Identifying Lipid Membrane Packing Defects in 3-Dimension: When Depth Does Matter.

Madhusmita Tripathy* and Anand Srivastava†

Molecular Biophysics Unit, Indian Institute of Science Bangalore, Bangalore 560012, India

Abstract

Being extremely difficult to characterize experimentally, lipid membrane packing defects and their role in the binding of peripheral proteins are generally investigated in computational studies. Such studies, while have been immensely successful in unraveling the key steps of the binding process, analyze the packing defects using their 2-dimensional projection, thus ignoring the crucial information on their depths. Here we present a simple yet computationally efficient algorithm, which identifies these defects in 3-dimension. We employ the algorithm to understand the nature of packing defects in flat bilayer membranes exhibiting liquid-ordered (L_o) and liquid-disordered (L_d) phases. Our results indicate the presence of shallower and smaller defects in the L_o phase membranes as compared to the L_d and mixed ones, in accordance to their distinct topological arrangements and temporal evolutions. Such analyses can elucidate the molecular scale picture behind the preferential binding of certain proteins to either of the liquid phases or their interface. Moreover, on the methodology front, our analyses suggest that the projection based 2-dimensional calculation of packing defects might result in inaccurate quantification of their sizes, thus advocating the importance of the 3-dimensional calculation.

Keywords

Lipid membrane packing defects, hydrophobic defects, defect localization, membrane-protein association, liquid ordered-liquid disordered phases.

*madhusmita@iisc.ac.in
†anand@iisc.ac.in

1 Introduction

Lipid membranes, a major cellular component, not only provide shape and stability to the cell and its organelles, but also play vital roles in many physiological functions, such as cell signaling and trafficking [1–4]. Many of these cellular functions are driven by membrane-protein interaction. One of the known feature that drives such interaction is the membrane curvature, where peripheral proteins preferentially adsorb to curved membranes [5, 6]. Another feature, which has recently been of great interest, is a highly transient aspect of lipid membranes: the lipid packing defects [7, 8]. These are identified as the surface of the hydrophobic lipid tails that are transiently exposed to the environment, thus also termed as hydrophobic defects. It has been shown that amphipathic motifs of peripheral proteins can sense these packing defects and subsequently bind to the lipid membrane by inserting their bulky hydrophobic residues into them [7–13].

Owing to their transient nature and small size, experimental characterization of lipid packing defects is extremely difficult. The same is indirectly inferred from experiments, where peripheral proteins preferentially partition into specific domains of the lipid membranes [10,12,14]. The contribution of computational tools and methods to characterize lipid packing defects, however, is remarkable. They have proven to be very useful in elucidating the process of binding of amphipathic motifs and their subsequent folding into α-helix in presence of packing defects, where the motif folding and defect coalescence has been shown to share a dynamic cooperativity [7, 8]. In this line, the scientific contributions from two groups, Voth and co-workers [7, 10] and Antonny and co-workers [8, 9, 11], need a special mention. The former group was among the first to characterize lipid packing defects in a model atomistic lipid membrane, exhibiting various state of curvature. They used a solvent accessible surface area (SASA) based method to identify the packing defects, which were then projected to a plane for subsequent analysis. The later group, on the other hand, used a 2-dimensional cartesian grid based method that scanned the lipid surface to identify exposed lipid tail atoms. Based on the nature of the atom encountered, they classified the defect to be geometric or chemical. Very recently, Antonny and co-workers optimized their algorithm and presented it in the form of a bioinformatic tool, named PackMem [15], where they also characterize packing defects to be shallow or deep. However, these elegant schemes still analyze the lipid packing defects in 2-dimension and present them as area, while in reality, such defects are 3-dimensional voids on the lipid membrane surface. Analyzing the defect projection is a technical simplification, where there is a possibility that some crucial information might get lost. The basic reason behind such simplification is the vast number of lipid atoms present in the membrane, which make a 3-dimensional grid search for atom overlap extremely time consuming. A scheme which can efficiently take care of the grid search should be able to fix this issue and identify the packing defects as voids in 3-dimension.

In this work, we present a simple, yet robust, algorithm to identify the lipid membrane packing defects in 3-dimension. Rooted from a free-volume based calculation, the algorithm identifies the de-

fect pockets by scanning a local (grid) volume around each atom in the lipid membrane. This is in contrast to the usual grid-based comparison, where each atom is compared to each grid point, in the reference box, to check for overlapping. The local search reduces the computational cost, thus making the algorithm efficient. We use this protocol to identify the packing defects in three different ternary lipid membrane systems, DPPC/DOPC/CHOL, PSM/DOPC/CHOL, and PSM/POPC/CHOL, each exhibiting pure L_o, pure L_d, and mixed L_o/L_d phases at various lipid compositions [16, 17]. While the nature of packing defects on curved membranes is well understood [7, 9], it is interesting to probe their nature on flat lipid membranes with co-existing liquid phases, where the distinct local topological rearrangements of lipids lead to characteristic nature of the ordered and disordered phases [18, 19]. We compare the spatial distributions of packing defects and find that the defects in the ordered and disordered phase of the membrane have distinctive features, an observation that can be attributed to the characteristics topological arrangement of lipids in the two phases. The distribution of defect pocket sizes turns out to be quite (remarkably) different from the 3-dimensional calculation, proposed here, and a projection based 2-dimensional calculation, raising questions on the validity of quantitative analyses from 2-dimensional calculation of lipid packing defects. The distribution of size and depths across these nine systems demonstrate some interesting features of lipid membranes exhibiting co-existing ordered and disordered phases.

The rest of the paper is organized as follows. The details of the atomistic simulation trajectories analyzed in this work are discussed in Section 2, along with a thorough discussion on the proposed algorithm to identify hydrophobic packing defects in lipid membranes. In Section 3, we provide and discuss results on the spatial, size and depth distributions of lipid packing defects for all the nine model lipid membrane systems studied in this work. Subsequently, to rigorously assess the usefulness of the methodology, we discuss the scope and limitations of the algorithm in Section 4. We conclude in Section 5 with our important results and probable future directions.

