JCB Journal of Cell Biology

TOOLS

Unbiased MD simulations identify lipid binding sites in lipid transfer proteins

Sriraksha Srinivasan[1*], Daniel Álvarez[1,2*], Arun T. John Peter[1], and Stefano Vanni[1,3]

The characterization of lipid binding to lipid transfer proteins (LTPs) is fundamental to understand their molecular mechanism. However, several structures of LTPs, and notably those proposed to act as bridges between membranes, do not provide the precise location of their endogenous lipid ligands. To address this limitation, computational approaches are a powerful alternative methodology, but they are often limited by the high flexibility of lipid substrates. Here, we develop a protocol based on unbiased coarse-grain molecular dynamics simulations in which lipids placed away from the protein can spontaneously bind to LTPs. This approach accurately determines binding pockets in LTPs and provides a working hypothesis for the lipid entry pathway. We apply this approach to characterize lipid binding to bridge LTPs of the Vps13-Atg2 family, for which the lipid localization inside the protein is currently unknown. Overall, our work paves the way to determine binding pockets and entry pathways for several LTPs in an inexpensive, fast, and accurate manner.

Introduction

Eukaryotic cells are organized into separate membrane-bound organelles, each with a characteristic lipid composition necessary for its proper functioning (van Meer et al., 2008). To achieve this complex homeostatic balance, lipids need to be selectively mobilized between different organelles in response to both extra- and intracellular stimuli (Fagone and Jackowski, 2009).

Intracellular lipid trafficking is largely achieved by two mechanisms—vesicular and non-vesicular transport. Non-vesicular lipid trafficking is mediated by a large group of lipid transfer proteins (LTPs), which encapsulate lipids in the hydrophobic cavity of a soluble lipid transfer domain (LTD) and transfer them between membranes of different organelles within the cell (Wong et al., 2017, 2019). Many LTPs have been discovered in the last few years (Gatta et al., 2015; Levine, 2022b; Kim et al., 2022), suggesting that intracellular lipid transport is more widespread and prominent than originally thought. Concomitantly, with the identification of new LTPs, new mechanisms of lipid transport are also being proposed (Wong et al., 2019; Dall'Armellina et al., 2023). Notably, certain LTPs have been suggested to transfer lipids between organelles by physically bridging them at membrane contact sites and by establishing long hydrophobic tunnels between the two organellar membranes (Wong et al., 2017, 2019; Reinisch and Prinz, 2021; Wozny et al., 2023).

Despite this plethora of new data on LTP cellular functions (Wong et al., 2017, 2019; Chiapparino et al., 2016), a detailed understanding of their mechanistic mode of lipid transport remains limited. This knowledge gap likely originates from the extreme and inherent molecular complexity of the process (Wong et al., 2017, 2019; Wong and Levine, 2016). Specifically, LTPs must recognize target donor and acceptor organellar membranes with high selectivity, and they must overcome non-negligible free energy barriers to extract and release lipids into these membranes. In addition, how transport directionality is achieved, as well as the auxiliary role of accessory proteins that bind LTPs and potentially affect the lipid transport cycle, remains largely unclear (Wong and Levine, 2016; Wong et al., 2017, 2019).

While structural studies are extremely helpful in this context, they are insufficient to paint a complete picture of the lipid transport mechanism. For example, even when a high-resolution structure of the LTP is available, the co-transported lipids are often missing, possibly as a consequence of the dynamic behavior of lipids inside the protein cavity. In addition, only few high-resolution structures of LTPs are available and, especially for those proposed to work via a bridge-like mechanism, AlphaFold (AF)-derived models are often used to provide a mechanistic interpretation of the functional data (Castro et al., 2022; Paul et al., 2022; Cai et al., 2022; Hanna et al., 2023; Guillén-Samander et al., 2022; Cai et al., 2022; van Vliet et al.,

[1]Department of Biology, University of Fribourg, Fribourg, Switzerland; [2]Departamento de Química Física y Analítica, Universidad de Oviedo, Oviedo, España; [3]Swiss National Center for Competence in Research Bio-inspired Materials, University of Fribourg, Fribourg, Switzerland.

*S. Srinivasan and D. Álvarez contributed equally to this paper. Correspondence to Stefano Vanni: stefano.vanni@unifr.ch.

2022; Neuman et al., 2022; Levine, 2022a; Dall'Armellina et al., 2023; Jumper et al., 2021). Finally, basic mechanistic features such as the exact entry/exit pathways for the loading and unloading of the lipids from the protein remain largely unknown.

To overcome these limitations, computational approaches hold promise to identify lipid binding poses for LTPs. A classical strategy in this regard is molecular docking (Bender et al., 2021). While this technique is widely used to determine ligand binding pockets in proteins (Forli et al., 2016), the high flexibility of lipid chains, possibly also inside the protein binding cavity, results in a major challenge in the case of LTPs. Recently, promising attempts to use artificial intelligence (AI)–derived tools to obtain the structure of protein–ligand complexes by directly building the protein structures around the corresponding molecules have been proposed (Krishna et al., 2023, *Preprint*), but these methods still perform quite poorly for low-affinity ligands (Krishna et al., 2023, *Preprint*). Alternatively, both unbiased atomistic molecular dynamics (MD) simulations, in which the ligand of interest is placed in the bulk solvent and eventually binds to the protein (Dror et al., 2011), as well as enhanced sampling MD methods (Niitsu et al., 2019; Limongelli et al., 2013), have been used to determine binding pathways for small molecules (Shan et al., 2022; Kuzmanic et al., 2020). However, since these processes occur over relatively long timescales, these approaches are inherently computationally expensive and thus difficult to extend to high-throughput investigations.

To alleviate these computational bottlenecks, a promising computational protocol based on unbiased coarse-grain (CG) MD simulations to determine protein–ligand interactions for small drug-like molecules has been recently proposed (Souza et al., 2020). There, the authors elegantly demonstrate that chemical-specific CG simulations (using the MARTINI force field [Souza et al., 2021]) can be used to simulate protein–ligand binding using a brute force approach, paving the way for potentially screening numerous drugs and proteins in a high-throughput fashion (Souza et al., 2020).

In this work, we explore the possibility of extending this protocol to the complex cases of lipids in LTPs. By tuning the computational parameters, we succeed in observing the spontaneous binding of lipids to LTPs and their eventual entry into the experimentally determined lipid binding pocket whenever this structural information is available. Our approach identifies a dominant entry pathway of lipids into LTPs, sheds light on the lipid-binding activity of poorly characterized proteins, and provides a structural view of the lipid density inside bridge-like LTPs (BLTPs), such as Vps13 and Atg2, that have been proposed to act as bulk lipid transporters but for which the lipid localization inside the cavity has remained uncharacterized so far. Our results pave the way for an improved characterization of the molecular mode of action of LTPs, and potentially for the in silico identification of new ones, in a high-throughput fashion at minimal computational cost.

Results

Brute force unbiased CG-MD simulations can reproduce crystallographic poses of lipids bound to LTPs

To determine whether unbiased CG-MD simulations can correctly identify lipid–protein interactions for LTPs, we first adapted a protein–ligand protocol recently proposed for small drug-like molecules using the CG MARTINI force field (Souza et al., 2020). In this protocol, drug-like molecules are placed in bulk solvent in the presence of a target protein, and during unbiased MD simulations, they can spontaneously relocate into the protein binding pocket due to the smoothness of the free energy profile in the CG representation (Cascella and Vanni, 2016). We reasoned that a similar protocol could work for lipid binding to LTPs even if lipid molecules carry extremely hydrophobic moieties: in the absence of alternative hydrophobic binding partners (such as a lipid bilayer) the lipid of interest could potentially be able to find the protein cavity of LTPs within reasonable time scales.

To test this hypothesis, we first identified several proteins that have been co-crystallized with a single lipid they are proposed to transport (Table 1, rows 1–13). These include ceramide-1-phosphate transfer protein (CPTP), fatty acid binding protein-1 (FABP1), mitoguardin-2 (Miga2), maintenance of lipid asymmetry protein C (MlaC), phosphatidyl-choline transfer protein (PCTP), neurofibromin-1 (NF1), oxysterol binding protein homolog-6 (Osh6), oxysterol binding protein homolog-4 (Osh4), phosphatidyl-inositol transfer protein-α (PITP), chromatin structure-remodeling complex subunit Sfh1, and StART domain-containing protein-11 (StARD11, also called ceramide transfer protein [CERT]), apolipoprotein M (ApoM), and Niemann-Pick disease type C2 protein (NPC2).

Starting from the x-ray structures, we next set up the following protocol: each protein was stripped off its co-crystallized ligand(s), placed in a box of water with 0.12 M NaCl with a lipid randomly inserted into the bulk solvent, and 1-µs-long unbiased MD simulations were performed (Fig. 1 A and Video 1). For each protein, five independent replicas were carried out. Hydrophobic tails of all simulated lipids were initially modeled as oleoyl tails (18:1), while headgroups were chosen to be similar to the ones in the corresponding crystal structure of the protein (Table 1). In case parameters for a given lipid in the crystal structure were unavailable in the MARTINI force field, structurally similar headgroups were chosen for comparison. For example, 1,2-dioleoyl-sn-glycero-3-phosphate (DOPA) was used instead of ceramide-1-phosphate.

To assess lipid binding, we used two distinct metrics. First, we collected the time traces of the minimum distance between the protein and the lipid (Data S1 A). Second, to discriminate between bona fide binding events inside the lipid-binding pocket versus futile lipid-binding events on the external surface of the protein, we used as a metric the solvation number of the lipid tails (defined as the number of water molecules within 5 Å of the lipid acyl chains) in the last 100 ns of the trajectory of each replica (Fig. 1 B). We selected this property since a lipid bound to an external surface of the protein would be surrounded by a larger number of water molecules as opposed to lipids bound to a hydrophobic cavity. Direct comparison between lipid solvation, the minimum distance between lipid and protein, and the center of mass distance between the CG lipid and the x-ray ligand (Data S1 A) confirms that lipid solvation is a good metric for lipid binding, marking a clear distinction between externally bound lipids and lipids that bind within the hydrophobic cavity.

Table 1. **Details of the LTPs tested in this study**

LTP	Structure source	Residues	Target Membrane	Proposed mechanism	Ligand in crystal structure	Lipids tested in MD simulations
CPTP	4K8N (Simanshu et al., 2013)	8–214	Golgi, plasma membrane, endosome, nucleus	Shuttle	Ceramide-1-phosphate	DOPA, DOPC, DPPA, DPSM, Chol, TO
MlaC	5UWA (Ekiert et al., 2017)	24–208	Bacterial outer membrane, bacterial inner membrane	Shuttle	DOG	DOPE, DOPS, DPPE, DPSM, Chol, TO
PCTP	1lN1 (Roderick et al., 2002)	8–210	-	Shuttle	DLPC	DOPC, DOPS, DPPC, DPSM, Chol, TO
NF1	2E2X (Welti et al., 2007)	1,567–1,816	Plasma membrane, nucleus, nucleolus	Shuttle	POPE	DOPE, DOPC, DPPE, DPSM, Chol, TO
PITP	1T27 (Yoder et al., 2001)	2–270	Nucleus	Shuttle	DOPC	DOPC, DOPS, DPPC, DPSM, Chol, TO
Sfh1	3B74 (Schaaf et al., 2008)	101–274	Nucleus	Shuttle	DOPE	DOPE, DOPS, DPPE, DPSM, Chol, TO
FABP1	1O8V (Jakobsson et al., 2003)	1–133		Shuttle	Palmitic acid	DOPE, OLAC
Osh4	1ZHY (Im et al., 2005)	2–354	Golgi, ER	Shuttle	Chol	Chol
NPC2	5KWY (Li et al., 2016)	20–151	ER, lysosome	Shuttle	Chol sulfate	DOPC, DPPC, DPSM, Chol, TO
ApoM	2YG2 (Christoffersen et al., 2011)	38–188	-	Shuttle	S1P	OLAC, DOPC, DPPC, DPSM, Chol, TO
Miga2	8EDV (Hong et al., 2022)	307–567	Mitochondria	Shuttle	DPPE	DOPE, OLAC, DPPE, DPSM, Chol, TO
StARD11	2E3R (Kudo et al., 2008)	364–598	Golgi, ER	Shuttle	C18-Ceramide	DOPA, DOPC, DPPC, DPSM, Chol, TO
Osh6	4B2Z (Maeda et al., 2013)	1–434	Plasma membrane, ER	Shuttle	DSPS	DOPC, DOPS
FABP1 Δ6	1O8V	7–133	-	-	Palmitic acid	DOPE, OLAC, DPPE, DPSM, Chol, TO
Osh6-Δ69	4B2Z	69–434	-	-	DSPS	DOPC, DOPS, DPPS, DPSM, Chol, TO
Osh4-Δ29	1ZHY	30–354	-	-	Chol	DOPC, DPPC, DPSM, Chol, TO
JHBP	2RQF (Suzuki et al., 2011)	1–225	-	-	Juvenile hormone III	DOPC, DOPS
Takeout	3E8T (Hamiaux et al., 2009)	5–220	-	-	Ubiquinone-8	DOPC, DOPS
Svf1	AF-Q05515-F1	1–481	Golgi, ER, nucleus	?	-	DOPC
SNX25	AF-Q9H3E2-F1	1–840	Endosome	?	-	DOPC
AsmA	AF-P28249-F1	27–577	Plasma membrane	?	-	DOPC
Mdm31	AF-P38880-F1	141–579	Mitochondrion	?	-	DOPC
Lec1	AF-Q06839-F1	1–1,073	Mitochondrion	?	-	DOPC
Nvj3	AF-Q03983-F1	71–463	Nucleus, vacuole, ER	?	-	DOPC
Atg2	6A9J (Osawa et al., 2019)	2–379	ER, preautophagosomal membrane	Bridge	DOPE	DOPE, DOPS
Vps13	6CBC (Kumar et al., 2018)	1–329	Nucleus, vacuole, mitochondrion, endosome, peroxisome, prospore membrane	Bridge	-	DOPC, DOPE
E-Syt2	4P42 (Schauder et al., 2014)	191–370[a]	Plasma membrane, ER	Shuttle	DOG, detergent	DOPA, DOPC
Atg2	AF-P53855-F1	1–1,344	ER, preautophagosomal membrane	Bridge	-	DOPC

Table 1. **Details of the LTPs tested in this study (*Continued*)**

LTP	Structure source	Residues	Target Membrane	Proposed mechanism	Ligand in crystal structure	Lipids tested in MD simulations
Csf1	AF	1–2,958 Fragments: 1–1,000 1,001–2,000 2,001–2,958	Plasma membrane, ER, mitochondrion	Bridge	-	DOPC
Vps13	AF	1–1,859 Fragments: 1–1,048 948–1,859	Endosome, mitochondrion, prospore membrane, vacuole, nucleus, peroxisome, Golgi	Bridge	-	DOPC

Abbreviations: Chol, cholesterol; OLAC, oleic acid; DOG, dioleyl-glycerol; DLPC, dilauroyl-phosphatidylcholine; POPE, 1-Palmitoyl-2-oleoyl-phosphatidylethanolamine; DSPS, distearoyl-phosphatidylserine; S1P, sphingosine 1-phosphate. The four-letter structure IDs are from the RCSB PDB while the ones starting with "AF" are from AlphaFold.
[a]Both chains A and B of the LTP were considered for the simulation.

