1. Lipid Composition. The relative proportions of sphingolipids, glycerophospholipids and sterols (free and conjugated) in the PM of *Arabidopsis thaliana* leaves were determined by Uemura et al.[44] to be 7, 47, 46%, repectively. Conjugated sterols were excluded from our model because of the low fraction (8%). The same study was used to inform the overall glycerophospholipids composition, from which five major ones were selected (see Table 1). While the leaflet asymmetry has not been determined by any experiment at this point to the best of our knowledge, we assumed an asymmetric distribution of glycerophospholipids between the two leaflets that is similar to other eukaryotic cells, where the cytosolic leaflet is enriched in PS and PI.[24] Since PG is also negatively charged, we presumed that it mainly resides in the cytosolic leaflet. To make the model representative of the real membrane and to facilitate computational analysis, PS, PI and PG lipids were excluded from the apoplastic leaflet and only incorporated into the cytosolic leaflet. The ratio between different glycerolipids were kept as close as possible to the original values reported by Uemura et al.,[44] which can be found in Table S1. For all glycerophospholipids, only the predominant species (tails) were used based on Uemura et al.,[44] which can be found in Table 1.

Table 1. Lipid Definition Abbreviations for glycerophospholipids are consistent with naming used in the CHARMM-GUI *Membrane Builder*.

lipid name	abbrev.	*sn*-1/ fatty acid chain	*sn*-2/ sphingosine chain
1-palmitoyl-2-linoleoyl-sn-glycero-3-phosphocholine	PLPC	16:0	18:2
1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine	DLiPE	18:2	18:2
1-palmitoyl-2-(α-linolenoyl)-sn-glycero-3-phospho-(1'-rac-glycerol)	PNPG	16:0	18:3
1-palmitoyl-2-linoleoyl-sn-glycero-3-phosphoinositol	PLPI	16:0	18:2
1,2-dilinoleoyl-sn-glycero-3-phospho-L-serine	DLiPS	18:2	18:2
D-glucose-(α1→4)-D-glucuronic acid-(α1→2)-inositol-phosphoryl-(N-palmitoyl-4R-hydroxy-D-erythro-sphingosine)	GIPC	16:0	18:0
β-D-glucosyl-N-palmitoyl-4R-hydroxy-D-erythro-sphingosine	GluCer	16:0	18:1
sitosterol	SITO		

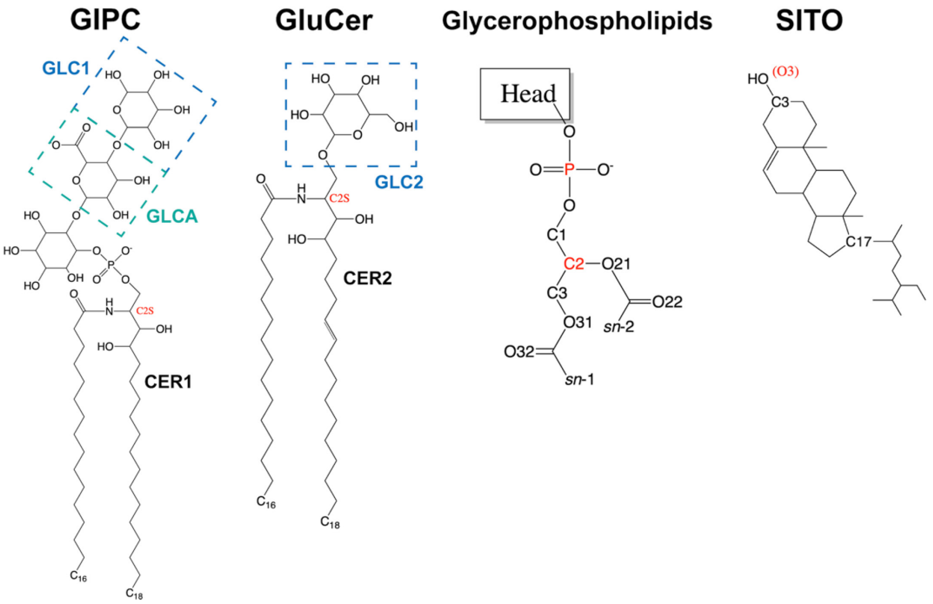


Figure 1. Chemical structures of lipid types involved in this study.

It is well known that the predominant sterols in plants are β-sitosterol (sitosterol hereafter), stigmasterol, and campesterol,[45] and Grosjean et al.[46] observed an almost even distribution of these species in tobacco suspension cells. Since these sterols are similar to each other in structure, only sitosterol was used in our model where it accounts for 50% of the outer leaflet and 42% of the inner leaflet.

Additional caution was imposed when determining the molecule species of sphingolipids and their precise structure. It has been reported by Markham et al.[47] that GIPC, glucosylceramide (GluCer), and ceramides accounts for 64, 34, and 2% of the total sphingolipids, so that ceramides were excluded from our model and the molar ratio of GluCer (1): GIPC (2) was used. The chemical structure of the predominant GIPC was also determined in the same study to be hexose-hexuronic-inositol-phosphoceramide for *Arabidopsis thaliana*. The linkage conformations between the six-membered rings determined in two other publications[48, 49] were used (see Table 1 for details). In terms of tails, long chain base (LCB) of eighteen carbon-long bearing two hydroxyl groups at carbon 3 and 4 has been identified as the predominant sphingosine chain, and a double bond may appear at carbon 8.[44, 47, 50] The fatty acid was determined to be mostly 16:0 by Uemura et al.[44] Within the PM of eukaryotic cells, sphingolipids are primarily located in the outer leaflet, so that they were excluded from the inner leaflet in our models. Table 2 lists the lipid composition for each model.

Table 2. Model Composition.

lipid	inner	outer	plant inner	plant1 outer	plant2 outer	plant3 outer
SITO	25	30	25	33	38	33
PLPC	3	17	3	18	15	19
DLiPE	12	5	12	5	5	6
PNPG	7		7			
PLPI	8		8			
DLiPS	5		5			
GluCer		3		3	3	4
GIPC		6		7	6	4
Total	60	61	60	66	67	66

2.2 Modeling Building and Simulation Details. The inner and outer models were constructed using the CHARMM-GUI *Membrane Builder*.[51, 52] For lipids not included in the CHARMM-GUI library, similar lipids were used as placeholders and mutated afterwards to the real ones using CHARMM.[53] A hydration number of 35 water molecules per lipid was used for the inner model and 50 was used for the outer model. NAMD[54] was used to perform the equilibrium and production runs for the two symmetric models (3 replicas for each) with the CHARMM36 (C36) lipid force field[55] and the modified TIP3 water model.[56, 57] The total area of the simulation box tangential to the membrane surface (the x-y plane henceforce) and the lipid-wise (component) area for each lipid species were computed after the systems reached equilibrium (see Section 2.3 for details). The asymmetric membranes were constructed by taking one leaflet (including the water molecules bound to that leaflet when the system is recentered around the center of the bilayer) from each model at the end of the simulation and joining them using an area-match method through CHARMM. Since the simulated total area of the inner model was larger than the outer model, the outer model was expanded in the x-y plane using periodic images and the excess lipids were selectively cutoff to make the sum of the total area of the primary cell and the lipid-wise areas (based on the component area from the images as closest to the total area of the inner model as possible. Since the two leaflets from the symmetric models were interdigitated at the end of simulation, the initial structures for the asymmetric models were also interdigitated. However, any steric clash between the two leaflets was expected to resolve quickly during the equilibration. The resulting lipid composition for each outer leaflet in the asymmetric model are listed in Table 2, while the inner leaflet has the same composition as the symmetric model. The different lipid compositions for the outer leaflet come from the randomness of lipid placement at the end of simulation for the outer model and they mimic the natural fluctuation of local lipid content in a realistic membrane. In terms of system size, the inner model measured at ~ $5.5\ (x) \times 5.5\ (y) \times 8.0\ (z)$ (nm); the outer model measured at ~ $5.3 \times 5.3 \times 10.0$ (nm); and the asymmetric models measured at ~ $5.5 \times 5.5 \times 9.0$ (nm).

The same software and force field was used to simulate the asymmetric membranes. For all the simulations in this study, the isothermal-isobaric (NPT) ensemble with a constant temperature of 298.15 K and a constant pressure of 1 atmosphere was used. Langevin dynamics was used to maintain the constant temperature and the Nosé-Hoover-Langevin piston[58, 59] was used to maintain the constant pressure, which allows the size of the simulation cell to change semi-isotropically ($x=y\neq z$). The parameters for the NPT control can be found in Table S2. The bond length between each hydrogen and the heavy atom to which it is bonded was constrained to the nominal length in the C36 parameter files through the SHAKE algorithm[60] and a displacement tolerance of 10^{-8} Å was used. The Particale Mesh Ewald (PME) method[61, 62] with an interpolation order of 6 and a maximum grid spacing of 1 Å was used for the long range electrostatics and the Lennard-Jones formulation with a force-swithching function[63] ranging from 10 to 12 Å was used for the van der Waals interactions. The production run was 400 ns for each replica of the inner model and 500 ns for the outer model. For each of the three asymmetric models, two replicas were simulated for 500 ns.

In addition to the "standard" models listed in Table 2, three replicas of large outer model were built using the coordinates (lipids, water molecules, and ions) from the 20th ns of the standard model simulations. The small systems were duplicated by three times and translated in the *x*-*y* plane by the amount of the cell size to form systems which are four times bigger in terms of total lipids. These larger membranes were used to study the size effect on the compositional phase separation. Simulation parameters for these large systems remained the same as the standard systems and the simulation time is 500 ns for each replica.

2.3. Analysis. Analysis for all model membranes were based on the equilibrated portion of the simulation, which was idendified through the time evolution of surface area per lipid (A_{lip}) and sterol-sterol radial distribution function. Other properties studied include component surface area, area compressibility modulus (K_{A}), chain deuterium order parameter (S_{CD}), lipid tilt, electron density profile (EDP), two-dimentional radial distribution function (2D-RDF) between lipids, lipid clustering, hydrogen bonding, compositional phase separation, electrostatic potential and pressure profile.