2 Methods

2.1 Atomistic Simulation Trajectories of Model Lipid Membranes

The set of atomistic simulation trajectories of lipid membrane system, which are used to benchmark the proposed algorithm, are borrowed from Edward Lymans group at the University of Delaware. The set consists of three different ternary lipid systems: DPPC/DOPC/CHOL, PSM/DOPC/CHOL, and PSM/POPC/CHOL, each simulated at three carefully chosen compositions at which the system exhibits pure L_o, pure L_d, and co-existing L_o/L_d phases. The simulations were carried out using NAMD on Anton supercomputer [20, 21] . CHARMM36 forcefield [22] was used to model the lipids, TIP3P [23] to model water and a forcefield from an earlier work [24] was used to model cholesterol. The entire trajectories

correspond to μs long simulation runs, where the system snapshots were collected at every 240 ps. Some of the important details of the trajectories are summarized in Table 1. Further details on the technicality of the system set up and simulation protocol can be found in the original references [16, 17].

Lipid system	Composition	Phase	Temperature (K)	Total Lipids	Analysis window (snapshots)
DPPC/DOPC/CHOL	0.55/0.15/0.30	Lo	298	512	500
	0.29/0.60/0.11	Ld	298	400	500
	0.37/0.36/0.27	Lo/Ld	298	380	500
PSM/DOPC/CHOL	0.64/0.03/0.33	Lo	295	560	193
	0.15/0.82/0.03	Ld	295	574	165
	0.43/0.38/0.19	Lo/Ld	295	452	313
PSM/POPC/CHOL	0.61/0.08/0.31	Lo	295	560	279
	0.23/0.69/0.08	Ld	295	560	500
	0.47/0.32/0.21	Lo/Ld	295	506	285

Table 1: Simulation details of the nine atomistic membrane systems analyzed in this work.

2.2 Calculating Hydrophobic Defects

To identify the packing defects on a membrane surface, we follow a grid-based free volume method. The defects are defined as the surface of the hydrocarbon tails of the lipids that are exposed to a probe, and are identified in a two-step calculation. In the first step, we identify the excluded or free volume of the surface of a membrane leaflet and in the second, the defect pockets are located. In the following sections, we discuss these two steps in details.

2.2.1 Identifying Membrane Surface Free Volume

To begin with, we consider one leaflet of the lipid bilayer lying on the XY plane. We define a cutoff Z value and select all the atoms of the monolayer that lie above the cutoff. The Z cutoff is arbitrary and can be efficiently chosen based on a simple visual inspection of the trajectory to be analyzed. A very low cutoff will lead to selection of a large number of atoms and increase the computational cost, while a very high cutoff may miss out important atoms over the trajectory. Therefore, a Z value is chosen which selects sufficient atoms from the monolayer, while making sure that all the lipids lie well above the Z cutoff over the entire trajectory. For simulation trajectories where the origin of the coordinate system is fixed at the center of the bilayer, Z=0 can serve as a decent choice. This selection renders a flat bottom to the membrane monolayer. The monolayer is then mapped to a 3-dimensional box with the flat bottom of the monolayer at the bottom face of the box. Periodic boundary conditions are applied along the X and Y directions, thus fixing the lateral boundaries as per the corresponding simulation box dimensions. The application of periodic boundary condition ensures precise identification of defects that are located near/at the boundaries of the simulation box. The height of the box is chosen to be sufficiently high, so as to keep all the surface atoms well within the box. The coordinates of all the atoms in the box

are then shifted such that the whole reference box lies in the positive XYZ quadrant, with origin at the corner of the bottom face. Subsequently, the box is mapped to a 3-dimensional grid of 1 Å resolution, and we identify and remove grid points that lie within one van der Waal (VDW) radius of any atom *plus* a probe distance. We use the VDW radius of atoms as specified by the forcefield and a probe distance of 1.4 Å, which is approximately the radius of a water molecule.

The general protocol for this part of the calculation is to compare the position of each grid point to that of each atom in the reference box to check for overlap. This comparison is usually achieved in two nested loops with either the grid index as the outer loop or the atom index and thus, is computationally expensive. We use a protocol that avoid such expensive nested loops over all atoms/grid points and rather compares an atom to only its neighboring grid points to check for overlapping. The idea is to consider a local 3-dimensional grid around each atom, in the selection, and compare the grid points within this local grid for possible overlap. As the corner of the reference box is kept at the origin and the grid spacing is 1 Å, we can identify the nearest neighboring grid points of any atom quite easily: it is the integer value of the position vector of the atom in any direction. In case the grid spacing was something different that 1 Å, it would be the integer value of the ratio of the position vector and the grid spacing. Thus, considering one atom of the monolayer, we can easily locate the grid points that lie within its VDW radius plus one probe distance in all directions. These grid points will constitute a local 3-dimensional grid around the central atom, under consideration. We now compare the distance between the position of the central atom and all the grid points within the local 3-dimensional grid around it. If this distance is found to be less than its VDW radius *plus* the probe distance, we remove the grid point as it overlaps with the central atom. We repeat the same procedure for all the atoms in the monolayer. The algorithm for this part of the calculation is given below.

```
initialize all grid points (of the entire box) to 0
for atom i in the box
    atom type[i] → t
    radius[t] → VDW radius[t] + probe distance
    lower boundary of the local box → (int)(r − radius[t])
    upper boundary of the local box → (int)(r + radius[t])
    check for overlapping of atom i position and grid points within the local box
    if grid point is overlaping
        grid point value → 1
    endif
endfor
```

At the end of this analysis, grid points with *grid point value* 0 constitute the free volume of the monolayer selection under consideration, i.e., the excluded volume of all the atoms in the box (selected from one monolayer). Starting from the rugged surface of the monolayer, this excluded volume extends all the way to the top of the reference box. One can visualize this excluded volume as a mould or cast of the monolayer under consideration. With the excluded volume information available, we next identify the defect pockets on the membrane surface.

2.2.2 Identifying Membrane Defect Pockets

The flat bilayer membranes that are studied in computer simulations are not essentially flat and exhibit transverse fluctuations that are much smaller in length scales as compared to the membrane's intrinsic curvature. These fluctuations can play a very important role in the formation of the defect pockets in such flat membranes. We, therefore, look for the defect pockets surrounding each lipid in its local neighborhood.