For the 13 proteins we tested, we observed two distinct behaviors: in some cases (8/13: CPTP, MlaC, NF1, PCTP, PITP, Sfh1, ApoM, NPC2; Fig. 1 B), the lipid could always bind spontaneously and irreversibly to the protein of interest, as indicated by the time traces of the minimum distances between the protein and the lipid and of the lipid tail solvation (Data S1 A). Solvation analysis indicates that the lipid tails are desolvated at the end of the trajectory, thus residing inside a hydrophobic binding pocket (Data S1, A and B, light purple).

In other cases (4/13: FABP1, StARD11, Osh6, Osh4), no entry of the lipid was observed (Data S1, A and B, gray), and the lipid could not find the binding pocket within the 1-μs elapsed simulation time. In one case (Miga2), the lipid entered the binding pocket only in three of the five replicas (Data S1 A and Fig. 1 B).

Next, we focused our attention on the false negative results (Fig. 1 B). We reasoned that the origin of such behavior could be related to high free energy barriers originating from protein dynamics, conformational changes potentially associated with lipid entry in physiological conditions and that are not well-reproduced by our CG simulations, or the inability of the lipid to find the entry pathway during the simulation time because of too limited sampling time. To qualitatively test these hypotheses, we performed three tests. First, we extended our simulations from 1 to 3 μs. This approach improved our results for Osh6 (Data S1, A and B), with two replicas displaying lipid binding in the extended simulation, and for Miga2, in which we observed lipid binding in one additional replica (Data S1, A and B). Next, we tested our protocol on lid-less Osh6 (Osh6-Δ69) and Osh4 (Osh4-Δ29) since oxysterol-related proteins (ORP) domains, like other LTDs, are known to have a lid-like helical region at the N-terminus that regulates lipid entry into the protein (Lipp et al., 2019; Moser von Filseck et al., 2015). Indeed, while we observed lipid entry for full-length Osh6 in only two replicas (in 3 μs), and no lipid binding for full-length Osh4, all five replicas displayed lipid entry for both Osh6-Δ69 and Osh4-Δ29 (Fig. 1 B and Data S1 A). Interestingly, in two of the replicas (see Data S1 A), the cholesterol molecule bound to Osh4-Δ29 exits the hydrophobic cavity, coming back to the bulk solvent. This phenomenon can be related to the protective role that the lid may have on the cavity, preventing the ligand from interacting with the solvent and thus unbinding. Analogously, the first six residues of FABP1 could also act as a lid, closing the entry to the hydrophobic cavity. However, the removal of the small lid did not lead to binding, although one replica underwent binding and unbinding of oleic acid (see Data S1 A).

Third, we opted to decrease the elastic network force constant that keeps the protein conformation close to its crystallographic structure in CG simulations, thus promoting increased flexibility. Indeed, upon reduction of the force constant (from 1,000 to 300 kJ mol^{-1} nm^{-2}), we could observe spontaneous lipid entry in the cavity of StARD11 (Fig. 1 B and Data S1 A).

Finally, to test whether the desolvated lipid-binding pose we observed is consistent with the one determined by x-ray crystallography, we computed the distance between the center of mass of the hydrophobic tails in the experimental structure and the corresponding one of the CG lipids, exclusively for MD frames, with a low lipid tail solvation (<2 water molecules, consistent with the mean value over the positive LTPs of 1.5 molecules on average) (Fig. 1 C).

Even though our simulations display a large variability, in all cases, we could observe values very close to the x-ray distance (Fig. 1, C and D), thus identifying binding poses very close to the crystallographic structure, with the sole exception of StARD11, where CG-MD simulations suggest a lipid binding pose with the lipid tails located more deeply inside the protein cavity than in the corresponding structure from x-ray crystallography (Fig. 1 D). The observation that our protocol provides additional desolvated lipid poses with respect to the crystal structure is potentially interesting. On the one hand, this observation could be simply attributed to the inaccuracy of our CG approach. On the other hand, this observation could originate from the ability of our protocol to identify additional lipid-binding poses. These conformations might be important along the lipid entry/exit pathway, and their presence is consistent with the hypothesis that the lipid must not be trapped in highly stable states inside the pocket to facilitate its uptake and release from membranes by the LTPs.

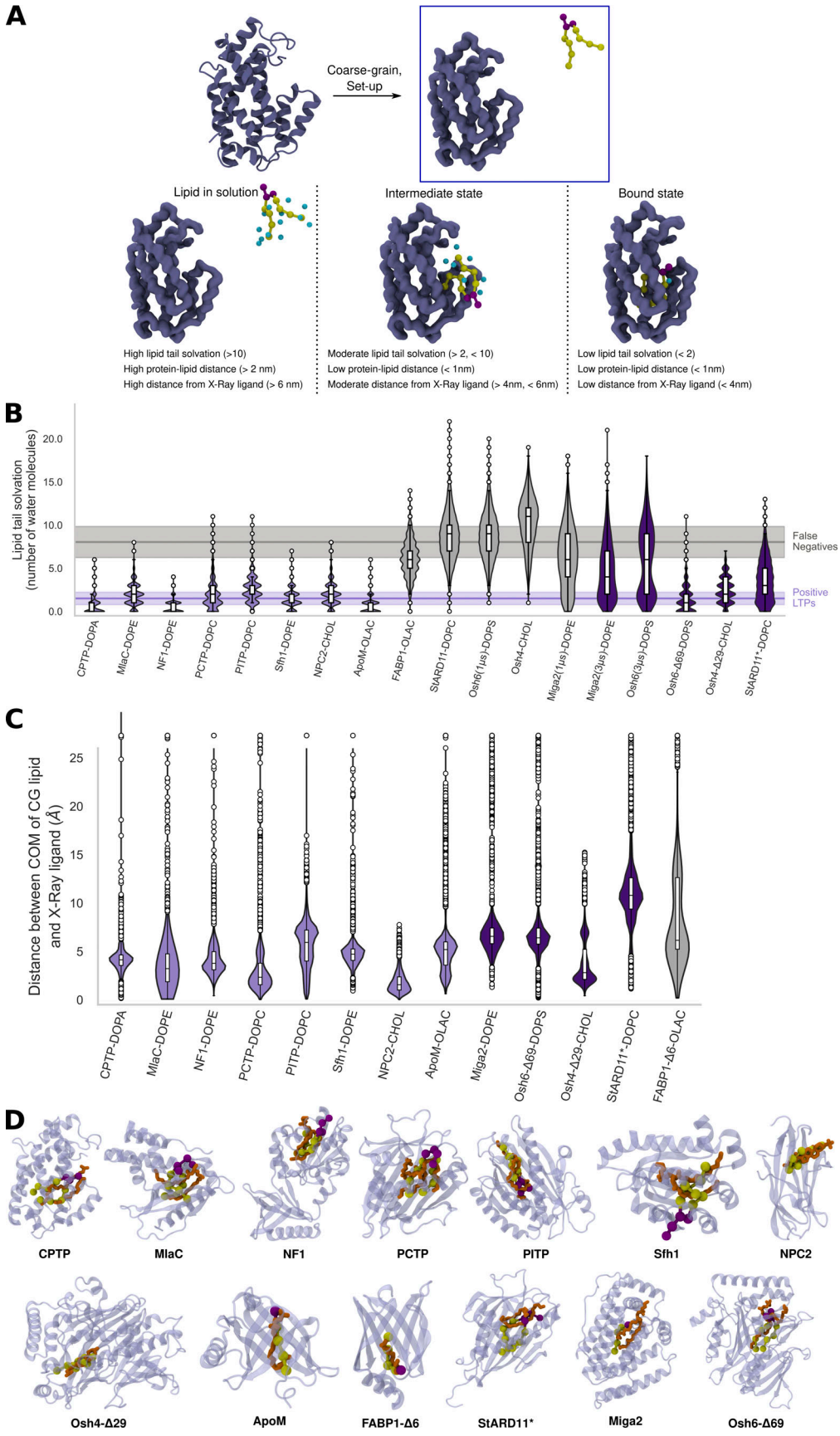

Figure 1. **CG simulations reproduce the correct lipid-binding pose to LTPs. (A)** Unbiased CG-MD simulation protocol used in this study. Atomistic structures of the protein were coarse-grained and set up with a lipid randomly placed in a bulk solvent far away from the protein, followed by MD simulations.

The lipid tail solvation (water beads within 5.0 Å of the lipid tails, represented as blue spheres) decreases as the lipid moves from bulk solvent to the hydrophobic binding cavity. **(B)** Lipid tail solvation in the last 100 ns of the CG-MD trajectories. Light purple indicates true positives, gray indicates false negatives, and dark purple indicates true positives after refinement of the simulation parameters. The two dotted lines indicate the average lipid tail solvation for the true positives (purple, 8/13 proteins) and the false negatives (gray, 5/13 proteins). **(C)** Distance between the center of mass (COM) of the hydrophobic tails of the experimental ligands and the lipids tested in the simulations. Only the frames with a lipid tail solvation <2 were considered. **(D)** Illustrative examples of agreement between lipid positions arising from our protocol (lipid tails: yellow, headgroups: purple) and the lipid/ligand position in x-ray structures (orange). The use of a different force constant for the elastic network of StARD11 is indicated by a *. Plots in B and C contain box plots and kernel density estimates (violin plots) of lipid tail solvation and distance, respectively. Box plots show the interquartile range as a white box, containing a horizontal line that depicts the median as well as minimum and maximum values shown as whiskers, which match the range displayed in the violin plots. Outliers are represented as white circles.

CG-MD simulations cannot reproduce the experimentally determined lipid specificity of LTPs

Next, we investigated whether our assay could reproduce the experimentally determined sensitivity for specific lipids. To do so, we performed identical simulations with multiple lipids that are not known to bind to the tested LTPs. In detail, for all proteins we tested a lipid with an identical polar head but with two saturated chains (dipalmitoyl-phostidylX [DPPX]), a dioleoyl phosphatidyl (DOPX) lipid with a different headgroup, cholesterol, triolein (TO), and a sphingolipid (N-stearoyl-D-erythro-sphingosylphosphorylcholine [DPSM]). In most cases, we found a similar binding behavior (Fig. S1), indicating that our protocol is generally unable to discriminate between different lipids. However, for some proteins (CPTP, MlaC, Osh6-Δ69, or those that bind smaller fatty acids, such as ApoM and Miga2 [Hong et al., 2022; Sevvana et al., 2009]), the use of a non-natural lipid substrate led to significantly worse binding prediction (Fig. S1), as was the case for almost all LTPs in the presence of TO, likely because of its bigger size.

Finally, analysis of the polar head placement in our simulations reveals a large variability in the binding poses between replicas (Fig. S2) as a consequence of the lack of sensitivity for the different headgroups. Rather, the polar heads can be stabilized by both interaction with the protein and by the surrounding solvent in our CG simulations.

Furthermore, to investigate the ability of our protocol to discriminate between different lipids, we applied it to two proteins, the takeout protein and the juvenile hormone binding protein (JHBP), that possess Synaptotagmin-like mitochondrial-lipid-binding (SMP) domains that are structurally similar to that of known LTPs. The takeout protein and the JHBP do not transport lipids, but they transport hydrophobic, lipid-like molecules: ubiquinone-8 and juvenile hormone III, respectively. In the presence of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) lipids, we observed spontaneous binding and very modest lipid solvation for both proteins (Fig. S3). Taken together, these results indicate that while our approach succeeds in the identification of hydrophobic lipid-binding pockets for LTPs, it has limitations for what concerns its ability to recapitulate the experimental selectivity for specific lipids and thus cannot be used to interrogate lipid specificity of LTPs.

Unbiased CG-MD simulations can discriminate between bona fide LTPs and negative control proteins that contain non-lipid-specific large cavities

The observation that our protocol is sometimes unable to reproduce the experimentally determined binding of lipids to LTPs suggests that our approach might not be biased toward false positive results. To further stress-test our protocol in this direction, we next investigated whether we could correctly flag proteins that do not transport lipids despite the presence of a large cavity in their structure. To do so, we selected 10 proteins that are not known to solubilize lipids but possess a cavity of large volume (Chwastyk et al., 2020) (Table 2), which is either hydrophobic or hydrophilic, as negative controls. Lipid binding to these negative control proteins was assessed using the same protocol described above, using DOPC as a model lipid.

For these negative controls, the lipids bound stably to the proteins during the CG simulations in many instances (Data S1 B). However, upon computing the spatial density of the lipids during the last 100 ns of the simulation trajectory, a superficial and outspread occupancy map was observed for the negative control proteins (Fig. 2 A). Similarly, time traces of lipid solvation show that in negative control proteins, the lipid remains highly solvated along the entire trajectory, even when bound to the protein (Data S1 B). This observation suggests that non-specific interactions of the lipid with different exterior surfaces of the proteins, rather than specific interactions in a hydrophobic pocket, characterize protein–lipid interactions with the negative controls in our assay. To quantify this observation, calculation of the solvation number of the lipid indicates that lipid solvation was significantly higher in the case of the negative controls (mean value = 6.4, Fig. 2 B) than that in the positive-result LTPs (mean value = 1.5, Fig. 2 B). This indicates that, for negative controls, lipids bound almost exclusively to

Table 2. **Details of the negative control proteins tested in this study**

Negative control	PDB ID	Nature of cavity	Lipid tested in MD simulations
Ferric-citrate transporter	1PO0	Hydrophilic	DOPC
Sheep lactoperoxidase	2IKC	Hydrophilic	DOPC
Hydroxylase-regulatory protein complex	2INP	Hydrophilic	DOPC
Sensory rhodopsin II	1JGJ	Hydrophobic	DOPC
Na^+/H^+ antiporter NhaA	1ZCD	Hydrophobic	DOPC
T4-Lysozyme	181L	Neutral	DOPC
CDK2	1AQ1	Hydrophobic	DOPC
Concanavalin A	3ENR	Hydrophilic	DOPC
Thyroid hormone receptor	1NAX	Hydrophobic	DOPC
Transcription factor T	1XBR	Hydrophilic	DOPC