A_{lip} was computed by dividing the area of the simulation box in the *x*-*y* plane by the number of lipids in each leaflet. For the asymmetric models, this led to different results for the two leaflets. Pymbar[64] was used to detect the starting points of area equilibrium, which are listed in Table S3.

K_{A} was computed from the fluctuation of the total area in the *x*-*y* plane through

$$K_{\mathrm{A}}=\frac{k_{\mathrm{B}}T\langle A\rangle}{\sigma_{\langle A\rangle}^{2}} \qquad (1)$$

where k_{B} is the Boltzmann's constant, T is the temperature (in Kelvin), $\langle A\rangle$ is the average total area per leaflet and $\sigma_{\langle A\rangle}^{2}$ is the mean square fluction of that area.

Component surface area was obtained by extracting the *x*, *y* coordinates of representative atoms for each lipid and feeding them into Qhull[65] to construct polygons, from which the total area for each lipid was computed.

The S_{CD} was calculated from

$$S_{\mathrm{CD}} = \left| \langle \frac{3}{2} cos^2\theta - \frac{1}{2} \rangle \right| \qquad (2)$$

where θ is the instantaneous angle formed by the vector of the C-H bond and the bilayer normal (z direction), and the average is for both time and different lipids.

The EDP was obtained by first centering the bilayer at $z = 0$, and then computing the atomic densities along the z axis. This calculation was only performed for the last 50 ns of each simulation. Functional groups were defined in Section 3 to interpret the EDP in a chemically meaningful way. Based on EDPs, the head-to-head distance D_{HH} (distance between the two peaks of the total EDP), the overall bilayer thickness D_{B} (distance between the half maximum points of the water EDP), and the hydrophobic distance $2D_{\mathrm{C}}$ (distance between the half maximum points of the acyl chain EDP) were determined. Chain interdigitation between the two leaflets were calculated based on eq 1 of Das et al.,[66] which utilizes the EDP of individual leaflets.

The 2D-RDF in the x-y plane was calculated based on the coordinates of representative atoms defined for each lipid species, which can be found in Table S4.

Hydrogen bonding is analyzed using CHARMM for headgroup atoms. the H-acceptor distance is restricted to be less than 2.4 Å and the minimum donor-hydrogen-acceptor angle is set to be 150°.

Lateral lipid clustering was computed using the density-based spatial clustering of applications with noise (DBSCAN) algorithm implemented in scikit-learn.[67, 68] While the cutoff distance in DBSCAN is a uniform value specified by the user, lipid-lipid distances from the trajectory can be preprossed (scaled) to reflect the different sizes of various lipids. This scaling was made possible by analyzing the 2D-RDFs between different lipid species and was shown to provide a more robust, and perhaps more importantly, an unbiased analysis of lipid clustering. More information about this will be presented in Section 3.

Compositional phase separation was analyzed using the method developed by Sodt et al.,[38] which utilized a hidden Markov model (HMM) to detect the "hidden" states for individual lipids. In this method, the local composition in the vicinity of each lipid determined by the six nearest lipids (including itself) is used as the emission signals. Asuming the two "hidden" states are the liquid-ordered (L_o) and the liquid-disordered (L_d) phases, it is reasonable to group the lipids into three major types, i.e., sitosterol, lipids with one or less double bonds (GIPC and GluCer), and lipids with two or more double bonds (all other lipids). The local composition is then encoded with the number of each lipid type, leading to 28 different local compositions. The sampling frequency for this analysis (0.2 ns) is much lower than the other analysis considering lipid diffusion is at the order of 10^{-7} cm^2/sec (thus it only covers an area of 2.4 Å^2 during 0.2 ns, on average), and this frequency is close to what was used by Sodt et al. (0.239 ns).

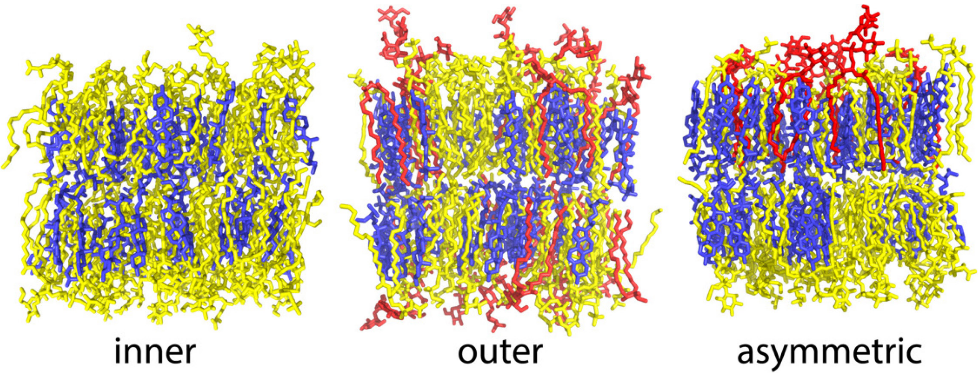


Figure 2. Example snapshots at the end of simulations. Yellow: glycerophospholipids; red: glycolipids; blue: sitosterols.

Electrostatic potential drop from the center of bilayer was calculated through

$$\phi(z) - \phi(0) = -\frac{4\pi}{\varepsilon_0}\int_0^z dz' \left[\int_0^{z'} dz''\rho_c(z'') + E_0\right] \qquad (3)$$

where $\phi(0)$ and $\phi(z)$ are the potential at the bilayer center (identified as the lowest point of the total EDP) and position z respectively, ε_0 is the dielectric constant of vaccum, ρ_c is the charge density. E_0 is the electric field at the bilayer center, which can be solved by using the boundary condition that the electrostatic potential at the two edges of the simulation box should be equal. For a symmetric bilayer, the leaflet-symmetrized, time-averaged charge density was used, so that E_0 is zero. For an asymmetric bilayer, the time-averaged charge density was used and E_0 was solved by imposing the following restraint

$$\int_{-0.5L}^{-0.4L} \phi(z)dz = \int_{0.4L}^{0.5L} \phi(z)dz \qquad (4)$$

where L the average length of the simulation box in the z direction. The averages of the first and the last 10% of the potential were used when determining E_0 instead of the potentials at $z = \pm L$ because of the noise in the potential profile (see Section 3.8).

Pressure profiles were generated by restarting the simulaitons in CHARMM from checkpoints containing both coordinates and velocities. A total of 100 checkpoints (every 2 ns from the last 200 ns of the NAMD simulation) were used for each inner membrane replica, and a total of 125 checkpoints (every 2 ns from the last 250 ns of the NAMD simulation) were used for each outer membrane replica and each asymmetric membrane replica. The resampling simulations were carried out with a developmental version of the CHARMM program, where the pressure profile is calculated using Harasima contour. A time step of 1 fs was used for these simulations and the simulation length is 0.1 ns for each checkpoint. NPT ensemble was maintained using the extended system barostat and thermostat. Electrostatic interactions were treated using the PME method with a real space cutoff of 12 Å, and the the Lennard-Jones formulation with a force-swithching function ranging from 10 to 12 Å was used for the van der Waals interactions. The local pressure was calculated and recorded at a frequency of 100 timesteps. While the total surface tension is zero in an NPT simulation, the contribution from each leaflet can be calculated by

$$\gamma_{lower} = \int_{-\frac{L}{2}}^{0}(P_N - P_T(z))dz \quad \text{or} \quad \gamma_{upper} = \int_{0}^{\frac{L}{2}}(P_N - P_T(z))dz \qquad (5)$$

where P_N (1 atm) and P_T are the components of the pressure tensor normal and tangential to the bilayer surface. Previous work has shown that P_T exhibits huge oscillations between negative values at the water/hydrocarbon interface and positive values at the bilayer center.[69] One can also calculate the first derivative of the bending free energy per lipid at zero curvature for each leaflet, $\bar{F}'(0)$, by

$$\bar{F}'(0) = \frac{d\bar{F}(c)}{dc}|_{c=0} = -\int_{0}^{\frac{L}{2}} z[P_T(z) - P_N]\,dz = -K_c^m c_0 \qquad (6)$$

where c stands for the curvature, K_c^m is the monolayer (leaflet) bending constant, and c_0 is the spontaneous curvature of the monolayer (leaflet). This integral provides the curvature tendency of the leaflet, although the extraction of the exact value of c_0 needs K_c^m to be determined, which is beyond the scope of this work.

3. Results

In this section, results will be presented for all the membrane models. Since the three asymmetric models are slightly different from each other in terms of their outer leaflet composition, we will use “plant 1”, “plant 2” and “plant 3” to distinguish them in the figures and tables. For all simulations, flip-flop of sitosterol was ruled out by checking the positions of the O3 atoms using the whole trajectory recentered around the bilayer center. When the average of the three asymmetric models is presented, the label “plant” will be used. The total surface area and the sitosterol 2D-RDF were used to estimate the starting points of equilibrium. The starting points of area equilibrium are reported in Table S3, and sample plots of total area versus time are provided in Figure S1. Sample plots of block averaged 2D-RDF for sitosterol are shown in Figure S2 to demonstrate general convergence of properties. Based on these results, the last 200 ns of the simulation data of the inner membrane model, and the last 250 ns of the simulation data of the outer membrane model and asymmetric membrane models were used for the analyses.

3.1. Surface Area and Area Compressibility. A_{lip} is reported in Table 3 and the A_{lip} is 48.35 ± 0.01 Å^2 for the symmetric inner model and 43.79 ± 0.02 Å^2 for the symmetric outer model. This difference is majorly contributed by the higher sitosterol proportion and the inclusion of GluCer and GIPC in the outer model. While the component areas (Table 4) for these lipids are lower than the glycerophospholipids, the higher concentration of sitosterol and saturated tails also produced an ordering effect, which slightly reduced the component areas for the other lipids in the system. In the asymmetric models, A_{lip} for each leaflet roughly matches its corresponding symmetric model. Plant-2 has a lower A_{lip} compared to the other two asymmetric models, especially for the outer leaflet. This likely originated from the higher number of sitosterols in the outer leaflet. It is interesting that the A_{lip} for the inner leaflet (which corresponds to the inner model) in all the asymmetric models are significantly lower than the A_{lip} for the symmetric inner model. While a possible origin is a small area mismatch between the two leaflets when they were joined together, a much lower A_{lip} for the lower leaflet in plant-2 indicates the higher sterol concentration in the outer leaflet might also modulate the A_{lip} of the opposite leaflet. A previous simulation study of soybean hypocotyl and root plasma membranes[70] generated A_{lip} of 52.7 ± 0.2 Å^2 and 51.9 ± 0.1 Å^2, respectively, which are larger than the A_{lip} for both leaflets of the *A. thaliana* plasma membrane.