For this part of the analysis, we define the bottom carbon atom of the glycerol group, from where the two tails bifurcate, of lipids and the hydroxyl group oxygen atom of cholesterols as the reference sites. At this point, it should be noted that in earlier works on hydrophobic defects, the lipid membranes were composed of either a single lipid component or a mixture without cholesterol [7,8,11]. As cholesterols lie deeper within the membrane as compared to lipids, neglecting them will introduce inaccuracy in defect pocket calculation. Also, the interaction between water molecules and the lipids can get sufficiently modified in the presence of nearby cholesterols. Therefore, we also consider cholesterols while locating defect pockets. Intuitively, all the reference atoms never fall on a plane and, and so, do the local defects. We define defect pockets as local excluded volume, which we obtain from the earlier calculation, that lie in the neighborhood of the reference sites and extend below them along the Z-direction, thus exposing the hydrophobic tails of the lipids and the aromatic rings of the cholesterols. The neighborhood is defined around the reference sites in the XY plane, within a local 3-dimensional grid of length 3 *times* the probe distance *plus* the VDW radius of the reference atom. With these conditions, the free volume grid points that lie above the membrane surface are automatically discarded and we are left with local defect pockets lying on the membrane.

$initialize\ defect\ grid\ points\ to\ 0$

for $reference\ atom\ j\ in\ the\ box$

$atom\ type[j] \rightarrow tt$

$neighborhood[tt] \rightarrow VDW\ radius[tt]\ +\ (3\ \times probe\ distance)$

$lower\ boundary\ of\ the\ local\ box\ \rightarrow\ (int)(r_{ref}\ -\ neighborhood[tt])$

$upper\ boundary\ of\ the\ local\ box\ \rightarrow\ (int)(r_{ref}\ +\ neighborhood[tt])$

check for overlapping of reference atom j position and grid points within the local box

if *grid point lies within neighborhood*[*tt*] *of the reference site in XY plane and below* z_{ref}

defect grid point value $\rightarrow 1$

endif

endfor

The grid points with *defect grid point value* 1, constitute the defect pockets. The algorithms discussed in this and the last subsections provide the coordinates of grid points as output, which constitute the defect pockets on the monolayer at a given snapshot in the trajectory. This data is collected over the entire trajectory, which can be subsequently used to analyze the defect structure. Figure 1 shows the membrane defects identified using this method in a PSM/DOPC/CHOL bilayer system showing L_o phase. The figures have been rendered using the visualization tool, VMD [25].

As one can expect, this calculation can (and does) also identify grid points that lie well within the membrane as defects. One can employ conditions to avoid such occurrence, such as the defect pocket should always originate from the reference atom plane and extend below. However, as the lipid tails are very closely packed, such ambiguous defect pockets are statistically insignificant in number and does not affect the final results.

3 Results and Discussion

3.1 Spatial Distribution of Packing Defects

To understand the spatial distribution of defects in mixed phase (L_o/L_d) lipid membranes, we analyze their spatial maps, which were generated in the following steps. The defect pockets were first projected onto the XY plane of the grid. Ten such projections were collected from ten consecutive system snapshots. Subsequently, the set of projected defects were binned in XY plane with bin size 1 Å (the original grid spacing) and averaged. Thus, the colorbar in the map indicates the probability that a grid point is a part of a defect. A high probability, therefore, indicates the persistence of a defect pocket in the XY plane, i.e., a localized defect. It should be noted that the system snapshots, considered here, are collected 240 *ps* apart in time. Therefore, considering a large number of snapshots will be pointless as the defects are highly transient structures.

The ten steps averaged spatial maps of defects, calculated for the L_o/L_d mixed phase of the three systems, are shown in Fig. 2. Also shown are the top view (in the XY plane) of the corresponding membrane configurations, generated at the middle (the fifth) snapshots of the chosen sets. The L_o regions in all these mixed phase systems are rich in saturated lipids, DPPC in case of DPPC/DOPC/CHOL system

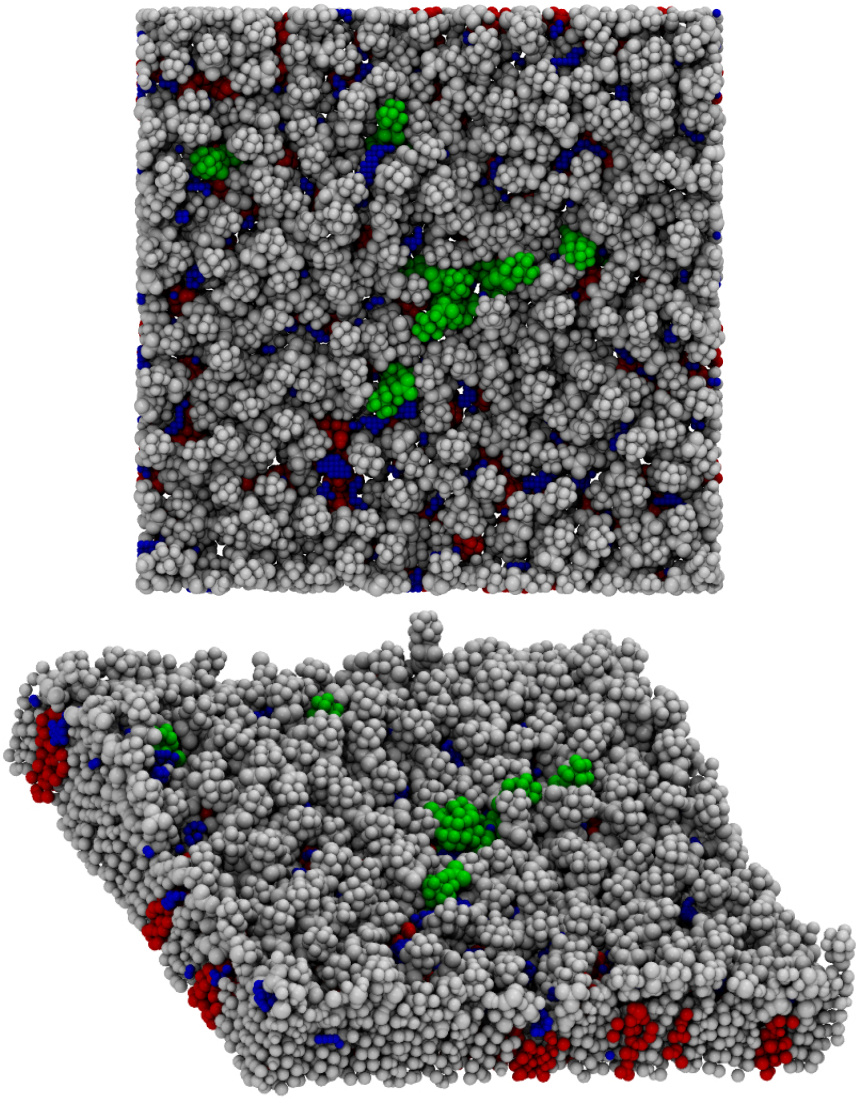

Figure 1: A top (top) and side (bottom) view of one leaflet of PSM/DOPC/CHOL bilayer system, exhibiting L_o phase, showing the membrane packing defects. Lipids and defects are shown using VDW representation: PSM in white, DOPC in green, cholesterol in red, and defects in blue.