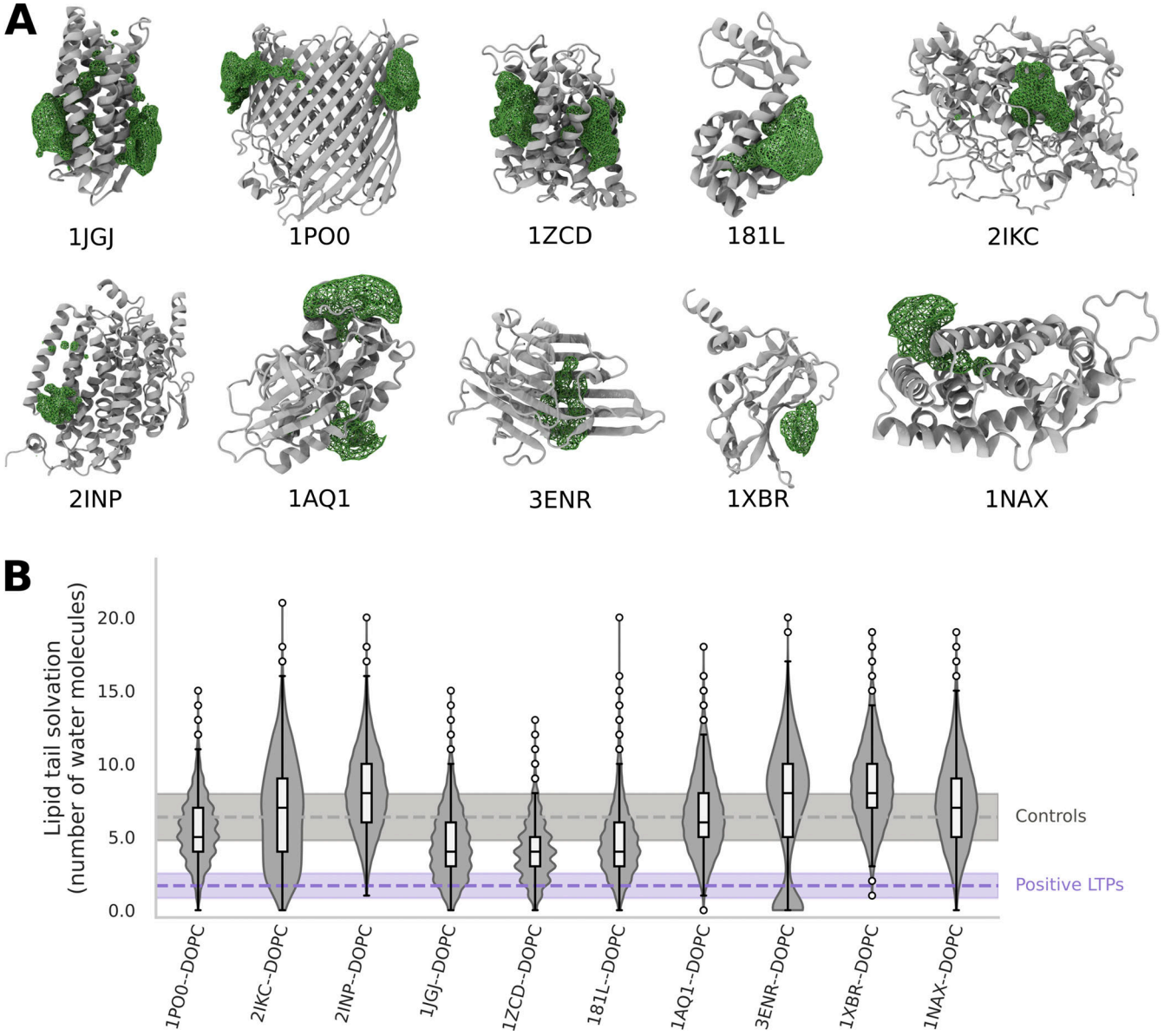

Figure 2. **CG simulations do not identify lipid-binding pockets for non-LTPs. (A)** Spatial density maps of lipid occupancy in the last 100 ns of unbiased CG simulations with DOPC for non-LTP negative controls. **(B)** Lipid solvation of non-LTPs in the last 100 ns of the trajectory.

external surfaces of the proteins despite the presence of a cavity. Precisely, out of 50 control simulations (five for each protein), we observed a single exception in one of the replicas of concanavalin A (PDB ID: 3ENR, Fig. 2 B, and Data S1 B), further suggesting that our approach is quite robust against false positive results.

Entry of the lipid into the protein cavity occurs via a dominant pathway

We next focused our attention on the path of lipid entry inside the LTP cavity as this process is critical to understand their mechanism of action. Since our protocol does not describe the physiological conditions of lipid entry/exit, with the lipid initially placed in bulk solvent rather than in a lipid bilayer, we wanted to investigate whether lipid entry was following a single dominant, possibly physiological, pathway or if rather the protein structure provided several entry points for fully water-solvated lipids.

Visualization of the simulation trajectories revealed the presence of a dominant pathway of lipid entry for all LTPs. By mapping the center of mass of the lipid during the trajectories and measuring the fraction of replicas in which the corresponding pathway was observed, we noticed that this pathway was unique for most LTPs (7 out of 12 positive results, Fig. 3 A) and observed with a high probability for the other LTPs. In almost all cases (including CPTP [Rogers and Geissler, 2023], Osh6 [Moser von Filseck et al., 2015], and PCTP [Khelashvili et al., 2019]), the pathway observed in these simulations corresponds to the one that has been proposed in literature. For all proteins, the averaged contact frequency between the LTPs residues and

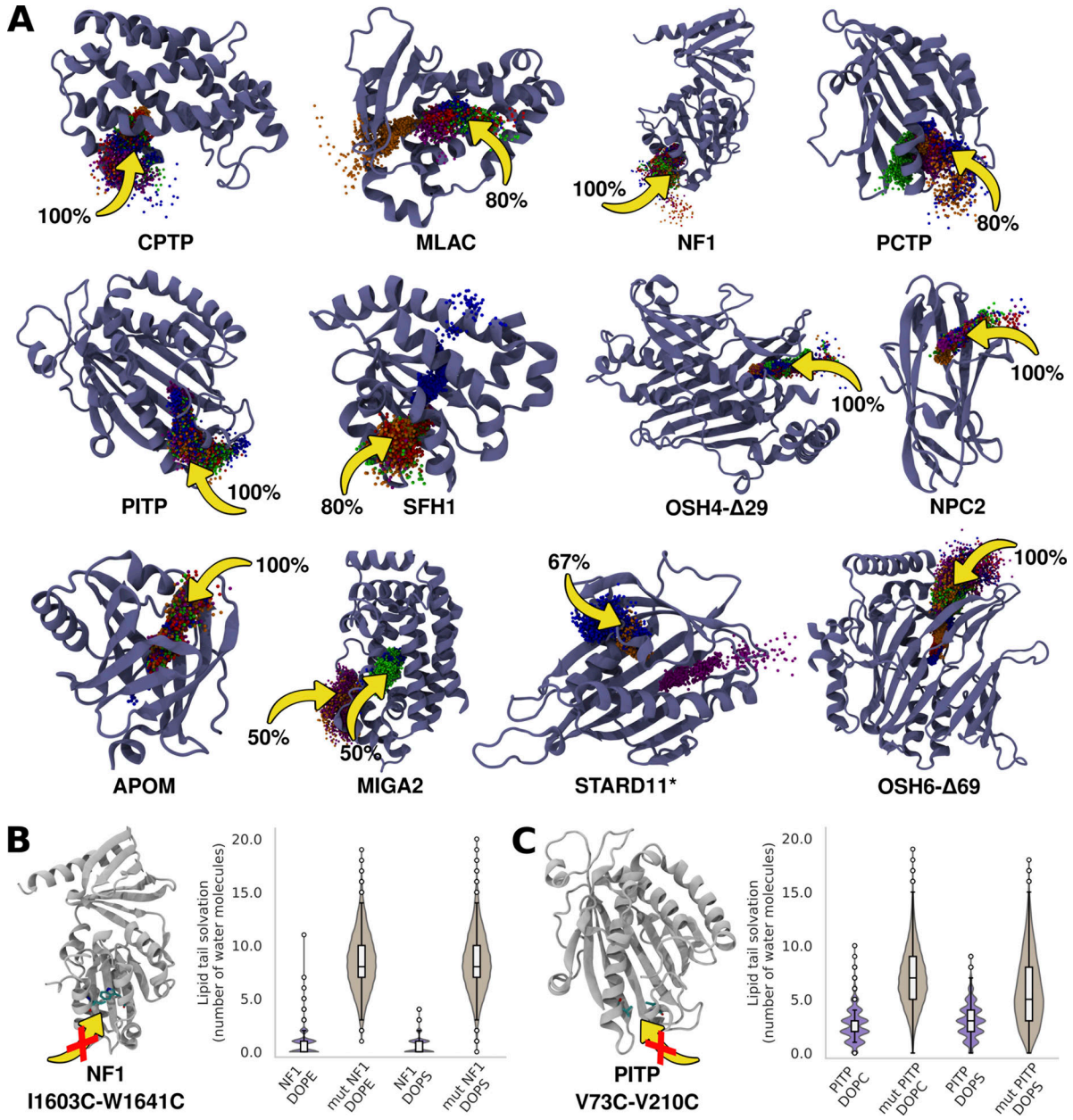

Figure 3. **Lipid entry pathways. (A)** Entry pathway of lipids into the LTD. The percentage values beside the arrows indicate the fraction of replicas in which the pathway was observed. Proteins are represented in cartoon representation and the center of mass of the lipid along the trajectory, neglecting the frames in which the lipid is in bulk solvent, represented as spheres. Each color of the sphere corresponds to a different replica with lipid binding. **(B and C)** Mutation of residues into disulfide bridges along the lipid-entry path in NF1 and PITP abolishes lipid entry into the binding pocket as indicated by the lipid solvation plots. The use of a different force constant for the elastic network of StARD11 is indicated by a *.

the simulated lipids during the last 100 ns of the simulations (Data S1 C) as well as the time traces of the contacts between the relevant residues and simulated lipids for ApoM, NF1, Osh4-Δ29, and PITP (Data S1 D) are shown in supplementary information to highlight the molecular resolution of our protocol.

Finally, to test the robustness of the observed entry pathway, we adopted two strategies. First, for the case of ApoM, we compared our pathway with that proposed in all-atom (AA) simulations (Zhang et al., 2016). Our data on the observed binding pathway (see Fig. S4 A) as well as lipid–protein residue contacts (see Fig. S4 B and Fig. S7), even with the intrinsic limitation of the lack of conformational changes in our CG simulations, is in excellent agreement with the one suggested by AA simulations.

Second, we designed mutations to block the entry process observed in the CG simulations. To do so, we performed identical simulations after introducing disulfide bridge mutations along the lipid entry pathway—I1603C-W1641C in the case of NF1 and V73C-210C in the case of PITP (Fig. 3, B and C). The residues were chosen such that the distance between Cα atoms of the residues is <8 Å. We observed that in both cases, the presence of the disulfide bridge abolished lipid entry into the

protein entirely, as shown by the lipid solvation in Fig. 3, B and C.

Brute-force CG-MD simulations can characterize the lipid-binding cavity of poorly characterized LTPs

As potential applications of our protocol, we first focused on proteins that have been proposed to transport lipids, but for which no experimental lipid-bound structure was determined, or for which direct experimental evidence of lipid binding is lacking. These include, for example, the recently characterized yeast ceramide transporter survival factor 1 (Svf1) (Limar et al., 2023), sorting nexin-25 (SNX25) (Paul et al., 2022), a protein that has been proposed to belong to a new class of LTPs with phox homology-associated (PXA) and C-terminal phox homology-associated (PXC) domains that form a hydrophobic channel, nucleus vacuole junction-3 (Nvj3), which is also expected to belong to this class (Paul et al., 2022), lipid droplet ergosterol cortex 1 (Lec1)/Ypr097w, which has been proposed as a LTP due to its superficial similarity to the previously mentioned class of LTPs (Paul et al., 2022), and the repeating β groove (RBG) proteins mitochondrial distribution and morphology protein-31 (Mdm31) and AsmA (Neuman et al., 2022). Using the AF structure as a starting point, our protocol was indeed able to propose a lipid binding pose for these proteins (Fig. 4, A–E).

The identified lipid-binding pocket of Svf1 is located between the two lipocalin domains (blue and violet in Fig. 4 A), which is in good agreement with what has been recently suggested using blind docking (Limar et al., 2023). For SNX25, our protocol suggests that lipid binding, which occurs in four out of five replicas (see Data S1 E), takes place in a long and conserved hydrophobic pocket formed by the PXA and PXC domains (red and orange in Fig. 4 B, respectively), in agreement with what was recently proposed based on structural considerations (Paul et al., 2022). The same results were obtained for the related Nvj3 protein, with lipid binding in all replicas along the hydrophobic channel between the predicted PXA-PXC domains (red and orange in Fig. 4 C, respectively). AsmA and Mdm31 are two prokaryotic RBG proteins distantly related to the eukaryotic BLTP superfamily, which includes well-studied LTPs such as vacuolar protein sorting 13 (Vps13) and autophagy-related protein 2 (Atg2). Our protocol displays lipid binding in all replicas for AsmA and in three out of five replicas for Mdm31 (Data S1 E) within the hydrophobic cavity formed by the RBG domains in a similar manner as their eukaryotic counterparts (Fig. 4, D and E). On the other hand, Lec1 displays lipid binding in only two out of five replicas (Data S1 E), with one more replica binding to the external surface of the protein. However, it is noteworthy that the two binding events occur within the highly conserved hydrophobic cavity formed by the PXYn (N-terminal phox homology-associated domain from yeast) and PXYc (C-terminal phox homology-associated domain from yeast) domains (green and orange in Fig. 4 F, respectively) (Paul et al., 2022).

In summary, our protocol can identify potential hydrophobic pockets of LTPs that are poorly characterized at the structural level, and our results are in good agreement with alternative methods such as docking or structural analysis.

LTPs that can bind multiple lipids in their cavity possess several continuous lipid interacting regions

Next, we applied our protocol to BLTPs (Braschi et al., 2022; Neuman et al., 2022). Recently, many LTPs have been proposed to transport lipids via a bridge-like mechanism by establishing a continuous hydrophobic conduit between membrane organelles (Neuman et al., 2022; Wong et al., 2019). Within this model, LTPs could bind many lipids concomitantly, and they could contribute to bulk lipid transport between organelles, such as in autophagosome formation (Osawa et al., 2019; Ghanbarpour et al., 2021) or lysosomal repair (Radulovic et al., 2022, *Preprint*; Tan and Finkel, 2022). Yet, available high- or medium-resolution structures of BLTPs are scarce and entirely devoid of lipids, with the sole exception of one single phosphatidylethanolamine molecule in the structure of a small region of Atg2 (Osawa et al., 2019).

To investigate lipid occupancy in BLTPs, we adapted our protocol to work in the presence of multiple lipids. To do so, we performed lipid-addition simulations iteratively to avoid lipid–lipid interactions (leading to micelle formation) in the solvent before binding to the protein. As a first test to validate our protocol, we investigated the SMP domain of extended synaptotagmin-2 (E-Syt2) dimer (Schauder et al., 2014; Saheki and De Camilli, 2017; Fernández-Busnadiego et al., 2015). While this protein is proposed to work via a shuttle-like mechanism (Schauder et al., 2014; Zhang et al., 2022), it has been co-crystallized with multiple lipids in its cavity (two diacylglycerols and two detergents), providing a natural positive control for our protocol.

Indeed, we observed that with the iterative lipid addition process, at least five lipids can be accommodated in the cavity of the SMP domain as indicated by the low solvation of the lipid tails (Fig. 5 D). Intriguingly, the average spatial density of lipids in E-Syt2 spans the entire hydrophobic conduit rather than a specific lipid binding pocket (Fig. 5 A).

We next applied our protocol to Atg2 and Vps13 as these proteins have been proposed to bind multiple lipids at once. We initially restricted our analysis to the respective chorein domains of these proteins, as these regions have been solved at high resolution (2.7 and 3.0 Å, respectively [Osawa et al., 2019; Kumar et al., 2018]). Again, our protocol shows that the proteins are able to bind multiple lipids and form a continuous lipid-filled tunnel spanning the entirety of the hydrophobic cavity formed by the chorein domain (Fig. 5, B and C, and individual time traces for Atg2 in Fig. S5).

Taken together, these data indicate that our protocol is able to predict the simultaneous binding of multiple lipids to BLTPs, and it confirms that the lipid binding mode for these proteins is distinct from those proposed to work in a shuttle-like fashion. Of note, in all cases, we observed that the presence of multiple lipids further decreases their solvation (Fig. 5, D–F). This suggests that the presence of lipids inside the binding cavity could promote sequestration of lipids from the lipid bilayer and, hence, lipid transport.