This is expected since the soybean membrane models contain more unsaturated tails and less sterols.

Table 3. A_{lip} and Compressibility Modulus (K_A). Errors for the inner/outer symmetric models are standard errors based on three replicas. Average K_A based on all three asymmetric models is reported with standard errors from all six replicas to reduce statistical error.

model	A_{lip} outer leaflet (Å^2)	A_{lip} inner leaflet (Å^2)	K_A (N/m)
inner	48.35 ± 0.01		0.68 ± 0.05
outer	43.79 ± 0.02		0.99 ± 0.05
plant 1	43.66 ± 0.03	48.03 ± 0.03	0.94 ± 0.04
plant 2	42.53 ± 0.04	47.50 ± 0.04	
plant 3	43.67 ± 0.04	48.03 ± 0.04	

K_A for all model membranes are listed in Table 3. K_A for the inner model is significantly lower than the outer model, although both are remarkably higher than a single-component glycerophospholipid bilayer,[69] reinforcing the role of sterols in the modulation of membrane stiffness.[71, 72] K_A for the asymmetric model is closer to the outer model, indicating the area compressibility is non-additive and biased to the leaflet with more compression resistance.

Table 4. Component area (Å^2).

lipid	inner	outer	plant inner	plant outer
SITO	31.3 ± 0.4	29.6 ± 0.3	31.3 ± 0.6	30.1 ± 0.3
PLPC	60.1 ± 1.4	58.6 ± 0.6	59.4 ± 2.4	59.3 ± 0.7
DLiPE	60.0 ± 0.9	58.4 ± 1.2	59.9 ± 1.1	59.4 ± 1.8
PNPG	61.9 ± 1.1		60.4 ± 1.6	
PLPI	60.1 ± 1.1		60.8 ± 1.9	
DLIPS	60.4 ± 1.6		60.8 ± 1.4	
GluCer		53.3 ± 1.1		53.1 ± 1.7
GIPC		54.1 ± 0.9		53.9 ± 1.3

3.2. Electron Density Profile and Membrane Thickness. The overall EDPs are shown in Figure 3, with the water EDP and the EDP of the hydrophobic core shown separately. The decreased peak of the overall EDP for the outer leaflet/model is associated with the replacement of a portion of glycerophospholipids with glycolipids, whose EDPs decrease at ~ ± 23 Å (Figure S3). EDPs for individual lipid types are presented in Figure S3. There is a small degree of chain interdigitation between the two leaflets, as shown in Table S5. Membrane thicknesses computed based on the EDPs are listed in Table 5. Here, only the averages of the asymmetric models are reported. D_{HH} for the three different systems are statistically the same. The overall thickness, D_{B}, is highest in the outer model and lowest in the inner model. Normally D_{B} can be used to reflect the degree of water penetration and a higher D_{B} means less water penetration. However, in our case, the major reason for the higher D_{B} in the outer model is the glycolipids taking up extra space beyond the membrane surface (see Figure 2). The hydrophobic thickness, $2D_{\mathrm{C}}$, is also higher in the outer model, probably caused by the higher percentage of saturated tails and sitosterol. For the asymmetric model (avg.), D_{B} and $2D_{\mathrm{C}}$ are approximately the average of the inner and outer model. Compared with the soybean model membranes,[70] the asymmetric membrane is slightly thicker, perhaps as a consequence of more saturated tails and higher sterol content.

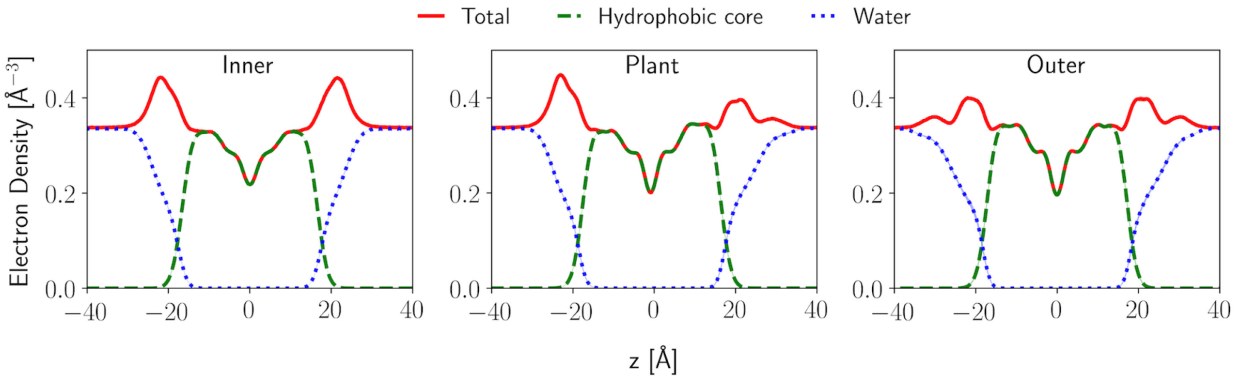


Figure 3. Total EDP, EDP of the hydrophobic core and water EDP.

Table 5. Membrane thicknesses for model membranes. Standard errors are based on replicas.

model	D_{B} (Å)	D_{HH} (Å)	$2D_{\mathrm{C}}$ (Å)
inner	40.1 ± 0.1	43.5 ± 0.1	33.4 ± 0.1
outer	41.6 ± 0.4	42.8 ± 0.5	34.8 ± 0.2
asymmetric (avg.)	40.8 ± 0.2	43.6 ± 0.4	34.3 ± 0.1

To gain more insights into the influence of lipid composition on the density distribution, leaflet-molar-fraction-normalized EDPs for individual lipid species were generated (Figure S4). For the asymmetric membrane, an error function was used to smooth the transition between two leaflets as it is necessary to account for the different number of lipids in each leaflet. Ideally, this smoothing could be avoided by calculating the density profile for each leaflet separately. However, the smoothing method only introduced negligible artifacts at the leaflet interface, which is not the focus of this analysis. Lipids in this set of plots can be divided into three groups: (1) lipids shared by both the inner and outer models; (2) lipids only presented in the inner model; (3) lipids only presented in the outer model. For lipids in the first group, the normalized density profiles for the outer model have higher peaks compared to their inner model counterparts, agreeing with the smaller A_{lip} and the following results for chain order parameters. Density profile for each leaflet in the asymmetric model overlaps with the corresponding symmetric model, indicating minimal change in lipid packing.

3.3. Order Parameter. S_{CD} is a metric used to measure chain order/disorder, and the order increases as the S_{CD} value increases. Figure 4 plots the S_{CD} for DLiPE and PLPC. DLiPE has two tails with double bonds at carbons 9/10 and 12/13, so that the S_{CD} dropped twice around that region for both tails. PLPC has only one tail (*sn*-2) containing double bonds at carbons 9/10 and 12/13, thus a similar double-dip is observed. It is clear from Figure 4 that the lipid tails are more ordered in the outer model or the outer leaflet of the asymmetric model, especially for the *sn*-1 chain of PLPC. This is expected because there are more saturated tails and sitosterols in the outer model/leaflet. The ordering effect of GIPC and phytosterols on glycerophospholipids were studied in a recent experiment by Mamode Cassim et al.,[73] where they measured the ^{2}H-NMR powder spectra of deuterated palmitoyl-oleoyl phosphatidylcholine containing 31 atoms of deuterium on the palmitoyl chain (POPC-(2)H_{31}) mixed with GIPC and/or phytosterols (sitosterol or stigmasterol). Adding phytosterols to the POPC-(2)H_{31} sample led the gel-to-fluid phase transition of the POPC bilayer at (−2.5 ± 2.4) °C to be abolished and highly ordered aliphatic chain was seen for all measured temperatures (-10 to 40 °C). By averaging the NMR data at 20 °C and 30 °C, a S_{CD} of 0.34 was obtained for the palmitoyl chain in a membrane composed of sitosterol and POPC-(2)H_{31} (1:1 mol). The ternary system composed of GIPC/sitosterol/POPC-(2)H_{31} (1:1:1 mol) also demonstrated a rigidifying effect of GIPC and sitosterol, with an estimated S_{CD} of 0.26 at 25 °C. Since the S_{CD} estimated from the NMR experiment should be interpreted as the average of carbon positions having significantly higher values compared to others, average S_{CD} of carbon 4 to carbon 11 of the palmitoyl (*sn*-1) chain (PLPC) in our models were computed. These averages are 0.30 for the inner model, 0.31 for the inner leaflet of the asymmetric model, and 0.36 for both the outer model and the outer leaflet of the asymmetric model, which compare favorably with the experiment considering the presence of other lipids and the different proportion of GIPC/sitosterol.

The S_{CD} profiles for the glycolipids are shown in Figure S5, which are basically identical between the outer model and asymmetric model. The only difference exsits in the top part of the sphingosine chain (carbon 2 to 7). However, it should be noted that S_{CD} at this region has relatively large errors, which is likely a consequence of different hydrogen bond patterns (see Section 4).

