JCB Journal of Cell Biology

SPOTLIGHT

Making lipids very unhappy to discover how they bind to proteins

Christopher Stefan[1] and Roberto Covino[2]

Membrane lipid composition is maintained by conserved lipid transfer proteins, but computational approaches to study their lipid-binding mechanisms are limiting. Srinivasan et al. (https://doi.org/10.1083/jcb.202312055) develop a clever molecular dynamics simulations assay to accurately model lipid-binding poses in lipid transfer proteins.

Advances in computational approaches that use machine learning to accurately predict structures of proteins and protein complexes, such as AlphaFold and trRosetta (1, 2), have transformed molecular biology. Yet even the most recent developments in these powerful structural prediction tools are still in the early stages of modeling protein interactions with small molecules, including lipids. Likewise, molecular docking tools often struggle to fit highly flexible lipids into binding pockets in proteins, and while new applications permit greater flexibility in binding states, the outcome is a static picture of the lipid-bound protein. Thus, despite tremendous recent advances in protein structure prediction and molecular docking tools, computational modeling of transient protein–lipid associations remains challenging. Atomistic molecular dynamics (MD) simulations offer deep insight into protein–lipid associations in rich molecular detail (3) but can be computationally costly. To address these computational challenges, Srinivasan et al. (4) have developed an elegant strategy that employs unbiased coarse-grained (CG) MD simulations to accurately predict lipid-binding sites in lipid transfer proteins.

Non-vesicular lipid transport pathways essential for membrane biogenesis and homeostasis are mediated by conserved lipid transfer proteins which extract and deliver lipids between cellular membranes. Lipid transfer proteins are generally classified into two groups: shuttle proteins that carry lipids between membranes and bridge proteins that form long hydrophobic tunnels between membranes. Knowledge of their molecular mechanisms of lipid transport is limited and primarily based on static structures of lipid transfer domains bound to lipids (5, 6). Yet, binding of lipid molecules in the hydrophobic pocket of a lipid transfer protein is an inherently dynamic process. The process occurs on very small temporal and spatial scales, challenging its experimental characterization. As mentioned, standard computational docking methods to predict the binding pose are inadequate because they do not consider dynamics. MD simulations are uniquely suited to overcome this challenge. However, lipids are typically embedded in bilayers, micelles, and other stable supramolecular assemblies. Sampling the spontaneous transition of a lipid from a bilayer to the pocket of a transfer protein requires crossing large free-energy barriers, which is usually impossible in equilibrium MD simulations. In their study, Srinivasan et al. (4) employ a simple yet very effective strategy. Instead of initializing a simulation in an equilibrium configuration, where lipids are happily part of a bilayer that they will not easily leave, the system is set up far out of equilibrium, with a deeply unhappy fully solvated lipid randomly spaced from the lipid transfer protein (Fig. 1). The lipid is highly dynamic and explores solutions to minimize the large free-energy penalty experienced from being solvated. Therefore, in this system, the lipid rapidly finds and associates with the lipid transfer protein's hydrophobic binding pocket (Fig. 1). For lipid transfer proteins that accommodate multiple lipids in a large cavity or channel, the lipid-binding process may be repeated successively until the cavity is fully occupied. Importantly, because the simulations are dynamic, they can provide information on lipid entry gates in the protein as well as stable binding sites.

To determine whether CG-MD simulations can accurately predict lipid-binding pockets, Srinivasan et al. (4) utilized crystal structures of 13 known lipid transfer domains bound to distinct lipid classes (glycerophospholipids, sphingolipids, or cholesterol). The atomistic protein structures were stripped of lipid, converted to CG models, and MD simulations (using the MARTINI 3 force field over μs timescales) were performed in the presence of a solvated lipid (Fig. 1). Remarkably, the authors could recover all known lipid-binding poses of all 13

[1]Laboratory for Molecular Cell Biology, University College London, London, UK; [2]Institute of Computer Science, Goethe University Frankfurt, and Frankfurt Institute for Advanced Studies, Frankfurt, Germany.