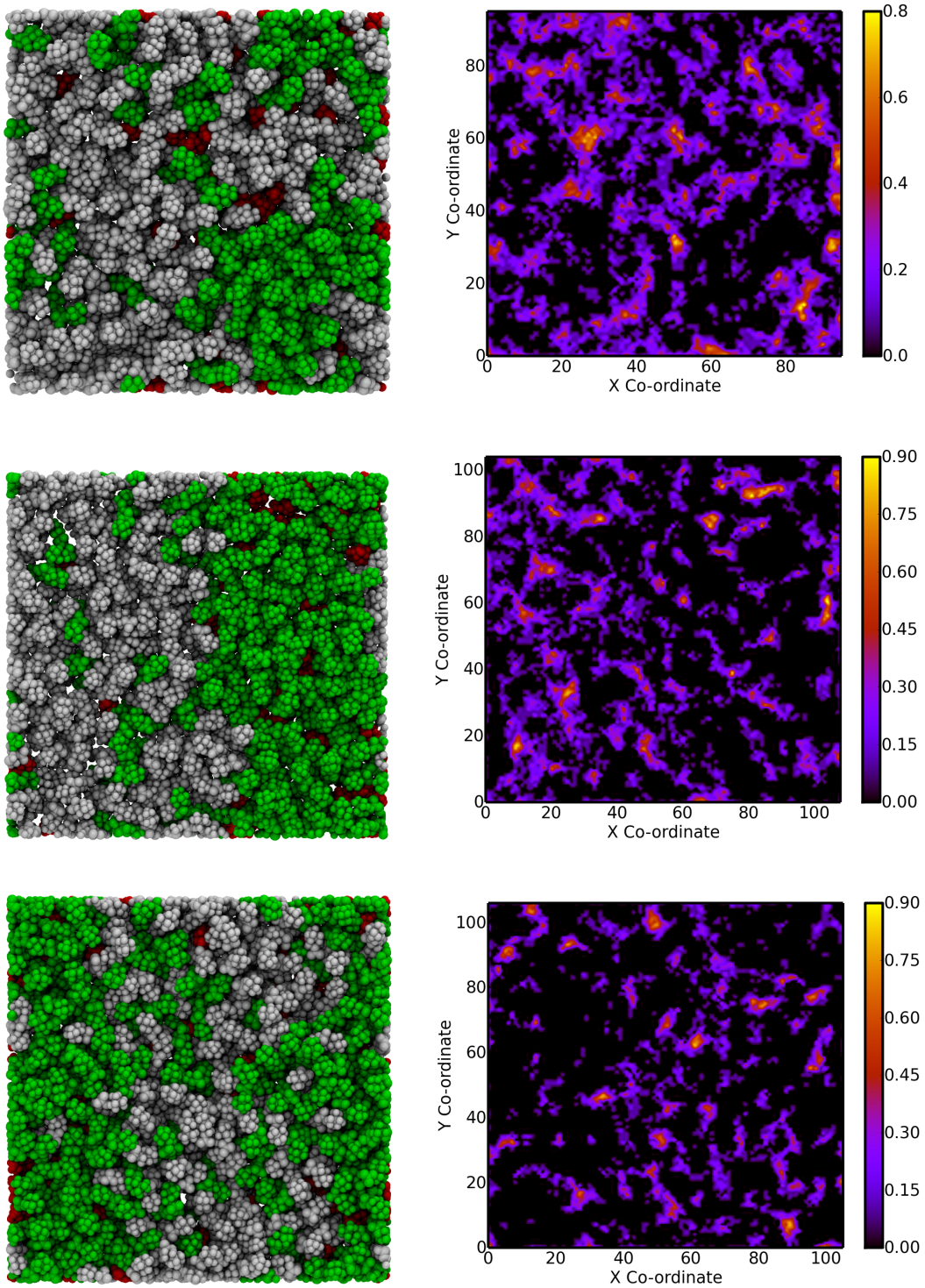

Figure 2: Membrane configurations (left panel) and spatial distribution maps of defects therein (right panel) for DPPC/DOPC/CHOL (top), PSM/DOPC/CHOL (middle), and PSM/POPC/CHOL (bottom) systems, each exhibiting L_o/L_d phases. All lipids are shown in VDW representation with saturated lipids (DPPC and PSM) in green, unsaturated ones (DOPC and POPC) in white, and cholesterols in red. The colorbars on right panel show the defect distribution count or occurrence probability. The axes shown are in Å units.

and PSM in case of PSM/DOPC/CHOL and PSM/POPC/CHOL systems. The straight configuration of the lipid tails is what leads to the compact ordered phase of these saturated lipids upon aggregation. As evident in the figure, the L_o regions of the membranes are found to be scarce in defect pockets: a direct consequence of the compact arrangement of the lipids therein, which does not allow wiggle space between the lipids such that the hydrophobic tails can get exposed. (The small patch on the bottom right in case of DPPC/DOPC/CHOl, right half in PSM/DOPC/CHOl, and near both boundaries along the X-axis in PSM/POPC/CHOL system.) On the other hand, the L_d regions of the membranes show a high density of defects. Also, given that the spatial maps are averaged over ten consecutive snapshots, one cannot fail to notice that the defects in the L_d regions have extended over the time, where as even the largest defects in the L_o regions seem to be localized, exhibiting high probability of defect occurrence. The large defects on top right corner in PSM/DOPC/CHOL system and those on the top left corner in PSM/POPC/CHOL system are few examples of such defects. Similar nature of defects are noticed in pure L_o and pure L_d membranes analyzed in this work (not shown here): while highly localized hydrophobic defects are associated with the ordered phase, the disorderedness of lipids in the L_d phase leads to a high density of defects and make them highly mobile. It is therefore evident that our algorithm is able to capture the distinct nature of defects in the ordered and disordered phases of a lipid membrane, which can be useful in understanding the preferential membrane-protein association at distinct phases of the membrane or at their interface [12, 14, 26].