Next, we investigated multiple proteins that belong to the BLTP superfamily, cold sensitive for fermentation protein 1 (Csf1, also known as BLTP1 or Tweek), Vps13, and Atg2 in their

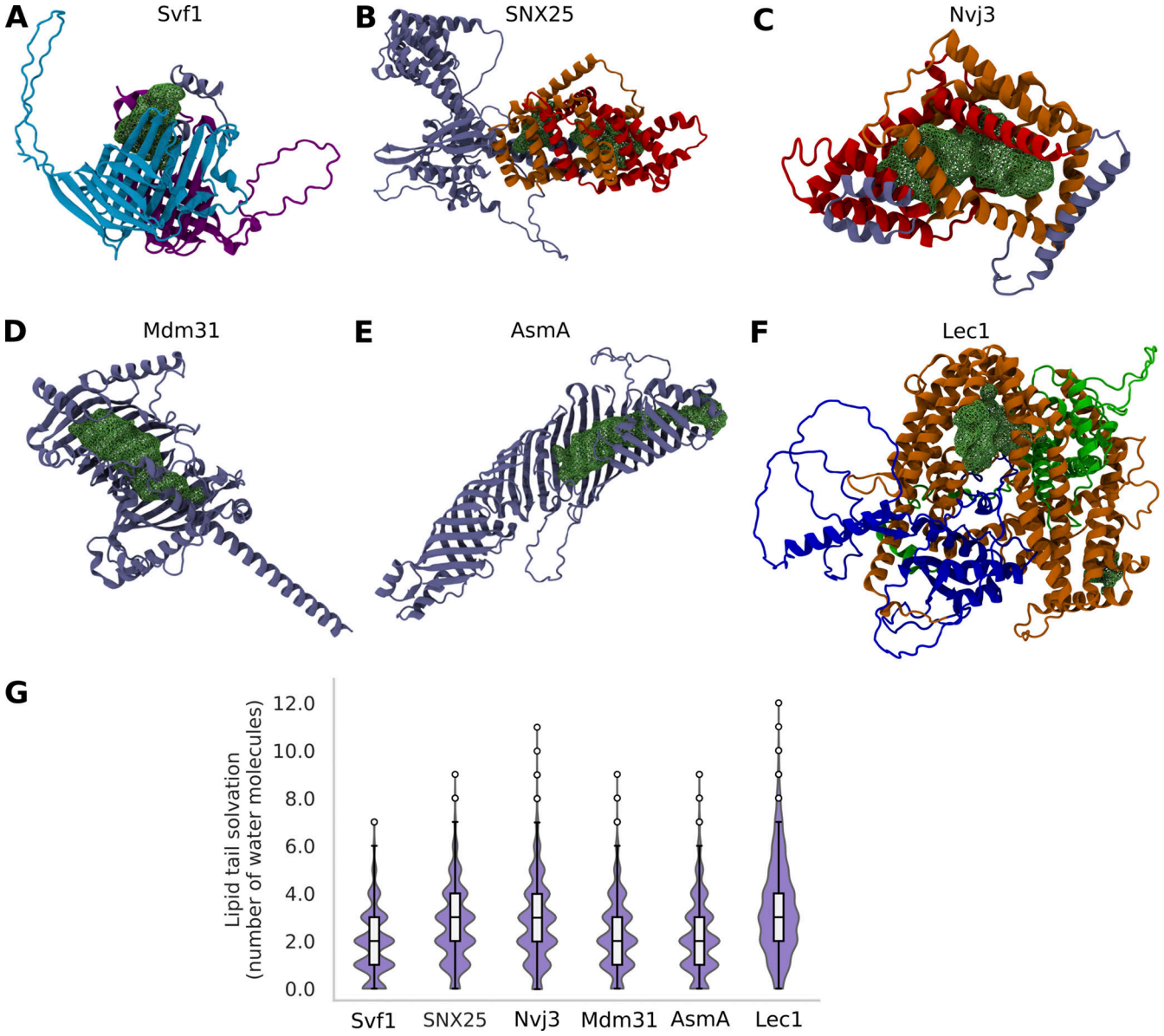

Figure 4. **Identification of the hydrophobic cavity of potential LTPs. (A–F)** Spatial density maps (A–F) of (A) Svf1, (B) SNX25, (C) Nvj3, (D) Mdm31, (E) AsmA, and (F) Lec1. The two lipocalin domains of Svf1 are shown in blue and violet. The PXA and PXC domains of SNX25 and Nvj3 are displayed in red and orange, respectively. The PXYn, PX, and PXYc domains of Lec1 are displayed in green, blue, and orange, respectively (notation and colors as in Paul et al., 2022). Averaged lipid tail densities are shown in green. **(G)** Lipid tail solvation for the six potential LTPs; only the replicas with lipid binding were considered for the analysis.

entirety (Fig. 6, A–C). For Atg2, we used the AF structure (identifier AF-P53855-F1) and residues 1–1,344 as the C-terminal region consists of several helices with low predicted local distance difference test scores that are not part of the hydrophobic cavity. In the case of Csf1, the structure of three different fragments was predicted using AF and then aligned to get the whole protein structure (2,958 residues), which was considered entirely. Finally, for Vps13, a similar protocol was employed, but residues 1,860–3,144 were discarded as they do not belong to the extended chorein domain that forms the hydrophobic cavity.

Similar to the results described above, using an adapted iterative lipid addition process (see Materials and methods and Video 2), we observed that the proteins can accommodate multiple lipids at the same time. Precisely, we observed the binding of 15 lipids in Atg2, 53 in Csf1, and 49 in Vps13 with low tail solvation. Despite the large number of lipids in the cavity, the use of an elastic network to restrain the secondary structure prevents large protein conformational changes, thus indicating that the presence of multiple lipids is compatible with the initial AF models. A similar number of lipids (15) has been proposed for ATG2A, a human ortholog of Atg2 with the same length, using structural analysis (Wang et al., 2023, *Preprint*). The average solvation data of the lipid tails (Fig. 6 F) are consistent with our previous analysis on shuttle LTPs (average <1). Interestingly, even though the cavities are almost completely filled with lipids, the headgroups are arranged in such a way that they face the

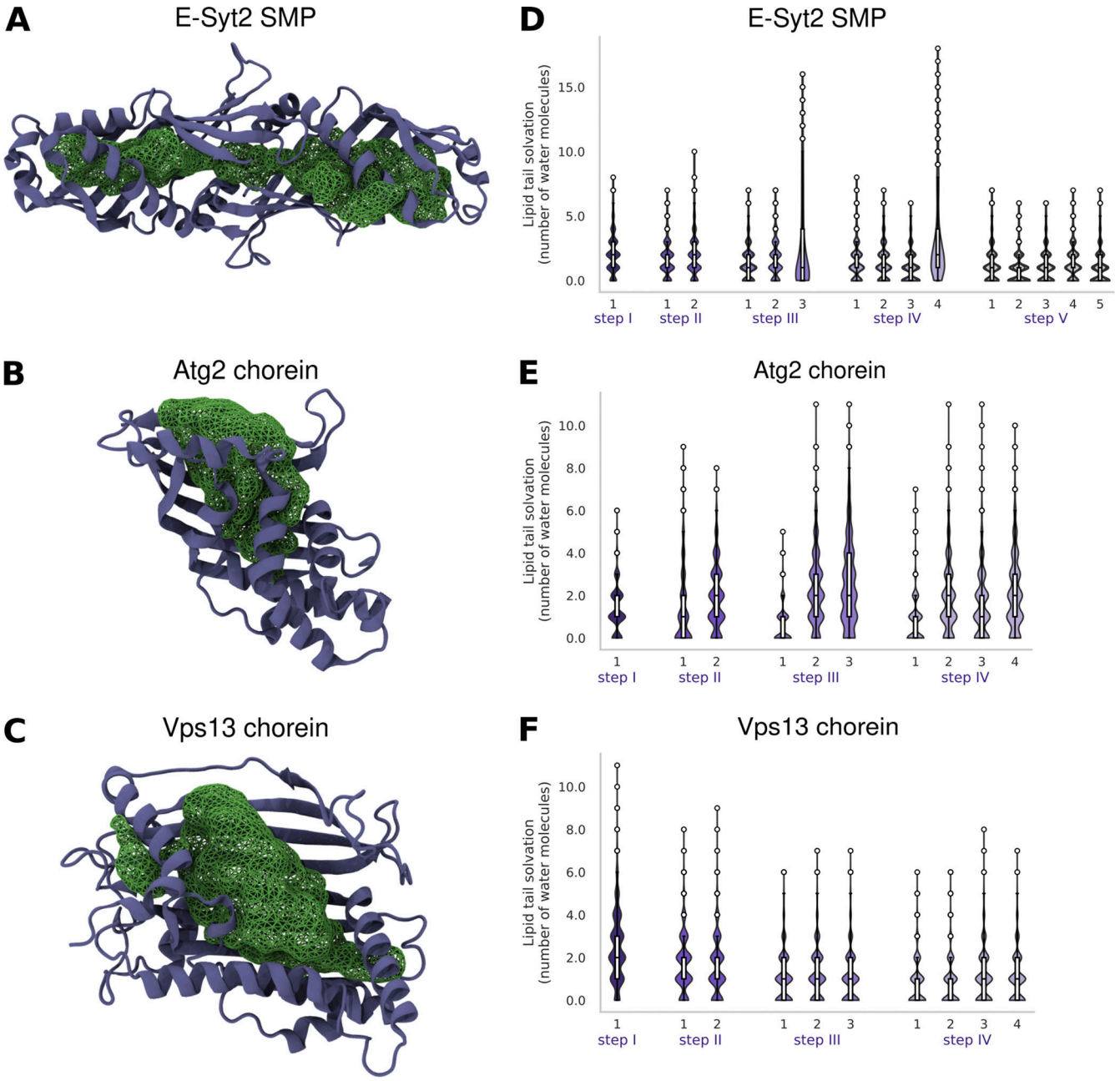

Figure 5. **LTDs with multiple lipid binding regions. (A–F)** (A–C) Spatial density maps of lipids and (D–F) corresponding solvation for each lipid in the multi-step iterative protocol for the (A and D) SMP domain of E-Syt2, (B and E) chorein motif of Atg2, and (C and F) chorein motif of Vps13. The protein is shown in purple; the averaged lipid density is in green.

solvent (Fig. S6 and Fig. S7). Further, the spatial density maps indicate a long hydrophobic conduit, but they also suggest the possible presence of bottlenecks in Atg2 and Csf1, which could be related to regulatory mechanisms.

Finally, to highlight the practical applications of our protocol, we designed mutations that block the hydrophobic cavity of Atg2, similar to those tested in ATG2A experiments (Valverde et al., 2019; Tan and Finkel, 2022). In these experiments, a small ring of hydrophobic residues was mutated to charged ones, resulting in impaired lipid transport in vitro. To mimic this approach, we selected four hydrophobic residues (F88, L180, L208, L311) close to the chorein motif of Atg2, where the lipid spatial density is large and uniform (Fig. 6 D), and we mutated them to charged residues (arginine and glutamate). We next performed a 500-ns-long unbiased CG-MD simulation with the 15 lipids already inside the Atg2 mutant and computed their spatial density map at the end of the simulation. The resulting map (Fig. 6 D) demonstrates that the new ring of charged residues generates a bottleneck where the lipids can no longer reside. A similar strategy has also been experimentally tested on Vps13 (Li et al., 2020), showing that mutations in the middle of the hydrophobic cavity (V690D/L692R/L694E/I715K/A717D/M720K/I722D/I761R/

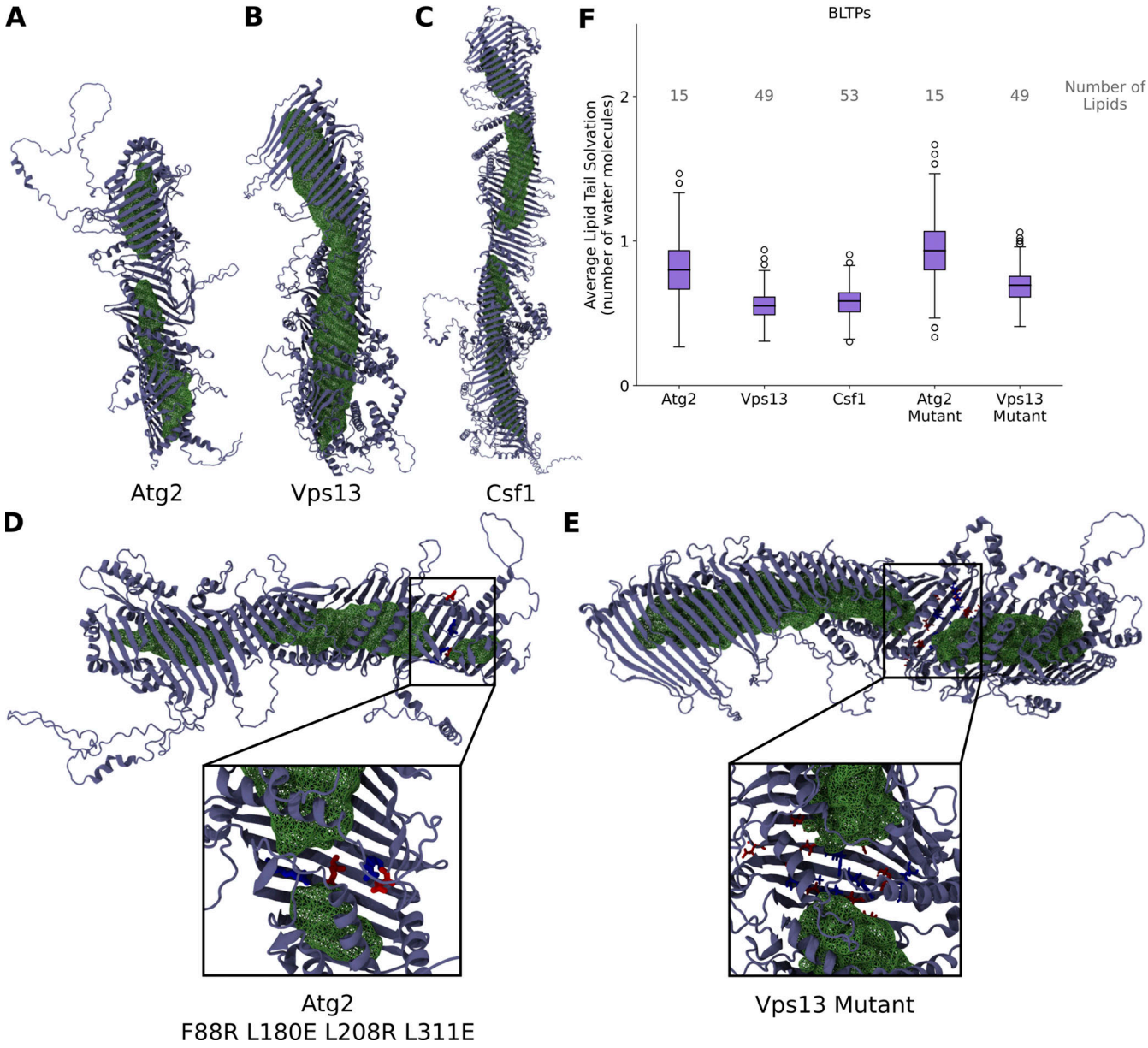

Figure 6. **BLTPs possess a long hydrophobic cavity that can be filled with multiple lipids. (A–F)** (A–E) Spatial density maps and (F) lipid tail solvation (per lipid) using the iterative lipid addition process for BLTPs: (A) Atg2, (B) Vps13, (C) Csf1, and mutants of (D) Atg2 and (E) Vps13. Proteins are shown in purple; averaged lipid tail densities are in green. The mutated residues are displayed in blue (arginine, lysine) or red (glutamate, aspartate).