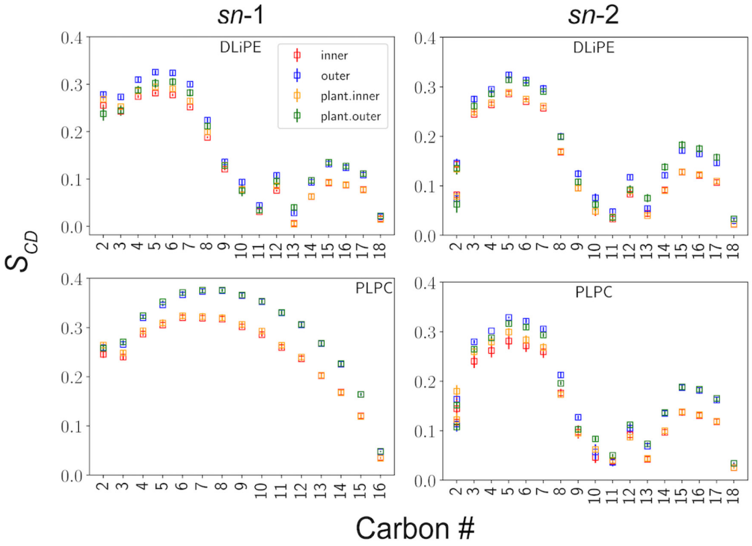


Figure 4. S_{CD} profiles for DLiPE and PLPC. Errors for the symmetric models are based on three replicas; errors for the symmetric model (plant.inner and plant.outer) are based on all three models listed in Table 2 (6 replicas in total).

3.4. Radial Distribution Functions. Two-dimensional radial distribution functions (2D-RDFs) were generated using representative atoms (Figure 1 and Table S1). These are the phosphorus atoms of the glycerophospholipids, the center carbon atoms in the amino alcohol of the glycolipids, and the oxygen atoms of sitosterols. Figure 5 presents the interactions between two major lipid types – sitosterol and glycolipid. There is a strong peak for the sitosterol-sitosterol 2D-RDF in all the model systems, perhaps due to the high concentration and the rigid structure of sitosterol. Despite of the higher sitosterol concentration in the outer model or the outer leaflet (in the asymmetric model), the height of this peak is lower compared to the inner model or the inner leaflet. This can be explained by the presence of glycolipids. Sitosterol shows preferential association with GluCer and GIPC at short range, and this might have diminished the sitosterol-sitosterol association. To determine whether this preferential association with sitosterol is specific to glycolipids, 2D-RDFs between sitosterol and two glycerophospholipids (DLiPE and PLPC) were generated (Figure S6). In addition, potential of mean force (PMF) based on the 2D-RDFs are shown in Figure S7. It is clear from these plots that sitosterol does not favor glycerophospholipids, which is expected since the glycerophospholipids have more unsaturated tails compared to the glycolipids. Another interaction of interest is the repulsion between glycolipids as shown in Figure 5 and Figure S7. This repulsion is likely caused by the negative net charge on GIPC (-2 *e*) as well as the steric effect of its bulky head.

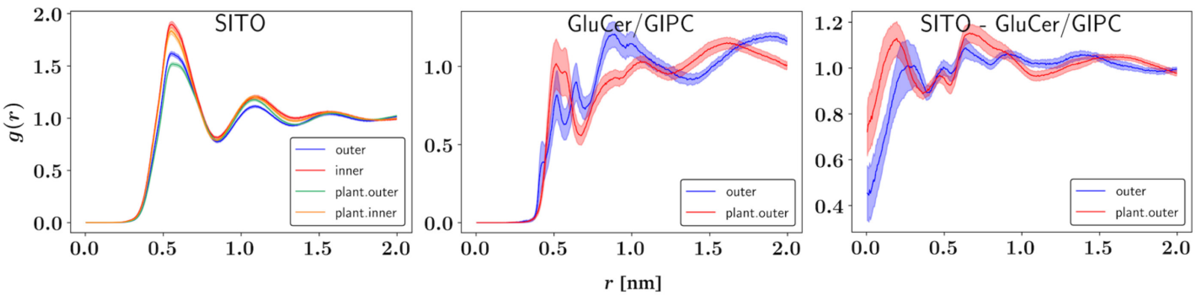


Figure 5. 2D-RDFs for selected lipid pairs. SITO: between any two sitosterols; GluCer/GIPC: between any two glycolipids; SITO – GluCer/GIPC: between sitosterols and glycolipids.

3.5. Lipid Clustering. Analysis of lipid clustering and compositional phase separation (Section 3.6) can provide information about lateral lipid organization within a leaflet. Clustering analysis was made possible by the DBSCAN algorithm, which uses distances between points of interest to compute the clusters. Here, the points are the same representative atoms used to compute the 2D-RDFs. In this set of analysis, two distance cutoff schemes were utilized. In the first, a uniform cutoff distance (6 Å) between lipids was used for all lipid pairs, regardless of their sizes. In the second, the distance was scaled by the average distance of two neighboring lipids of specific types estimated from the 2D-RDFs. This led to different cutoffs for different lipid pairs (Table S6). A lower cutoff means the two lipids need to be closer to form a cluster. Table 6 reports the tendency to form clusters for each lipid type. The different cutoff schemes did not change the tendency significantly for most lipids except sitosterol. This is because the modifications to the pairwise distance were mostly related to sitosterol and were all in the direction of decreasing the cutoff. The tendency for DLiPE increases when the pairwise distances are used, because it is a potent hydrogen bond donor as will be shown in Section 3.7. Moving from the symmetric membranes to the asymmetric membrane did not change the tendency to form clusters for all lipids, which means there was little inter-leaflet coupling in terms of clustering.

Table 6. Lipid % in cluster using different D_{cut} scheme. The minimum number of lipids to form a cluster is set to be 4. Membrane composition (%) also provided for comparison.

lipid	SITO	PLPC	DLIPE	PNPG	PLPI	DLIPS	GIPC	GluCer
	inner							
composition	42	5	20	12	13	8		
uniform D_{cut}	49 ± 1	4 ± 1	19 ± 1	9 ± 1	12 ± 1	7 ± 1		
customized D_{cut}	42 ± 1	4 ± 1	24 ± 1	8 ± 1	13 ± 1	9 ± 1		
	outer							
composition	49	28	8				10	5
uniform D_{cut}	57 ± 2	22 ± 1	7 ± 1				9 ± 1	5 ± 1
customized D_{cut}	53 ± 2	24 ± 1	9 ± 1				9 ± 1	5 ± 2
	plant inner							
composition	42	5	20	12	13	8		
uniform D_{cut}	48 ± 1	4 ± 1	19 ± 1	9 ± 1	12 ± 1	7 ± 1		
customized D_{cut}	42 ± 1	4 ± 1	24 ± 1	8 ± 1	13 ± 1	9 ± 1		
	plant outer							
Composition (avg.)	52	26	8				9	5
uniform D_{cut}	57 ± 1	22 ± 1	7 ± 1				9 ± 1	5 ± 1
customized D_{cut}	53 ± 2	24 ± 1	9 ± 1				9 ± 1	5 ± 1

The influences of the minimum number of samples (lipids)[67] and the overall scaling factor of the cutoff distance are reported in Figure S8. It turns out lipid clustering is not sensitive to these settings once the pairwise distances are used. In almost all cases, sitosterol and DLiPE have a slightly higher chance to form a cluster compared to the other lipids, perhaps due to the ability of sterol to fit into voids between lipid tails and the ability of DLiPE to form hydrogen bonds with neighboring lipids.

To study the stability and time scale of these lipid clusters, the residence time in cluster for each lipid type was calculated (Figure S9). Different minimum cluster sizes and cutoffs were used, and it was found that smaller minimum cluster size and larger cutoffs lead to longer residence time. Sitosterol has the shortest residence time compared to the other lipids, indicating a potential faster translational diffusion and is consistent with the relatively short lifetimes of hydrogen bonds involving sitosterol (see Section 3.7). In addition, sitosterol does not act as a hydrogen bond acceptor since there is no hydrogen bond donor around its polar hydroxyl group, which means it has more freedom for rotational diffusion, leading to a less stable interaction with neighboring lipids. DLiPE and DLiPS have relatively longer residence time because of their ability to form hydrogen bonds (see Section 3.7). Unfortunately, the large standard errors for GIPC and GluCer obstructed unambiguous interpretations of their residence times.

Clustering analysis can also be utilized to study the affinity between lipids. Table 7 presents the propensity of two lipids to coexist in a cluster, with a focus on sitosterol and GluCer/GIPC. These numbers measure the probability to find the second lipid (the accompanying lipid) in the cluster where the first lipid (the kernel lipid) is presenet. Percentage of the accompanying lipid in the composition is also provided for comparison. When the kernel lipid is sitosterol, the percentages of other lipids are not significantly modified compared to the composition, though a slight decrease for CluCer/GIPC is observed. When the kernel lipid is GluCer/GIPC, sitosterol is more favorable than the other lipids and GluCer/GIPC is less favorable, which is consistent with the repulsion between the glycolipids discovered by the 2D-RDF.

Table 7. Percentages of accompanying lipids in cluster where a kernel lipid is present. D_{cut} are those in Table S6 and the minimum number of lipids to form a cluster is set to 4.

model/leaflet	kernel lipid	accompanying lipid(s)	composition (%)	same cluster (%)
inner	SITO	SITO	42	43 ± 1
outer	SITO	SITO	49	51 ± 1
	SITO	GluCer/GIPC	15	12 ± 1
	GluCer/GIPC	SITO	49	55 ± 1
	GluCer/GIPC	GluCer/GIPC	15	8 ± 1
plant inner	SITO	SITO	42	42 ± 1
plant outer	SITO	SITO	52	54 ± 1
	SITO	GluCer/GIPC	14	11 ± 1
	GluCer/GIPC	SITO	52	57 ± 3
	GluCer/GIPC	GluCer/GIPC	14	9 ± 2

3.6. Compositional Phase Separation. As stated in the Method section, HMM was used to assign states to individual lipids based on the local composition. The two putative states are the L_o and L_d phases, considering experimental obsevations and simulation results of similar systems.[21, 38] The representative atoms to locate a lipid in the local composition calculation are same as those used in the 2D-RDF and clustering analyses (Table S6). In this analysis, we try to answer two questions: (1) Does certain type of lipid prefer to stay in one state (phase)? (2) Does HMM provide additional or different information regarding lipid organization compared to the clustering analysis? Since the focus here is potential phase separation, only the outer model and the outer leaflet from the asymmetric membrane were investigated because they contain both saturated and unsaturated tails.