In our earlier studies [18, 19], we observed that the compact topological rearrangement of lipids in the L_o regions make them dynamically steady, which should be the reason for the localization of the defects therein. On the other hand, comparatively faster evolution of lipids in the L_d phase can make the defects more dynamic. We are currently working on establishing any such possible correlation between the evolution of lipids in the membrane and the defect pockets therein, which will be discussed in our future publication.

3.2 Size Distribution of Packing Defects

The most general way to characterize the packing defects is to analyze the distribution of their sizes. The existing protocols to analyze packing defects, which deal with the projection of the defects on a grid plane, represent the defect size as defect area, which is the area of the square grids identified as part of defects. However, here we extract the packing defects as 3-dimensional voids, and thus the calculation of defect size becomes straightforward. We represent these voids as collection of spheres of diameter 1 Å (positioned at each grid point) and calculate defect size as the total number of grid points within it. We count this number by using a simple distance based clustering algorithm. where the grid points, identified from the $defect\ grid$ data (obtained from the algorithms discussed in the last section), which lie within a diagonal distance of each other are merged as a single cluster, and thus form a defect pocket.

Mathematically, the exact volume of the void and the size, calculated this way, differs only by a constant factor and therefore, the statistics should be comparable.

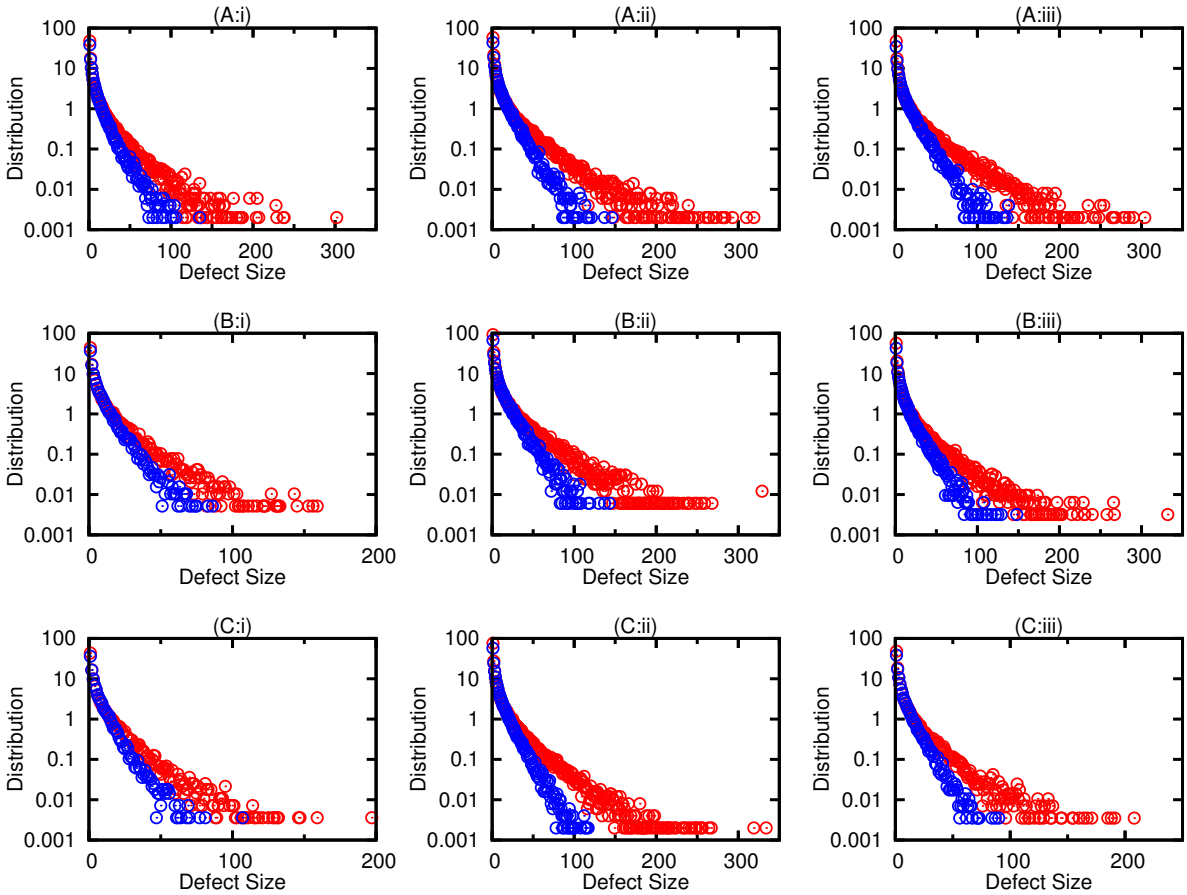

Figure 3: Distribution of defect pocket sizes in 3-dimension (red) and 2-dimension(blue) for (A) DPPC/DOPC/CHOL, (B) PSM/DOPC/CHOL, and (C) PSM/POPC/CHOL systems exhibiting pure L_o (i), pure L_d (ii), and mixed L_o/L_d (iii) phases. Defect size is calculated as the total number of grid points in a defect pocket.

In defect size distribution plots, the data points corresponding to very small defects, such as those less than 20, can be conveniently ignored as such small defects appear in all lipid membrane systems irrespective of their nature. On the other hand, very large defects are rare and thus have very low probability of occurrence. Therefore, only the defects with size in the intermediate range are statistically relevant. In general, the defect size distribution, in this range, is found to be exponential and so, the slope of the distribution, plotted in a semi-log scale, indicates the average defect size.

To understand the nature of packing defects in flat lipid membranes, exhibiting pure and mixed liquid phases, we calculate the distribution of their sizes. We first compare these distributions to those obtained from a 2-dimensional defect calculation. To calculate the defect size distribution in 2-dimension, we project the 3-dimensional defects onto the XY-plane and calculate their sizes using our clustering algorithm. In Fig. 3, we compare these distributions for all the nine systems studied in this work. As evident, the defect size distributions, in the range of interest, differ substantially for the two cases and

subsequently, the estimated value of average defect size. [[calculate]] Such an observation raises doubt on the validity of calculations that are based on a 2-dimensional analysis of defects. We, therefore, emphasize that studies intending to rigorously quantify lipid packing defects should identify them in 3-dimension rather than analyzing a projection of the same.