I768E/F790D/M796D/L798R/V802E/I816R/G820D/L827E) do not impair the lipid-binding ability of Vps13 but do abrogate the Vps13 function in sporulation. As for the previous case, our protocol clearly indicates the formation of a significant bottleneck in the Vps13 mutant that could be responsible for the observed loss-of-function (Fig. 6 E).

Discussion

In the last few years, lipid transport by proteins has emerged as a central process in membrane and organelle homeostasis, but its molecular mechanisms remain largely unclear. To bridge this gap, we present here a computational protocol based on brute-force unbiased CG-MD simulations, adapted from a recently described strategy to determine protein–ligand interactions for hydrophilic small drug-like molecules (Souza et al., 2020), to propose a structural model for the binding pose of lipids inside LTPs. Our approach provides a physics-based hypothesis on the structure of the lipid–LTP complex that could contribute toward a mechanistic interpretation of in vitro and cellular experiments.

Our protocol is easy to reproduce (requiring only publicly available open-source software) and computationally cheap, as most simulations take only a few hours to perform in a standalone graphics processing unit–accelerated workstation. In addition, it does not require previous knowledge of the binding pocket, and, by taking advantage of physicochemical properties of both lipids and proteins, can easily distinguish between polar

and hydrophobic interactions, thus fully considering the amphipathic nature of most lipids. In addition, our iterative approach allows to propose a binding mode for LTPs binding multiple lipids in a dynamic way, with each lipid able to rearrange its localization inside the binding pocket upon the binding of subsequent lipids. This provides a clear advantage over docking methods, where the binding pose of each individual lipid is mostly static and can't be easily modified upon the docking of subsequent molecules.

On the other hand, we acknowledge that our protocol has three main limitations. First, it is mostly unable to discriminate between different lipids. As such, it can't be used to investigate lipid specificity of LTPs. We foresee that future improvements in CG force fields or the use of multiscale strategies (e.g., by backmapping to all-atom simulations), possibly coupled with free-energy calculations, could help in this direction.

Second, our protocol can generate false negative results, as we could not straightforwardly identify any lipid-binding pose for a few well-characterized LTPs, rather requiring a posteriori refinements in the simulation setup. This is very likely a consequence of the conformational plasticity of LTPs, which can adopt different conformations in their apo and holo states (Srinivasan et al., 2023, *Preprint*). Hence, if the initial protein structure is in a "closed" conformation (for example, in the presence of a lid that precludes lipid entry), our protocol is unable to identify the correct binding pathway and pose since our CG approach can't reproduce protein conformational flexibility as the protein structure is restricted by an elastic network. We expect that enhancing conformational sampling to generate a diverse set of starting protein structures, e.g., by all-atom MD simulations (Srinivasan et al., 2023, *Preprint*) or by machine-learning-based approach (Janson et al., 2023; Degiacomi, 2019) could help mitigate this issue in the future.

Third, since in our simulations the lipid is initially placed in the surrounding aqueous solution, our protocol does not reproduce the physiological process of lipid entry, in which the lipid enters the LTP's cavity directly from a donor membrane, nor that of lipid desorption. However, the good agreement between the observed lipid entry pathway and that proposed for multiple LTPs (Osh4, Osh6, ApoM, PITP…) indicates that the pathway proposed by our protocol warrants further experimental investigation, possibly by site-directed mutagenesis studies.

In contrast, a strong advantage of our method is its robustness towards false positives. This suggests that whenever a desolvated lipid-binding pose is found, there is a high likelihood that this interaction has meaningful consequences for protein function and that this warrants further experimental investigations. Specifically, we expect that our method will be extremely useful in at least three main areas. First, it will allow for generation of mechanistic hypotheses for subsequent experimental validation for what concerns the molecular details of lipid transport, such as the identification of regulatory mechanisms, including posttranslational modifications or potential bottlenecks along the transport pathway. Second, it will provide lipid-bound structures to potentially investigate the mechanism of lipid release by LTPs into lipid bilayers in silico, thus providing new avenues toward the characterization of the lipid transport mechanism in more realistic conditions. Third, we expect it will become a powerful tool for the direct identification of novel LTPs in silico, with the quality of the AI-based structural predictions as the main limiting factor.

Materials and methods

Software details

MD simulations of all systems were performed with the GROMACS (v 2021.x) (Van Der Spoel et al., 2005) package using the MARTINI 3 force field (Marrink et al., 2007). Molecular images were rendered using Visual Molecular Dynamics (VMD) (Humphrey et al., 1996).

Protein selection

LTPs were selected to demonstrate the protocol's flexibility across different families, folds/structures, and bound lipids. LTPs that function via both—shuttle and tunnel mechanisms—were chosen for the study. The non-LTPs with RCSB PDB IDs 1PO0, 2IKC, 2INP, 1JGJ, and 1ZCD tested in the study were chosen to represent proteins that have large cavities and can geometrically accommodate one or more molecules that resemble lipids in size. These controls were chosen from Table 1 of Chwastyk et al. (2020), which contains 50 structures with the largest hydrophilic and hydrophobic cavities. T4 Lysozyme was included in the negative controls to test a protein with a neutral cavity while CDK2 was included to test an enzyme. The other three non-LTPs were chosen to represent proteins that bind diverse molecules such as sugars (hence concanavalin A [PDB ID 3ENR]—a carbohydrate-binding protein), DNA (hence transcription factor T [PDB ID 1XBR]), and hormones (hence thyroid hormone receptor [PDB ID 1NAX]).

System setup

The atomistic structures of the protein were obtained from the RCSB PDB (Berman et al., 2000) and were converted to CG models using the martinize (de Jong et al., 2013) script. An additional elastic network with a force constant of 1,000 kJ mol^{-1} nm^{-2} was used to restrain the secondary structure of the protein, with a 0 nm elastic bond lower cutoff and 0.8 nm upper cutoff. A force constant of 300 kJ mol^{-1} nm^{-2} was also tested for StARD11 and FABP1. Side chain corrections were applied to all proteins. The CG proteins were then placed in the center of a cubic box in which the distance between the protein and the edge of the box was at least 2.0 nm, and one lipid molecule was randomly placed in the bulk solvent in each of the replicas without any distance restrictions. The system was then solvated, considering a Van der Waals (VdW) distance of 0.21 nm, and ionized with 0.12 M NaCl. For BLTPs, the structure was predicted using AF version 2.0, and, in the case of Vps13 and Csf1, different fragments were predicted and then aligned using common residues to get the structure of the entire protein. For LTPs that could bind multiple lipids in their cavity, the lipid addition procedure was repeated iteratively, such that the structure of the protein with n lipids bound in it served as the starting structure for the addition of (n+1)th lipid, solvating and

ionizing again every time a new lipid was added. This protocol was concluded once no more lipid entry into the hydrophobic cavity was observed. Rather, excess lipids would bind at the extremities of the cavity and remain solvated. In the case of Vps13 and Csf1, due to their large size, the protocol was applied separately to different fragments (two and three fragments, respectively) with similar size (see Table 1). Once filled, they were aligned to the complete structure and simulated.

Simulation details for protein in water/LTP entry setup

Five independent replicas of 1 µs each were simulated for each protein–lipid system, with the exception of FABP1, FABP1 Δ6, Miga2, StARD11, Osh6, Osh6-Δ69, Osh4, and Osh4-Δ29, for which five independent replicas were simulated for 3 µs each. Initial equilibration was carried out by performing energy minimization using the steepest descent algorithm, followed by a short MD run of 125 ps. Production runs were performed at 310 K using a velocity-rescale thermostat (Bussi et al., 2007), with separate temperature coupling for protein and non-protein particles. The md integrator was used for the production runs, with a time step of 25 fs. The Parrinello-Rahman barostat (Parrinello and Rahman, 1981) was used to maintain the pressure at 1 bar, along with an isotropic pressure coupling scheme and a nstpcouple parameter set to 10. The Coulombic terms were calculated using reaction field, with an epsilon (dielectric constant) of 15 and a cut-off distance of 1.1 nm. A cut-off scheme was used for the VdW terms, with a cut-off distance of 1.1 nm and the Verlet cut-off scheme for the potential shift (de Jong et al., 2016). The non-bonded interactions were calculated by generating a pair-list using the Verlet scheme with a buffer tolerance of 0.005. The system setup and simulation parameters are in line with the recently proposed protocol for studying protein–ligand binding with the MARTINI force field (Souza et al., 2020).

The Osh6-Δ69 protein was modeled by removing residues 36–69 that correspond to the lid of the ORP domain. Similarly, Osh4-Δ29 was modeled by removing residues 2–29, and FABP1-Δ6 was modeled by removing residues 1–6. Disulfide bridges in mutants NF1 1603C-W1641C and PITP V73C-210C were added using the CHARMM-GUI PDB Reader and Manipulator (Jo et al., 2008) web server before coarse-graining the structure using the martinize script. The BLTP mutants were modeled with Pymol v2.3.0 (Schrodinger, 2015) and aligned to the wild-type BLTPs filled with lipids after coarse-graining with the martinize script.

Analysis

The minimum distance between the protein and the lipid was computed using the gmx mindist tool. An in-house tcl script was used to calculate the solvation number of the lipid, counting the number of water molecules within 5.0 Å of the lipid tail beads (headgroup and backbone were excluded) for each frame of the trajectory, for each replica. Only the last 100 ns of the trajectories were used for the boxplot analysis. In the case of cholesterol, all beads but the "ROH" one, which corresponds to the polar hydroxyl group, were considered for the analysis.

Spatial density maps for lipids were computed with the Volmap plugin of VMD (https://www.ks.uiuc.edu/Research/vmd/plugins/volmapgui/) with a resolution of 0.5 Å, a uniform weight of one (the number density), and the default atom size by averaging over a concatenated trajectory containing the last 100 ns of each replica.

To compute the distance to the lipid tails of the crystallographic ligands, the backbone beads of the CG-MD trajectories were aligned to the N atoms of the protein crystal structure using the fit option of gmx trjconv. Then, a tcl script was used to compute the distance between the center of mass of the hydrophobic beads of the lipid and the center of mass of the lipid hydrophobic tails in the crystal structure for each frame of every replica. A similar procedure was used to compute the distance to the headgroups of the ligands.

To determine the lipid entry pathways, the center of mass of the lipid along the aligned trajectories was computed and saved in a separate PDB file using gmx traj tool. Then, the resulting files were visualized using VMD, representing each center of mass as a sphere, with different colors for different replicas, and removing the points that correspond to the lipid in bulk solvent. The percentage of occurrence of each pathway was calculated by dividing the number of replicas in which the lipid entered the LTP via the dominant pathway over the total number of replicas in which the lipid entered the protein cavity.

The analysis of the residues involved in protein–lipid contacts during the last 100 ns of the simulations was performed using gmx select, and a contact was identified if any of the beads of the protein residue was within 5.0 Å of any of the lipid beads in that frame. The time trace of contacts for selected residues was computed using gmx mindist using the same criteria for the contacts.

All scripts for system preparation and analysis are available, with instructions, at https://github.com/danialv4/Unbiased_simulations_characterize_lipid_binding/, along with the input files and gro files for the lipids used in the simulations.

Final frames of all protein–lipid pairs considered in this work are also provided in the same public repository.

Online supplemental material

Data S1 shows all the time traces of lipid tail solvation, minimum distance between protein and lipid, and distance to the x-ray ligand of all the protein–lipid pairs considered in Fig. 1; the time traces of lipid tail solvation and protein–lipid distance of all the control proteins (Fig. 2) with DOPC; an analysis of protein–lipid contacts per residue for several LTPs during the last 100 ns of simulations; the time traces of protein–lipid contacts per residue, upon binding, for few LTPs; and the time traces of lipid tail solvation and protein–lipid minimum distance for the protein–lipid pairs considered in Fig. 4. Fig. S1 shows the lipid tail solvation of the simulations with LTPs and alternative lipids. Fig. S2 illustrates the distance between the headgroups of the simulated lipids and the ones of the x-ray ligands. Fig. S3 contains the time traces of solvation and distance as well as the plots of lipid tail solvation for the simulations with Takeout and JHBP proteins. Fig. S4 illustrates the results of the entry pathway and specific residues present in the hydrophobic cavity of ApoM for comparison with all-atom studies. Fig. S5 collects the time traces of lipid tail solvation for the iterative addition of four lipids to the chorein domain of Atg2. Fig. S6 demonstrates the

differences in solvation between the headgroups and lipid tails of BLTPs, and Fig. S7 is an explicit representation of their disposition. Video 1 displays the application of our protocol, as in Fig. 1 A. Video 2 shows the iterative addition of 15 lipids to one of the BLTPs (Atg2).

Data availability

All scripts for system preparation and analysis are available, with instructions, at https://github.com/danialv4/Unbiased_simulations_characterize_lipid_binding/, along with the input files and gro files for the lipids used in the simulations. Final frames of all protein–lipid pairs considered in this work are also provided in the same public repository.

Acknowledgments

This work was supported by the Swiss National Science Foundation (PP00P3_194807 and PP00P3_163966). This work was also supported by grants from the Swiss National Supercomputing Centre under project ID s1030, s1132, and s1221 and has also received funding from the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement no. 803952). D. Alvarez acknowledges support from the Margarita Salas program 2021–2023 funded by Ministerio de Universidades (MU-21-UP2021-030-53773022). S. Vanni and A.T. John Peter acknowledge support from the Novartis Forschungsstiftung via a FreeNovation grant.

Author contributions: S. Srinivasan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing—original draft, Writing—review & editing, D. Alvarez: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review & editing, A.T. John Peter: Methodology, S. Vanni: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing.

Disclosures: The authors declare no competing interests exist.

Submitted: 13 December 2023
Revised: 29 May 2024
Accepted: 16 July 2024

References

Bender, B.J., S. Gahbauer, A. Luttens, J. Lyu, C.M. Webb, R.M. Stein, E.A. Fink, T.E. Balius, J. Carlsson, J.J. Irwin, and B.K. Shoichet. 2021. A practical guide to large-scale docking. *Nat. Protoc.* 16:4799–4832. https://doi.org/10.1038/s41596-021-00597-z

Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. 2000. The protein Data Bank. *Nucleic Acids Res.* 28:235–242. https://doi.org/10.1093/nar/28.1.235

Braschi, B., E.A. Bruford, A.T. Cavanagh, S.D. Neuman, and A. Bashirullah. 2022. The bridge-like lipid transfer protein (BLTP) gene group: Introducing new nomenclature based on structural homology indicating shared function. *Hum. Genomics*. 16:66. https://doi.org/10.1186/s40246-022-00439-3

Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101. https://doi.org/10.1063/1.2408420

Cai, S., Y. Wu, A. Guillén-Samander, W. Hancock-Cerutti, J. Liu, and P. De Camilli. 2022. In situ architecture of the lipid transport protein VPS13C at ER-lysosome membrane contacts. *Proc. Natl. Acad. Sci. USA*. 119: e2203769119. https://doi.org/10.1073/pnas.2203769119

Cascella, M., and S. Vanni. 2016. Toward Accurate Coarse-Graining Approaches for Protein and Membrane Simulations. Chemical Modelling. Vol. 12. The Royal Society of Chemistry, London, UK.