Correspondence to Christopher Stefan: c.stefan@ucl.ac.uk; Roberto Covino: covino@fias.uni-frankfurt.de.

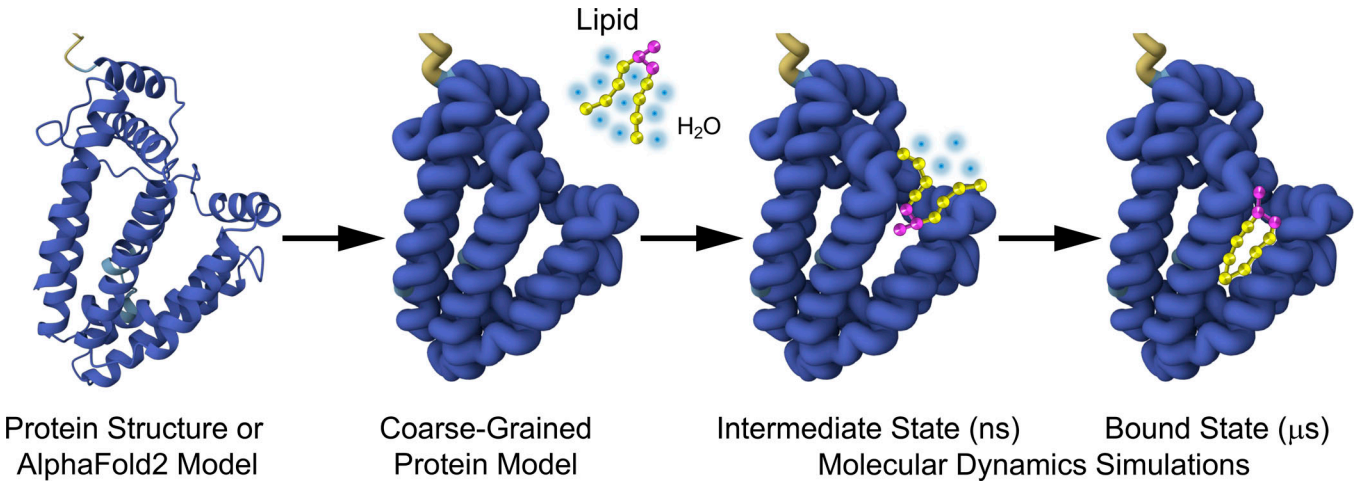

Figure 1. **Illustrative schematic of the unbiased CG-MD simulations used in the study by Srinivasan et al. (4).** The atomistic structure of a lipid transfer protein, either experimentally derived or predicted by AlphaFold2 (MIGA2 residues 307–567 AlphaFold2 model: blue), is converted to a CG model devoid of lipid and set up with a fully solvated lipid (lipid tails: yellow, lipid headgroup: magenta, H_2O: light blue spheres) randomly placed away from the protein. MD simulations are then performed over μs timescales to reveal lipid trajectories, the lipid entry gate into the protein (intermediate state, ns), and the lipid-binding pose in the hydrophobic pocket of the protein (bound state, μs). The fully solvated lipid is highly dynamic in the simulations and explores solutions to minimize the large free-energy penalty from being solvated. Thus, in this assay, the lipid tail solvation decreases as the lipid moves from the bulk solvent into the lipid transfer protein's hydrophobic binding pocket.

experimentally characterized lipid transfer proteins with the position of the CG lipid in the binding pocket closely resembling the position of the lipid determined in the experimental structures. Moreover, mapping of the lipid trajectories during the simulations revealed specific entry gates for each of the lipid transfer proteins tested. The CG-MD simulation protocol not only identified true positives (known lipid transfer proteins) but also true negatives (proteins that do not bind or transfer lipids). For each of the negative control proteins, the lipid remained highly solvated even if non-specifically associated with the protein surface. However, some limitations remain, as the CG-MD simulations cannot accurately predict the lipid-binding specificity of a lipid transfer protein, probably due to the CG representations of the proteins and lipids used for MD simulations. For example, the known cholesterol-binding protein NPC2 readily associated with glycerophospholipids and the known glycerophospholipid-binding proteins PCTP, PITP, and Sfh1 readily associated with cholesterol. Also, in a few cases the pipeline did not work right away and identifying the lipid-binding poses required an iterative "ad hoc" approach. The authors expressly point out these caveats in the manuscript.