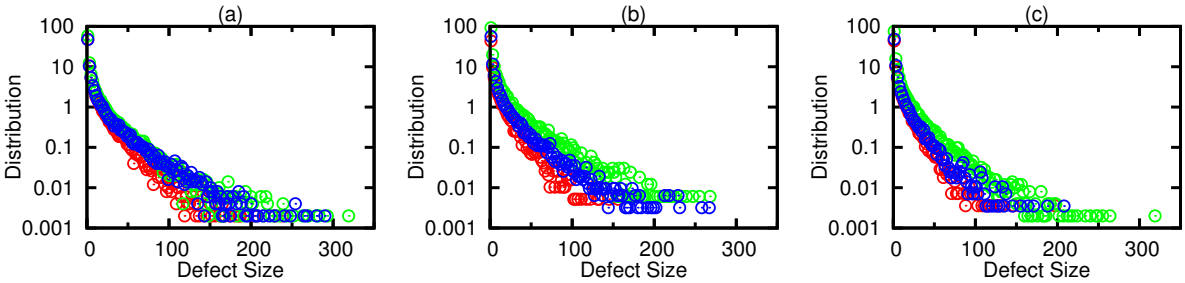

Figure 4: Distribution of defect pocket sizes for (a) DPPC/DOPC/CHOL, (b) PSM/DOPC/CHOL, and (c) PSM/POPC/CHOL systems exhibiting pure L_o (red), pure L_d (green), and mixed L_o/L_d (blue) phases. Every alternate data point is plotted for visual clarity.

To access the nature of the packing defects across the three lipid mixtures under consideration, we compare their size distributions in Fig 4. In all cases, the L_o phase is found to comprise smaller defect pockets as compared to the L_d phase. This observation can be attributed to the compact arrangement of lipids in an L_o phase that does not allow significant amount of voids to persist on the membrane. The mixed phase exhibit characteristics of both the ordered and disordered phases and therefore, the L_o/L_d phase in PSM/DOPC/CHOL systems is found to exhibit defect sizes that are intermediate to both the pure phases. However, the defect size distribution in the mixed phase DPPC/DOPC/CHOL membrane is found to be similar to that of the pure L_d phase, while that in PSM/POPC/CHOL membrane is similar to that of the pure L_o phase. Therefore, we believe that the mixed phase in DPPC/DOPC/CHOL system is dominated by the L_d subphase, while that in the PSM/POPC/CHOL system is dominated by the corresponding L_o subphase. The spatial distribution of defects shown in Fig. 2 also indicates the same behavior. However, rigorous analyses will be necessary to support this claim. Few of our current results support this view, that the nature of defects in a mixed phase membrane system is influenced by the richer phase, and will be discussed further in our upcoming publication.

3.3 Depth Distribution of Packing Defects

To further explore the importance of analyzing packing defects in 3-dimensional, we calculate the distribution of their depths. The depth of a defect is defined as its net extension along the Z-direction, which is again straightforward to calculate from the *defect grid* data. In Fig. 5, we present these distributions for all the nine systems under consideration.

As earlier, we ignore very shallow ($<$ 6 Å in depth) and very deep ($>$ 12 Å in depth) defects, and focus on the defects in the intermediate range of depths. For all the three lipid mixtures, the

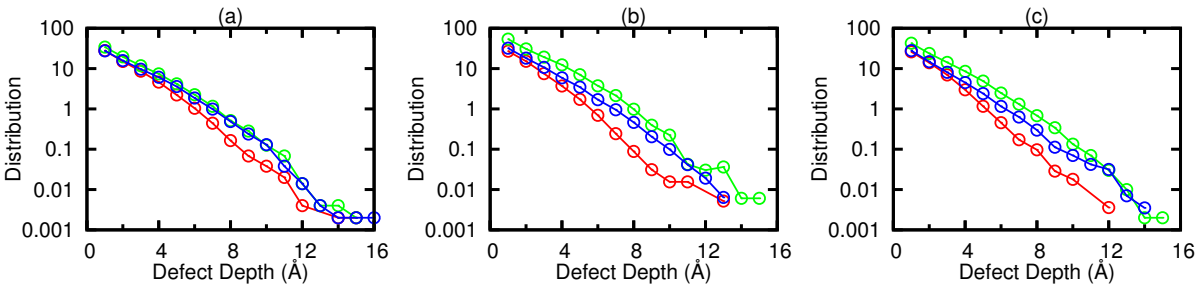

Figure 5: Distribution of defect pocket depths for (a) DPPC/DOPC/CHOL, (b) PSM/DOPC/CHOL, and (c) PSM/POPC/CHOL systems exhibiting pure L_o (red), pure L_d (green), and mixed L_o/L_d (blue) phases. Lines are guide to the eyes.

membrane exhibiting L_o phase comprise shallower defects than those in L_d and mixed phase L_o/L_d membranes. For example, in PSM/DOPC/CHOL and PSM/POPC/CHOL systems, the defect pockets with depth between 8 and 10 Å differ close to one order of magnitude in their distributions (occurrence probability) between the L_o and L_d cases. In DPPC/DOPC/CHOL systems, this difference is somewhat smaller, but nonetheless pronounced. The defects pockets in mixed L_o/L_d phase membrane are found to be comparable in depth to those in L_d phase membrane for DPPC/DOPC/CHOL systems, while for other two cases, mixed phase membranes exhibit shallower defect pockets as compared to disordered membranes.

The difference in defect depths across the L_o, L_d and L_o/L_d systems is again as expected. In our recent study [18, 19], we have observed the spatio-temporal evolution of lipids in the ordered and disordered regions of the membrane to be significantly different. This, along with the compact packing of lipids in the ordered regions, does not allow deeper defects to persist. On the other hand, fast moving disordered lipids in the L_d phases leads to larger and deeper defect pockets. Such information on depth distribution of membrane packing defects can be immensely important in the study of membrane-protein association [7–10], where deeper defects can facilitate protein anchoring and subsequent binding. It is needless to mention that such differences in terms of defect depths cannot be identified in a 2-dimensional projection based defect calculation. Though the defect analysis tool from Antonny and co-workers [15], PackMem, can identify defects to be shallow or deep, it cannot provide the exact measure of their depths.