Figure 6 plots the end-of-simulation states for replica 1 of the outer model. Lipid clustering is also shown for comparison. As shown in the plot, all glycolipids were assigned to the same state while sitosterols and glycerophospholipids appeared in both states. In the clustering analysis, however, only five glycolipids were identified as “in cluster”. End-of-simulation states for all replicas and leaflets are plotted in Figure S10 and Figure S11. The glycolipid-enriched state is filled with color in these subplots. It is interesting that a stripe-like pattern is found for almost all the replicas/leaflets of the outer model except the upper leaflet from replica 1. In the asymmetric membrane, this pattern is maintained. Besides, the glycolipid-enriched state also contains much more sitosterols than glycerophospholipids. To quantify this differentiation, the percentage of each lipid type that appeared in the glycolipid-enriched state was calculated (Table S8). On average, for the outer model, 98.0 % of the glycolipids and 58.4 % of the sitosterol appeared in the glycolipid-enriched state, and it is only 41.2 % for glycerophospholipids. These averages are 94.9 %, 54.7 % and 37.0 % for the asymmetric models, respectively. It should be noted that the stripe-like pattern was not observed when visualizing lipid clustering (not shown except the example in Figure 6). This is not surprising because different metrics were used. In the clustering analysis, the cutoff distances determined the assignment of state; in the phase separation analysis, however, the local composition played an important role. With glycolipids being less populated in the model, using local composition as the metric could potentially introduce bias into HMM so that the lipids around glycolipids are more likely to be assigned to one state. More about this topic will be discussed in Section 4. When extending this analysis to a larger system, the stripe-like pattern becomes less evident, and a web-like pattern is seen where lipids in the same state interconnect with each other (Figure S12). However, the preference of being in the glycolipid-enriched state for each lipid type has not changed significantly (95.1 % for glycolipids, 55.9 % for sitosterol, and 42.0 % for glycerophospholipids). More discussion about this is available in Section 4.

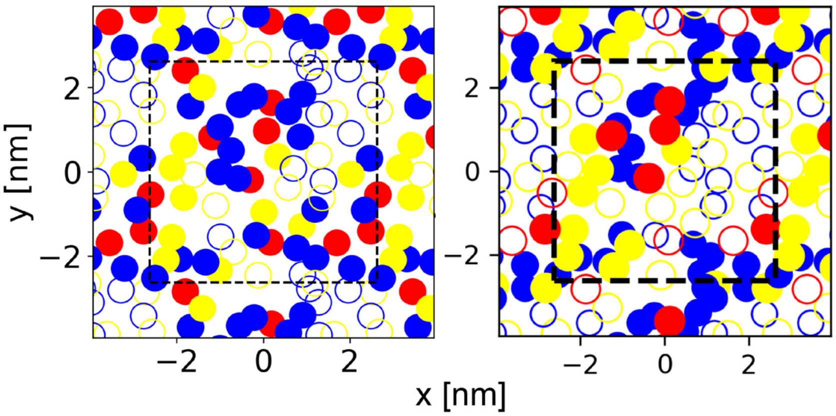


Figure 6. (Left) End-of-simulation states for lipids in the upper leaflet from replica 1 of the outer model. Glycolipids in pink, glycerophospholipids in orange, and sitosterol in green. Lipids in the glycolipid-enriched state are filled. (Right) End-of-simulation lipid clustering. Glycolipids in red, glycerophospholipids in yellow, and sitosterol in blue. Lipids in clusters are filled. The center boxes indicate the boundaries of the simulation cells.

3.7. Hydrogen Bonding. The probability to form inter-lipid hydrogen bond is shown in Figure 7 and the lifetime of hydrogen bond is shown in Table S9 and Table S10. In the inner model or the inner leaflet of the asymmetric membrane, DLiPE is the most active hydrogen bond donor, which has a high probability to form hydrogen bonds with DLiPS and PNPG. The driving force could be the positive charge carried by the amine group. Further analysis has shown that, among the DLiPE – DLiPS hydrogen bonds, about half were formed between the amine group on DLiPE and the phosphate group on DLiPS, and the other half were formed between the amine group and the carboxyl group on DLiPS. In DLiPE – PNPG hydrogen bonds, half were between amine and the phosphate on PNPG and the other half were between amine and the terminal glycerol group on PNPG. Since DLiPS also has an amine group, it can act as a donor. However, the negative net charge decreased the chance for DLiPS to form hydrogen bonds with other DLiPS and PNPG.

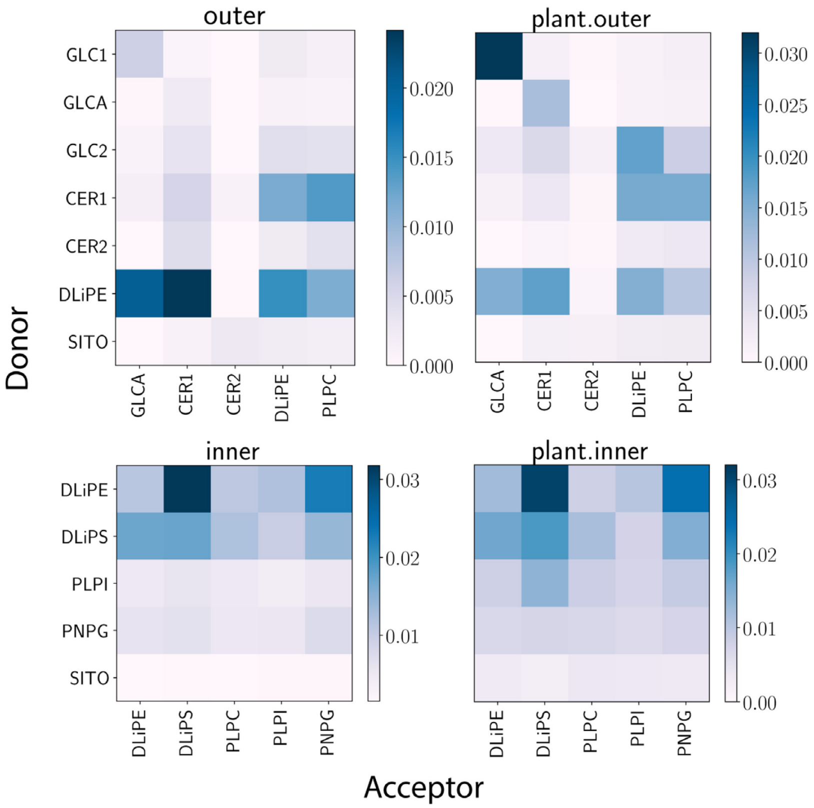


Figure 7. The probability to form inter-lipid hydrogen bond between lipid pairs. The probability is calculated as the average number of hydrogen bonds divided by the number of lipid pairs. Lipid and group (GLCA, CER1, CER2) definitions can be found in Figure 1.

The inter-lipid hydrogen bond pattern is slightly different between the outer model and the outer leaflet, and the differences are mostly related to the glycolipids. This is not surprising since the number of glycolipids in the outer model/leaflet is low, causing randomness in lipid distribution. For example, the probability to form hydrogen bond between GLCA (donor) and DLiPE (acceptor) is about 0.016 for the asymmetric membrane, but the error based on all the six replicas is 0.008. In both the outer and inner models, sitosterol formed little hydrogen bonds with other lipids, indicating the favorable interaction between sitosterol and GluCer/GIPC was not driven by hydrogen bonds.

There were also several lipids forming intra-lipid hydrogen bonds frequently (Table 8). The frequency to form intra-lipid hydrogen bonds for DLiPE and PNPG is pretty conserved between different systems and replicas. DLiPE can form intra-lipid hydrogen bonds between its amine hydrogens and the phosphate oxygens or the carbonyl oxygens. The two hydroxyl groups in PNPG formed most of its intra-hydrogen bonds with the phosphate oxygens. In most cases, GIPC formed intra-lipid hydrogen bonds between the inositol hydrogens and the amide oxygens. However, in the third asymmetric model, many intra-lipid hydrogen bonds (> 0.4 / lipid) were formed between the upper sphingosine hydroxyl and the lower phosphate ester oxygen. While the distance between the hydroxyl hydrogen and the ester oxygen is less than 2.4 Å without exception, the donor-hydrogen-acceptor angle could be influenced by the local lipid content. After plotting out the 2D-RDFs between GIPC and other lipid types, we found more sitosterols around GIPCs in the third asymmetric model (Figure S13), which could be an effect of the lower GIPC concentration. The higher local sitosterol concentration might have altered the conformation of the sphingosine backbone and made the donor-hydrogen-acceptor angle more favaroble. However, this difference between models should be considered as a deficiency of the small system size, which led to the unbalanced GIPC population between different models.

Table 8. Intra-lipid hydrogen bond per lipid ($N_{HB\text{-}intra}$ / lipid) for GIPC, DLiPE and PNPG.

lipid,	model	$N_{HB\text{-}intra}$ / lipid
GIPC	outer	0.094 ± 0.006
	plant.outer	0.239 ± 0.120
DLiPE	outer	0.069 ± 0.001
	plant.outer	0.067 ± 0.004
	inner	0.075 ± 0.002
	plant.inner	0.071 ± 0.002
PNPG	inner	0.479 ± 0.005
	plant.inner	0.476 ± 0.006

3.8. Electrostatic potential drop. The average electrostatic potential profile (EPP) for the symmetric and asymmetric membranes are shown in Figure 8. While the three asymmetric models differ slightly in their compositions, the EPPs for them are close to each other (Figure S14) so that only the average (plant) is reported in the main text. The total potential drop from the bilayer center to the water phase for the asymmetric membrane is closer to the outer model. Polarization density of water is also shown in Figure 8. For $z > 20$ Å and $z < -20$ Å, the average z component of water dipole is opposite to the bilayer normal, which must have contributed to the potential drop in this region. The potential drop within -20 Å < z < 20 Å could be related to the orientation of the lipid tail and thus the CH vectors. The water polarization density profile of the asymmetric membrane overall agrees with the symmetric models for the corresponding leaflets. However, due to different sizes of the simulation cell (in *z* direction), there are noticeable mismatches toward the two edges of the cell, and the influence on the EPP is not clear in the current study. It should be noted that the C36 lipid force field overestimated the potential drop of the dipalmitoylphosphatidylcholine (DPPC) bilayer by about two-fold,[55] so that similar overestimations should be expected for the model membranes studied here.