Castro, I.G., S.P. Shortill, S.K. Dziurdzik, A. Cadou, S. Ganesan, R. Valenti, Y. David, M. Davey, C. Mattes, F.B. Thomas, et al. 2022. Systematic analysis of membrane contact sites in *Saccharomyces cerevisiae* uncovers modulators of cellular lipid distribution. *Elife*. 11:e74602. https://doi.org/10.7554/eLife.74602

Chiapparino, A., K. Maeda, D. Turei, J. Saez-Rodriguez, and A.C. Gavin. 2016. The orchestra of lipid-transfer proteins at the crossroads between metabolism and signaling. *Prog. Lipid Res.* 61:30–39. https://doi.org/10.1016/j.plipres.2015.10.004

Christoffersen, C., H. Obinata, S.B. Kumaraswamy, S. Galvani, J. Ahnström, M. Sevvana, C. Egerer-Sieber, Y.A. Muller, T. Hla, L.B. Nielsen, and B. Dahlbäck. 2011. Endothelium-protective sphingosine-1-phosphate provided by HDL-associated apolipoprotein M. *Proc. Natl. Acad. Sci. USA*. 108:9613–9618. https://doi.org/10.1073/pnas.1103187108

Chwastyk, M., E.A. Panek, J. Malinowski, M. Jaskólski, and M. Cieplak. 2020. Properties of cavities in biological structures—A survey of the Protein Data Bank. *Front. Mol. Biosci.* 7:591381. https://doi.org/10.3389/fmolb.2020.591381

Dall'Armellina, F., M. Stagi, and L.E. Swan. 2023. In silico modeling human VPS13 proteins associated with donor and target membranes suggests lipid transfer mechanisms. *Proteins*. 91:439–455. https://doi.org/10.1002/prot.26446

de Jong, D.H., S. Baoukina, H.I. Ingólfsson, and S.J. Marrink. 2016. Martini straight: Boosting performance using a shorter cutoff and GPUs. *Comput. Phys. Commun.* 199:1–7. https://doi.org/10.1016/j.cpc.2015.09.014

de Jong, D.H., G. Singh, W.F.D. Bennett, C. Arnarez, T.A. Wassenaar, L.V. Schäfer, X. Periole, D.P. Tieleman, and S.J. Marrink. 2013. Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theor. Comput.* 9:687–697. https://doi.org/10.1021/ct300646g

Degiacomi, M.T. 2019. Coupling molecular dynamics and deep learning to mine protein conformational space. *Structure*. 27:1034–1040.e3. https://doi.org/10.1016/j.str.2019.03.018

Dror, R.O., A.C. Pan, D.H. Arlow, D.W. Borhani, P. Maragakis, Y. Shan, H. Xu, and D.E. Shaw. 2011. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. USA*. 108:13118–13123. https://doi.org/10.1073/pnas.1104614108

Ekiert, D.C., G. Bhabha, G.L. Isom, G. Greenan, S. Ovchinnikov, I.R. Henderson, J.S. Cox, and R.D. Vale. 2017. Architectures of lipid transport systems for the bacterial outer membrane. *Cell*. 169:273–285.e17. https://doi.org/10.1016/j.cell.2017.03.019

Fagone, P. and S. Jackowski. 2009. Membrane phospholipid synthesis and endoplasmic reticulum function. *J. Lipid Res.* 50 Suppl:S311–S316. https://doi.org/10.1194/jlr.R800049-JLR200

Fernández-Busnadiego, R., Y. Saheki, and P. De Camilli. 2015. Three-dimensional architecture of extended synaptotagmin-mediated endoplasmic reticulum-plasma membrane contact sites. *Proc. Natl. Acad. Sci. USA*. 112:E2004–E2013. https://doi.org/10.1073/pnas.1503191112

Forli, S., R. Huey, M.E. Pique, M.F. Sanner, D.S. Goodsell, and A.J. Olson. 2016. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* 11:905–919. https://doi.org/10.1038/nprot.2016.051

Gatta, A.T., L.H. Wong, Y.Y. Sere, D.M. Calderón-Noreña, S. Cockcroft, A.K. Menon, and T.P. Levine. 2015. A new family of StART domain proteins at membrane contact sites has a role in ER-PM sterol transport. *Elife*. 4: e07253. https://doi.org/10.7554/eLife.07253

Ghanbarpour, A., D.P. Valverde, T.J. Melia, and K.M. Reinisch. 2021. A model for a partnership of lipid transfer proteins and scramblases in membrane expansion and organelle biogenesis. *Proc. Natl. Acad. Sci. USA*. 118: e2101562118. https://doi.org/10.1073/pnas.2101562118

Guillén-Samander, A., Y. Wu, S.S. Pineda, F.J. García, J.N. Eisen, M. Leonzino, B. Ugur, M. Kellis, M. Heiman, and P. De Camilli. 2022. A partnership between the lipid scramblase XK and the lipid transfer protein VPS13A at the plasma membrane. *Proc. Natl. Acad. Sci. USA*. 119:e2205425119. https://doi.org/10.1073/pnas.2205425119

Hamiaux, C., D. Stanley, D.R. Greenwood, E.N. Baker, and R.D. Newcomb. 2009. Crystal structure of Epiphyas postvittana takeout 1 with bound ubiquinone supports a role as ligand carriers for takeout proteins in insects. *J. Biol. Chem.* 284:3496–3503. https://doi.org/10.1074/jbc.M807467200

Hanna, M., A. Guillén-Samander, and P. De Camilli. 2023. RBG motif bridge-like lipid transport proteins: Structure, functions, and open questions. *Annu. Rev. Cell Dev. Biol.* 39:409–434. https://doi.org/10.1146/annurev-cellbio-120420-014634

Hong, Z., J. Adlakha, N. Wan, E. Guinn, F. Giska, K. Gupta, T.J. Melia, and K.M. Reinisch. 2022. Mitoguardin-2-mediated lipid transfer preserves mitochondrial morphology and lipid droplet formation. *J. Cell Biol.* 221: e202207022. https://doi.org/10.1083/jcb.202207022

Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: Visual molecular dynamics. *J. Mol. Graph.* 14:33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Im, Y.J., S. Raychaudhuri, W.A. Prinz, and J.H. Hurley. 2005. Structural mechanism for sterol sensing and transport by OSBP-related proteins. *Nature*. 437:154–158. https://doi.org/10.1038/nature03923

Jakobsson, E., G. Alvite, T. Bergfors, A. Esteves, and G.J. Kleywegt. 2003. The crystal structure of Echinococcus granulosus fatty-acid-binding protein 1. *Biochim. Biophys. Acta*. 1649:40–50. https://doi.org/10.1016/S1570-9639(03)00151-1

Janson, G., G. Valdes-Garcia, L. Heo, and M. Feig. 2023. Direct generation of protein conformational ensembles via machine learning. *Nat. Commun.* 14:774. https://doi.org/10.1038/s41467-023-36443-x

Jo, S., T. Kim, V.G. Iyer, and W. Im. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865. https://doi.org/10.1002/jcc.20945

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596: 583–589. https://doi.org/10.1038/s41586-021-03819-2

Khelashvili, G., N. Chauhan, K. Pandey, D. Eliezer, and A.K. Menon. 2019. Exchange of water for sterol underlies sterol egress from a StARkin domain. *Elife*. 8:e53444. https://doi.org/10.7554/eLife.53444

Kim, H., S. Lee, Y. Jun, and C. Lee. 2022. Structural basis for mitoguardin-2 mediated lipid transport at ER-mitochondrial membrane contact sites. *Nat. Commun.* 13:3702. https://doi.org/10.1038/s41467-022-31462-6

Krishna, R., J. Wang, W. Ahern, P. Sturmfels, P. Venkatesh, I. Kalvet, G.R. Lee, F.S. Morey-Burrows, I. Anishchenko, I.R. Humphreys, et al. 2023. Generalized biomolecular modeling and design with RoseTTAFold all-atom. *bioRxiv*. https://doi.org/10.1101/2023.10.09.561603 (Preprint posted October 09, 2023).

Kudo, N., K. Kumagai, N. Tomishige, T. Yamaji, S. Wakatsuki, M. Nishijima, K. Hanada, and R. Kato. 2008. Structural basis for specific lipid recognition by CERT responsible for nonvesicular trafficking of ceramide. *Proc. Natl. Acad. Sci. USA*. 105:488–493. https://doi.org/10.1073/pnas.0709191105

Kumar, N., M. Leonzino, W. Hancock-Cerutti, F.A. Horenkamp, P. Li, J.A. Lees, H. Wheeler, K.M. Reinisch, and P. De Camilli. 2018. VPS13A and VPS13C are lipid transport proteins differentially localized at ER contact sites. *J. Cell Biol.* 217:3625–3639. https://doi.org/10.1083/jcb.201807019

Kuzmanic, A., G.R. Bowman, J. Juarez-Jimenez, J. Michel, and F.L. Gervasio. 2020. Investigating cryptic binding sites by molecular dynamics simulations. *Acc. Chem. Res.* 53:654–661. https://doi.org/10.1021/acs.accounts.9b00613

Levine, T.P. 2022a. Sequence analysis and structural predictions of lipid transfer bridges in the repeating beta groove (RBG) superfamily reveal past and present domain variations affecting form, function and interactions of VPS13, ATG2, SHIP164, hobbit and Tweek. *Contact*. 5: 251525642211343. https://doi.org/10.1177/25152564221134328

Levine, T.P. 2022b. TMEM106B in humans and Vac7 and Tag1 in yeast are predicted to be lipid transfer proteins. *Proteins*. 90:164–175. https://doi.org/10.1002/prot.26201

Li, P., J.A. Lees, C.P. Lusk, and K.M. Reinisch. 2020. Cryo-EM reconstruction of a VPS13 fragment reveals a long groove to channel lipids between membranes. *J. Cell Biol.* 219:e202001161. https://doi.org/10.1083/jcb.202001161

Li, X., P. Saha, J. Li, G. Blobel, and S.R. Pfeffer. 2016. Clues to the mechanism of cholesterol transfer from the structure of NPC1 middle lumenal domain bound to NPC2. *Proc. Natl. Acad. Sci. USA*. 113:10079–10084. https://doi.org/10.1073/pnas.1611956113

Limar, S., C. Körner, F. Martínez-Montañés, V.G. Stancheva, V.N. Wolf, S. Walter, E.A. Miller, C.S. Ejsing, V.V. Galassi, and F. Fröhlich. 2023. Yeast Svf1 binds ceramides and contributes to sphingolipid metabolism at the ER cis-Golgi interface. *J. Cell Biol.* 222:e202109162. https://doi.org/10.1083/jcb.202109162

Limongelli, V., M. Bonomi, and M. Parrinello. 2013. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. USA*. 110: 6358–6363. https://doi.org/10.1073/pnas.1303186110

Lipp, N.F., R. Gautier, M. Magdeleine, M. Renard, V. Albanèse, A. Čopič, and G. Drin. 2019. An electrostatic switching mechanism to control the lipid transfer activity of Osh6p. *Nat. Commun.* 10:3926. https://doi.org/10.1038/s41467-019-11780-y

Maeda, K., K. Anand, A. Chiapparino, A. Kumar, M. Poletto, M. Kaksonen, and A.C. Gavin. 2013. Interactome map uncovers phosphatidylserine transport by oxysterol-binding proteins. *Nature*. 501:257–261. https://doi.org/10.1038/nature12430

Marrink, S.J., H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. de Vries. 2007. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B*. 111:7812–7824. https://doi.org/10.1021/jp071097f

Moser von Filseck, J., A. Čopič, V. Delfosse, S. Vanni, C.L. Jackson, W. Bourguet, and G. Drin. 2015. INTRACELLULAR TRANSPORT. Phosphatidylserine transport by ORP/Osh proteins is driven by phosphatidylinositol 4-phosphate. *Science*. 349:432–436. https://doi.org/10.1126/science.aab1346

Neuman, S.D., T.P. Levine, and A. Bashirullah. 2022. A novel superfamily of bridge-like lipid transfer proteins. *Trends Cell Biol.* 32:962–974. https://doi.org/10.1016/j.tcb.2022.03.011

Niitsu, A., S. Re, H. Oshima, M. Kamiya, and Y. Sugita. 2019. De novo prediction of binders and nonbinders for T4 Lysozyme by gREST simulations. *J. Chem. Inf. Model.* 59:3879–3888. https://doi.org/10.1021/acs.jcim.9b00416

Osawa, T., T. Kotani, T. Kawaoka, E. Hirata, K. Suzuki, H. Nakatogawa, Y. Ohsumi, and N.N. Noda. 2019. Atg2 mediates direct lipid transfer between membranes for autophagosome formation. *Nat. Struct. Mol. Biol.* 26:281–288. https://doi.org/10.1038/s41594-019-0203-4

Parrinello, M., and A. Rahman. 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52:7182–7190. https://doi.org/10.1063/1.328693

Paul, B., S. Weeratunga, V.A. Tillu, H. Hariri, W.M. Henne, and B.M. Collins. 2022. Structural predictions of the SNX-RGS proteins suggest they belong to a new class of lipid transfer proteins. *Front. Cell Dev. Biol.* 10: 826688. https://doi.org/10.3389/fcell.2022.826688

Radulovic, M., E.M. Wenzel, S. Gilani, L.K.K. Holland, A.H. Lystad, S. Phuyal, V.M. Olkkonen, A. Brech, M. Jäättelä, K. Maeda, et al. 2022. Lysosome repair by ER-mediated cholesterol transfer. *bioRxiv*. https://doi.org/10.1101/2022.09.26.509457 (Preprint posted September 28, 2022).

Reinisch, K.M., and W.A. Prinz. 2021. Mechanisms of nonvesicular lipid transport. *J. Cell Biol.* 220:e202012058. https://doi.org/10.1083/jcb.202012058

Roderick, S.L., W.W. Chan, D.S. Agate, L.R. Olsen, M.W. Vetting, K.R. Rajashankar, and D.E. Cohen. 2002. Structure of human phosphatidylcholine transfer protein in complex with its ligand. *Nat. Struct. Biol.* 9: 507–511. https://doi.org/10.1038/nsb812

Rogers, J.R., and P.L. Geissler. 2023. Ceramide-1-phosphate transfer protein enhances lipid transport by disrupting hydrophobic lipid-membrane contacts. *PLoS Comput. Biol.* 19:e1010992. https://doi.org/10.1371/journal.pcbi.1010992

Saheki, Y., and P. De Camilli. 2017. The extended-synaptotagmins. *Biochim. Biophys. Acta Mol. Cell Res.* 1864:1490–1493. https://doi.org/10.1016/j.bbamcr.2017.03.013

Schaaf, G., E.A. Ortlund, K.R. Tyeryar, C.J. Mousley, K.E. Ile, T.A. Garrett, J. Ren, M.J. Woolls, C.R. Raetz, M.R. Redinbo, and V.A. Bankaitis. 2008. Functional anatomy of phospholipid binding and regulation of phosphoinositide homeostasis by proteins of the sec14 superfamily. *Mol. Cell*. 29:191–206. https://doi.org/10.1016/j.molcel.2007.11.026

Schauder, C.M., X. Wu, Y. Saheki, P. Narayanaswamy, F. Torta, M.R. Wenk, P. De Camilli, and K.M. Reinisch. 2014. Structure of a lipid-bound extended synaptotagmin indicates a role in lipid transfer. *Nature*. 510: 552–555. https://doi.org/10.1038/nature13269

Schrodinger, L. 2015. The PyMOL Molecular Graphics System [Online]. Available at: https://pymol.org/2/.