4 Scope and Limitations

The algorithm proposed in this work, while efficiently handles the job of overlap check, does not impose any fixed parameters. Rather, users can choose various parameters as per their requirement. Firstly, by suitably choosing the Z-cutoff, the number of atoms in the leaflet under consideration can be greatly reduced, thus reducing the computation time. Secondly, for membranes with various species of biomolecules other than lipids, such as sterols, peptides, and proteins, one can choose the reference sites based on the

nature of these biomolecules, so as to correctly identify the packing defects around them. As discussed earlier, neglecting such molecules from the calculation can introduce inaccuracy in the defect pocket identification. Lastly, one also has the flexibility of choosing the probe distance and thus identify packing defects of specific nature. In our calculations, we have used a probe distance of 1.4 Å, the approximate radius of a water molecule, so that the defects are the hydrophobic surface available to water. One can also choose the same to be the radius of a bulky hydrophobic residue on a peptide (approximately 3 Å), as used in earlier analyses by other groups [7, 8].

The lipid membranes considered in our analyses are mostly flat and so, both leaflets are essentially identical in nature, in terms of curvature and composition. While we have analyzed the defects in only one of the leaflets, the same analyses can be performed on the other as well with a suitable choice of Z-cutoff. In fact, it is recommended to perform the analyses on both leaflets of the bilayer membrane, so as to improve the statistics of the distribution. Being a benchmarking exercise, we analyze defects in only one leaflet of the bilayer for all the systems, so the comparison is unbiased.

For curved membranes, where lipid packing can be drastically different between the inner and outer leaflets, the characteristics of defects on the two leaflets can be very distinct. For such curved membranes, a Z-cutoff cannot be used to select a subset of atoms from one of the leaflets. In stead, all atoms from one leaflet have to be chosen for analysis and the leaflet can be suitably moved to lie at the bottom boundary of the reference box. It should be mentioned, at this point, that for such a curved membrane, the first part of the proposed algorithm will identify excluded volume that will lie beneath the leaflet itself and so, can be identified as defects in the second part of the calculation. Therefore, to handle such ambiguity, we can impose additional constraints, such as, we can locate carbon atoms near the end of each lipid tail and remove all grid points below it. However, the proposed protocol, of checking overlapping between an atom and its neighboring grid points lying within a local 3-dimensional box, can be used to efficiently identify defect pockets in membranes exhibiting various state of curvature.

5 Conclusions

In this work, we provide a coherent framework to systematically characterize lipid membrane packing defects in 3-dimension. The protocol allows to choose a subset of atoms from a monolayer to analyze surface defect pockets and the algorithm follows a local grid search method to identify overlaps, thus making the framework computationally very efficient. For all the nine systems considered in this work, with very different system sizes as summarized in Table 1, it takes less than 1 second to identify defect pockets for a given snapshot of the trajectory. The algorithm takes care of periodic boundary conditions in the lateral directions, so that the defect pockets lying at and near the boundaries are correctly identified. We show that in comparison to a 2-dimensional defect calculation, based on the projection of defect

pockets on a plane, the proposed algorithm yields disparate distributions of defect sizes, thus raising questions about the reliability of the average defect sizes calculated from 2-dimensional calculations. Being able to capture the real 3-dimensional structure of the defect pockets, the algorithm also allows the calculation of defect depths. Such information can be crucial while studying membrane-protein association, which is well known to be facilitated by packing defects [7–10]. The framework is best suited for a flat lipid membrane. However, the same can be extended to curved membranes with a little additional cost of computation. Finally, the algorithm is very simple and thus allows modifications based on user requirements.

We used this method to identify defect pockets in 3 systems of flat lipid bilayers: DPPC/DOPC/CHOL, PSM/DOPC/CHOL, and PSM/POPC/CHOL, each exhibiting pure L_o, pure L_d, and mixed L_o/L_d phases. The L_o phase membranes were found to exhibit smaller and shallower defects than those in the L_d and L_o/L_d phases, which can be attributed to the compact arrangement of lipids in a L_o system. The mixed L_o/L_d phases were found to exhibit defects that are intermediate in size and depth to both the pure phase membranes. The data from the defect calculation are thus in line with the current understanding of membranes exhibiting pure and mixed liquid phases [16–19], and validates the proposed framework. We are currently in the process of mapping out the connection between the nature of defect pockets in such membranes and the spatio-temporal evolution of the membrane lipids. Such a connection can shed light on the molecular origin of the spatial heterogeneities observed in biological membranes that act as platforms for peripheral proteins to bind.

6 Acknowledgement

The authors thank Edward Lyman (University of Delaware, USA) for sharing the simulation trajectories and acknowledge Anton supercomputer facility for making the trajectories available. The financial support from Indian Institute of Science Bangalore are greatly acknowledged. AS thanks the startup grant provided by the Ministry of Human Resource Development, India and MT acknowledges the research fellowship from the Department of Biotechnology, India.

References

[1] Kai Simons and Elina Ikonen. Functional rafts in cell membranes. *Nature*, 387(6633):569, 1997.

[2] Katharina Gaus, Elena Chklovskaia, Barbara Fazekas de St Groth, Wendy Jessup, and Thomas Harder. Condensation of the plasma membrane at the site of t lymphocyte activation. *J Cell Biol*, 171(1):121–131, 2005.

[3] Gerrit van Meer, EH Stelzer, Roel W Wijnaendts-van Resandt, and Kai Simons. Sorting of sphingolipids in epithelial (madin-darby canine kidney) cells. *The Journal of cell biology*, 105(4):1623–1635, 1987.

[4] Katharina Gaus, Soazig Le Lay, Nagaraj Balasubramanian, and Martin A Schwartz. Integrin-mediated adhesion regulates membrane order. *The Journal of cell biology*, 174(5):725–734, 2006.

[5] Nikos S Hatzakis, Vikram K Bhatia, Jannik Larsen, Kenneth L Madsen, Pierre-Yves Bolinger, Andreas H Kunding, John Castillo, Ulrik Gether, Per Hedegård, and Dimitrios Stamou. How curved membranes recruit amphipathic helices and protein anchoring motifs. *Nature Chemical Biology*, 5:835, sep 2009.

[6] Vikram Kjøller Bhatia, Nikos S. Hatzakis, and Dimitrios Stamou. A unifying mechanism accounts for sensing of membrane curvature by bar domains, amphipathic helices and membrane-anchored proteins. *Seminars in Cell and Developmental Biology*, 21(4):381 – 390, 2010. Membrane Curvature Developmental Vertebrate Palaeontology.