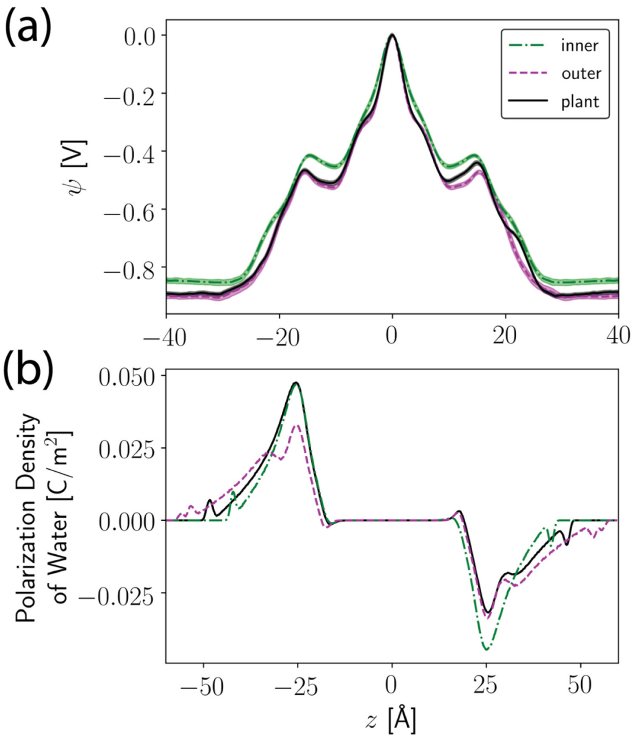


Figure 8. (a) Electrostatic potential profiles; (b) Polarization density of water in *z* direction.

3.9. Pressure profile. The tangential (to the bilayer normal) component of the pressure tensor is shown in Figure 9, which exhibits huge oscillations between positive and negative values. Compared to the pressure profiles obtained for pure phosphoglycerolipids,[69] there are additional peaks at $z = \sim \pm 10$ Å, which might be caused by sitosterols. The outer membrane has larger overall oscillations compared to the inner membrane, which is consistent with the larger compressibility modulus it exhibits. In the asymmetric membrane, pressure profile for each leaflet follows their corresponding symmetric models, though slightly modulated by the opposite leaflet.

For the asymmetric membrane, the surface tension of each leaflet was computed based on eq 5. The average surface tension for the inner leaflet is 6.0 ± 2.0 dyn/cm and the average for the outer leaflet is -6.0 ± 2 dyn/cm, which are small compared to the compressibility modulus, indicating a good match between leaflet areas. $K_c^m c_0$ is -0.15 ± 0.03 kcal/mol/Å for each leaflet in the inner model and -0.30 ± 0.03 kcal/mol/Å in the outer model, both indicating a positive spontaneous curvature of the monolayer. These numbers did not change drastically in the asymmetric model, as $K_c^m c_0$ values for the inner and outer leaflet are -0.21 ± 0.03 kcal/mol/Å and -0.29 ± 0.04 kcal/mol/Å, respectively.

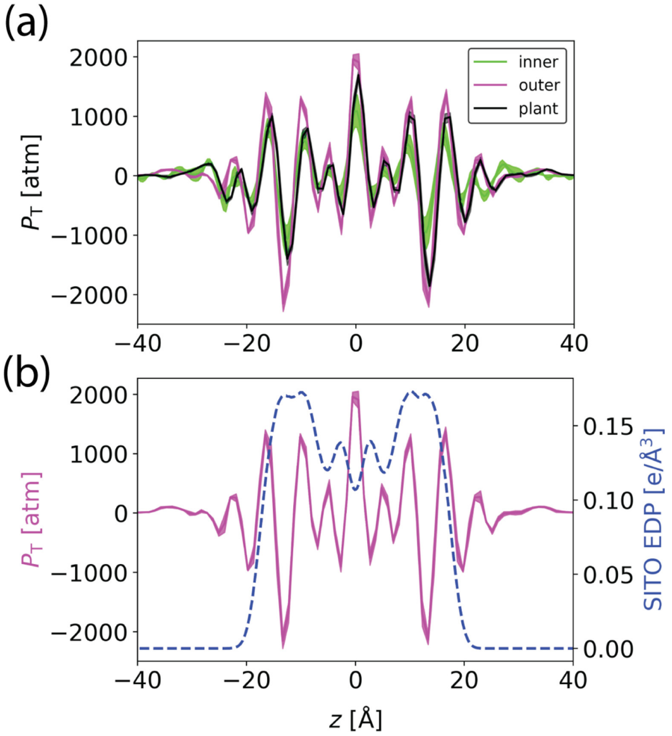


Figure 9. (a) the tangential pressure profile for the three model membranes. (b) The tangential pressure profile for the outer model aligned with the sitosterol EDP.

4. Discussion

In this study, we modeled the *Arabidopsis thaliana* plasma membrane using the CHARMM36 all-atom lipid force field. Two symmetric bilayer models representing the inner and outer leaflets of the membrane were simulated before merged to form the asymmetric membrane model. The major difference between the two leaflets in our model is that the outer leaflet contains glycolipids while the inner leaflet contains more glycerophospholipids but no glycolipid. The lipid composition was primarily determined based on Uemura et al.,[44] where the percentage of sphingolipids was determined to be 7 mol%. It should be noted that the proportion of sphingolipids varies among different plants. For example, Cacas et al.[20] measured GIPCs represent up to 40 mol % of total tobacco PM lipids, a number much higher than *Arabidopsis thaliana* reported by Uemura et al. Apart from the species gap which must have contributed to this difference, the lipid extraction method used by Uemura et al. might have caused incomplete solubilization of GIPCs as pointed out later by Markham et al.[47] Despite these uncertainties from experiments, MD simulation can provide valuable insights about how different lipid types interact with each other and how the lipid composition influences membrane properties.

For all the models simulated in this study, a high area compressibility modulus was observed, which is in line with our previous simuation study of soybean plasma membrane.[70] These results imply that plant membranes might be rigid in general, though more evidence is needed to corroborate this proposition. Compared to the inner model, the outer model is more rigid, which is caused by the higher sterol concentration and more saturated tails. K_A for the asymmetric model is closer to the outer model, indicating the non-additive nature of area compressibility. The glycolipids present in the outer model/leaflet likely also contributed to the higher K_A since they formed a fair amount of inter-lipid hydrogen bonds. The higher K_A for the outer model also agrees with the more condensed environment inferred from the component areas, the EDPs and the chain order parameters.

Due to the existence of sitosterol, pressure profiles for the model membranes exhibit more oscillations at z = ~ ± 10 Å compared to bilayers composed of pure glycerolphospholipids or sphingolipids.[69] As discussed by Ollila et al.,[74] these oscillations are likely utilized by the membrane to modulate the structure of membrane proteins since the associated energetics is at the order of a few k_BT. The calculated spontaneous curvature is negative for both the inner and outer leaflets, though the exact numbers were not determined. While a direct comparison with experiment or other simulation study is not available, this result is qualitatively in agreement with the result for DOPC (80%)/CHOL (20%) bilayer simulated by Ollila et al.[74] using a united-atom force field. Using X-ray diffraction, Chen et al.[75] observed that 5 mol% cholesterol in DOPC induced the inverted hexagonal (H_{II}) phases of very large dimension, and that increasing the level of cholesterol resulted in reduced lattice dimensions hence more negative spontaneous curvatures.

In addition, using small-angle X-ray scattering (SAXS), Kollmitzer et al.[76] measured the (partial) spontaneous curvature of cholesterol in the H_{II} phase to be −4.94 ± 0.13 Å, which was significantly more negative compared to PC, PE and sphingolipid. These results indicate that sterols may also introduce negative spontaneous curvature in a lamaller phase. Since the asymmetic membrane studied here is relatively rigid so that a high K_c^m is expected, which means a small $|c_0|$. Even though, one might expect the net spontaneous curvature to be positive in order to form a cell. One reasonable argument is that membrane proteins can modulate the curvature of the bilayer; another possible reason is that the percentage of GIPC in the outer leaflet might be higher than what Uemura et al.[44] measured, which likely has a positive (partial) spontaneous curvature because $K_c^m c_0$ for palmitoylsphingomyelin (PSM) is positive in the lamaller phase.[69]

Chain order parameters were computed for glycerophospholipids and the ceramide moiety in glycolipids. Statistical errors for the order parameter are generally small except for the sphingosine chain in ceramide. To investigate the relationship between the noisy S_{CD}'s and hydrogen bonding, Pearson correlation coefficients between these S_{CD}'s and each hydrogen bond pair involving CER1 (GIPC) were calculated based on replica averages. It turns out that S_{CD}'s of C2 and C3 do not have strong correlation with any hydrogen bond. However, the S_{CD}'s of C4-C6 have correlation coeffients (absolute value) over 0.7 with one or more hydrogen bonds (Table S11). These hydrogen bonds are mostly associated with the amide group of the ceramide, which is three bonds away from the C4 atom of the sphingosine chain.

The DBSCAN algorithm with pairwise cutoffs was used to study lipid clustering, which showed stability against different settings. While there was no lipid type showing significantly higher probability to form clusters, additional analysis has shown that sitosterol favors glycolipid-containing clusters. This is consistent with the 2D-RDFs, where sitosterol showed high affinity with glycolipids. HMM was used to detect potential compositional phase separation by grouping all the lipids into three major types (i.e., glycerophospholipids, glycolipids and sitosterol). Through this analysis, we have found that sitosterols prefer the glycolipid-enriched state, which agrees with the 2D-RDFs. However, the high probability (> 90 %) for glycolipids to be in the glycolipid-enriched state is concerning. Because of the small number of glycolipids in the model, one can imagine that lipids close to any glycolipid would have a much higher chance to be assigned to the glycolipid-enriched state. This could have biased the HMM process such that the glycolipid positions become definitive and other information is overlooked. Two different system sizes were used to investigate the size dependence of the compositional phase separation, and a change from the stripe-like pattern to the web-like pattern was observed when increasing the system size by four times. It is possible that the stripe-like pattern in the smaller system is a biased result caused by the periodic box condition, but it could also be that the stripes are just portions of the larger webs. Ideally, systems of tens of nanometers in both x and y directions should be used to study raft formation because the size of natural lipid rafts is at this order.[21] Nevertheless, these patterns and the affinity between sitosterol and glycolipids support a potential phase separation which could eventually lead to lipid rafts. Interestingly, we have found that glycolipids have unfavorable interactions with each other. In order to form lipid rafts, sitosterols must act as bridges connecting the glycolipids.