Sevvana, M., J. Ahnström, C. Egerer-Sieber, H.A. Lange, B. Dahlbäck, and Y.A. Muller. 2009. Serendipitous fatty acid binding reveals the structural determinants for ligand recognition in apolipoprotein M. *J. Mol. Biol.* 393:920–936. https://doi.org/10.1016/j.jmb.2009.08.071

Shan, Y., V.P. Mysore, A.E. Leffler, E.T. Kim, S. Sagawa, and D.E. Shaw. 2022. How does a small molecule bind at a cryptic binding site? *PLoS Comput. Biol.* 18:e1009817. https://doi.org/10.1371/journal.pcbi.1009817

Simanshu, D.K., R.K. Kamlekar, D.S. Wijesinghe, X. Zou, X. Zhai, S.K. Mishra, J.G. Molotkovsky, L. Malinina, E.H. Hinchcliffe, C.E. Chalfant, et al. 2013. Non-vesicular trafficking by a ceramide-1-phosphate transfer protein regulates eicosanoids. *Nature*. 500:463–467. https://doi.org/10.1038/nature12332

Souza, P.C.T., R. Alessandri, J. Barnoud, S. Thallmair, I. Faustino, F. Grünewald, I. Patmanidis, H. Abdizadeh, B.M.H. Bruininks, T.A. Wassenaar, et al. 2021. Martini 3: A general purpose force field for coarse-grained molecular dynamics. *Nat. Methods*. 18:382–388. https://doi.org/10.1038/s41592-021-01098-3

Souza, P.C.T., S. Thallmair, P. Conflitti, C. Ramírez-Palacios, R. Alessandri, S. Raniolo, V. Limongelli, and S.J. Marrink. 2020. Protein-ligand binding with the coarse-grained Martini model. *Nat. Commun.* 11:3714. https://doi.org/10.1038/s41467-020-17437-5

Srinivasan, S., A.D. Luca, A.T.J. Peter, C. Gehin, M.A. Lone, T. Hornemann, G. D'Angelo, and S. Vanni. 2023. Conformational dynamics of lipid transfer domains provide a general framework to decode their functional mechanism. *bioRxiv*. https://doi.org/10.1101/2023.04.11.536463 (Preprint posted April 13, 2023).

Suzuki, R., Z. Fujimoto, T. Shiotsuki, W. Tsuchiya, M. Momma, A. Tase, M. Miyazawa, and T. Yamazaki. 2011. Structural mechanism of JH delivery in hemolymph by JHBP of silkworm, Bombyx mori. *Sci. Rep.* 1:133. https://doi.org/10.1038/srep00133

Tan, J.X., and T. Finkel. 2022. A phosphoinositide signalling pathway mediates rapid lysosomal repair. *Nature*. 609:815–821. https://doi.org/10.1038/s41586-022-05164-4

Valverde, D.P., S. Yu, V. Boggavarapu, N. Kumar, J.A. Lees, T. Walz, K.M. Reinisch, and T.J. Melia. 2019. ATG2 transports lipids to promote autophagosome biogenesis. *J. Cell Biol.* 218:1787–1798. https://doi.org/10.1083/jcb.201811139

Van Der Spoel, D., E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen. 2005. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718. https://doi.org/10.1002/jcc.20291

van Meer, G., D.R. Voelker, and G.W. Feigenson. 2008. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* 9:112–124. https://doi.org/10.1038/nrm2330

van Vliet, A.R., G.N. Chiduza, S.L. Maslen, V.E. Pye, D. Joshi, S. De Tito, H.B.J. Jefferies, E. Christodoulou, C. Roustan, E. Punch, et al. 2022. ATG9A and ATG2A form a heteromeric complex essential for autophagosome formation. *Mol. Cell*. 82:4324–4339.e8. https://doi.org/10.1016/j.molcel.2022.10.017

Wang, Y., S. Dahmane, R. Ti, X. Mai, L. Zhu, L.-A. Carlson, and G. Stjepanovic. 2023. Structural basis for lipid transfer by the ATG2A-ATG9A complex. *bioRxiv*. https://doi.org/10.1101/2023.07.08.548186 (Preprint posted July 08, 2023).

Welti, S., S. Fraterman, I. D'Angelo, M. Wilm, and K. Scheffzek. 2007. The sec14 homology module of neurofibromin binds cellular glycerophospholipids: Mass spectrometry and structure of a lipid complex. *J. Mol. Biol.* 366:551–562. https://doi.org/10.1016/j.jmb.2006.11.055

Wong, L.H., A. Čopič, and T.P. Levine. 2017. Advances on the transfer of lipids by lipid transfer proteins. *Trends Biochem. Sci.* 42:516–530. https://doi.org/10.1016/j.tibs.2017.05.001

Wong, L.H., A.T. Gatta, and T.P. Levine. 2019. Lipid transfer proteins: The lipid commute via shuttles, bridges and tubes. *Nat. Rev. Mol. Cell Biol.* 20:85–101. https://doi.org/10.1038/s41580-018-0071-5

Wong, L.H., and T.P. Levine. 2016. Lipid transfer proteins do their thing anchored at membrane contact sites… but what is their thing? *Biochem. Soc. Trans.* 44:517–527. https://doi.org/10.1042/BST20150275

Wozny, M.R., A. Di Luca, D.R. Morado, A. Picco, R. Khaddaj, P. Campomanes, L. Ivanović, P.C. Hoffmann, E.A. Miller, S. Vanni, and W. Kukulski. 2023. In situ architecture of the ER-mitochondria encounter structure. *Nature*. 618:188–192. https://doi.org/10.1038/s41586-023-06050-3

Yoder, M.D., L.M. Thomas, J.M. Tremblay, R.L. Oliver, L.R. Yarbrough, and G.M. Helmkamp Jr. 2001. Structure of a multifunctional protein. Mammalian phosphatidylinositol transfer protein complexed with phosphatidylcholine. *J. Biol. Chem.* 276:9246–9252. https://doi.org/10.1074/jbc.M010131200

Zhang, H., K. Pluhackova, Z. Jiang, and R.A. Böckmann. 2016. Binding characteristics of sphingosine-1-phosphate to ApoM hints to assisted release mechanism via the ApoM calyx-opening. *Sci. Rep.* 6:30655. https://doi.org/10.1038/srep30655

Zhang, Y., J. Ge, X. Bian, and A. Kumar. 2022. Quantitative models of lipid transfer and membrane contact formation. *Contact*. 5:1–21. https://doi.org/10.1177/25152564221096024

Supplemental material

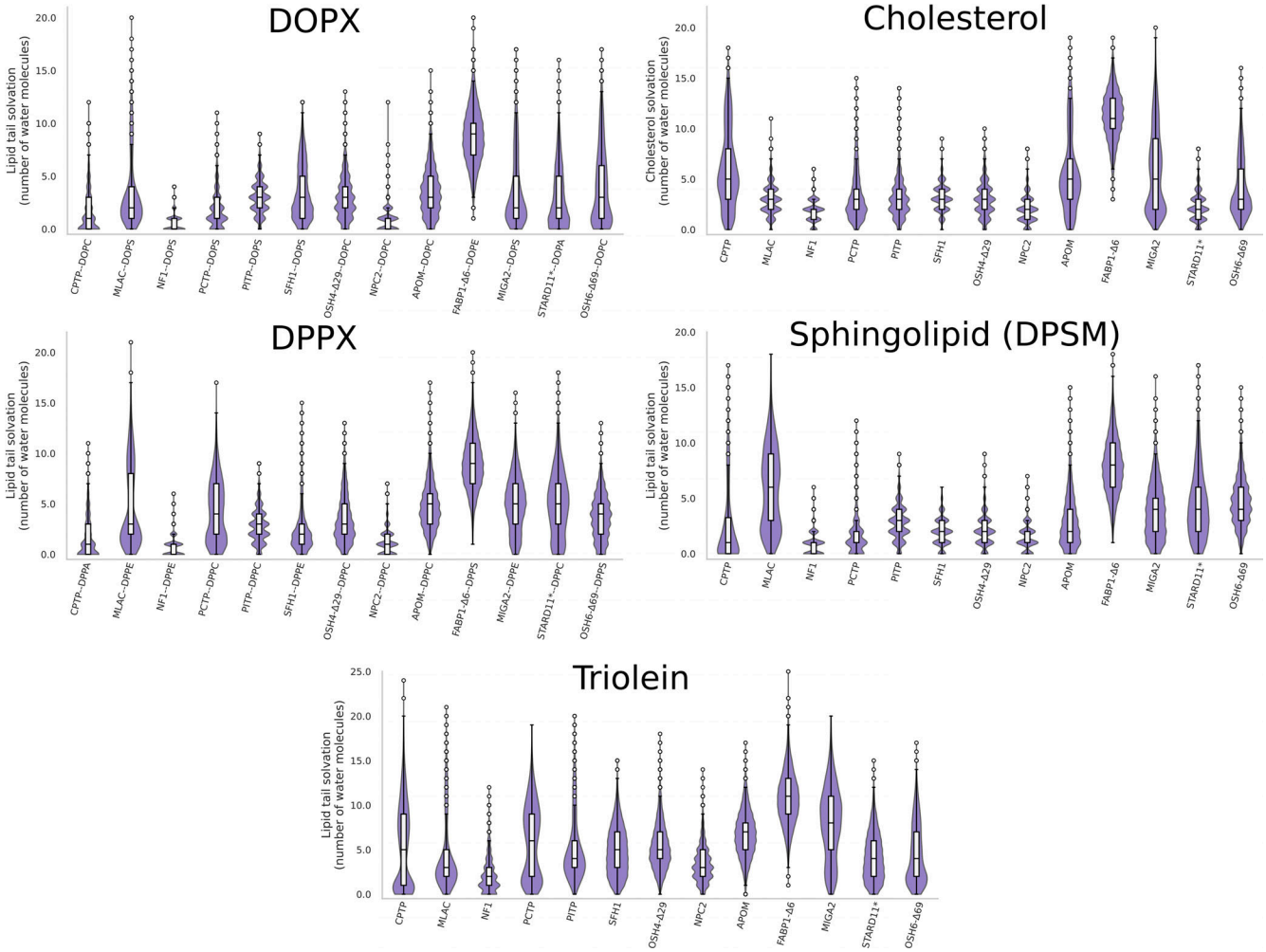

Figure S1. **Lipid tail solvation in the last 100 ns of all the CG-MD trajectories of LTPs with alternative lipids.** Monounsaturated phospholipids with different headgroups (18:1 PX lipids, DOPX, specific polar groups in the figure x-axis), phospholipids with saturated chains (16:0 PX lipids, DPPX), cholesterol, a sphingolipid (d18:1/18:0 sphingomyelin, DPSM), and triolein (18:1 triglyceride).

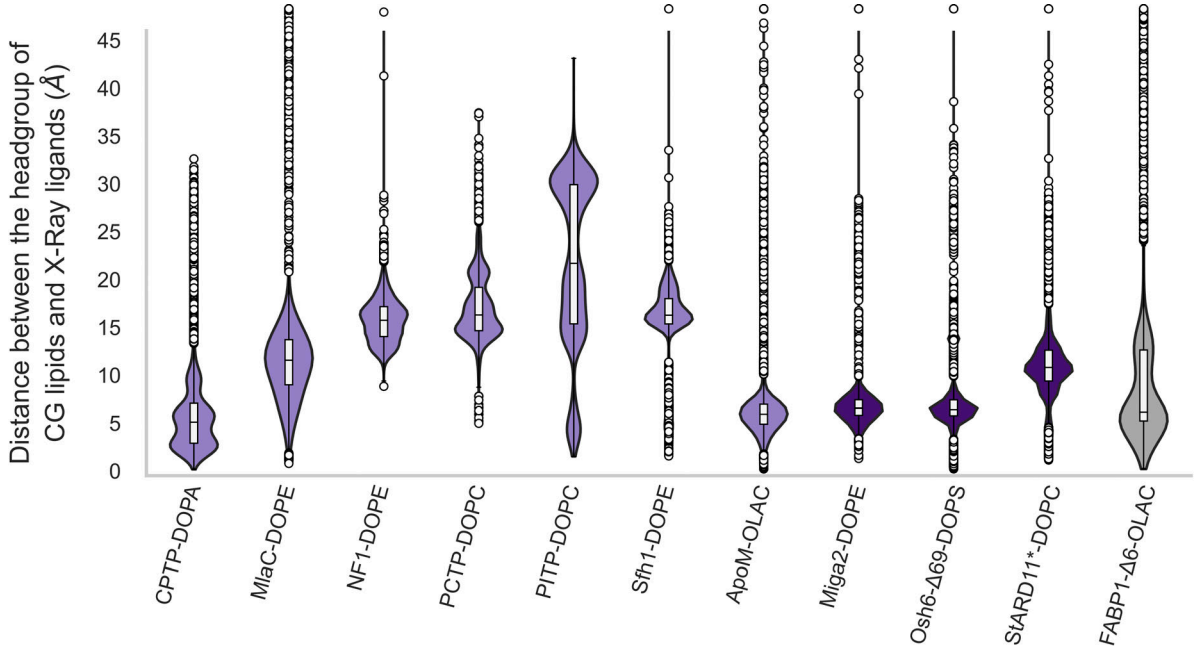

Figure S2. **Distance between the center of mass of the headgroup of the experimental ligands and the PO4 bead of the lipids tested in the simulations.** Only the frames with a lipid tail solvation of <2 were considered.