The authors then built upon the CG-MD methods to identify large lipid-binding cavities in poorly characterized lipid transfer proteins and to gain further mechanistic insight into the bridge family of lipid transfer proteins, including Atg2 and Vps13. Previous studies have employed AlphaFold2 and trRosetta (1, 2) to reveal a large hydrophobic cavity in the human SNX25 protein as well as the budding yeast Nvj3 and Lec1/Ypr097w proteins that are formed by folding of the PXA and PXC domains in SNX25/Nvj3 and the so-called PXYn and PXYc domains in Lec1 (7, 8). These findings suggest that SNX25, Nvj3, and Lec1 may serve as lipid transfer proteins but experimental evidence for lipid binding and transfer is lacking. Indeed, iterative CG-MD simulations performed in the Srinivasan et al. study (4) indicate that the hydrophobic cavity in SNX25, Nvj3, and Lec1 may accommodate multiple lipids, supporting their proposed functions as lipid transfer proteins. Next, the authors examined the bridge-like lipid transfer proteins Atg2 and Vps13 that are proposed to bind multiple lipids within a long continuous hydrophobic channel formed by the chorein domains of these proteins (9, 10). Crystal structures of bridge-like lipid transfer proteins bound to multiple lipids are currently lacking. Following sequential rounds of lipid addition, the MD simulations predict that these bridge-like lipid transfer proteins can simultaneously bind multiple lipids (15 in Atg2 and 49 in Vps13) that span the entire length of a continuous hydrophobic channel. Finally, mutant forms of Atg2 and Vps13 bearing substitutions in hydrophobic residues were examined as a proof-of-concept that the CG-MD simulations can model lipid trajectories within the transfer protein. Strikingly, the MD simulations revealed bottlenecks in the mutant proteins where lipids did not reside, likely resulting in impaired lipid transfer.

The clever MD simulation assay by Srinivasan et al. (4) provides a powerful new approach to investigate the mechanisms of lipid entry, binding, and delivery by lipid transfer proteins. The CG representation used in the MD simulations may cause some limitations in the chemical precision of the predictions. However, these predictions can guide the design of more precise experiments, broadening our understanding of crucial aspects of membrane lipid dynamics in the cell.

Acknowledgments

C. Stefan is supported by UK Research and Innovation Biotechnology and Biological Sciences Research Council funding, award reference BB/X017184/1. R. Covino is supported by the Goethe University Frankfurt and the Frankfurt Institute for Advanced Studies.

References

1. Jumper, J., et al. 2021. *Nature*. https://doi.org/10.1038/s41586-021-03819-2
2. Yang, J., et al. 2020. *Proc. Natl. Acad. Sci. USA*. https://doi.org/10.1073/pnas.1914677117
3. Nishimura, T., et al. 2023. *Sci. Adv*. https://doi.org/10.1126/sciadv.adh1281
4. Srinivasan, S., et al. 2024. *J. Cell Biol.* https://doi.org/10.1083/jcb.202312055
5. Reinisch, K.M., and W.A. Prinz. 2021. *J. Cell Biol.* https://doi.org/10.1083/jcb.202012058
6. Wong, L.H., et al. 2019. *Nat. Rev. Mol. Cell Biol.* https://doi.org/10.1038/s41580-018-0071-5
7. Paul, B., et al. 2022. *Front. Cell Dev. Biol.* https://doi.org/10.3389/fcell.2022.826688
8. Castro, I., et al. 2022. *Elife*. https://doi.org/10.7554/eLife.74602
9. Levine, T., 2022. *Contact*. https://doi.org/10.1177/25152564221134328
10. Lees, J., et al. 2020. *Curr. Opin. Cell. Biol.* https://doi.org/10.1016/j.ceb.2020.02.008