5. Conclusion

In conclusion, an asymmetric membrane model was built for the plasma membrane of *Arabidopsis thaliana*. The major difference between the two leaflets is that the outer leaflet contains glycolipids while the inner leaflet only has glycerophospholipids and sitosterol. The saturated tails from the glycolipids and the increased amount of sitosterol lead to a tighter packing of the outer leaflet. Clustering and compositional phase separation analyses have confirmed that sitosterol has a high affinity with glycolipids and potential phase separations were detected using HMM. The current models, especially model 1 for the asymmetric membrane which has the closest composition to the experiment, would aid in future simulation studies related to plant membranes, such as membrane partitioning and trans-membrane transport of small molecules.

6. Supporting Information Available

Tables of additional details of methods, experimental composition, additional similation parameters, starting points of area equilibrium, chain interdigitation, percentage of lipid in the glycolipid-abundant phase, hydrogen bond lifetimes, correlation coefficients between S_{CD} (CER1) and hydrogen bond; and figures of area vs. time, time evolution of SITO-SITO 2D-RDF, EDPs, S_{CD} for ceramides, 2D-RDFs between lipids of interest, potential of mean force between lipid pairs, clustering analysis, electrostatic potential profiles for asymmetric models.

7. Acknowledgments

The authors thank Soohyung Park and Richard M. Venable for the pressure tensor code in CHARMM and their helpful discussions on the pressure profile calculation. The authors also thank Alexander J. Sodt for sharing the HMM code. This research was supported in part by the Intramural Research Program of the NIH (National Heart, Lung, and Blood Institute, NHLBI), NSF (MCB-1951425), and the use of the high-performance computational capabilities at the Maryland Advanced Research Computing Center (MARCC), Deepthought2 at U. Maryland-College Park, and NHLBI LoBoS at the National Institutes of Health, Bethesda, MD.

8. References:

1. Koornneef, M.; Meinke, D., The Development of Arabidopsis as a Model Plant. *The Plant Journal* **2010,** *61* (6), 909-921.

2. Lloyd, A. M.; Barnason, A. R.; Rogers, S. G.; Byrne, M. C.; Fraley, R. T.; Horsch, R. B., Transformation of Arabidopsis Thaliana with Agrobacterium Tumefaciens. *Science* **1986,** *234* (4775), 464.

3. Valvekens, D.; Van Montagu, M.; Van Lijsebettens, M., Agrobacterium Tumefaciens-Mediated Transformation of Arabidopsis Thaliana Root Explants by Using Kanamycin Selection. *Proc. Natl. Acad. Sci. U. S. A.* **1988,** *85* (15), 5536-5540.

4. Koncz, C.; Martini, N.; Mayerhofer, R.; Koncz-Kalman, Z.; Körber, H.; Redei, G. P.; Schell, J., High-Frequency T-DNA-Mediated Gene Tagging in Plants. *Proc. Natl. Acad. Sci. U. S. A.* **1989,** *86* (21), 8467.

5. Feldmann, K. A.; David Marks, M., Agrobacterium-Mediated Transformation of Germinating Seeds of Arabidopsis Thaliana: A Non-Tissue Culture Approach. *Molecular and General Genetics MGG* **1987,** *208* (1), 1-9.

6. Leutwiler, L. S.; Hough-Evans, B. R.; Meyerowitz, E. M., The DNA of Arabidopsis Thaliana. *Molecular and General Genetics MGG* **1984,** *194* (1), 15-23.

7. Cheng, C.-Y.; Krishnakumar, V.; Chan, A. P.; Thibaud-Nissen, F.; Schobel, S.; Town, C. D., Araport11: A Complete Reannotation of the Arabidopsis Thaliana Reference Genome. *The Plant Journal* **2017,** *89* (4), 789-804.

8. Meyerowitz, E. M. a. S. C. R. e., *Arabidopsis*. Cold Spring Harbor Press: New York, 1994; p 1300 pp.

9. The Arabidopsis Information Resource. https://www.arabidopsis.org (accessed Oct 10, 2020).

10. Jones, A. M.; Chory, J.; Dangl, J. L.; Estelle, M.; Jacobsen, S. E.; Meyerowitz, E. M.; Nordborg, M.; Weigel, D., The Impact of Arabidopsis on Human Health: Diversifying Our Portfolio. *Cell* **2008,** *133* (6), 939-943.

11. Jones, J. D. G.; Dangl, J. L., The Plant Immune System. *Nature* **2006,** *444* (7117), 323-329.

12. Parry, G.; Estelle, M., Auxin Receptors: A New Role for F-Box Proteins. *Curr. Opin. Cell Biol.* **2006,** *18* (2), 152-156.

13. Deng, X. W.; Caspar, T.; Quail, P. H., Cop1: A Regulatory Locus Involved in Light-Controlled Development and Gene Expression in Arabidopsis. *Genes Dev.* **1991,** *5* (7), 1172-1182.

14. Yadeta, K. A.; Elmore, J. M.; Coaker, G., Advancements in the Analysis of the Arabidopsis Plasma Membrane Proteome. *Frontiers in plant science* **2013,** *4*, 86-86.

15. Singer, S. J.; Nicolson, G. L., The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* **1972,** *175* (4023), 720.

16. Engelman, D. M., Membranes Are More Mosaic Than Fluid. *Nature* **2005,** *438* (7068), 578-580.

17. Kusumi, A.; Fujiwara, T. K.; Chadda, R.; Xie, M.; Tsunoyama, T. A.; Kalay, Z.; Kasai, R. S.; Suzuki, K. G. N., Dynamic Organizing Principles of the Plasma Membrane That Regulate Signal Transduction: Commemorating the Fortieth Anniversary of Singer and Nicolson's Fluid-Mosaic Model. *Annual Review of Cell and Developmental Biology* **2012,** *28* (1), 215-250.

18. 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. Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 3-30.

19. Clouse, S. D., Arabidopsis Mutants Reveal Multiple Roles for Sterols in Plant Development. *The Plant Cell* **2002,** *14* (9), 1995.

20. Cacas, J.-L.; Buré, C.; Grosjean, K.; Gerbeau-Pissot, P.; Lherminier, J.; Rombouts, Y.; Maes, E.; Bossard, C.; Gronnier, J.; Furt, F.; Fouillen, L.; Germain, V.; Bayer, E.; Cluzet, S.; Robert, F.; Schmitter, J.-M.; Deleu, M.; Lins, L.; Simon-Plas, F.; Mongrand, S., Revisiting Plant Plasma Membrane Lipids in Tobacco: A Focus on Sphingolipids. *Plant Physiol.* **2016,** *170* (1), 367.

21. 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.

22. Stanislas, T.; Bouyssie, D.; Rossignol, M.; Vesa, S.; Fromentin, J.; Morel, J.; Pichereaux, C.; Monsarrat, B.; Simon-Plas, F., Quantitative Proteomics Reveals a Dynamic Association of Proteins to Detergent-Resistant Membranes Upon Elicitor Signaling in Tobacco. *Molecular & Cellular Proteomics* **2009,** *8* (9), 2186.

23. Titapiwatanakun, B.; Blakeslee, J. J.; Bandyopadhyay, A.; Yang, H.; Mravec, J.; Sauer, M.; Cheng, Y.; Adamec, J.; Nagashima, A.; Geisler, M.; Sakai, T.; Friml, J.; Peer, W. A.; Murphy, A. S., Abcb19/Pgp19 Stabilises Pin1 in Membrane Microdomains in Arabidopsis. *The Plant Journal* **2009,** *57* (1), 27-44.

24. van Meer, G.; Voelker, D. R.; Feigenson, G. W., Membrane Lipids: Where They Are and How They Behave. *Nature Reviews Molecular Cell Biology* **2008,** *9* (2), 112-124.

25. Saliba, A.-E.; Vonkova, I.; Gavin, A.-C., The Systematic Analysis of Protein–Lipid Interactions Comes of Age. *Nature Reviews Molecular Cell Biology* **2015,** *16* (12), 753-761.

26. Saliba, A.-E.; Vonkova, I.; Ceschia, S.; Findlay, G. M.; Maeda, K.; Tischer, C.; Deghou, S.; van Noort, V.; Bork, P.; Pawson, T.; Ellenberg, J.; Gavin, A.-C., A Quantitative Liposome Microarray to Systematically Characterize Protein-Lipid Interactions. *Nat. Methods* **2014,** *11* (1), 47-50.

27. Li, X.; Gianoulis, T. A.; Yip, K. Y.; Gerstein, M.; Snyder, M., Extensive in Vivo Metabolite-Protein Interactions Revealed by Large-Scale Systematic Analyses. *Cell* **2010,** *143* (4), 639-650.

28. Maeda, K.; Anand, K.; Chiapparino, A.; Kumar, A.; Poletto, M.; Kaksonen, M.; Gavin, A.-C., Interactome Map Uncovers Phosphatidylserine Transport by Oxysterol-Binding Proteins. *Nature* **2013,** *501* (7466), 257-261.

29. Laufer, C.; Fischer, B.; Huber, W.; Boutros, M., Measuring Genetic Interactions in Human Cells by Rnai and Imaging. *Nat. Protoc.* **2014,** *9* (10), 2341-2353.

30. Isakoff, S. J.; Cardozo, T.; Andreev, J.; Li, Z.; Ferguson, K. M.; Abagyan, R.; Lemmon, M. A.; Aronheim, A.; Skolnik, E. Y., Identification and Analysis of Ph Domain-Containing Targets of Phosphatidylinositol 3-Kinase Using a Novel in Vivo Assay in Yeast. *The EMBO Journal* **1998,** *17* (18), 5374-5387.