[7] Haosheng Cui, Edward Lyman, and Gregory A. Voth. Mechanism of membrane curvature sensing by amphipathic helix containing proteins. *Biophysical Journal*, 100(5):1271 – 1279, 2011.

[8] Stefano Vanni, Lydie Vamparys, Romain Gautier, Guillaume Drin, Catherine Etchebest, PatrickF.J. Fuchs, and Bruno Antonny. Amphipathic lipid packing sensor motifs: Probing bilayer defects with hydrophobic residues. *Biophysical Journal*, 104(3):575 – 584, 2013.

[9] Stefano Vanni, Hisaaki Hirose, Hélène Barelli, Bruno Antonny, and Romain Gautier. A sub-nanometre view of how membrane curvature and composition modulate lipid packing and protein recruitment. *Nature Communications*, 5:4916 EP –, 09 2014.

[10] Coline Prvost, Morris E. Sharp, Nora Kory, Qingqing Lin, Gregory A. Voth, Robert V. Farese, and Tobias C. Walther. Mechanism and determinants of amphipathic helix-containing protein targeting to lipid droplets. *Developmental Cell*, 44(1):73 – 86.e4, 2018.

[11] Lydie Vamparys, Romain Gautier, Stefano Vanni, W.F. Drew Bennett, D. Peter Tieleman, Bruno Antonny, Catherine Etchebest, and PatrickF.J. Fuchs. Conical lipids in flat bilayers induce packing defects similar to that induced by positive curvature. *Biophysical Journal*, 104(3):585 – 593, 2013.

[12] Guillaume Kulakowski, Hugo Bousquet, Jean-Baptiste Manneville, Patricia Bassereau, Bruno Goud, and Lena K. Oesterlin. Lipid packing defects and membrane charge control rab gtpase recruitment. *Traffic*, 19(7):536–545, 2018.

[13] Leili Zhang, Manohary Rajendram, Douglas B. Weibel, Arun Yethiraj, and Qiang Cui. Ionic hydrogen bonds and lipid packing defects determine the binding orientation and insertion depth of reca on multicomponent lipid bilayers. *The Journal of Physical Chemistry B*, 120(33):8424–8437, 2016.

[14] Sung-Tae Yang, Volker Kiessling, and Lukas K. Tamm. Line tension at lipid phase boundaries as driving force for hiv fusion peptide-mediated fusion. *Nature Communications*, 7:11401 EP –, 04 2016.

[15] Romain Gautier, Amélie Bacle, Marion L. Tiberti, Patrick F. Fuchs, Stefano Vanni, and Bruno Antonny. Packmem: A versatile tool to compute and visualize interfacial packing defects in lipid bilayers. *Biophysical Journal*, 115(3):436 – 444, 2018.

[16] Alexander J Sodt, Richard W Pastor, and Edward Lyman. Hexagonal substructure and hydrogen bonding in liquid-ordered phases containing palmitoyl sphingomyelin. *Biophysical journal*, 109(5):948–955, 2015.

[17] Alexander J Sodt, Michael Logan Sandar, Klaus Gawrisch, Richard W Pastor, and Edward Lyman. The molecular structure of the liquid-ordered phase of lipid bilayers. *Journal of the American Chemical Society*, 136(2):725–732, 2014.

[18] Sahithya S. Iyer, Madhusmita Tripathy, and Anand Srivastava. Fluid phase coexistence in biological membrane: Insights from local nonaffine deformation of lipids. *Biophysical Journal*, 115(1):117 – 128, 2018.

[19] Madhusmita Tripathy, Sahithya S. Iyer, and Anand Srivastava. Chapter three - molecular origin of spatiotemporal heterogeneity in biomembranes with coexisting liquid phases: Insights from topological rearrangements and lipid packing defects. volume 28 of *Advances in Biomembranes and Lipid Self-Assembly*, pages 87 – 114. Academic Press, 2018.

[20] David E. Shaw, Ron O. Dror, John K. Salmon, J. P. Grossman, Kenneth M. Mackenzie, Joseph A. Bank, Cliff Young, Martin M. Deneroff, Brannon Batson, Kevin J. Bowers, Edmond Chow, Michael P. Eastwood, Douglas J. Ierardi, John L. Klepeis, Jeffrey S. Kuskin, Richard H. Larson, Kresten Lindorff-Larsen, Paul Maragakis, Mark A. Moraes, Stefano Piana, Yibing Shan, and Brian Towles. Millisecond-scale molecular dynamics simulations on anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, SC '09, pages 39:1–39:11, New York, NY, USA, 2009. ACM.

[21] D. E. Shaw, J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, A. Forte, J. Gagliardo, G. Gill, B. Greskamp, C. R. Ho, D. J. Ierardi, L. Iserovich, J. S. Kuskin, R. H. Larson, T. Layman, L. S. Lee, A. K. Lerer, C. Li, D. Killebrew,

K. M. Mackenzie, S. Y. H. Mok, M. A. Moraes, R. Mueller, L. J. Nociolo, J. L. Peticolas, T. Quan, D. Ramot, J. K. Salmon, D. P. Scarpazza, U. B. Schafer, N. Siddique, C. W. Snyder, J. Spengler, P. T. P. Tang, M. Theobald, H. Toma, B. Towles, B. Vitale, S. C. Wang, and C. Young. Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 41–53, Nov 2014.

[22] RichardM. Venable, AlexanderJ. Sodt, Brent Rogaski, Huan Rui, Elizabeth Hatcher, AlexanderD. MacKerell, RichardW. Pastor, and JefferyB. Klauda. Charmm all-atom additive force field for sphingomyelin: Elucidation of hydrogen bonding and of positive curvature. *Biophysical Journal*, 107(1):134 – 145, 2014.

[23] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

[24] Michael C. Pitman, Frank Suits, Alexander D. MacKerell, and Scott E. Feller. Molecular-level organization of saturated and polyunsaturated fatty acids in a phosphatidylcholine bilayer containing cholesterol. *Biochemistry*, 43(49):15318–15328, 2004. PMID: 15581344.

[25] William Humphrey, Andrew Dalke, and Klaus Schulten. Vmd: Visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[26] Xubo Lin, Alemayehu A Gorfe, and Ilya Levental. Protein Partitioning into Ordered Membrane Domains: Insights from Simulations. *Biophysical Journal*, 114(8):1936–1944, apr 2018.