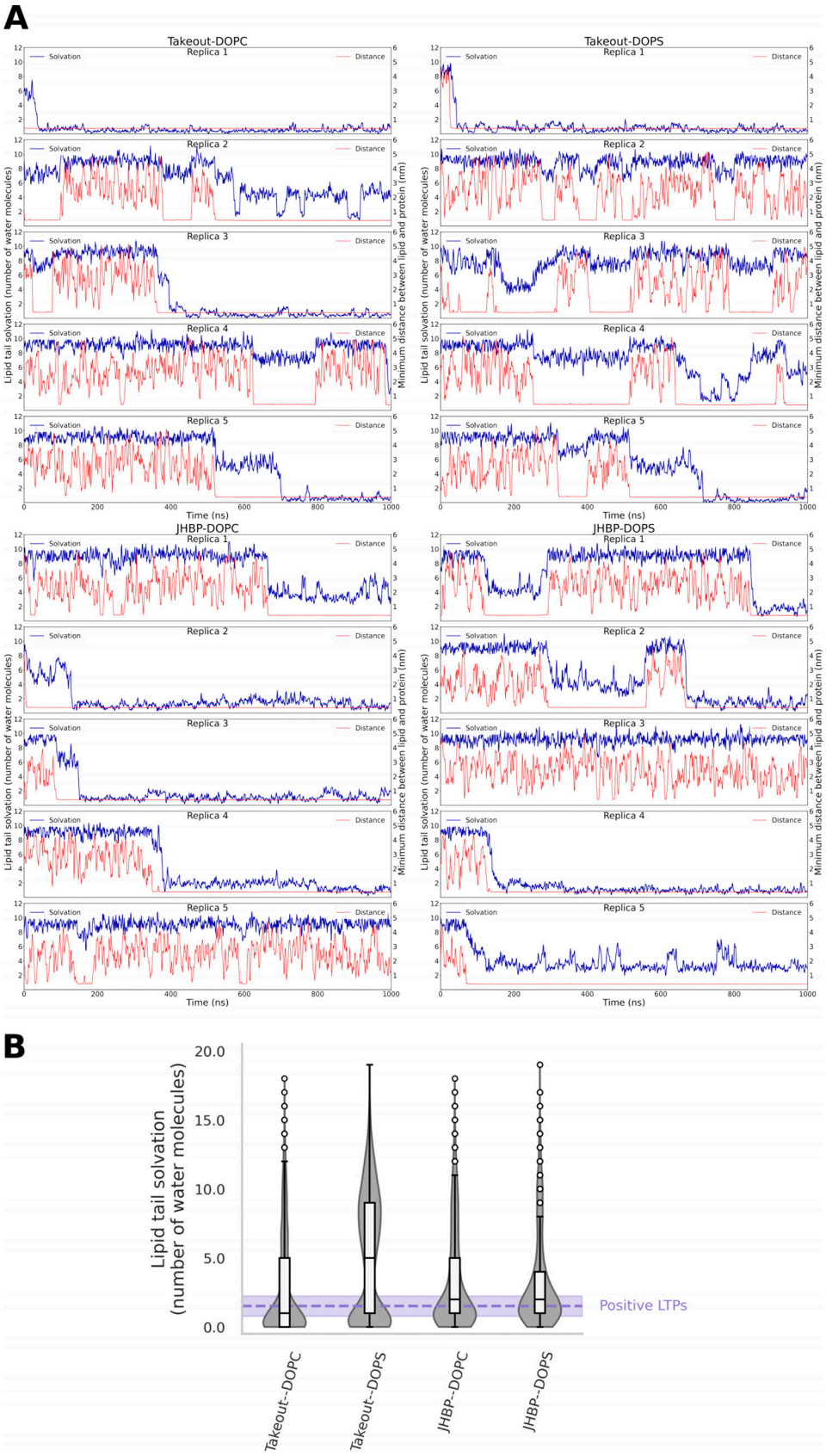

Figure S3. **Lipid binding to Takeout and JHBP proteins. (A)** Time trace of minimum distance values between the lipids and the Takeout and JHBP proteins (red) along with the lipid tail solvation (blue). Both plots have been block-averaged every 10 frames. **(B)** Lipid tail solvation in the last 100 ns of the trajectories, mean value for positive LTP is shown as a dotted line.

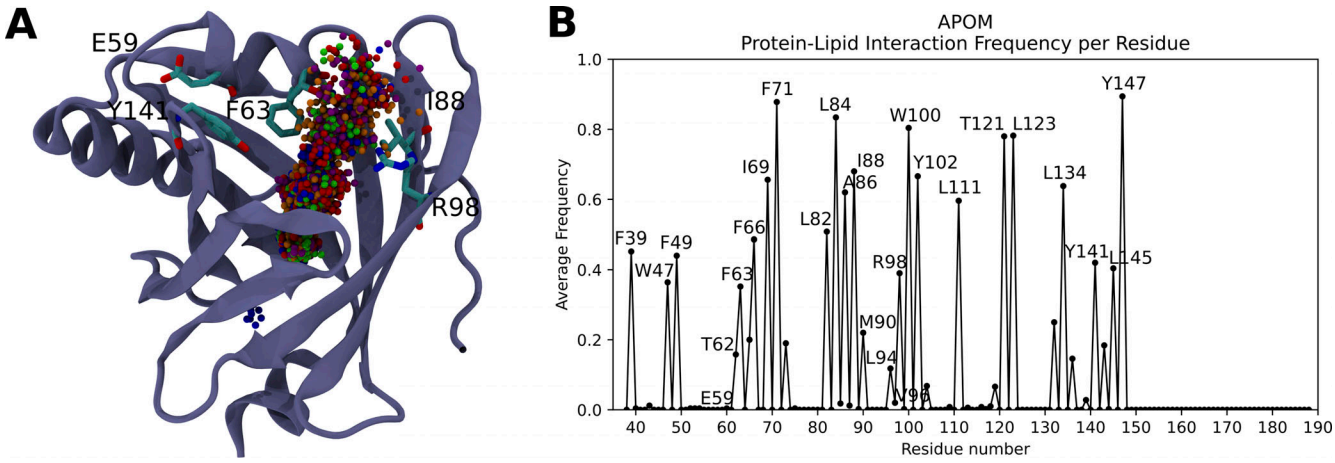

Figure S4. **Analysis of lipid entry pathway of ApoM. (A)** Mapping as spheres of the center of mass of the CG lipid long the binding process; each color represents a different replica. Some residues have been highlighted as in Fig. 6 of reference Zhang et al. (2016). **(B)** Averaged contact frequency between the protein residue and the simulated lipid during the last 100 ns of simulations.

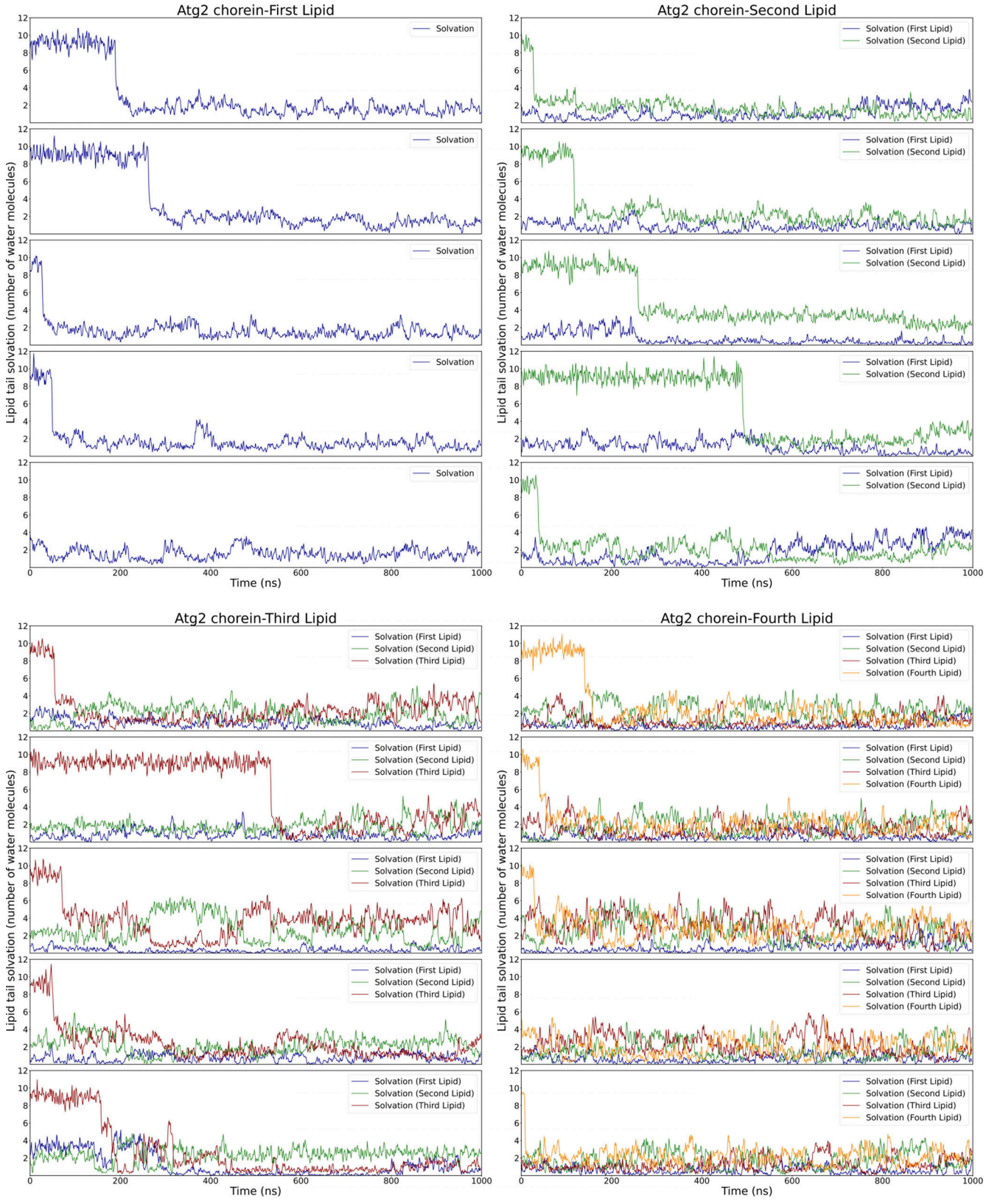

Figure S5. **Time trace of lipid tail solvation for the iterative addition of four DOPE lipids to Atg2 (chorein motif).** Each individual plot displays the results of a different replica. All the plots have been smoothened using block averages every 10 frames.

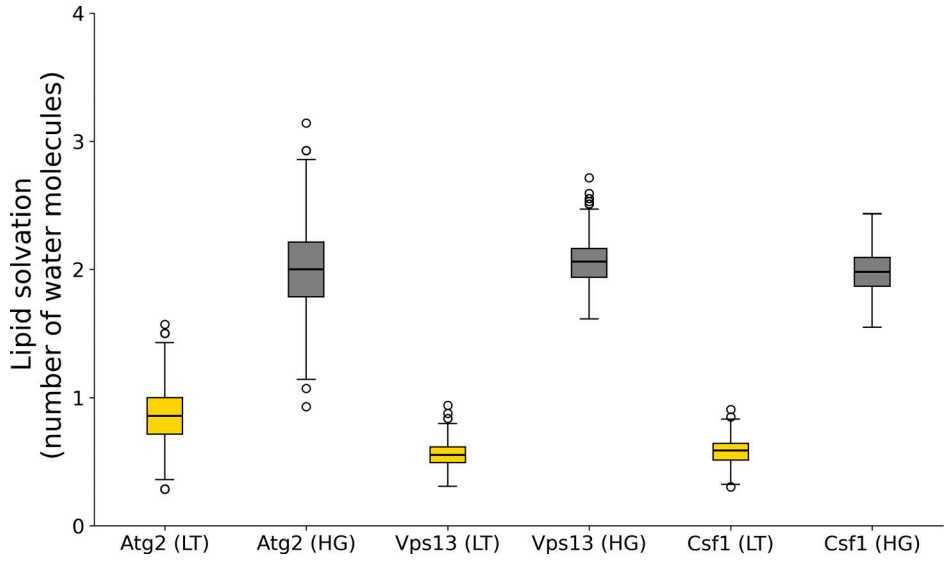

Figure S6. **Comparison of the lipid solvation number of the lipid tails (LT) and the headgroups (HG) for the BLTPs considered in this study (Atg2, Vps13, Csf1).**

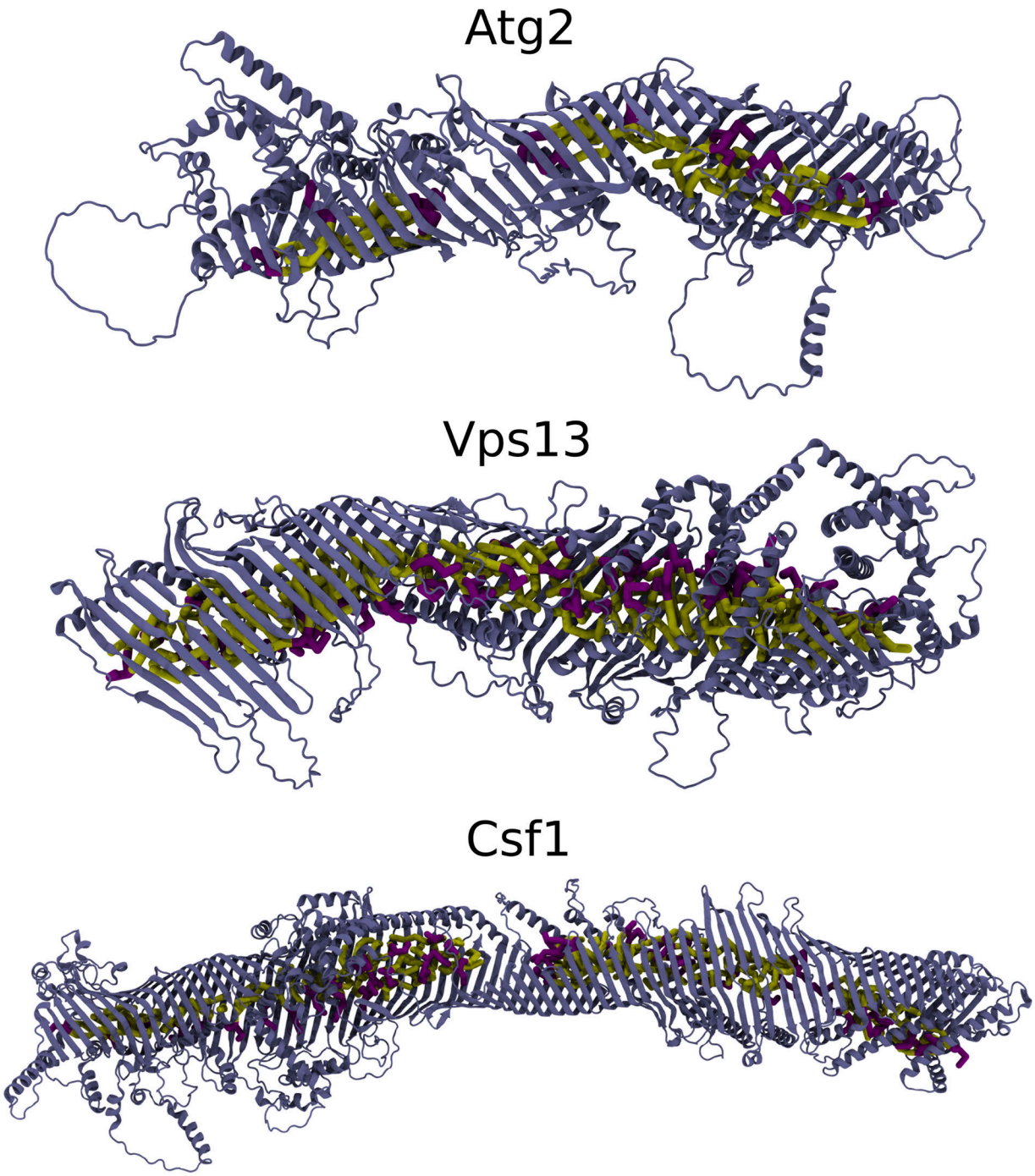

Figure S7. **Explicit representation of the DOPC lipids inside the long hydrophobic cavities of RBG BLTPs.** Headgroups (purple) face the solvent, while lipid tails (yellow) face the hydrophobic interior.

Video 1. **The application of our protocol, as in** Fig. 1 A**.** Video plays at 30 frames per second.

Video 2. **The iterative addition of 15 lipids to one of the BLTPs (Atg2).** Headgroups are represented in purple and lipid tails are colored in yellow. Video plays at 30 frames per second.

Provided online is Data S1, which shows a time trace of minimum distance values between LTPs and lipids (red), lipid tail solvation (blue), and distance between the center of mass of the CG lipid and the bound lipid in the x-ray structure (gray); a time trace of minimum distance values between negative control proteins and DOPC (red), and lipid tail solvation (blue); the averaged contact frequency between the LTP residues and the simulated lipids during the last 100 ns of the simulations; a time trace of the contacts for the relevant residues of ApoM, NF1, Osh4 Δ29, and PITP with the simulated lipids (same lipids as in Fig. 1 B); and a time trace of minimum distance values between potential LTPs and DOPC (red), and lipid tail solvation (blue).