31. Contreras, F. X.; Ernst, A. M.; Haberkant, P.; Björkholm, P.; Lindahl, E.; Gönen, B.; Tischer, C.; Elofsson, A.; von Heijne, G.; Thiele, C.; Pepperkok, R.; Wieland, F.; Brügger, B., Molecular Recognition of a Single Sphingolipid Species by a Protein's Transmembrane Domain. *Nature* **2012,** *481* (7382), 525-529.

32. Hulce, J. J.; Cognetta, A. B.; Niphakis, M. J.; Tully, S. E.; Cravatt, B. F., Proteome-Wide Mapping of Cholesterol-Interacting Proteins in Mammalian Cells. *Nat. Methods* **2013,** *10* (3), 259-264.

33. Haberkant, P.; Holthuis, J. C. M., Fat & Fabulous: Bifunctional Lipids in the Spotlight. *Biochim. Biophys. Acta* **2014,** *1841* (8), 1022-1030.

34. Khakbaz, P.; Klauda, J. B., Probing the Importance of Lipid Diversity in Cell Membranes Via Molecular Simulation. *Chem. Phys. Lipids* **2015,** *192*, 12-22.

35. Monje-Galvan, V.; Klauda, J. B., Modeling Yeast Organelle Membranes and How Lipid Diversity Influences Bilayer Properties. *Biochemistry* **2015,** *54* (45), 6852-6861.

36. Wang, E.; Klauda, J. B., Examination of Mixtures Containing Sphingomyelin and Cholesterol by Molecular Dynamics Simulations. *J. Phys. Chem. B* **2017,** *121* (18), 4833-4844.

37. Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E., Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics* **2012,** *41* (1), 429-452.

38. Sodt, A. J.; Sandar, M. L.; Gawrisch, K.; Pastor, R. W.; Lyman, E., The Molecular Structure of the Liquid-Ordered Phase of Lipid Bilayers. *J. Am. Chem. Soc.* **2014,** *136* (2), 725-732.

39. Dror, R. O.; Green, H. F.; Valant, C.; Borhani, D. W.; Valcourt, J. R.; Pan, A. C.; Arlow, D. H.; Canals, M.; Lane, J. R.; Rahmani, R.; Baell, J. B.; Sexton, P. M.; Christopoulos, A.; Shaw, D. E., Structural Basis for Modulation of a G-Protein-Coupled Receptor by Allosteric Drugs. *Nature* **2013,** *503* (7475), 295-299.

40. Wu, Emilia L.; Engström, O.; Jo, S.; Stuhlsatz, D.; Yeom, Min S.; Klauda, Jeffery B.; Widmalm, G.; Im, W., Molecular Dynamics and Nmr Spectroscopy Studies of e.Coli lipopolysaccharide Structure and Dynamics. *Biophys. J.* **2013,** *105* (6), 1444-1455.

41. 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.

42. 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. *The FASEB Journal* **2010,** *24* (4), 1128-1138.

43. Gilmour, S. J.; Hajela, R. K.; Thomashow, M. F., Cold Acclimation in Arabidopsis Thaliana. *Plant Physiol.* **1988,** *87* (3), 745.

44. 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.

45. Guo, D.-a.; Venkatramesh, M.; Nes, W. D., Developmental Regulation of Sterol Biosynthesis in Zea Mays. *Lipids* **1995,** *30* (3), 203.

46. Grosjean, K.; Mongrand, S.; Beney, L.; Simon-Plas, F.; Gerbeau-Pissot, P., Differential Effect of Plant Lipids on Membrane Organization: Specificities of Phytosphingolipids and Phytosterols*. *J. Biol. Chem.* **2015,** *290* (9), 5810-5825.

47. Markham, J. E.; Li, J.; Cahoon, E. B.; Jaworski, J. G., Separation and Identification of Major Plant Sphingolipid Classes from Leaves. *J. Biol. Chem.* **2006,** *281* (32), 22684-22694.

48. Hsieh, T. C. Y.; Kaul, K.; Laine, R. A.; Lester, R. L., Structure of a Major Glycophosphoceramide from Tobacco Leaves, Psl-I: 2-Deoxy-2-Acetamido-D-Glucopyranosyl(A1 → 4)-D-Glucuronopyranosyl(A1 → 2)Myoinositol-1-O-Phosphoceramide. *Biochemistry* **1978,** *17* (17), 3575-3581.

49. Hsieh, T. C.; Lester, R. L.; Laine, R. A., Glycophosphoceramides from Plants. Purification and Characterization of a Novel Tetrasaccharide Derived from Tobacco Leaf Glycolipids. *J. Biol. Chem.* **1981,** *256* (15), 7747-7755.

50. Cacas, J.-L.; Furt, F.; Le Guédard, M.; Schmitter, J.-M.; Buré, C.; Gerbeau-Pissot, P.; Moreau, P.; Bessoule, J.-J.; Simon-Plas, F.; Mongrand, S., Lipids of Plant Membrane Rafts. *Prog. Lipid Res.* **2012,** *51* (3), 272-299.

51. Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W., Charmm-Gui Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* **2009,** *97* (1), 50-58.

52. 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.

53. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., Charmm: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009,** *30* (10), 1545-1614.

54. Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E., Scalable Molecular Dynamics on Cpu and Gpu Architectures with Namd. *J. Chem. Phys.* **2020,** *153* (4), 044130.

55. 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* (23), 7830-7843.

56. Durell, S. R.; Brooks, B. R.; Ben-Naim, A., Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem.* **1994,** *98* (8), 2198-2202.

57. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79* (2), 926-935.

58. Martyna, G. J.; Tobias, D. J.; Klein, M. L., Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994,** *101* (5), 4177-4189.

59. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R., Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995,** *103* (11), 4613-4621.

60. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977,** *23* (3), 327-341.

61. 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* (19), 8577-8593.

62. 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.

63. Steinbach, P. J.; Brooks, B. R., New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. *J. Comput. Chem.* **1994,** *15* (7), 667-683.

64. Chodera, J. D., A Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **2016,** *12* (4), 1799-1805.

65. Barber, C. B.; Dobkin, D. P.; Huhdanpaa, H., The Quickhull Algorithm for Convex Hulls. *ACM Trans. Math. Softw.* **1996,** *22* (4), 469–483.

66. Das, C.; Noro, M. G.; Olmsted, P. D., Simulation Studies of Stratum Corneum Lipid Mixtures. *Biophys. J.* **2009,** *97* (7), 1941-1951.

67. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X., A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, AAAI Press: Portland, Oregon, 1996; pp 226–231.

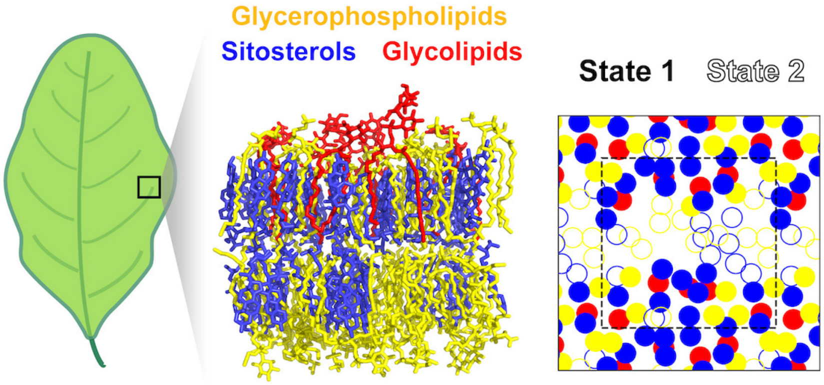
68. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É., Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011,** *12* (null), 2825–2830.

69. Venable, R. M.; Brown, F. L. H.; Pastor, R. W., Mechanical Properties of Lipid Bilayers from Molecular Dynamics Simulation. *Chem. Phys. Lipids* **2015,** *192*, 60-74.

70. 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.

71. Gracià, R. S.; Bezlyepkina, N.; Knorr, R. L.; Lipowsky, R.; Dimova, R., Effect of Cholesterol on the Rigidity of Saturated and Unsaturated Membranes: Fluctuation and Electrodeformation Analysis of Giant Vesicles. *Soft Matter* **2010,** *6* (7), 1472-1482.
72. Hofsäß, C.; Lindahl, E.; Edholm, O., Molecular Dynamics Simulations of Phospholipid Bilayers with Cholesterol. *Biophys. J.* **2003,** *84* (4), 2192-2206.
73. Mamode Cassim, A.; Navon, Y.; Gao, Y.; Decossas, M.; Fouillen, L.; Grélard, A.; Nagano, M.; Lambert, O.; Bahammou, D.; Van Delft, P.; Maneta-Peyret, L.; Simon-Plas, F.; Heux, L.; Jean, B.; Fragneto, G.; Mortimer, J. C.; Deleu, M.; Lins, L.; Mongrand, S., Biophysical Analysis of the Plant-Specific Gipc Sphingolipids Reveals Multiple Modes of Membrane Regulation. *J. Biol. Chem.* **2021,** *296*.
74. Samuli Ollila, O. H.; Róg, T.; Karttunen, M.; Vattulainen, I., Role of Sterol Type on Lateral Pressure Profiles of Lipid Membranes Affecting Membrane Protein Functionality: Comparison between Cholesterol, Desmosterol, 7-Dehydrocholesterol and Ketosterol. *J. Struct. Biol.* **2007,** *159* (2), 311-323.
75. Chen, Z.; Rand, R. P., The Influence of Cholesterol on Phospholipid Membrane Curvature and Bending Elasticity. *Biophys. J.* **1997,** *73* (1), 267-276.
76. Kollmitzer, B.; Heftberger, P.; Rappolt, M.; Pabst, G., Monolayer Spontaneous Curvature of Raft-Forming Membrane Lipids. *Soft Matter* **2013,** *9* (45), 10877-10884.

TOC Graphic:


Glycerophospholipids
Sitosterols
Glycolipids
State 1
State 2